**Software Requirements Specification Document**

|  | Tower of Babble, Team 6:- Arup Biswas, Arvind Srinivasan, Dhriti Chintakunta, Keshav Kumar Manjhi, Abhinav Gupta |
| --- | --- |

# **Brief problem statement**

In this project, we provide a proposed method to address the existing requirement for translation and transcription of communications within global corporations, where individuals from diverse linguistic backgrounds are involved. At present, individuals are required to upload their audio or video content onto other servers, which may give rise to a potential breach of privacy. A significant number of current systems likewise operate on a subscription-based model, with the specific methodologies employed being undisclosed. Our team offers a free, open-source web application that provides a layer of abstraction for the backend infrastructure, also allowing it to be hosted on the organization's own servers.

# **System requirements**

MERN Stack - The combination of MongoDB, ExpressJS, ReactJS and NodeJS that powers a web full stack

Whisper.cpp - Our desired transcription service - provides both speed and accuracy

Fairseq - Our translation model - built for non-english translation

Docker - Provides the abstractions and dependency resolution to allow the project to run smoothly

# **Users profile**

Companies who wish for privacy and self-host the website so as to prevent sharing of data to third-party servers are our primary target audience.

We wish to simplify(abstract) the translation and transcription of audio for the end-users.

The users are not expected to know the underlying workings of the web-app, and would just be required to upload the audio/video file(s) to be transcribed/translated and they would be provided with the output, to save.