Requirements

Hardware Requirements:

- High-performance GPU servers for real-time processing.
- Scalable storage solutions, e.g., Amazon S3 or Google Cloud Storage.
- Streaming servers, e.g., NGINX with RTMP module for live stream ingestion.

Software Requirements:

- **Streaming Ingest Server**: NGINX with RTMP module, or Kinesis Video Streaming from AWS.
- **Real-time Processing**: Apache Kafka for message queuing, Apache Flink for stream processing.
- **Transcription**: Whisper from OpenAI (support for Arabic).
- **Translation**: AWS Translate.
- **Emotion Recognition**: BERT-based pre-trained model, or T-5.
- **Logo Recognition**: AWS Rekognition since it is already packaged with many famous faces, logos, and other things

Human resources Requirements:

- **Web Developers**: To build an interface for accessing and managing the live stream analytics.
- **Data Engineers**: To ensure efficient data pipeline setup and integration of various APIs and models.
- **IT Team**: To manage the infrastructure and ensure high availability and scalability.
- **Machine Learning Engineers:** Since we are going to use pre-trained models, mostly we will require a Machine Learning Engineer to make simple fine-tuning techniques, and to ensure the operationalization & Maintenance of all the models in the most efficient way.

Milestones and Time Estimates

1. **Project Initialization (2 weeks)**
    - Requirement analysis and finalization.
    - Setting up the project infrastructure.
2. **Real-time Processing Pipeline Setup (4 weeks)**
    - Setting up the NGINX RTMP server.
    - Configuring Apache Kafka and Flink.
3. **Audio Transcription Integration (3 weeks)**
    - Integrating real-time transcription using Whisper from OpenAI.
4. **Translation Integration (3 weeks)**
    - Integrating AWS Translate for real-time translation.
5. **Emotion Recognition System (3 weeks)**
    - Integrating BERT-based pre-trained model, or T-5.
6. **Logo Recognition System (4 weeks)**
    - Implementing AWS Rekognition
7. **Testing and Optimization (4 weeks)**

    - End-to-end testing of the entire pipeline.
    - Performance optimization and scalability testing.
8. **Deployment (2 weeks)**
    - Deployment to production environment.
    - Documentation and training for the client.
9. **Operationalization & Maintenance (2 weeks)**
    - Setting the needed metrics and alarms for the models
    - Apply Machine Learning Operation technique if the models degraded

Selected Technologies

- **Streaming Ingest Server**: NGINX with RTMP module, or Kinesis Video Streaming from AWS.
- **Real-time Processing**: Apache Kafka and Apache Flink.
- **Transcription**: Whisper from OpenAI.
- **Translation**: AWS Translate.
- **Emotion Recognition**: OpenCV, TensorFlow, PyTorch.
- **Logo Recognition**: AWS Rekognition, YOLO models.

Questions for the Client

1. **What are the expected peak concurrent viewers for the live stream?**
2. **Do you have any preferred cloud service provider (AWS, Google Cloud, Azure)?**
3. **Are there any specific requirements for the user interface and reporting?**
4. **What are the acceptable latencies for real-time processing and analytics?**
5. **Are there any specific logos that need to be prioritized for recognition?**
6. **Do you require multi-language support for the emotion recognition output?**

Pre-trained Models and Custom Training

- **Transcription**: Utilize Google Cloud Speech-to-Text or Deepgram API, which supports Arabic.
- **Translation**: Use Google Cloud Translation API or AWS Translate for multilingual support.
- **Emotion Recognition**: BERT-based pre-trained model, or T-5.
- **Logo Recognition**: AWS Rekognition.