# Assignment 1

Shaikhutdinov Murat

## Team Members and Roles

Solo project by Shaikhutdinov Murat.
**AI Tools Used**: Grok for creating the LaTeX file from a text file and for grammar checking.

## Three Mom Principles

1. **Talk about their life instead of your idea**: Avoid pitching the tool or seeking validation for the idea. Focus on the interviewee's current behaviors, challenges, and needs.
2. **Ask about specifics in the past instead of generics or opinions about the future**: Prioritize questions about concrete past actions over hypothetical future behaviors or opinions.
3. **Talk less and listen more**: Craft questions that encourage detailed responses and avoid leading the interviewee to desired answers.

## Interview Script for Follow My Reading Project

### Introduction

Thank you for participating in this interview. I am exploring how people manage and share texts and audio recordings in their work or personal projects. This conversation will help me understand your current processes, challenges, and needs. I understand that I already have a project specification, but let's stick to the Mom-principled interview for deeper research quality. It should take about 20–30 minutes.

### Section 1: Project Background and Goals

Q1 Have you or your team used any tools or processes for uploading texts, recording audio, or sharing them with others in the past?

- Yes
- No
- (if Yes): Can you briefly describe one or two tools or methods you've used and what stood out about them?

Q2 Can you walk me through a recent situation where you needed to share or work with a text document, like uploading it for others to access or having someone read it aloud? What was the context, and what challenges came up?

Q3 When you've needed to create or use audio versions of texts in the past, how did you go about it? Can you describe a specific example, including what tools or steps you used?

Q4 Can you share a time when you or your team struggled to organize, find, or use texts or audio recordings effectively? What happened, and how did you handle it?

**Changes from Originally Thought Questions**

- **Q1 (Kept)**: Retained as it asks about past tool usage, aligning with Rule 2.
- **Q2 (Changed)**: Original asked about motivations for the tool, which risks focusing on the idea (Rule 1 violation) and future implementation (Rule 2 violation). Replaced with a question about a recent situation to focus on past behaviors.
- **Q3 (Changed)**: Original asked to prioritize features, which is hypothetical and idea-focused (Rules 1 and 2 violations). Replaced with a question about past audio creation to gather practical data.
- **Q4 (Changed)**: Original asked for a vision of the tool's fit, which is future-focused (Rule 2 violation). Replaced with a question about past struggles to stay grounded in specifics.

## Section 2: Operating in the Problem Space

Q5 Can you describe your current process for handling tasks like uploading, storing, or sharing text documents/audio? Walk me through the last time you did this.

Q6 How often do you or your users create audio recordings of texts?

- Daily
- Weekly
- Monthly
- Rarely
- Follow-up: Can you share an example of the last text you recorded and why you needed to record it?

Q7 Can you tell me about the last time you thought about/tried analyzing an audio recording to match it to text? What tools or methods did you used, and what was challenging?

Q8 Have you ever used a tool or method to synchronize text with audio playback?

- Yes
- No
- (if Yes): Can you describe how you used it and what worked or didn't?
- (if No): How have you managed without this capability?

Q9 Can you walk me through how you currently organize and find texts or audio recordings?

**Changes from Original Section 2**

- **Q5 (Kept)**: Already asks about current processes and past behaviors, aligning with Rules 1 and 2.
- **Q6 (Kept)**: Asks about frequency (past behavior) and includes a follow-up for specifics, aligning with Rule 2.
- **Q7 (Kept)**: Asks about past usage of synchronization tools, aligning with Rule 2.
- **Q8 (Kept)**: Focuses on past challenges with audio analysis, aligning with Rule 2.
- **Q9 (Changed)**: Original asked about desired parameters of web app, which is hypothetical and future-focused (Rule 2 violation). Replaced with a question about current organization methods to focus on past behaviors.

### Closing

Is there anything else about your experiences with text or audio tasks that you think is important for us to understand? Do we need to contact somebody else?

#### Changes from Original Closing

- **(Kept)**: Already open-ended and invites additional insights without focusing on the idea, aligning with Rules 1 and 3.

### End

Thank you for sharing your insights! Your feedback will help us better understand the needs of the app. Would you be open to future conversations as we continue our research?

## Product Research Board

`https://docs.google.com/document/d/1DavFaoV_xYFIjP7-G1u9jKdkVUVwZzvQJDbv1T9dDk4/edit?usp=sharing`

## Qualitative Analysis Table

Table 1: Qualitative Analysis of Competing Products

| Product | Text Extraction | Audio Recording | Audio Segmentation | Text-Audio Sync | User Accessibility | Notes |
|---|---|---|---|---|---|---|
| **Descript** | Uploads text or transcribes audio automatically. | Built-in high-quality recording. | Auto-segments by speaker/pauses; editable. | Syncs text with audio, highlights words. | Web-based, intuitive UI. | Strong for podcasting; may be overkill. |
| **Audacity** | Manual import via plugins; no native extraction. | Robust recording and editing. | Manual via labels; no auto word-level. | No native sync; needs external tools. | Free, but steep learning curve. | Best for editing, not syncing. |
| **Otter.ai** | Real-time audio-to-text; no text upload. | Records or imports audio. | Segments by speaker/time; not word-level. | Highlights text, not precise sync. | Mobile/web, user-friendly. | Focused on meetings, not custom sync. |
| **Speechify** | Uploads/scans text for extraction. | No user recording; synthetic audio. | No user audio segmentation. | Highlights text with synthetic audio. | Mobile/web, accessibility-focused. | Lacks user recording; text-to-speech only. |
| **ReadSpeaker** | Extracts from documents/web. | No user recording; synthetic speech. | No user audio segmentation. | Syncs synthetic audio with text. | Web-based, multi-language. | Enterprise-focused; no user audio. |

## List of Learned/Requiring Qualification Things from the Interview

(Theoretical as there was no interview in real life)

### Project Domain

Language learning, school utility/edtech, audiobook creation on later stages, speech analysis.

## Stakeholders

- **Voice Actors**: With integrating to other services.
- **Readers**: Interact with content, read aloud, and contribute recordings if implementing a public library.
- **Educators/Tutors**: Use the tool to assess pronunciation or fluency.

## Need to Estimate with Customer

### Market

- Who are the target users (age, profession, motivation)?
- What alternatives do they currently use?
- Are they even motivated to upload or read texts?

### Objective

- What does successful interaction look like for different users?
- How often would a typical user use the product?
- What's the primary motivation: learning, fun?

### Method

- How do they currently record themselves reading?
- How would users interact with audio playback and text highlighting?
- Would users be okay with recordings being public/shared?

## MVP

Simple Telegram text transcriber mapping the spoken word to text.