**Project Proposal: Bilingual AI Virtual News Anchor**

**Project Code :**

**Project Advisor**:

Mr. Farooq Javed

**Project Manager**:  
 Dr. Muhammad Ilyas

**Project Team**:

Mohsin Ali Roll. No: BSCS51F21S083 (TL)

Saim Sikandar Roll. No: BSCS51F21S101

Abdul Samad Roll. No: BSCS51F21S050

**Submission Date:**

23/ Oct/ 2024

**Table of Contents**

1. Abstract
2. Background and Justification
3. Research Methodology
4. Project Scope
5. High-Level Project Plan
6. References

**1. Abstract**

This project will develop an **AI-powered bilingual virtual news anchor** that presents news in **Urdu and English** using **Retrieval-Augmented Generation (RAG)** for accurate, real-time news summaries. The virtual anchor will deliver news via a video avatar, with lifelike speech and synced lip movements, offering an innovative, scalable way to deliver multilingual news content.

**2. Background and Justification**

With the rise of digital media, there is a need for **real-time multilingual news delivery**. Existing solutions often generate inaccurate information or lack support for both Urdu and English. This project addresses this gap by using **AI and RAG** to provide **accurate, bilingual news** through a **virtual avatar**. Previous works have limited language capabilities, while this project focuses on real-time, bilingual news delivery, leveraging AI for a better user experience.

**3. Research Methodology**

1. **Data Collection**: Use APIs and web scraping for news datasets in Urdu and English.
2. **News Generation**: Implement **Llama 3.1** and **RAG** for summarizing real-time news.
3. **Text-to-Speech (TTS)**: Use **TTS models** (e.g., Coqui TTS) for bilingual speech synthesis.
4. **Avatar and Lip-Sync**: Integrate model for accurate lip-syncing with TTS output.
5. **Interface**: Build an interface for user interaction and news selection and deployment on cloud.

**4. Project Scope**

This project will:

* Cover **bilingual news generation** (Urdu/English) using AI.
* Integrate **TTS and lip-syncing** for realistic avatar presentation.
* Use **RAG** to provide **real-time news retrieval and accuracy**.

**Out of Scope**:

* Development of custom hardware or 3D avatars.
* Support for languages beyond Urdu and English in this phase.

**5. High-Level Project Plan**

| **Phase** | **Activity** |  |
| --- | --- | --- |
| **Phase 1**: TTS & Lip-Syncing  **Phase 2**: Setup and Data Collection | Implement TTS and lip-sync with Wav2Lip.  Gather datasets and set up RAG workflow. |  |
|  |  |  |
| **Phase 3**: News Generation | Fine-tune Llama 3.1 for bilingual news summaries. |  |
| **Phase 4**: UI Development & Testing | Build the Streamlit interface, conduct testing. |  |
| **Phase 5**: Final Integration | Full system testing and refinement. |  |

**6. References**

* Coqui TTS: [coqui-ai/TTS: 🐸💬 - a deep learning toolkit for Text-to-Speech, battle-tested in research and production](https://github.com/coqui-ai/TTS)
* Wav2Lip: <https://github.com/Rudrabha/Wav2Lip>
* 3d avatar: [bornfree/talking\_avatar: A nice 3D avatar that can speak input text with facial expressions](https://github.com/bornfree/talking_avatar?tab=readme-ov-file)
* [azure ai-services speech-service | Microsoft Learn](https://learn.microsoft.com/pdf?url=https%3A%2F%2Flearn.microsoft.com%2Fen-us%2Fazure%2Fai-services%2Fspeech-service%2Ftoc.json)

**Signed By Supervisor**

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**