Infosys Springboard Virtual Internship 6.0 Completion Report

Team Details:

Batch Number : 3

Start date : 18 September 2025

Names:

- Shaikh Mohammad Affan Inayat
- NallaGonda Surya Vamsi
- Ravalika Tulagapu
- Chikkam Bala Satya Aditya
- Gowdham K
- Rishika Roy
- Rajeev Reddy
- Sanket S Mathpati
- Nandhana C H
- Sadhvika Bhaskar
- Jayahar J
- Thariga Shrinithi R
- K Sneha

Internship Duration: 8 Weeks

1. Project Title:

AI-Powered Real-Time Speech Translation For Multilingual Content

A multi-module AI-powered platform for multilingual speech translation — integrating offline, real-time, and web-based translation tools using open-source technologies

2. Project Objective:

To build a **modular, end-to-end AI speech translation system** capable of:

- Converting spoken audio into text (STT)
- Translating the text into 12+ Indian languages
- Generating translated speech (TTS)
- Providing both **batch** and **real-time** translation capabilities via CLI and Web UI

The project demonstrates how **AI can break language barriers** in real-world communication such as OTT platforms, education, and accessibility tools.

3. Project description in detail:

QalamAI is structured as a **multi-module monorepo**, containing four integrated modules:

Module	Folder	Purpose
Module 1	`speech-translator/module1`	Environment setup and basic translation sanity check
Module 2	`speech-translator/module2`	Batch audio translation (multi-language support)
Module 3	`speech-translator/module3`	Real-time OTT / CLI-based translation
Module 4	`speech-translator/module4`	Flask Web App with live mic and file upload translation

Technologies Used

- **Python 3.11+**
- **Flask** (Web UI)
- **SpeechRecognition / Whisper** (Speech-to-Text)
- **Google Translate / Deep Translator** (Translation)
- **gTTS / Edge TTS** (Text-to-Speech)
- **librosa, pydub, ffmpeg, moviepy** (Audio processing)
- **Hugging Face Transformers** (Language models)
- **HTML, CSS, JavaScript** (Frontend for Flask UI)

Real-World Impact

The project can be adapted for:

- **Educational platforms** — to provide real-time lecture translations.
- **OTT/YouTube** — for automated multilingual captions and dubbing.
- **Accessibility solutions** — to help differently-abled users access content in preferred languages.
- **Government / Public services** — enabling voice-based multilingual communication.

4. Timeline Overview:

Week	Activities Planned	Activities Completed
Week 1	Environment setup, Python library installation, and initial GitHub repo creation.	Virtual environment created, dependencies installed, and base repo initialized.
Week 2	Build Module 1: Environment verification & test translation pipelines.	Successfully executed `test_env.py` using Helsinki-NLP model; verified BLEU scores.
Week 3	Develop Module 2: Batch audio translation using SpeechRecognition & gTTS.	Implemented dataset fetcher, audio converter, and batch translator scripts.
Week 4	Integrate translation and TTS pipelines; generate CSV logs and outputs.	End-to-end pipeline completed; tested on multilingual audio samples.
Week 5	Build Module 3: Real-time OTT translation with Whisper / Google	Created `module3_ott_realtime.py` supporting Whisper and Google STT.

|		
	STT.	
Week 6	Test multilingual real-time translator; refine latency and audio handling.	Reduced processing time; validated support for 12 Indian languages.
Week 7	Build Module 4: Flask web application (mic + upload + YouTube).	UI implemented with file upload, mic input, and translation playback.
Week 8	Testing, documentation, and final presentation.	Completed full testing, screenshots, README, and project submission.

5a. Key Milestones:

| Milestone	Description	Date Achieved
Project Kickoff	Environment setup & team alignment	Week 1
Prototype / First Draft	Basic translation workflow (Module 1–2)	Week 3
Mid-Term Review	End-to-end demo of STT→Translate→TTS	Week 5
Final Submission	All modules completed & tested	Week 8
Presentation	Final showcase & report submission	Week 8

5b. Project execution details:

The execution followed an **Agile-like sprint structure** across 8 weeks.

- **Phase 1 (Setup & Testing):** Environment setup, dependency validation, and base translation tests using Hugging Face models.
- **Phase 2 (Batch Processing):** Built and tested batch translation pipeline supporting 12+ languages.
- **Phase 3 (Real-time Translation):** Developed CLI-based live translator with Whisper and Google STT.
- **Phase 4 (Web Integration):** Integrated Flask-based web app for live mic and file translation.
- **Phase 5 (Optimization & Reporting):** Improved processing speed, UI polish, documentation, and report generation.

6. Snapshots / Screenshots:
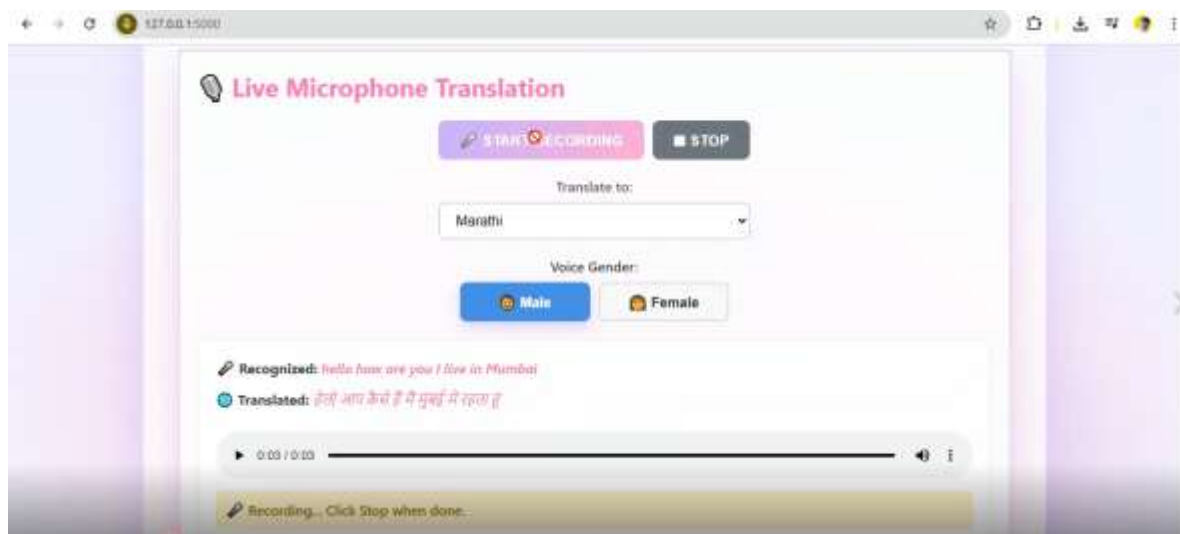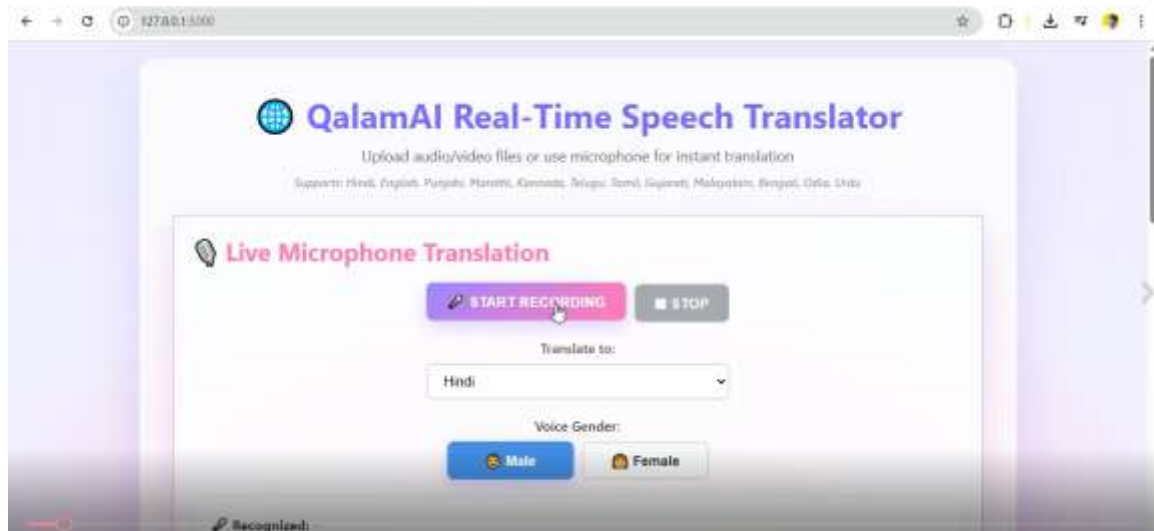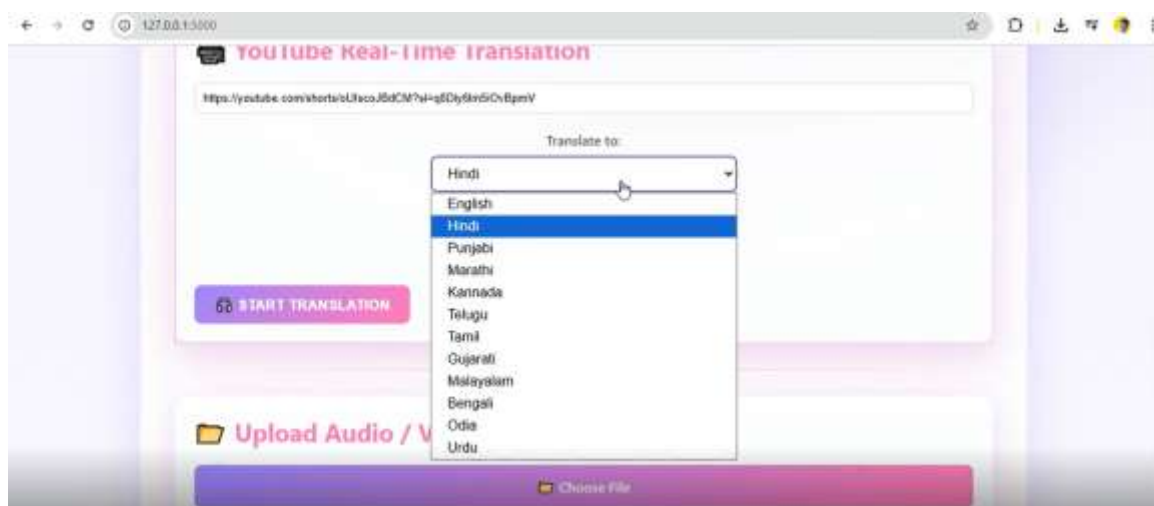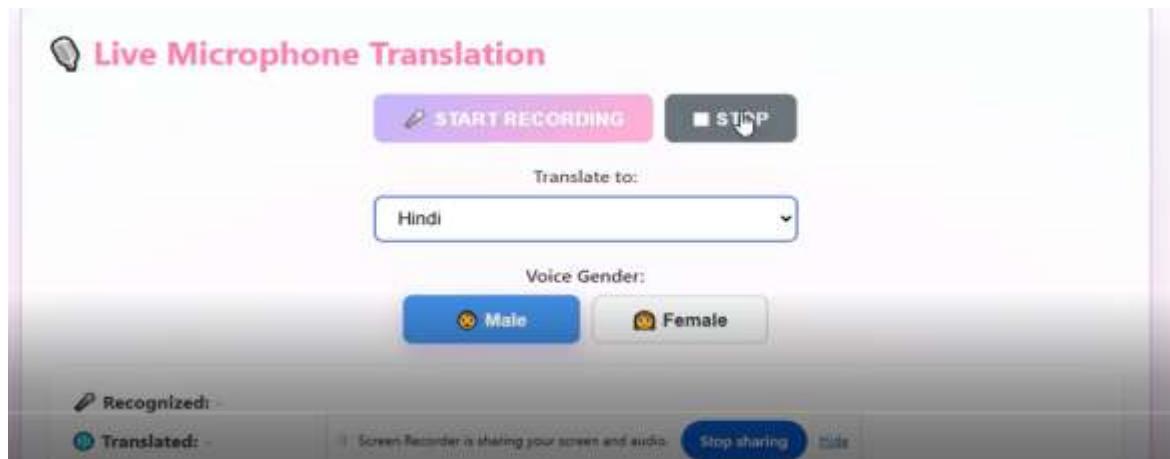
7. Challenges Faced:

Challenge	Description	Resolution
FFmpeg dependency errors	Audio conversion failed during early testing.	Installed FFmpeg via PATH and used `imageio_ffmpeg` fallback.
SpeechRecognition API timeouts	High latency for long audio files.	Implemented chunk-based processing and retry logic.
Translation inconsistencies	Google Translate API rate limits hit during batch mode.	Added delay handling and offline alternatives.
gTTS language limitations	Some Indian languages (Punjabi, Odia) unsupported.	Used Hindi fallback TTS output for missing languages.

8. Learnings & Skills Acquired:

Technical Skills

- Speech Processing (STT, TTS, Translation)
- Python (Flask, Whisper, transformers, pydub, ffmpeg)
- Web Development (HTML, CSS, JavaScript)
- RESTful API design and integration
- Dataset handling and logging with Pandas

Soft Skills

- Collaboration and version control using Git/GitHub
- Problem-solving and debugging complex pipelines
- Technical documentation and reporting
- Time management and team coordination

9. Testimonials from team:

"Working on QalamAI helped us understand how speech translation works end-to-end — from recognizing voices to generating natural translations in multiple languages."

"It was an enriching experience that improved both our technical and teamwork abilities."

"We feel proud to have built a functional prototype that could genuinely make communication more inclusive."

10. Conclusion:

This 8-week Infosys Springboard Virtual Internship provided hands-on exposure to **real-world AI project development**, covering model integration, backend logic, and frontend deployment.

QalamAI Speech Translator "**AI-Powered Real-Time Speech Translation For Multilingual Content"** is not just a prototype — it is a scalable foundation for AI-driven multilingual accessibility.
The project aligns with our **career goals in AI, NLP, and software engineering**, strengthening our confidence in end-to-end development.

11. Acknowledgements:

We would like to extend our heartfelt gratitude to **Neha Ma'am** for her continuous guidance, mentorship, and encouragement throughout our **Infosys Springboard Virtual Internship 6.0** journey. Her patient support, valuable insights, and efficient daily assistance motivated us to stay focused and perform every task with precision and enthusiasm.

We also take this opportunity to sincerely thank our amazing team members: **Shaikh Mohammad Affan Inayat, NallaGonda Surya Vamsi, Ravalika Tulagapu, Chikkam Bala Satya Aditya, Gowdham K, Rishika Roy, Rajeev Reddy, Sanket S Mathpati, Nandhana C H, Sadhvika Bhaskar, Jayahar J, Thariga Shrinithi R, and K Sneha** — for their collaboration, dedication, and excellent teamwork.

Each member contributed significantly in research, coding, report writing, and preparing the **technical presentation**, ensuring that our project, **"AI-Powered Real-Time Speech Translation For Multilingual Content"** was completed successfully and to the highest standard.

Finally, we sincerely thank **Infosys Springboard** for providing us with this incredible learning opportunity that has strengthened our technical expertise, teamwork, and professional growth.

.