

第一个 study(证明需求分析的有效性) (12 人) (10 min)

初始任务需求:

Task1: 模拟面试: 我想开发一个帮助用户准备面试的服务

Task: 自己思考一个 task 进行需求分析

任务要求:

参与者使用 Sapper v2 在 5-10min 内完成需求分析

完成任务后保存 json 文件/将澄清聊天记录截图

填写以下问卷

第二个 study(证明任务分解的有效性) (12+3 人)

任务要求

参与者将自己思考的任务使用任务分解 copilot 进行分解

下载为 json 文件/将分解结果截图

完成任务后进行填写问卷/进行打分 (表格)

1、Usefulness（有用性）：

- 1分： 极其不有用 - 工具几乎没有对您澄清需求提供帮助。
- 2分： 不太有用 - 工具提供了一些帮助，但效果非常有限。
- 3分： 中立 - 工具对您澄清需求提供了一些有用的帮助，但效果一般。
- 4分： 有用 - 工具在澄清需求方面提供了明显的帮助，对您有所裨益。
- 5分： 非常有用 - 工具极大地提高了您澄清需求的效率和效果。

2、Serendipity（意外收获）：

- 1分： 毫无意外 - 在使用工具时，没有任何令人惊喜或意外的发现。
- 2分： 微不足道的意外 - 有一些小的、无足轻重的意外发现。
- 3分： 中立 - 工具提供了一些意外的信息，但没有太多令人惊喜的。
- 4分： 有意外收获 - 工具提供了一些令人惊喜的信息或见解。
- 5分： 非常有意外收获 - 工具提供了极具价值的、意外的信息或见解，超出了您的期望。

3、Diversity（多样性）：

- 1分： 单一和局限 - 工具提供的信息或建议非常单一，没有多样性。
- 2分： 很少多样性 - 工具提供了一些不同的信息，但种类有限。
- 3分： 中立 - 工具提供了一些多样的信息，但仍然有改进的空间。
- 4分： 有多样性 - 工具提供了多种不同类型的信息或建议。
- 5分： 非常多样性 - 工具提供了广泛多样的信息或建议，让您可以从多个角度看待需求。

4、Ease-of-Use（易用性）：

- 1分： 非常难以使用 - 操作工具非常困难，容易导致错误。
- 2分： 很难使用 - 操作工具有一些困难，可能容易导致错误。
- 3分： 中立 - 操作工具既不特别容易，也不特别困难，没有明显的问题。
- 4分： 容易使用 - 操作工具相对容易，不太容易导致错误。
- 5分： 非常容易使用 - 工具非常直观和易于操作，不容易出错。

5、Relevance（相关性）：

- 1分： 完全不相关 - 工具提供的建议与您的需求完全不相关。
- 2分： 不太相关 - 工具提供的建议有一些相关性，但大部分不相关。
- 3分： 中立 - 工具提供的建议在一些方面与您的需求相关，但还需要改进。
- 4分： 相关 - 工具提供的建议在大部分情况下与您的需求相关。
- 5分： 非常相关 - 工具提供的建议与您的需求密切相关，帮助您更好地理解 and 满足需求。

6、与在没有使用这个工具时相比，你认为这个工具在提供多样的澄清建议方面有何不同？

7、这个工具是否提供了与你之前没有考虑到的方式或角度来澄清你的需求？如果是的话，可以分享一些例子吗？

8、您认为工具的哪些特定特性或方面需要改进或细化，为什么？

9、您还有其他意见吗？

Usefulness（有用性）：这个方面关注的是我们的 copilot 对于您澄清需求是否真正有帮助。在使用我们的工具后，您是否感到能够更清楚地理解您的需求？请考虑您使用工具后需求澄清的效果。

Serendipity（意外收获）：这一部分涉及到在使用工具时是否有令人惊喜的发现。您是否在使用工具时发现了一些意外的、有趣的信息或见解，这些信息可能超出了您最初的期望？

Diversity（多样性）：多样性方面考虑的是在使用工具时，您是否能够获取多种不同的见解或建议。工具是否提供了多样的方式或角度来澄清需求？请思考工具是否能够帮助您获得不同的信息。

Ease-of-Use（易用性）：这个方面涉及到您使用工具的难易程度。您是否发现工具易于操作，是否需要花费很多时间来学习或执行操作？请考虑您的用户体验，特别是与工具的使用相关的方面。

Relevance（相关性）：相关性方面关注的是工具提供的建议是否与您的需求相关。您是否认为工具提供的澄清建议与您的需求密切相关，或者是否存在无关或不相关的信息？请思考工具是否能够准确地理解和满足您的需求。

1、整体结构（Overall Structure）：

- 1分：非常不合理 - 任务分解与原始任务需求没有明显关联，结构混乱且难以理解。
- 2分：不太合理 - 任务分解与原始需求有一定关联，但结构较混乱或者缺少关键信息。
- 3分：中立 - 任务分解在一定程度上与需求相关，但还有改进空间。
- 4分：合理 - 任务分解与需求关联紧密，结构清晰，但仍有些许改进空间。
- 5分：非常合理 - 任务分解与需求高度相关，结构非常清晰，容易理解，没有遗漏关键信息。

2、完整性（Completeness）：

- 1分：非常不完整 - 任务分解严重缺乏重要信息，无法执行任务子步骤。
- 2分：不太完整 - 任务分解缺少重要信息，需要进一步细化和补充。
- 3分：中立 - 任务分解在某种程度上包含了需求的关键要点，但需要更多细节。
- 4分：完整 - 任务分解涵盖了需求的大部分要点，但仍可提供更多信息。
- 5分：非常完整 - 任务分解详尽地包含了需求的所有要点，没有遗漏。

3、易修改性（Ease-of-Modification）：

- 1分：非常难以修改 - 任务分解无法轻松修改，需要大量时间和努力。
- 2分：很难修改 - 任务分解可以修改，但需要耗费相当多的精力。
- 3分：中立 - 任务分解在修改方面提供了一定的灵活性，但仍有改进的空间。
- 4分：容易修改 - 任务分解相对容易修改，需要的时间和努力适中。
- 5分：非常容易修改 - 任务分解提供了极大的灵活性，可以轻松地根据需要进行修改，不会导致错误。

4、相对于没有使用我们的工具时，使用我们的工具是否能够帮助您更好地规划和组织任务的执行？请描述您的体验。

5、在进行任务分解后，您是否感到更自信和准备好开始任务？它是否让您更容易开始工作？

6、您认为工具的哪些特定特性或方面需要改进或细化，为什么？

7、您还有其他意见吗？

Overall-structure（总体结构）：这个方面关注的是你在使用我们的 copilot 进行任务分解时，是否感到任务的总体结构非常清晰。也就是说，你能否在分解任务时充分考虑到任务的核心条件和关键步骤，确保任务分解后的结构有助于更好地理解任务。

Completeness（完整性）：完整性方面考察的是你在使用我们的工具分解任务时是否感到任务的所有细节都得到了充分的考虑。这包括了任务的所有子任务、步骤或相关要素是否都被列举出来，以确保任务不会有遗漏。

Ease-of-modification（易修改性）：易修改性方面关注的是在分解任务后，你是否感到可以轻松地进行修改和调整。也就是说，你是否能够方便地添加、删除或编辑任务的子任务或步骤，以满足可能出现的变化或进一步优化任务结构的需要。