

The first study (demonstrating the effectiveness of requirements analysis) (10 minutes)

Initial Task Requirements:

- Task 1: Simulated Interview: I want to develop a service to help users prepare for interviews.
- Task: Think of a task for requirements analysis on your own.

Task Requirements:

- Participants should use Sapper v2 to complete requirements analysis within 5-10 minutes.
- After completing the task, save the JSON file or take a screenshot of the clarification chat.
- Fill out the following questionnaire:

https://qualtricsxml2d5csky4.qualtrics.com/jfe/form/SV_9BmYeEvp2FEGdkW

The second study (demonstrating the effectiveness of task decomposition)

Task Requirements:

Participants should use Task Decomposition Copilot to break down the task they have thought of and run the generate skeleton

- Download the results as a JSON file or take a screenshot of the decomposition.
- After completing the task, fill out the questionnaire and provide a rating (in a table format):

https://qualtricsxml2d5csky4.qualtrics.com/jfe/form/SV_5duw7bmcgLgahds

1. Usefulness:

1 point: Extremely unhelpful - The tool provided almost no assistance in clarifying your requirements.

2 points: Not very useful - The tool offered some help, but the effect was minimal.

3 points: Neutral - The tool provided some useful assistance in clarifying your requirements, but the effect was average.

4 points: Useful - The tool significantly helped in clarifying requirements and was beneficial to you.

5 points: Very useful - The tool greatly improved the efficiency and effectiveness of your requirements clarification.

2. Serendipity:

1 point: No serendipity - There were no surprising or unexpected discoveries while using the tool.

2 points: Minor serendipity - Some minor, insignificant unexpected findings occurred.

3 points: Neutral - The tool provided some unexpected information, but there were not many surprising insights.

4 points: Serendipitous - The tool provided some surprising information or insights.

5 points: Very serendipitous - The tool offered highly valuable and unexpected information or insights, exceeding your expectations.

3. Diversity:

1 point: Single and limited - The tool provided very uniform information or suggestions with minimal diversity.

2 points: Limited diversity - The tool provided some different information but with limited variety.

3 points: Neutral - The tool provided some diverse information, but there is room for improvement.

4 points: Diverse - The tool provided a variety of different types of information or suggestions.

5 points: Very diverse - The tool offered a wide range of diverse information or suggestions, allowing you to approach requirements from multiple angles.

4. Ease-of-Use:

1 point: Extremely difficult to use - Operating the tool was very challenging and prone to errors.

2 points: Very difficult to use - Using the tool was somewhat challenging and could lead to errors.

3 points: Neutral - Operating the tool was neither exceptionally easy nor difficult, with no significant issues.

4 points: Easy to use - The tool was relatively user-friendly and not prone to errors.

5 points: Very easy to use - The tool was highly intuitive and easy to operate, with minimal risk of errors.

5. Relevance:

1 point: Completely irrelevant - The tool's suggestions were entirely unrelated to your requirements.

2 points: Not very relevant - The tool's suggestions had some relevance, but most were unrelated.

3 points: Neutral - The tool's suggestions were relevant in some aspects of your requirements but needed improvement.

4 points: Relevant - The tool's suggestions were relevant in most cases to your requirements.

5 points: Very relevant - The tool's suggestions were closely related to your requirements, helping you better understand and meet them.

6. How is this tool different in providing diverse clarification suggestions compared to not using it?

7. Did this tool provide ways or perspectives to clarify your requirements that you had not considered before? If so, can you share some examples?

8. Which specific features or aspects of the tool do you believe need improvement or refinement, and why?

9. Do you have any additional feedback?

Usefulness: This aspect focuses on the extent to which our Copilot has been genuinely helpful in clarifying your requirements. After using our tool, did you feel that you could better understand your requirements? Please consider the impact of using the tool on the clarity of your requirements.

Serendipity: This section pertains to whether there were any surprising discoveries while using the tool. Did you come across unexpected and intriguing information or insights that exceeded your initial expectations during tool usage?

Diversity: The diversity aspect considers whether you were able to obtain a variety of different insights or suggestions while using the tool. Did the tool provide diverse ways or perspectives to clarify requirements? Please reflect on whether the tool helped you acquire different types of information.

Ease-of-Use: This aspect relates to the ease or difficulty of using the tool. Did you find the tool easy to operate, or did it require a significant amount of time to learn and use? Please think about your user experience, especially as it relates to the tool's usability.

Relevance: The relevance aspect focuses on whether the tool's suggestions were closely related to your requirements. Do you believe the tool provided clarification suggestions that were highly relevant to your needs, or did it include irrelevant or unrelated information? Please consider whether the tool accurately understood and addressed your requirements.

1. Overall Structure:

1 point: Extremely unreasonable - Task decomposition has no clear connection to the original task requirements, and the structure is chaotic and difficult to understand.

2 points: Not very reasonable - Task decomposition is somewhat related to the original requirements but lacks clarity or key information.

3 points: Neutral - Task decomposition is somewhat related to the requirements but has room for improvement.

4 points: Reasonable - Task decomposition is closely related to the requirements, the structure is clear, but there is still room for improvement.

5 points: Very reasonable - Task decomposition is highly relevant to the requirements, the structure is very clear, easy to understand, and no key information is missing.

2. Completeness:

1 point: Extremely incomplete - Task decomposition severely lacks important information and cannot execute task substeps.

2 points: Not very complete - Task decomposition lacks essential information and requires further refinement and supplementation.

3 points: Neutral - Task decomposition to some extent includes key points of the requirements but needs more details.

4 points: Complete - Task decomposition covers most of the requirements' key points but can provide more information.

5 points: Very complete - Task decomposition comprehensively includes all key points of the requirements, with nothing missing.

3. Ease-of-Modification:

1 point: Very difficult to modify - Task decomposition cannot be easily modified and requires a significant amount of time and effort.

2 points: Quite difficult to modify - Task decomposition can be modified but demands a considerable amount of effort.

3 points: Neutral - Task decomposition offers some flexibility in modification but still has room for improvement.

4 points: Easy to modify - Task decomposition is relatively easy to modify, requiring moderate time and effort.

5 points: Very easy to modify - Task decomposition provides great flexibility and can be easily modified as needed without causing errors.

4. Could using our tool help you better plan and organize the execution of tasks compared to not using it? Please describe your experience.

5. After performing task decomposition, did you feel more confident and prepared to start the task? Did it make it easier for you to get started?

6. Which specific features or aspects of the tool do you believe need improvement or refinement, and why?

7. Do you have any additional feedback?

Overall Structure: This aspect focuses on whether you found the overall structure of the task decomposition using our Copilot to be very clear. In other words, whether you were able to adequately consider the core conditions and key steps of the task during decomposition to ensure that the post-decomposition structure contributes to a better understanding of the task.

Completeness: Completeness aspect examines whether you felt that all details of the task were thoroughly considered when using our tool for task decomposition. This includes whether all sub-tasks, steps, or relevant elements of the task were listed to ensure that there are no omissions.

Ease-of-Modification: Ease-of-modification aspect concerns whether you felt that it was easy to make modifications and adjustments after task decomposition. In other words, whether you could conveniently add, delete, or edit sub-tasks or steps of the task to meet potential changes or further optimize the task structure.