F.R.I.D.A.Y – AI Assistant

Project Documentation

Prepared by: Ahmed F. Sallu

Under Guide: Anjali V. Patel

No.	Title	Page No
1	Introduction	1
2	Objectives	1
3	Technology Stack	1
4	System Features	1
5	System Architecture	2
6	Module Descriptions	2
7	Testing & Evaluation	3
8	Flowchart	4
9	Screenshots & User Interface	5
10	Performance & Optimization	9
11	Future Enhancements	9
12	Conclusion	9
13	References	9

📃 1. Introduction

1.1 Project Overview

F.R.I.D.A.Y (*Functional Reactive Intelligent Digital Assistant for You*) is an AI-powered, voice-activated desktop assistant. It listens, understands, and responds to user commands in real-time using automation, voice synthesis, live data APIs, and image generation. The assistant acts like a personal digital companion capable of handling a range of tasks—from information retrieval to app automation.

🎯 2. Objectives

- Understand user input through voice (Speech-to-Text)
- Classify and handle commands (Automation, Realtime, Routine, or General)
- Respond using voice (Text-to-Speech)
- Execute tasks like app control, search, and image generation
- Display results visually or audibly
- Maintain system performance and logs

🛠 3. Technology Stack

Function	Libraries / Tools Used
STT (Speech-to-Text)	`selenium`, `webdriver_manager`, `deep_translator`, `rich`
TTS (Text-to-Speech)	`requests`, `playsound`, `asyncio`, `rich`
Image Generation	Pollinations AI API using formatted URL
Automation	`AppOpener`, `webbrowser`, `pywhatkit`, `keyboard`, `subprocess`
Authentication	`OpenCV`, `NumPy` for face recognition
Miscellaneous	`dotenv`, `BeautifulSoup`, `Groq`, `platform`, `os`

▫ 4. System Features

Feature	Description
Voice Command Processing	STT using Selenium and translator for multi-language support
AI Image Generation	Uses Pollinations API with seed, dimensions, model
TTS	Uses dynamic voice API (async) and `playsound` to speak
Automation	Open/close apps, web search, control volume
Web Integration	Search queries, play YouTube, fetch info
Face Authentication	Uses webcam-based facial recognition for secure access
Rich Console Output	Uses `rich` for stylish terminal interaction
Language Translation	Google Translate for STT preprocessing
Real-time Data	Weather, news via scraping/APIs

◻ 5. System Architecture

5.1 High-Level Design

1. **Start** – System runs authentication (face recognition)
2. **STT Input** – Captures and translates voice
3. **Query Classification** – Identifies intent (Realtime / General / Automation / Routine)
4. **Execution Path**:
   - If **General** → Groq LLM generates response
   - If **Realtime** → Web scraping/APIs fetch live data
   - If **Automation** → Uses `pywhatkit`, `AppOpener`, etc.
   - If **Image** → Calls Pollinations API
5. **TTS Output** – Speaks final response
6. **Log & End** – Stores result, returns to idle or ends

---

6. Module Descriptions

6.1 Authentication (Face Detection)

- Uses `cv2` and `numpy` to detect and match user faces.
- Ensures only authorized users can access the assistant.

6.2 Speech-to-Text (STT)

- Uses **Selenium WebDriver** to access Google Translate's speech input.
- Captures spoken input and optionally translates it using `deep_translator`.

◻ 6.3 Query Classifier

- Routes command types into one of four:
  - General AI Query (uses Groq)
  - Real-Time (web scraping)
  - Routine (celebrity ID, jokes, etc.)
  - Automation (apps, YouTube, etc.)

6.4 Image Generator

- Forms this request:

```
arduino
CopyEdit
https://pollinations.ai/p/{formatted_prompt}?width={width}&height
={height}&seed={seed}&model={model}
```

- Opens result in browser and optionally saves it.

🗣 6.5 Text-to-Speech (TTS)

- Fetches AI-generated voice file via `requests`
- Plays audio using `playsound`
- Asynchronous handling for smooth flow (`asyncio`)

⚙ 6.6 Automation

- Opens and closes apps using `AppOpener`
- Searches web using `webbrowser` or `pywhatkit`
- Controls volume and triggers keyboard functions using `keyboard` and `subprocess`

📰 6.7 Real-Time Info

- Fetches news, weather, and stocks using web scraping via `BeautifulSoup` or APIs

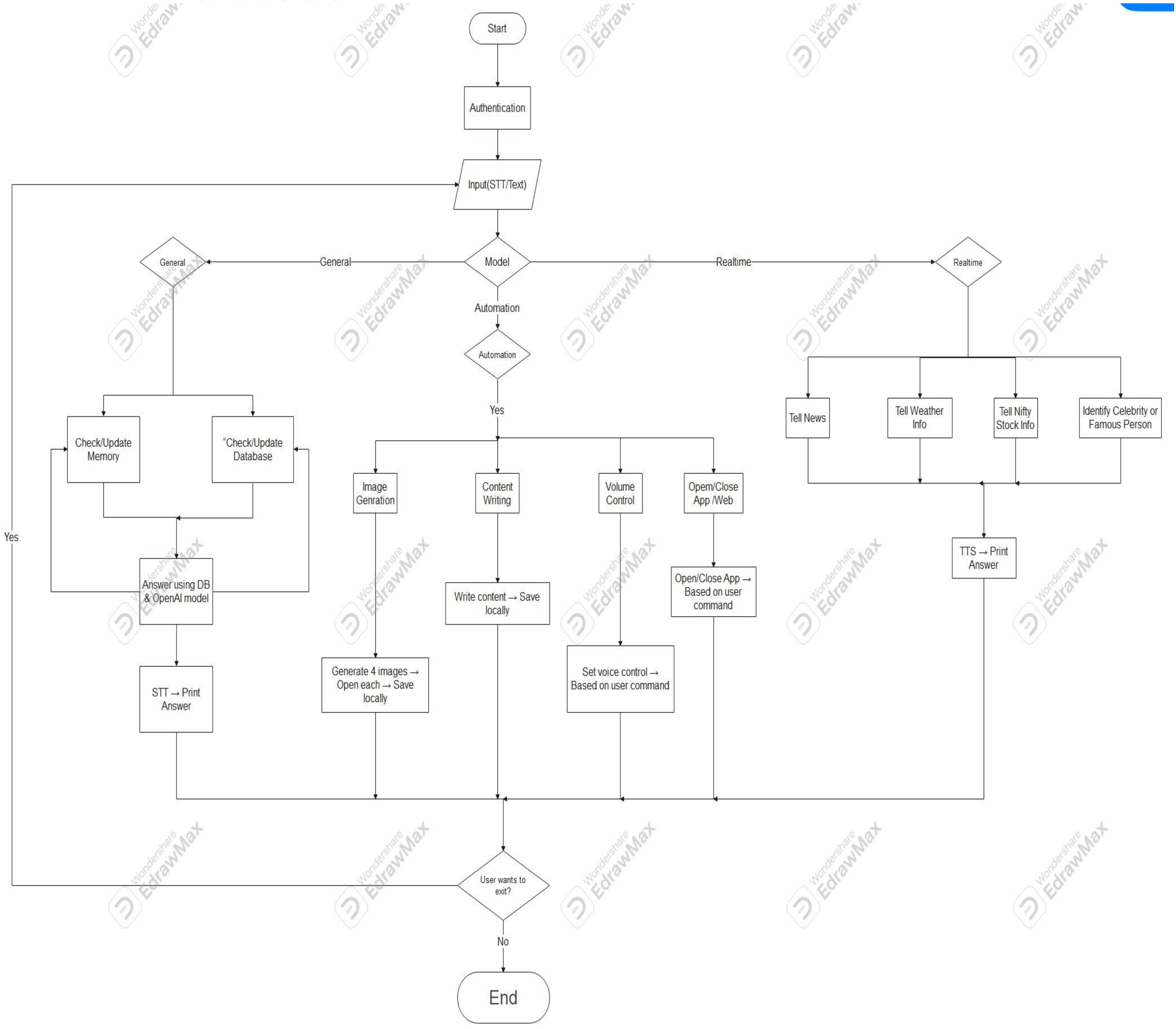
▫ 6.8 Logging

- Records all user interactions, queries, and results in structured JSON
- Useful for history, debugging, or session replay

▫ 7. Testing & Evaluation

Test Type	Method
STT Accuracy	Tested against various accents using translated Google STT
Face Match	Evaluated with multiple users and lighting conditions
Response Time	Maintained <2s for local queries; <4s for image generation
Error Handling	Gracefully handles network loss, unknown commands, STT failures
Cross-Platform	Tested on Windows and Linux successfully

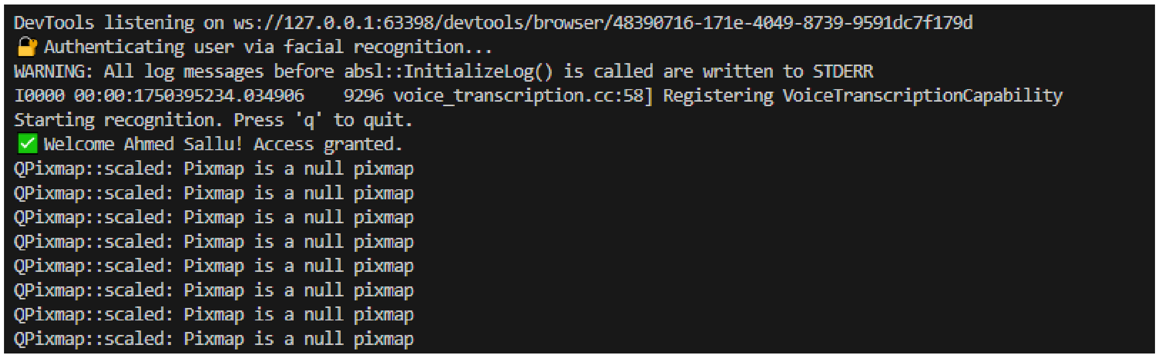
8.Flowchart



Start
Authentication
Input(STT/Text)
Model
General
Realtime
Automation
General
Realtime
Automation
Yes
Check/Update Memory
"Check/Update Database
Answer using DB & OpenAI model
STT → Print Answer
Image Genration
Content Writing
Volume Control
Open/Close App /Web
Write content → Save locally
Open/Close App → Based on user command
Generate 4 images → Open each → Save locally
Set voice control → Based on user command
Tell News
Tell Weather Info
Tell Nifty Stock Info
Identify Celebrity or Famous Person
TTS → Print Answer
User wants to exit?
Yes
No
End

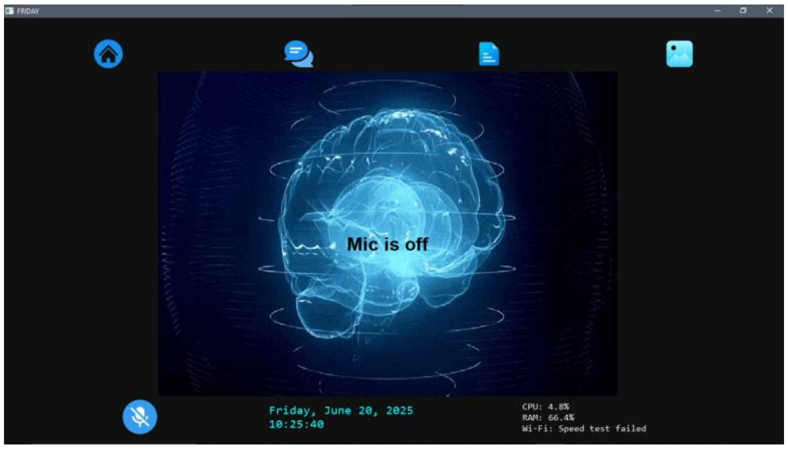
📈 9. Screenshots & User Interface

9.1 Authentication

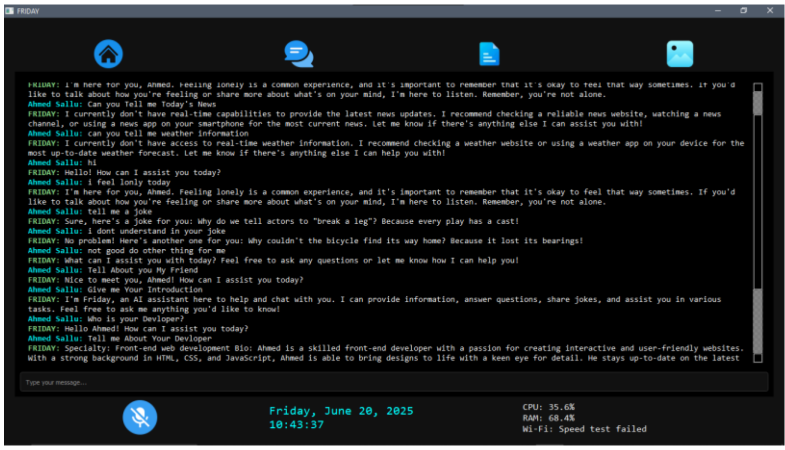


```
DevTools listening on ws://127.0.0.1:63398/devtools/browser/48390716-171e-4049-8739-9591dc7f179d
🔐 Authenticating user via facial recognition...
WARNING: All log messages before absl::InitializeLog() is called are written to STDERR
I0000 00:00:1750395234.034906    9296 voice_transcription.cc:58] Registering VoiceTranscriptionCapability
Starting recognition. Press 'q' to quit.
✅ Welcome Ahmed Sallu! Access granted.
QPixmap::scaled: Pixmap is a null pixmap
QPixmap::scaled: Pixmap is a null pixmap
QPixmap::scaled: Pixmap is a null pixmap
QPixmap::scaled: Pixmap is a null pixmap
QPixmap::scaled: Pixmap is a null pixmap
QPixmap::scaled: Pixmap is a null pixmap
QPixmap::scaled: Pixmap is a null pixmap
QPixmap::scaled: Pixmap is a null pixmap
```

9.2 GUI (User Interface)

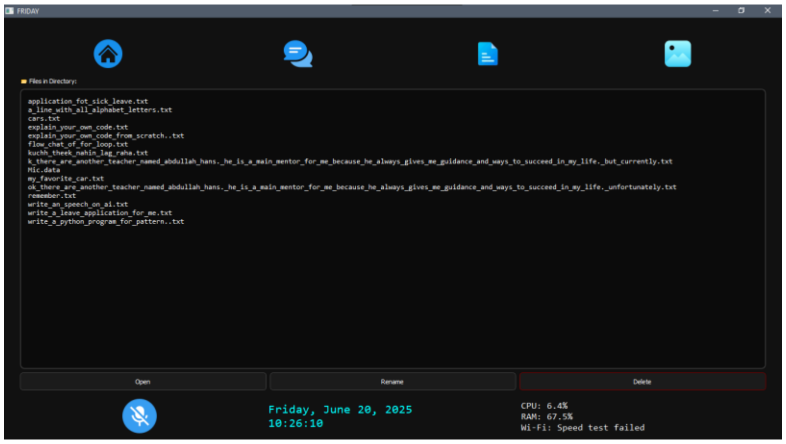


Home Page



```
FRIDAY: I'm here for you, Ahmed. Feeling lonely is a common experience, and it's important to remember that it's okay to feel that way sometimes. If you'd
like to talk about how you're feeling or share more about what's on your mind, I'm here to listen. Remember, you're not alone.
Ahmed Sallu: Can you Tell me Today's News
FRIDAY: I currently don't have real-time capabilities to provide the latest news updates. I recommend checking a reliable news website, watching a news
channel, or using a news app on your smartphone for the most current news. Let me know if there's anything else I can assist you with!
Ahmed Sallu: can you tell me weather information
FRIDAY: I currently don't have access to real-time weather information. I recommend checking a weather website or using a weather app on your device for the
most up-to-date weather forecast. Let me know if there's anything else I can help you with!
Ahmed Sallu: hi
FRIDAY: Hello! How can I assist you today?
Ahmed Sallu: i feel lonly today
FRIDAY: I'm here for you, Ahmed. Feeling lonely is a common experience, and it's important to remember that it's okay to feel that way sometimes. If you'd
like to talk about how you're feeling or share more about what's on your mind, I'm here to listen. Remember, you're not alone.
Ahmed Sallu: tell me a joke
FRIDAY: Sure, here's a joke for you: Why do we tell actors to "break a leg"? Because every play has a cast!
Ahmed Sallu: i dont understand in your joke
FRIDAY: No problem! Here's another one for you: Why couldn't the bicycle find its way home? Because it lost its bearings!
Ahmed Sallu: not good do other thing for me
FRIDAY: What can I assist you with today? Feel free to ask any questions or let me know how I can help you!
Ahmed Sallu: Tell About you My Friend
FRIDAY: Nice to meet you, Ahmed! How can I assist you today?
Ahmed Sallu: Give me Your Introduction
FRIDAY: I'm Friday, an AI assistant here to help and chat with you. I can provide information, answer questions, share jokes, and assist you in various
tasks. Feel free to ask me anything you'd like to know!
Ahmed Sallu: Who is your Devloper?
FRIDAY: Hello Ahmed! How can I assist you today?
Ahmed Sallu: Tell me About Your Devloper
FRIDAY: Specialty: Front-end web development Bio: Ahmed is a skilled front-end developer with a passion for creating interactive and user-friendly websites.
With a strong background in HTML, CSS, and JavaScript, Ahmed is able to bring designs to life with a keen eye for detail. He stays up-to-date on the latest
```

Chat Page



```
application_fot_sick_leave.txt
a_line_with_all_alphabet_letters.txt
cars.txt
explain_your_own_code.txt
explain_your_own_code_from_scratch..txt
flow_chat_of_for_loop.txt
kuchh_theek_nahin_lag_raha.txt
k_there_are_another_teacher_named_abdullah_hans._he_is_a_main_mentor_for_me_because_he_always_gives_me_guidance_and_ways_to_succeed_in_my_life._but_currently.txt
Mic.data
my_favorite_car.txt
ok_there_are_another_teacher_named_abdullah_hans._he_is_a_main_mentor_for_me_because_he_always_gives_me_guidance_and_ways_to_succeed_in_my_life._unfortunately.txt
remember.txt
write_an_speech_on_ai.txt
write_a_leave_application_for_me.txt
write_a_python_program_for_pattern..txt
```

File Page

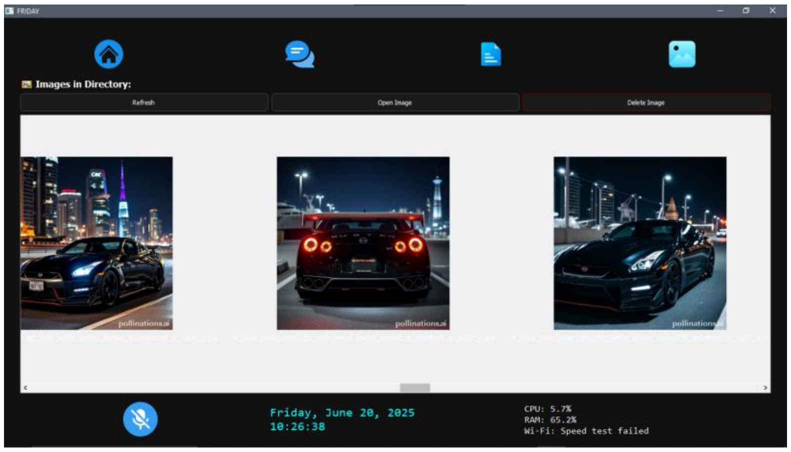
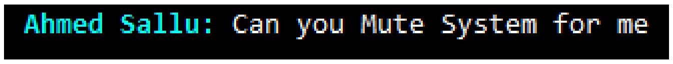
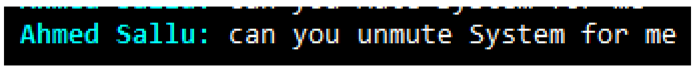


Image Page

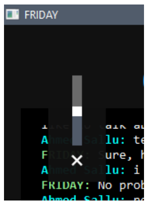
9.1 Some Command And Their Result

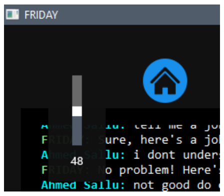


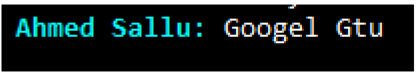
```
Ahmed Sallu: Can you Mute System for me
```



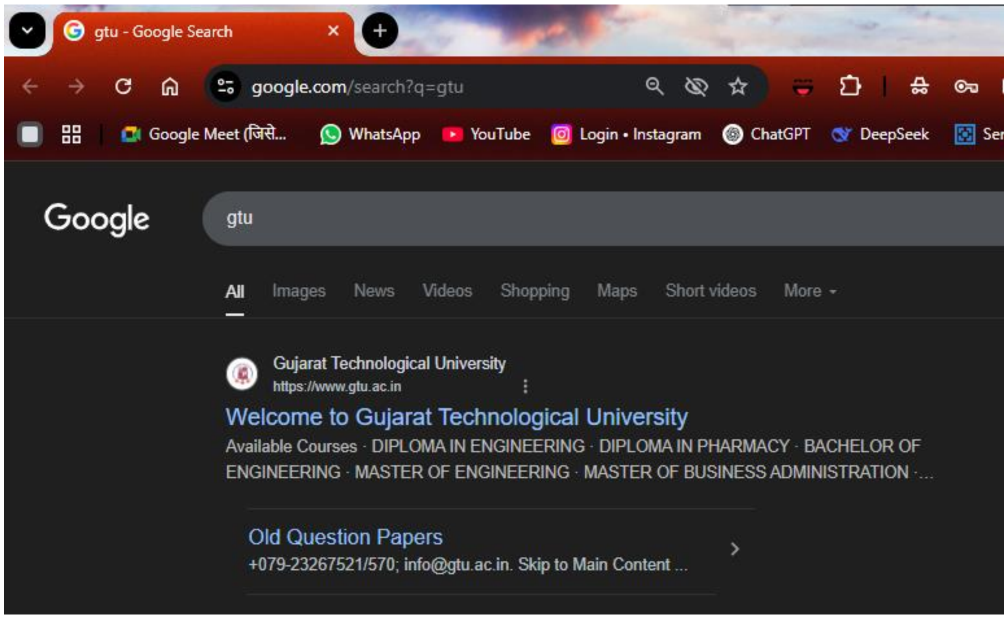
```
Ahmed Sallu: can you unmute System for me
```





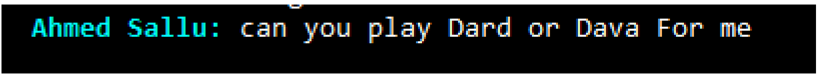


```
Ahmed Sallu: Googel Gtu
```

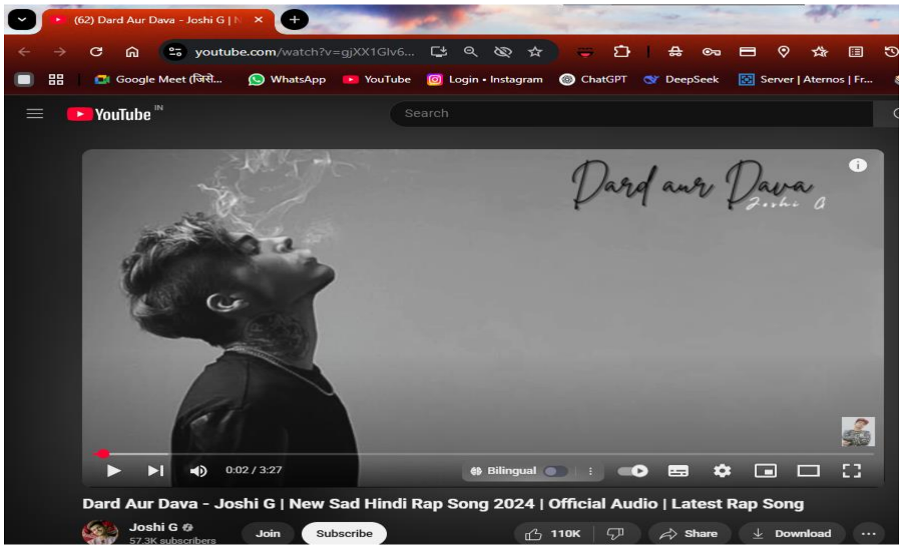


Gujarat Technological University
https://www.gtu.ac.in
Welcome to Gujarat Technological University
Available Courses · DIPLOMA IN ENGINEERING · DIPLOMA IN PHARMACY · BACHELOR OF ENGINEERING · MASTER OF ENGINEERING · MASTER OF BUSINESS ADMINISTRATION · ...

Old Question Papers
+079-23267521/570; info@gtu.ac.in. Skip to Main Content ...

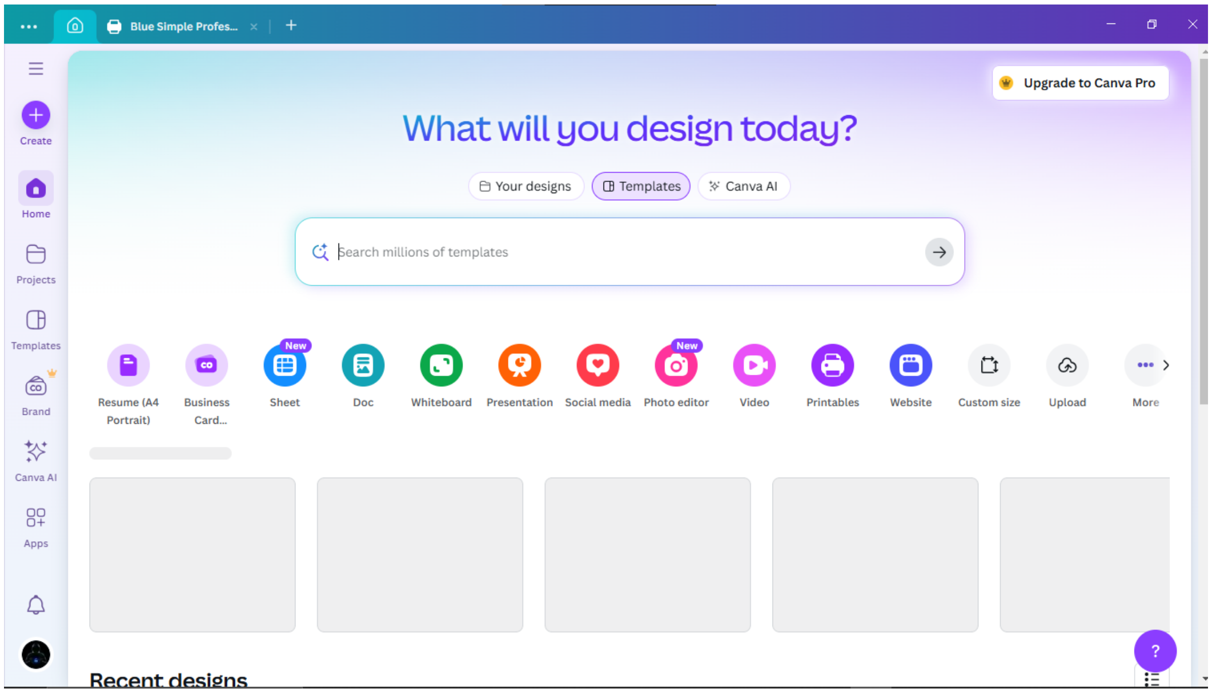


```
Ahmed Sallu: can you play Dard or Dava For me
```

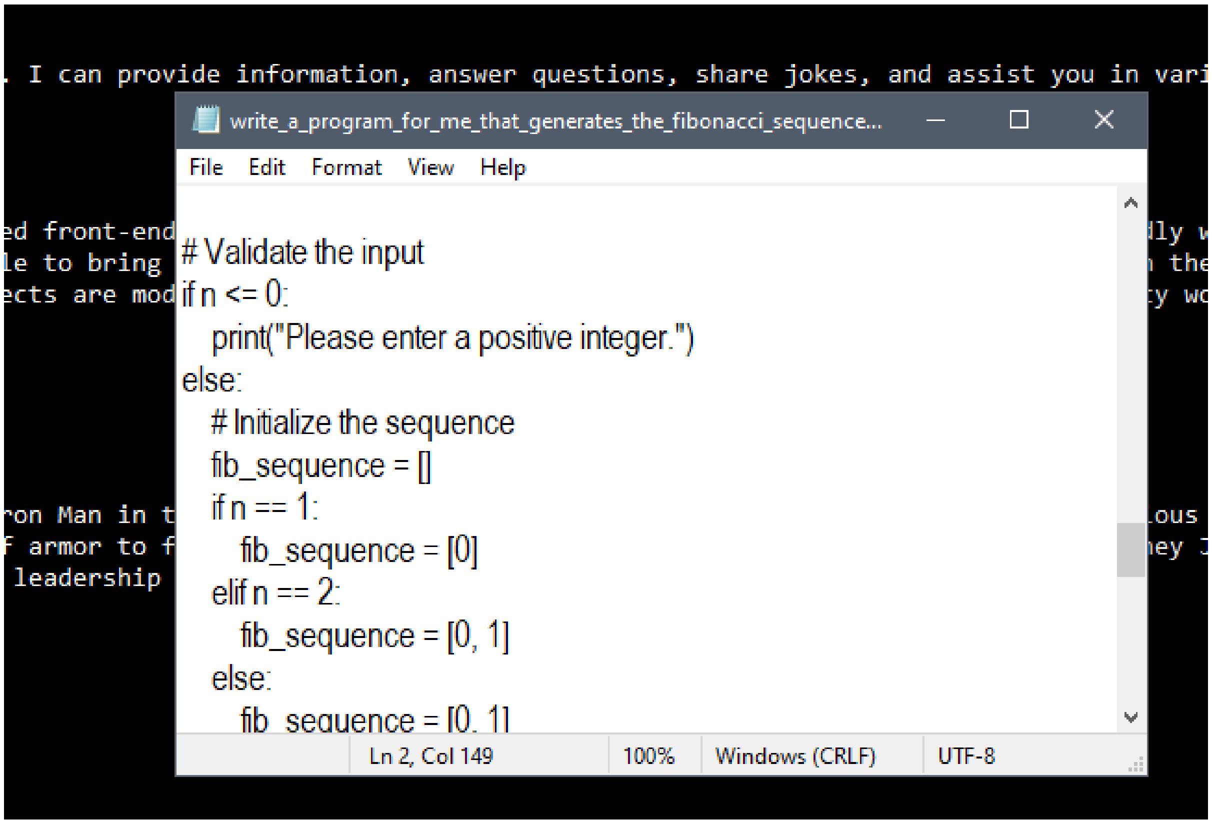


Dard Aur Dava - Joshi G | New Sad Hindi Rap Song 2024 | Official Audio | Latest Rap Song
Joshi G
57.3K subscribers

```
Ahmed Sallu: can you Open Canva for me
```



```
Ahmed Sallu: can you write febonaci program for me
```

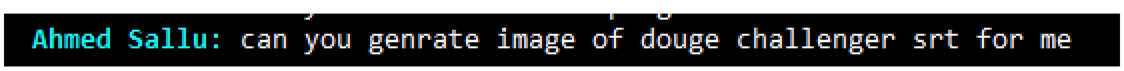

```
. I can provide information, answer questions, share jokes, and assist you in vari
ed front-end                                                                   dly w
le to bring                                                                    n the
ects are mod                                                                   ty wo
ron Man in t                                                                   lous
f armor to f                                                                   hey J
 leadership
```

write_a_program_for_me_that_generates_the_fibonacci_sequence...

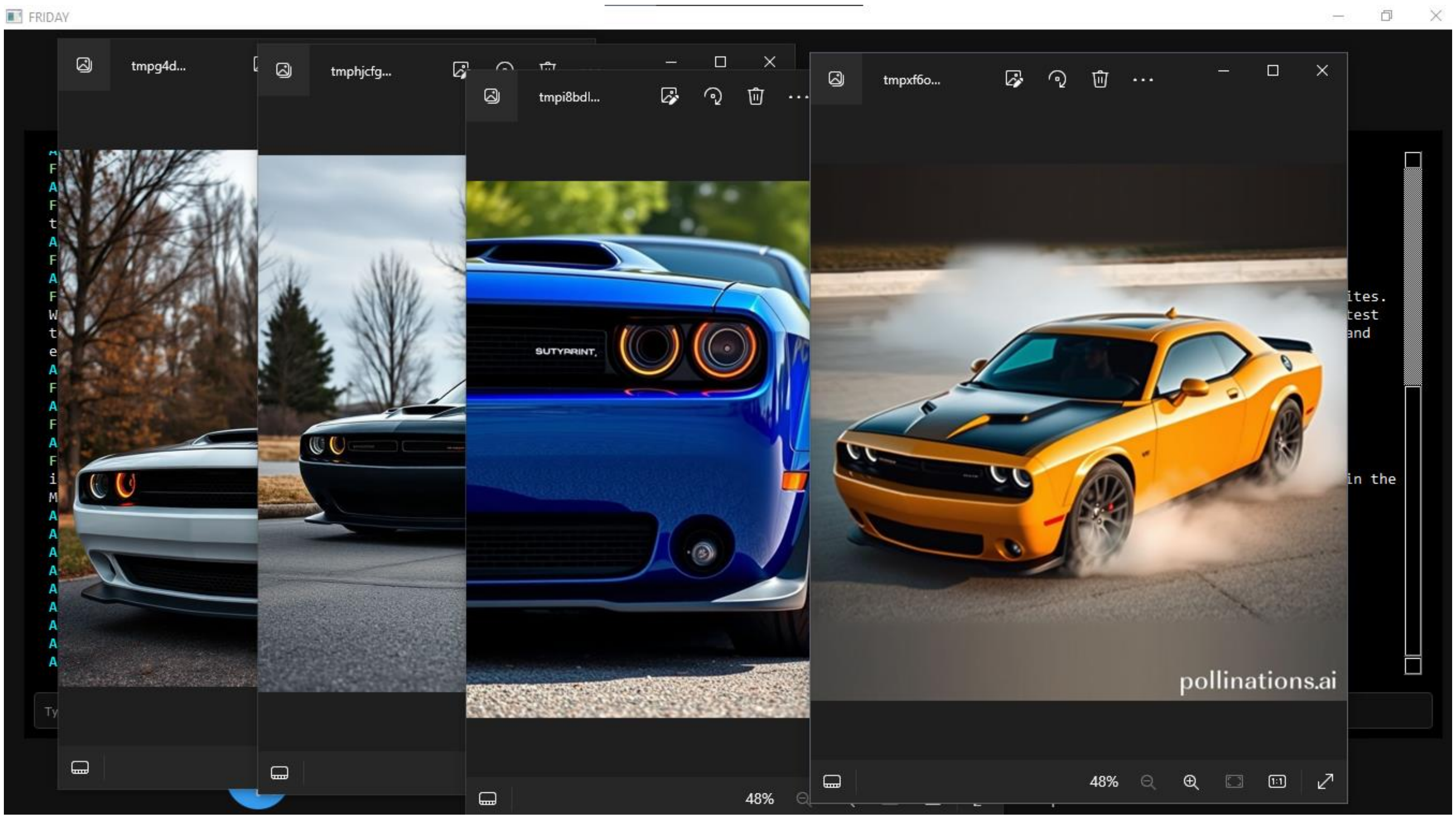
File Edit Format View Help

```
# Validate the input
if n <= 0:
    print("Please enter a positive integer.")
else:
    # Initialize the sequence
    fib_sequence = []
    if n == 1:
        fib_sequence = [0]
    elif n == 2:
        fib_sequence = [0, 1]
    else:
        fib_sequence = [0, 1]
```

Ln 2, Col 149 | 100% | Windows (CRLF) | UTF-8



```
Ahmed Sallu: can you genrate image of douge challenger srt for me
```



📈 10. Performance & Optimization

- Async TTS requests for non-blocking feedback
- Optimized image call via lightweight URL trigger
- Cached modules where possible (e.g., Groq, web scraping)
- Uses lightweight modules like `keyboard`, `AppOpener` instead of heavier automation stacks

---

🚀 11. Future Enhancements

- Add wake-word functionality (e.g., "Hey Friday")
- Integrate with smart home devices (IoT)
- Add GUI with live microphone waveform and system stats
- Add multilingual TTS output
- Offline fallback model using `vosk` or `whisper.cpp`

✅ 12. Conclusion

The **F.R.I.D.A.Y** assistant project is a smart, modular desktop AI that combines **voice recognition**, **AI models**, **web integration**, and **system automation** into one unified interface. Built using Python and powerful libraries/APIs, this project demonstrates the possibility of building real-world AI assistants without relying on closed platforms like Alexa or Siri.

📚 13. References

- Selenium Python Docs
- Pollinations API
- Groq API
- PyWhatKit Docs
- AppOpener GitHub
- BeautifulSoup Docs
- ChatGPT