Towards the acceptance of a Prompt Engineering-based Approach to Build Proto-personas during Product Discovery

Fernando Ribeiro Ayach[1], **Awdren Fontão (supervisor)**[1]

[1] Faculty of Computing – Federal University of Mato Grosso do Sul
PO Box 79.002 – 79.070-900 – Campo Grande – MS – Brazil

{fernando.ayach, awdren.fontao}@ufms.br

Abstract. *Product discovery approaches such as Lean Inception (LI) typically span five days (40 working hours). During LI, the participants create and refine proto-personas during four working hours to understand user needs. Proto-personas are preliminary representations of ideal users that guide initial design discussions. The accuracy of the generated proto-personas has been counter-intuitive due to the limited time. There is a gap in exploring the use of prompt engineering and proto-persona strategies to support the Product Discovery approaches. We report an exploratory case study where six participants used a prompt engineering-based approach to generate proto-personas in Product Discovery (LI). Participants accepted our approach well. Our approach used an average of 11 minutes of working hours (SD ≈ 2.24 minutes), traditionally this time in LI is four hours.*

1. Introduction

Product Discovery approaches stand as a pivotal stage in Requirements Engineering (RE), aiming to decipher user needs and delineate clear initial directions for the project [Trieflinger et al. 2023]. Lean Inception (LI) stands out among existing approaches, spanning five days (40 working hours). Within LI, participants create and refine proto-personas, allocating a focused four-hour window to grasp user requirements. Proto-personas can be understood as an initial representation of a target user used to kickstart product development before more detailed research [Gothelf 2013]. The accuracy of these proto-personas, crafted within the constraints of time and resource limitations, may be counterintuitive.

Ensuring high accuracy and relevance in proto-personas is crucial for effective product discovery. This necessity for rapid and effective Product Discoveries is also expressed by Jeff Gothelf - co-author of Lean UX and Sense & Respond - "*One of the most common questions that comes up after teams complete a Lean UX or product discovery training engagement with us, "How will we find the time to do this and deliver work?" Their rationale reflects the velocity-driven demands put on them by their stakeholders.*". LLMs (Large Language Models) have become tools that could support software engineering activities, especially in RE [Marques et al. 2024]. An important process for using LLMs like ChatGPT is formulating appropriate prompts. A prompt can be understood as a set of instructions provided to an LLM, programming it by customizing its capabilities [Liu et al. 2023]. It influences the output generated by an LLM by providing specific rules and guidelines. Prompt Engineering is how LLMs are systematically programmed via prompts to optimize their results.

Regarding the use of personas in RE, [Karolita et al. 2023] did not identify any studies involving the use of prompt engineering for the construction of proto-personas. The authors recommend new approaches to creating personas. [Marques et al. 2024], when exploring the use of ChatGPT in RE, did not identify studies in the context of proto-personas. The authors emphasize the importance of exploring techniques across various software development stages to enhance quality in prompt engineering approaches. We perform an exploratory case study involving six participants to advance towards a prompt engineering-based approach to build proto-personas in a Product Discovery approach (i.e., LI) based on limited input, facilitating initial ideation and validation.

2. Related Work

With the rapid advancement of Large Language Models (LLMs) like ChatGPT, leveraging LLMs for persona development has become increasingly significant across various domains such as software engineering [Arora et al. 2023], UI design [Atlas 2023] and personas based on LLMs [Zhang et al. 2024]. There is a lack of exploring the synergy between prompt engineering and proto-persona strategies. In the existing secondary studies on personas in Requirements Engineering [Karolita et al. 2023], Agile methodologies [Losana et al. 2021], use cases for design personas [Salminen et al. 2022], LLMs for Software Engineering [Fan et al. 2023] and ChatGPT in Requirements Engineering [Marques et al. 2024], no studies are addressing prompt engineering for the generation of proto-personas in the context of Product Discovery. Considering this, it is worth investigating due to the impacts on projects generated by failures in building proto-personas.

3. The approach

Our approach was modeled based on the Prompt Patterns cataloged by [White et al. 2023] and OpenAI[1]. The approach was refined through pilot studies by the authors using the ChatGPT 3.5. Four versions of the model were created. For each version, we performed a business process modeling and a formal description[2], containing details of each activity: name, responsible, prompt pattern used, goal, prompt to be inserted, and expected ChatGPT response. The current version, Figure 1, comprises two events, five activities, and three artifacts. The process begins with accessing the LLM (ChatGPT) and ends with building a set of proto-personas.

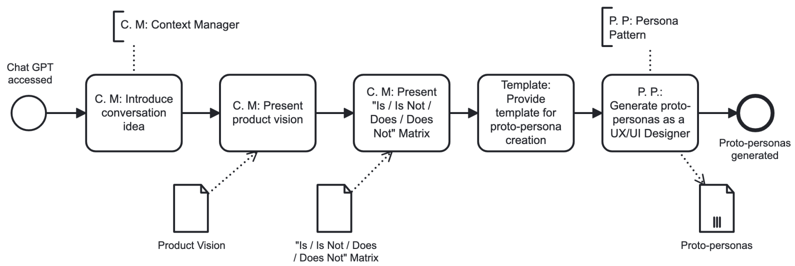

Figure 1. Our proposed approach

Activity 1: *Introduce conversation idea.* The first prompt aims to establish an information context for the LLM and limit its response for proto-persona generation in

[1]https://platform.openai.com/docs/guides/prompt-engineering.
[2]https://tinyurl.com/bdfc25dk.

a Lean Inception-based product discovery. ***The prompt pattern used was the Context Manager, which aims to restrict the AI analysis to a specific context.*** This pattern is crucial in this initial stage because ChatGPT is a generic text generator, and we need to guide the model's line of reasoning to work with information from our context, avoiding hallucinations and off-topic responses [Belzner et al. 2023]. This step uses the product discovery context, which will guide subsequent activities.

Activity 2: *Present product vision.* In this, we input the product vision as a prompt for the LLM, aiming to restrict the LLM's response to the context of the software product that includes the product vision and adds more application context. ***Once again, we utilize the Context Manager as the prompt pattern***, given its similarity to the previous topic. This step uses the product vision input for subsequent activities.

Activity 3: *Present "Is/Is Not/Does/Does Not" Matrix.* This is the final contextual activity of the model and aims to input the "Is/Is Not/Does/Does Not" LI matrix as a prompt for the LLM, restricting the LLM's response to the context of the software product. Similar to the previous activities, ***it also uses the Context Manager*** and is equally necessary for proto-persona generation in the last stage of the process, as notions of the product vision and its functionalities are essential for proto-persona conception. This step uses input on what the product does and does not do for subsequent activities.

Activity 4: *Provide the desired template for proto-persona creation.* After contextualization, we supply the LLM with the template of the proto-persona we want to generate, in this case, that of Lean Inception (profile, needs, and behaviors), to ensure that the generated response aligns with the LI's specification. ***We use the Template Pattern as the prompt pattern*** because it fulfills the function we desire for this activity: formatting the generated output into a specific format. This step prepares the proto-persona template to be used in the last activity of the process.

Activity 5: *Ask the LLM to act as a UI/UX Designer with experience in proto-persona creation and request that it generate the proto-personas for the software.* Finally, here we generate the artifacts, making the LLM act as an experienced professional in building software proto-personas. ***The prompt pattern used is the Persona Pattern***, aiming to incorporate a role into the LLM, causing it to assume (or at least attempt to assume) the line of thinking of the specified role. It can improve the focus of the model's line of reasoning. This last step yields the proto-personas of the specified product, based on the inputs of product vision, "Is/Is Not/Does/Does Not" LI matrix, and proto-persona template.

Reflections on the approach conception: The approach underwent four versions, with some activities, like the validation step, removed. We tried using minimal context for validation, based on a method that generated questions and improvement points via Reflexion [Zhang et al. 2024]. The outcome showed minimal deviation, suggesting that without additional inputs, a validation step is unnecessary. Also, prompt sensitivity affected results; for instance, using "Let's think step by step" led to only one proto-persona being generated, requiring repeated prompts.

4. Case Study Design

We used the guidelines of [Runeson and Höst 2009] to conduct and report an exploratory case study, which is applied in the search for new insights, generating ideas and hypotheses for research when there is little information about the phenomenon being studied.

4.1. Case, Unit of analysis and Rationale

The case analyzed in this study refers to using a prompt engineering-based approach to build proto-personas during the initial phases of a product discovery approach. ***It is characterized as a single case*** because we are exploring a phenomenon in a specific context. ***As the unit of analysis***, we used a mobile application project that supports the Solidarity Economy development. It involved a team of 12 software engineers and three specialists in the fields of administration and economics. As part of the product discovery phase, the Lean Inception [Caroli 2017] was used, generating proto-personas in this process. After 14 months of development, a usability test was conducted with 11 users, revealing a previously uncataloged persona during the Lean Inception and subsequent project progression and persona refinement. This discovery necessitated a comprehensive redesign of the application, including a UX workshop and modifications to workflows and interfaces. Such changes are exceedingly costly for any project and could have been mitigated with a precise initial definition of the system's personas. The presence of a single UX researcher in the unit of analysis led to task overload due to the high demand for revisions caused by the lack of persona mapping in the usability test. This highlights a broader issue of insufficient resources for persona formulation, likely affecting other projects as well.

Rationale: Given the case described above, the use of external tools such as LLMs (e.g. ChatGPT) can be useful to at least minimize such problems [Arora et al. 2023]. However, the casual use of tools like ChatGPT is not sufficient for the development of more complex tasks, such as persona generation [Marques et al. 2024]. Prompt engineering serves as a means to maximize the effectiveness of the approach's response [Liu et al. 2023, White et al. 2023], and our approach focuses precisely on this.

4.2. Research Question and Study's Goal

Our goal is to investigate the acceptance of a prompt engineering-based approach to build proto-personas during Product Discovery (i.e., Lean Inception). To support it we developed the following research question and metrics: ***RQ: What is the acceptance of "our approach" of proto-persona generation?*** **Metric 1:** Perceived Usefulness (PU). **Metric 2:** Perceived Ease of Use (PEoU). **Metric 3:** Attitude Toward Use (ATU). **Metric 4:** Behavioral Intention to Use (BItU).

4.3. Data Collection & Execution

At the end of each study session, we collected data regarding the set of generated proto-personas and the evaluation of the process. This evaluation was carried out by the participant during and after the study. During the study session, the researchers made notes related to the application of the think-aloud method. **Artifacts**: *Before the study*, we transposed the artifacts from the unit of analysis LI - "Product Vision" and "Is/Is Not/Does/Does Not" matrix - into text format for use in ChatGPT. They are essential inputs for the process execution. Natalia Arsand developed the persona description template used [Caroli 2017], which serves as a reference during the "Describe the persona" stage of Lean Inception. We chose to have the artifacts ready for use by the participants, as our evaluation did not aim to verify the collection of inputs before the process execution. Along the paper, we share the artifacts link. **The process evaluation questions** were based on the Technology Acceptance Model (TAM). For this study, we chose the most relevant variables for evaluating our approach using a five-point Likert Scale: Perceived Usefulness (PU), Perceived Ease of Use (PEoU), Attitude Toward Use (ATU), and Behavioral Intention to Use (BItU).

Participants selection: Participant selection was based on the following categories: *(a) Academic professional with experience in User Research* (*P3* - PhD. 16 years of industry experience. 11 years as a researcher, *P6* - PhD in progress. 24 years of industry experience); *(b) Industry professional with experience in User Research* (*P1* - Specialist. 6 years of industry experience); *(c) Newly integrated member of the unit of analysis* (*P4* - Undergraduate in progress. 4 months of experience); *(d) Experienced member of the unit of analysis* (*P2* - Undergraduate in progress. 1 and a half years of experience); and, *(e) Domain expert of the unit of analysis* (*P5* - PhD. 7 years of experience). The categories were chosen based on the experience variable, both about the project and knowledge of User Research. We opted for this type of sample stratification to reflect the usage scenarios of the process by different project members (following Rainer and Wohlin recommendations [Rainer and Wohlin 2022]): those with UX/UI expertise, with domain experts, or with members without experience in the previous two criteria. Additionally, we chose an industry and an academic User Research Specialist to gain insights into the process.

Study execution: The study took place both online and in person. Before conducting the study, we provided a form to the participant[3], which contained the consent form, a field to submit the proto-personas generated in the process, and the evaluation questions. We described the unit of analysis context briefly - concept of Solidarity Economy and the application's purpose - and then provided the participant with an execution script[4], a document that contains the prerequisites to start the process and the activities to be performed. Similarly, we prepared a script to support the researchers[5]. We conducted the execution of the process with one participant at a time and with two responsible researchers. One researcher guided the process, explaining the study proposal, while the other was designated for observation and note-taking. To enrich the observations, we used the think-aloud method, whereby we instructed the participant to narrate their actions, thoughts, and emotions during the process. We recorded audio, screen, and the total execution time. We instructed the participants to follow the execution script and evaluate the generated proto-personas[6] based on their UX/UI and project experience.

4.3.1. Data Analysis

To answer RQ, after assuring the similarity between the LLM-generated proto-persona and the project one during RQ1, we collected data from the TAM questionnaire and analyzed each statement. It aims to understand if there were trends toward agreement, disagreement, or neutrality. In the case of agreement, we verified if there was partiality or full agreement. In the case of disagreement, we discussed hypotheses that could generate the disagreement, based on qualitative data (think-aloud observations and participant

[3]https://tinyurl.com/mtnkb3ys.
[4]https://tinyurl.com/yessmtpz.
[5]https://tinyurl.com/jh6fpa78.
[6]Process-generated proto-personas samples: https://tinyurl.com/2r9jk6xe.

feedback on the process) collected during the studies, extracting refinement ideas for the process from this. For cases of neutrality, we analyzed the reasons behind the participant's non-positioning, also based on the other qualitative data mentioned. We also calculated the average execution time of the proto-persona generation process and compared it with the time for the same activity in the conventional LI activity. We also analyzed feedback from the process, the TAM questionnaire, and the proto-persona generation. We extracted transcription units, preliminary codes, and axial codes [Charmaz 2006] to understand the relationship between participant feedback and process metrics (e.g., proto-persona quality, ease of use). Interviews were conducted and coded by two or more authors over several iterative cycles.

4.4. Results' Analysis and Discussion

We extracted preliminary and axial codes[7] [Charmaz 2006] from think-aloud, TAM questions and proto-personas feedback from the collected qualitative data (Table 1). The analysis reinforced Likert scale results. We found that there was some little confusion about the process (as Table 1 exemplifies), but nothing that prevented the participants from executing with the facility. The axial codes from the proto-persona feedback revealed acceptance from the participants.

Transcription Unit	"Asking LLM to act as... what is LLM?"
Preliminary Code	Doubt about the definition of LLM.
Axial Code	Questions about definitions in the script

Table 1. Code extraction example - Participant 2

4.4.1. (RQ) What is the acceptance of "our approach" of generating proto-personas?

In this section, for each TAM topic, we present the results of the Likert scale and perform an analysis to understand the level of agreement.

Perceived Usefulness (PU): Using our approach makes it easier for you to build new proto-personas in your projects: Five participants (83%) agreed and one participant (17%) disagreed. **Analysis:** By analyzing the feedback provided by P1 who expressed disagreement regarding the process and outcome of the generation, additional efforts are required to review our approach results. Our approach may not consistently serve as a significant facilitator. P1 made an analogy of LLM with an intern, whose inputs need to be validated: "*So, at first, I have to see him as my intern who is creating for your persona, and then I will Check it out...*". P3 expressed a similar opinion regarding the response verification stage, although he/she agreed that it depends on a case-by-case basis: "*...we have to very critically analyze the responses received. So sometimes it can even make it difficult, a little bit...*".

We understand that our approach, especially in more complex projects, which require a more detailed specification of the proto-personas, may require work after generation since ChatGPT responses are not always 100% coherent with the given context.

[7]https://tinyurl.com/5n8hyhju.
https://tinyurl.com/3m8y8jw3.
https://tinyurl.com/yc62ps7t.

This review stage was already a hypothesis we had regarding our approach. In other versions of the process model, we tried to insert this validation via prompt, but we did not obtain any improvements in the results. Therefore, this activity after the generation of proto-personas is still an open topic to be explored in other works.

Using our approach speeded up the process of building proto-personas: four participants (66%) totally agreed and two (34%) partially agreed. **Analysis:** The average execution time of our approach is 11 minutes (SD ≈ 2.24). Compared to the 4 hours of the LI, it represents a considerable time saving: 95.42% faster. P5, who agreed, expressed: "*... I can't say it made it faster in terms of finalizing the process, but in terms of obtaining proto-personas it made it easier, yes.*". P1 and P3 expressed similar opinions.

Even if results are inconsistent, the time savings justify using generated proto-personas. They offer a solid starting point, allowing Lean Inception (LI) professionals to focus on refinement rather than creation. This accelerates persona development by providing quick initial proto-personas that can be refined or deemed sufficient to move on to the next LI activities, like developing user journeys.

<u>Perceived Ease of Use (PEoU):</u> **Understanding the instructions of our approach was easy for me:** five participants (83%) totally agreed and one (17%) partially agreed. **Analysis:** Our goal with the process was to make it as simple and clear as possible. Observing the qualitative data, the agreement given by the scale was reinforced. P3 highlights: "*Understanding this was very easy.*". P5 reinforces: "*Learning the instructions of our approach was easy for me, totally. The instructions themselves were easy.*". P1, the only one who partially agreed, defends: "*Again, there were the points that I mentioned, which seemed like a lot. I agree that I had the help, but totally not because it was quite extensive, you know...*". There were some doubts about understanding the process pointed out by P6, as described in "*Now the last activity, asking him to generate the proto-personas, again, right, here he is, right now, it's because it hasn't been requested yet, now he's going to ask.*", but as the transcription shows, the participant understood the instruction by himself/herself later.

<u>Attitude Toward Use (ATU)</u>: **I believe that using our approach is a great idea for generating proto-personas:** two participants (34%) agreed and four participants (66%) partially agreed. **Analysis:** Although some participants reported inconsistencies and indicated refinements in the process, everyone agreed that its use is a great idea for a proto-persona generation. P5 said: "*... highlighting that the process has to be refined to have some points, or at least add some related information, such as these warnings, etc.*" Another factor that led to a partial agreement was distrust in the quality of results in more complex contexts, as P4 said: "*So if it were a more complex project, maybe. That's why I don't know, I don't totally agree...*". P3 reinforces the uncertainty in the process: "*... It's a belief, it's not a certainty.*". The participants liked the idea and result of the generation, as reinforced by P6: "*... I think he did it right, I thought it looked cool.*", but many tests remained, and a single execution of the process was insufficient to ensure certainty.

I believe it is much better to use our approach, rather than a classic proto-persona generation process[8]: We observed a tendency towards neutrality: four partic-

[8]By the classic proto-persona generation process, we are primarily referring to the method proposed by [Caroli 2017]: a "diverge-converge" dynamic typically conducted over four hours.

ipants (66%) were neutral, one participant (17%) disagreed and another (17%) totally agreed. **Analysis:** It was the question with the most variations in responses from participants, most likely due to the weight of the statement. Our hypothesis about the four participants (66%) who presented neutrality is that, despite having liked the results of the process, the participants did not have experience using our approach in other projects, thus, they were not sure whether the quality of the results would hold in different cases. P5 expresses this idea: "*I think that depending on the way the process classic is conducted, we manage to generate better proto-personas.*". P3 was the one who disagreed with this statement, with the argument that our approach does not replace a classic process: "... *if I have something more complex, it can generate very simple results. So, sometimes, you can use it as an auxiliary and complementary platform in the creation process, and not just exchange one for the other.*". Thus, from the previous report, it is understood that our approach would be more useful as a support tool, an opinion shared by P1 and P5. P2 believes it is better than a classic activity: "*For other people, it's a bit boring to do a classic process.*". We can infer that our approach can include those people who do not like classic activity in LI, which contributes to generating proto-personas that they would not be able to carry out the LI activity, probably in a scenario with inexperienced members.

I like to use AI tools and processes to help with my UX activities: We observed a tendency towards partial agreement: one participant (17%) totally agreed and five (83%) partially agreed. **Analysis:** The partial agreement observed is due to the participants' view of AI tools as more of a support and with use with certain reservations. All six participants believe these tools and processes are welcome, but there should not be too strong an attachment to them, as tools like ChatGPT do not guarantee that the answers are 100% true. P4 expresses this frustration in: "*I just won't totally agree, because I'm still a little afraid of the tools.*". P1 has a similar opinion: "… *The ones we have on the market today, especially the design part, aren't that good yet, but the business has been evolving quite.*". The perception of AI as something more auxiliary (P3, P5).

Behavorial Intention to Use (BItU): I intend to use our approach whenever possible: five participants (83%) agreed, of which 3 (50%) totally agreed and 2 (33.33%) partially agreed. Furthermore, one participant (17%) was neutral. **Analysis**: even those who had their reservations regarding generation, for example, concerns in more complex cases, agreed with the statement, given the openness that the issue brings. P5 reinforces this idea: "*I agree. Given the way it is, I still have maybe a few saved, as I said, but it's easy and if I'm going to put it in whenever possible...* ". However, 1 participant (17%) was neutral. The participant's qualitative data indicate that he/she does not frequently perform the persona generation activity. It combined with P1's knowledge in generating proto-personas, suggests the hypothesis he/she is unsure whether he/she needs the process. Another statement from P1 highlights this uncertainty: "*We are clear that there was no need to carry out a new study. So, it depends, okay? So I neither agree nor disagree.*". P1 considers using the process in an academic scenario.

I would adopt new tools similar to our approach in the future: Five participants (83%) totally agreed and one (17%) partially agreed. **Analysis:** We observed that the participants liked our approach, which makes it consistent that the vast majority (83%) totally agreed with the statement. P3 highlights: "*I totally agree. If there are tools that are reliable that use AI, obviously I will want to use them.*". Some participants also highlighted that they would like to adapt our approach for their projects and were excited by the ease and speed it brings. P3 himself comments: "*I intend to use it and I even intend to adapt it.*". Even those who were more critical, like P1, expressed: "*I personally if I had the opportunity, I would even recommend this system of yours.*". P5, who commented on process refinements, said: "*Yes, good tools are always welcome. I would adopt new tools similar to our approach in the future, yes, for sure.*". Therefore, we analyzed that the proposal of our approach has the potential to be a good alternative for generating proto-personas in a Lean Inception. Just some little additions have to be made, as including pictures to the generated proto-personas, as suggested by P6: "*...it usually generates a photo, you could work on that from there, but I think it turned out pretty well.*".

5. Limitations

Given the specific nature of the unit of analysis and the participants involved, some findings are likely specific to this context. The application and its user interactions present unique characteristics that may not generalize to other types of applications or user groups. Additionally, the proto-persona generation process using Lean Inception and prompt engineering may yield different results in other settings. The reliance on self-reported data from participants also introduces the potential for bias.

6. Conclusion and Future Work

This research is part of a bigger work, published in the *38th Brazilian Symposium on Software Engineering* (SBES) 2024: "*A Prompt Engineering-based Process to Build Proto-personas during Lean Inception*". This work includes the research of two master's students focused on Product Discovery and UX/UI. **Our research** found that proto-personas developed using our approach better represented the target audience than those from LI, despite some technological inconsistencies. The approach was well accepted by participants, though there was neutrality due to limited testing. Limitations include handling critical inconsistencies or complex projects with brief proto-personas.

Implications to industry and research: **Faster Product Discoveries:** Our approach cuts proto-persona development in LI from 4 hours to 11 minutes. **Inception closer to solution:** Domain-based proto-personas better align with solution requirements, moving inception closer to the solution. **Context-Grounded Solutions:** Ross Mayfield - Header of Zoom - emphasized the importance of problem context[9]. Our approach enhances context compliance using the Context Manager prompt engineering pattern. Further research can explore different approaches and new applications of prompt engineering in other Product Discovery activities. **Future Work:** It is crucial to continue examining the use of LLMs in RE and Product Discovery, focusing on result conformity, similarity evaluation strategies, adapting the process for various inputs and activities, conducting tests, using other AI models, and refining the persona template. Additionally, comparing prompts, adopting rigorous validation, and applying the approach to more complex cases are important steps.

References

Arora, C., Grundy, J., and Abdelrazek, M. (2023). Advancing requirements engineering through generative ai: Assessing the role of llms. *arXiv preprint arXiv:2310.13976*.

[9]https://productcoalition.com/problems-with-product-discovery-31601fe95d9c.

Atlas, S. (2023). Chatgpt for higher education and professional development: A guide to conversational ai.

Belzner, L., Gabor, T., and Wirsing, M. (2023). Large language model assisted software engineering: prospects, challenges, and a case study. In *International Conference on Bridging the Gap between AI and Reality*, pages 355–374. Springer.

Caroli, P. (2017). Lean inception. *São Paulo, BR: Caroli. org*.

Charmaz, K. (2006). *Constructing grounded theory: A practical guide through qualitative analysis*. sage.

Fan, A., Gokkaya, B., Harman, M., Lyubarskiy, M., Sengupta, S., Yoo, S., and Zhang, J. M. (2023). Large language models for software engineering: Survey and open problems. *arXiv preprint arXiv:2310.03533*.

Gothelf, J. (2013). *Lean UX: Applying lean principles to improve user experience*. " O'Reilly Media, Inc.".

Karolita, D., McIntosh, J., Kanij, T., Grundy, J., and Obie, H. O. (2023). Use of personas in requirements engineering: A systematic mapping study. *Information and Software Technology*, page 107264.

Liu, P., Yuan, W., Fu, J., Jiang, Z., Hayashi, H., and Neubig, G. (2023). Pre-train, prompt, and predict: A systematic survey of prompting methods in natural language processing. *ACM Computing Surveys*, 55(9):1–35.

Losana, P., Castro, J. W., Ferre, X., Villalba-Mora, E., and Acuña, S. T. (2021). A systematic mapping study on integration proposals of the personas technique in agile methodologies. *Sensors*, 21(18):6298.

Marques, N., Silva, R. R., and Bernardino, J. (2024). Using chatgpt in software requirements engineering: A comprehensive review. *Future Internet*, 16(6):180.

Rainer, A. and Wohlin, C. (2022). Recruiting credible participants for field studies in software engineering research. *Information and Software Technology*, 151:107002.

Runeson, P. and Höst, M. (2009). Guidelines for conducting and reporting case study research in software engineering. *Empirical software engineering*, 14:131–164.

Salminen, J., Wenyun Guan, K., Jung, S.-G., and Jansen, B. (2022). Use cases for design personas: A systematic review and new frontiers. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, pages 1–21.

Trieflinger, S., Lang, D., Spies, S., and Münch, J. (2023). The discovery effort worthiness index: How much product discovery should you do and how can this be integrated into delivery? *Information and software technology*, 157:107167.

White, J., Fu, Q., Hays, S., Sandborn, M., Olea, C., Gilbert, H., Elnashar, A., Spencer-Smith, J., and Schmidt, D. C. (2023). A prompt pattern catalog to enhance prompt engineering with chatgpt. *arXiv preprint arXiv:2302.11382*.

Zhang, X., Liu, L., Wang, Y., Liu, X., Wang, H., Arora, C., Liu, H., Wang, W., and Hoang, T. (2024). Auto-generated personas: Enhancing user-centered design practices among university students. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*, pages 1–7.