Project Title: EchoMe AI

Real-Time Personalized Virtual Assistant with AI-Generated Avatar Based on Real People

Submitted by:

Team: Abdelrahman Mohamed, Ahmed ElSayed, Beshoy Emad, Beshoy Gamal, Nada Ahmed, Sama Mohsen
As part of the **Digital Egypt Pioneers** – Generative AI Track

1. Project Description

This project aims to develop **EchoMe AI**, a real-time personalized virtual assistant that replicates the **appearance, voice, personality, and conversational style** of real individuals — including users themselves. Unlike generic assistants such as Siri or Alexa, our system delivers **true human-like interaction through live video**, enabling emotional connection, cultural relevance, and personal identity in AI.

Users can upload:

- A **photo** (for visual appearance),
- A **voice sample** (to clone tone and accent),
- And short **text or recorded speech** (to define knowledge and speaking style).

From these inputs, the system generates an **interactive AI character** that appears as a **talking head in real-time video format**, listens via speech-to-text, processes queries using a fine-tuned LLM, and responds naturally — not just with words, but with **lipsynced facial expressions, gestures, and authentic vocal patterns**.

The assistant will be able to **answer questions, explain concepts, tell stories, and interact conversationally**, behaving like a digital twin of the user or any chosen persona. With strong support for **Modern Standard Arabic and regional dialects**, this solution empowers education, content creation, elderly companionship, and customer service across Arabic-speaking communities.

2. Group Members & Roles

Name	Role	Responsibilities
Sama Mohsen	**Team Leader – Project Coordination & Deployment Supervision**	Oversees all development phases; coordinates integration between modules; ensures timeline adherence; leads final deployment, testing, documentation, and presentation delivery.
Beshoy Gamal	**Voice Cloning & Audio Engineer**	Collects and preprocesses voice datasets; implements voice cloning using ElevenLabs/Coqui TTS; generates high-quality, expressive speech output; handles noise reduction, normalization, and format conversion.
Abdelrahman Mohamed	**Avatar Designer (2D/3D & Animation)**	Generates realistic 2D head-and-face avatars from photos; integrates D-ID API or Wav2Lip + GANs for precise lip-sync and subtle facial animation; ensures natural-looking video output synchronized with generated audio.
Nada Ahmed	**LLM Developer & Personality Modeling**	Implements and fine-tunes multilingual LLMs (e.g., AraBERT, Zephyr) to emulate specific personalities; designs prompts based on user-provided texts; evaluates response fluency, accuracy, and cultural appropriateness in Arabic and English.
Ahmed ElSayed	**Backend Developer & API Integration**	Builds RESTful endpoints in Flask; manages data flow between components; handles file uploads, session logic, background tasks, and error handling across services.
Beshoy Emad	**Backend Developer & API Integration**	Develops core Flask application structure; orchestrates external APIs (Hugging Face, Whisper, ElevenLabs, D-ID); ensures scalability, performance, and secure handling of secrets and media files.

3. Team Leader

Sama Mohsen

4. Objectives

1. **Build a Real-Time Interactive Avatar Engine**: Create a backend system that takes image, voice, and text inputs and returns a video of an AI-generated person responding intelligently.
2. **Enable Identity Replication**: Allow users to create a digital version of themselves or others — visually and vocally accurate.
3. **Support Multilingual Conversations**: Deliver fluent, dialect-aware responses in both **Arabic (MSA + common dialects)** and **English** using advanced NLP models.
4. **Achieve Natural Human-Like Interaction**: Generate emotionally expressive, context-aware replies with synchronized lip movements and facial cues.
5. **Ensure Ethical & Secure Design**: Implement consent mechanisms, watermarking, temporary storage, and data encryption to prevent misuse and deepfake abuse.

5. Tools & Technologies

Component	Technology Used
Backend Framework	Flask (Python)
Large Language Model (LLM)	Hugging Face Transformers: `AraBERT`, `Jais`, `Zephyr-7B`, `Phi-3-mini`
Voice Cloning	ElevenLabs API (high-fidelity, emotion-preserving), Coqui TTS (open-source alternative)
Avatar Generation	D-ID API / Wav2Lip + First Order Motion Model (FOMM) for lipsync and facial animation
Speech-to-Text (STT)	OpenAI Whisper (supports Arabic & English, robust transcription)
Text-to-Speech (TTS)	ElevenLabs (supports Arabic, dynamic prosody control)
Multilingual NLP	AraBERT, Camel-AI, langdetect, custom dialect datasets

Component	Technology Used
Data Processing	Pandas, NumPy, Librosa (audio), Pillow (image)
API Testing	Postman, cURL
Deployment	Docker, Git, Render.com or Hugging Face Spaces
Development Environment	VS Code / Jupyter Notebooks, Python 3.10+, `.env` for secrets

6. Milestones & Deadlines

Milestone	Description	Deadline
M0: Proposal Submission	Finalize and submit the project proposal document including team roles, objectives, tools, and initial plan	**October 12, 2025**
M1: Setup & Planning	Finalize roles, define scope, set up environment, create GitHub repo	October 15, 2025
M2: Flask Backend Core	Initialize Flask app, routing, config, logging, error handling	October 20, 2025
M3: LLM Integration	Integrate Arabic/English LLMs for generating responses via API endpoint	October 25, 2025
M4: Voice Pipeline (STT/TTS)	Add Whisper (STT) and ElevenLabs (TTS); test Arabic synthesis and voice cloning	October 28, 2025
M5: Avatar Generation Module	Connect D-ID/Wav2Lip to generate talking head videos from photo + audio	November 2, 2025
M6: Full API Integration	Chain modules into one workflow: POST → LLM → TTS → Avatar → Return URL	November 5, 2025
M7: Testing & Optimization	Fix bugs, optimize speed, evaluate Arabic accuracy, add error handling	November 8, 2025

Milestone	Description	Deadline
M8: Deployment & Documentation	Deploy online, write README, document API endpoints, record demo	November 10, 2025
M9: Final Presentation Preparation	Rehearse presentation, finalize slides, prepare live demo	November 25, 2025
M10: Project Completion & Final Submission	Submit all deliverables, ensure deployment stability, complete documentation	**November 30, 2025**

7. Key Performance Indicators (KPIs)

KPI	Customized Measurement for Our Project
1. Data Preparation Quality	Ensure uploaded images/audio/text are properly preprocessed (resize, normalize, clean). Validate tokenization for Arabic dialects using custom dataset sampling.
2. Model Performance & Accuracy	Evaluate LLM outputs using ROUGE/L for Arabic fluency; use human evaluation (1–5 scale) for realism of voice/avatar. Target ≥ 4.0 average score.
3. Pipeline Integration & Automation Level	End-to-end automation: User submits → system processes → returns video without manual steps. All APIs connected via Flask routes.
4. MLOps & Deployment Readiness	App containerized with Docker, hosted online (Render/HF Spaces), version-controlled on GitHub, secrets managed via `.env`.
5. Output Quality & Usability Score	Conduct internal user testing: “How natural was the conversation?” and “Did the avatar resemble the input?” Target ≥ 4/5 satisfaction.
6. Documentation & Presentation Quality	Complete documentation (README, comments), clear slide deck, recorded demo video, professional delivery during final presentation.