**Business Requirements Document (BRD)**

**Project: Live Arabic-to-Urdu Video Translation & Dubbing Pipeline**

**Prepared For**: [TV Organization Name]  
**Prepared By**: [Your Name / Team Name]  
**Date**: 2025-07-17  
**Version**: 1.0

**1. 🎯 Executive Summary**

[TV Organization Name] seeks to implement a real-time solution that receives a live video feed with **Arabic audio**, translates the speech into **Urdu**, and outputs a **live video stream with dubbed Urdu audio**. The system must ensure low latency, compatibility with HLS streaming protocols, and scalability to support additional output languages in the future. Additionally, the Arabic source text and corresponding translated Urdu text should be persisted for future evaluation and training purposes.

**2. 🧩 Business Objectives**

* Enable **real-time Arabic-to-Urdu dubbing** for live broadcasts.
* Ensure **low-latency, high-availability HLS streaming** for both input and output.
* Create a **modular and scalable architecture** to add other languages easily in the future (e.g., English, Turkish, French).
* Store **intermediate translation data** (Arabic transcription and Urdu translation) for evaluation and training.
* Support **24/7 broadcasting** operations with high reliability and fault tolerance.

**3. 📡 Scope**

**3.1 In-Scope**

* Accept live **HLS video stream** with Arabic audio as input.
* Perform **automatic speech recognition (ASR)** on Arabic audio.
* Translate Arabic text to Urdu using a machine translation (MT) model.
* Convert translated Urdu text to **natural-sounding speech** via Urdu TTS.
* Replace Arabic audio with Urdu audio in the output HLS stream.
* Store transcription + translation data with timestamps for later evaluation and model retraining.
* Scalable infrastructure design to support **multi-language output in future**.

**3.2 Out of Scope**

* Evaluation UI and manual feedback collection system (covered in a separate document).
* Lip-syncing of Urdu audio with video (not required at this stage).

**4. 📥 Input and 📤 Output**

**4.1 Input Specification**

* Format: **HLS (HTTP Live Streaming)**
* Content: **Live video feed with Arabic audio**
* Codec: To be determined based on current infrastructure (e.g., H.264 video, AAC audio)

**4.2 Output Specification**

* Format: **HLS stream**
* Content: **Same video with dubbed Urdu audio** (non-lipsynced)
* Same resolution, framerate, and segment duration as the input.
* Latency: Target **< 5 seconds** end-to-end

**5. 🔧 Functional Requirements**

**5.1 Audio Processing Pipeline**

* Extract Arabic audio stream from the HLS input in real time.
* Pass audio to an **Arabic ASR module** to get transcribed Arabic text.
* Translate transcribed Arabic text into **Urdu using MT module**.
* Convert translated Urdu text to speech using **Urdu TTS engine**.
* Recombine Urdu audio with original video segments to generate output HLS.

**5.2 Storage of Intermediate Data**

* For each time-aligned segment:
  + Store Arabic transcription.
  + Store Urdu translation (MT output).
* Persist in a structured format (JSON/NoSQL) for future access by the evaluation tool.
* Include metadata (timestamps, segment ID, confidence scores).

**5.3 Stream Reassembly & Output**

* Replace original Arabic audio with generated Urdu audio (non-lipsynced).
* Ensure smooth segment transition with correct HLS manifest update.
* Maintain video quality and alignment with original visuals.

**6. 🏗️ Non-Functional Requirements**

**6.1 Performance**

* Target end-to-end latency: **< 5 seconds**
* High throughput and uptime for continuous live streaming

**6.2 Scalability**

* Ability to add support for more output languages with minimal changes
* Plug-and-play architecture for ASR, MT, and TTS components

**6.3 Reliability**

* Automatic failover or buffering in case of translation/TTS failure
* Logging and alerting for failed segments

**6.4 Security**

* Secure handling of HLS streams
* Restricted access to stored transcription/translation data

**7. 🧱 Architecture Overview (High-Level)**

text

CopyEdit

+---------------------+

| Live HLS Input |

| (Arabic Audio) |

+----------+----------+

|

+----------v----------+

| Audio Extractor |

+----------+----------+

|

+----------v----------+

| Arabic ASR Module |

+----------+----------+

|

+----------v----------+

| Arabic → Urdu MT |

+----------+----------+

|

+----------v----------+

| Urdu TTS Engine |

+----------+----------+

|

+----------v----------+

| Video + Urdu Audio |

| HLS Stream Builder |

+----------+----------+

|

+----------v----------+

| Live Urdu HLS Output|

+---------------------+

+---------------------+

| Data Logger |

| (Store transcripts) |

+---------------------+

**8. 📦 Data Model: Translation Storage Format**

Each segment is stored in the following format:

json

CopyEdit

{

"segment\_id": "seg\_000124",

"start\_time": "00:05:24.000",

"end\_time": "00:05:29.000",

"arabic\_transcript": "مرحبا بكم في هذه النشرة الإخبارية",

"urdu\_translation": "اس خبرنامے میں خوش آمدید",

"confidence\_scores": {

"asr": 0.93,

"mt": 0.88

},

"language\_pair": "ar-ur"

}

Stored in a structured database or flat file system accessible by the evaluation tool.

**9. 🧪 Assumptions and Constraints**

* Assumes availability of trained ASR, MT, and TTS models for Arabic and Urdu.
* No requirement for real-time speaker identification or diarization.
* Latency depends on ASR and TTS processing capabilities.
* Assumes current infrastructure supports HLS ingest and output.

**10. 🧭 Future Roadmap**

* Support for multiple output languages (e.g., Urdu, English, French).
* Real-time subtitles (caption generation) for accessibility.
* Optional lipsyncing with AI voice modeling.
* Integration with manual evaluation and post-editing tools.

**11. ✅ Success Criteria**

|  |  |
| --- | --- |
| **KPI** | **Target** |
| Translation Latency | ≤ 5 seconds |
| ASR Accuracy | ≥ 90% (WER) |
| Translation Accuracy | Subject to evaluation via eval tool |
| Output Stream Stability | 99.99% uptime |
| Multi-language Expansion | Plug-in support for other language pairs |