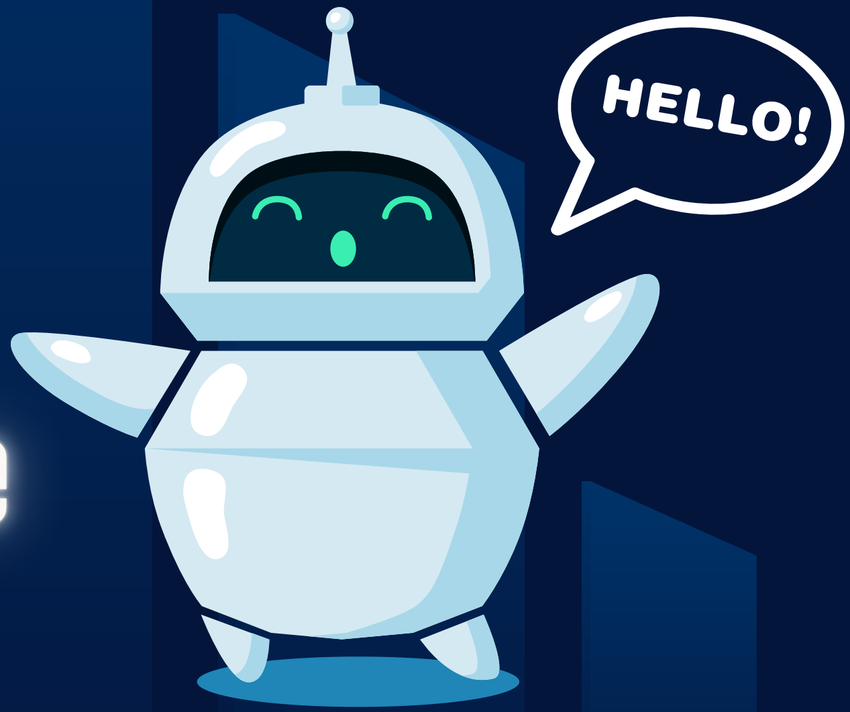
remoteBase.

@TeamAutonoMinds

Enhance Browsing Through AutoBrowse Agent



Team Members

Danish Karim

Usama Faheem Ahmed

Zehra Arshad

Muhammad Kashif Abdullah

The Problem

1. Most existing browser automation tool require programming knowledge and manual scripting, making them difficult for non-technical users.

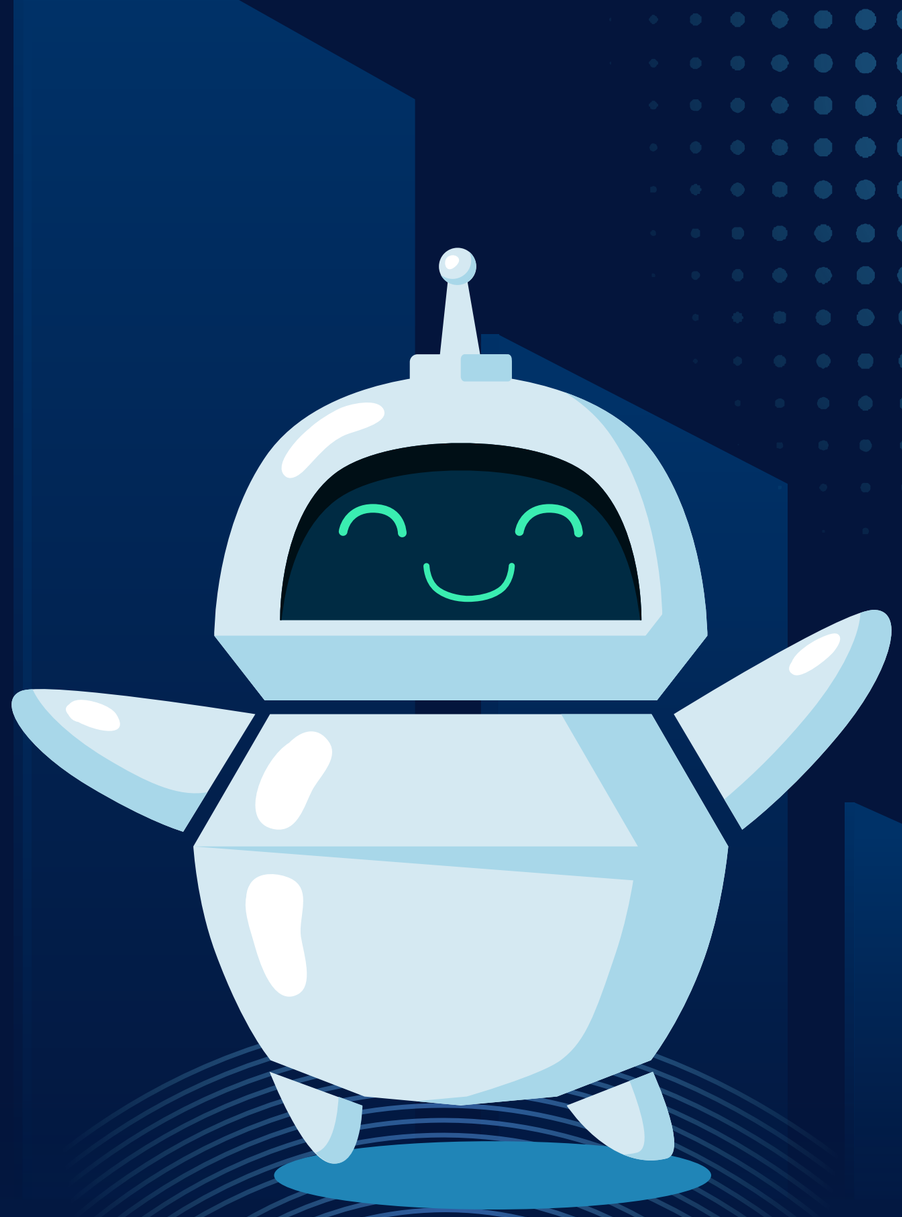
2. Current automation systems often lack natural, conversational interfaces, limiting their usability for everyday users.

3. Users want to give high-level, natural-language instructions rather than precise commands

4. Users often struggle to interact with their computers when input devices like the touchpad or keyboard fail

Introducing AutoBrowse

- AutoBrowse is a voice-controlled browser automation assistant designed to make web interaction seamless and hands-free.
- By understanding natural spoken language and converting it into intelligent browser actions.
- AutoBrowse empowers users to navigate the internet, perform complex tasks, and interact with web content using only their voice.
- This creates a more intuitive, accessible, and user-friendly way to browse — whether for convenience, accessibility, or as a backup when hardware fails.



Key Features

NATURAL LANUAGE UNDERSTAND

Enables users to control the browser using natural voice commands

UNDERSTAND COMPLEX INSTRUCTIONS

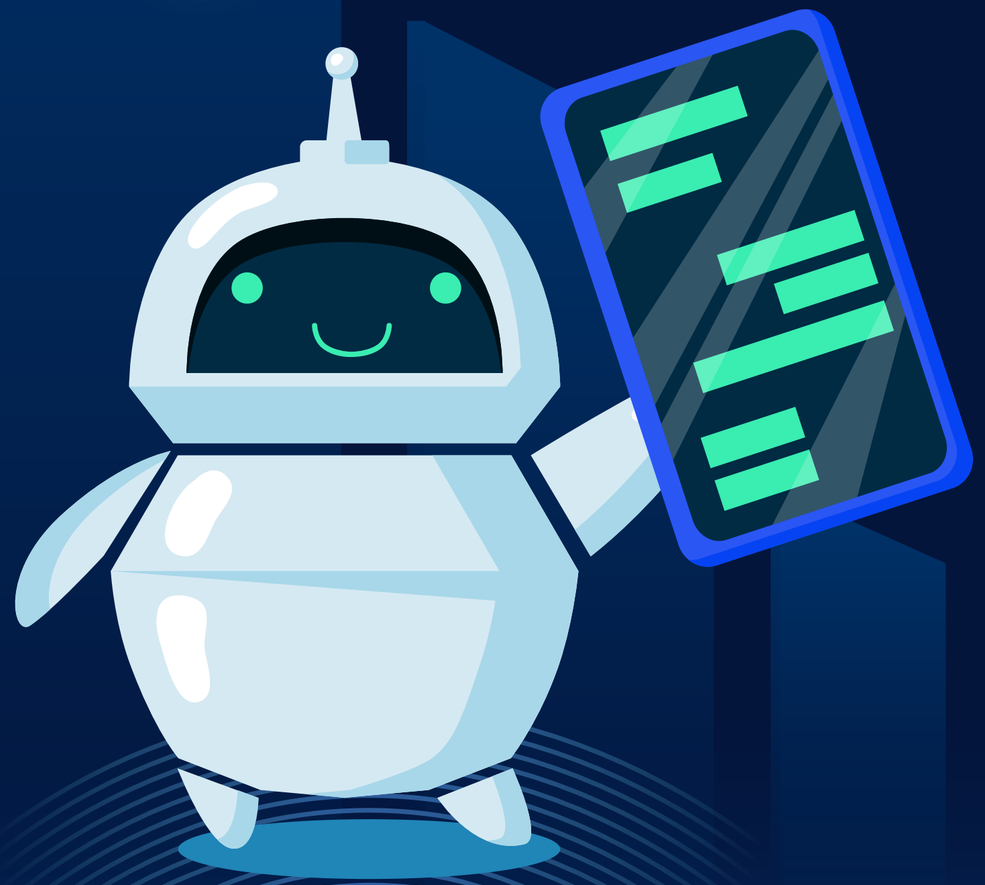
Understands complex instructions and interprets user intent intelligently

WEB TASKS

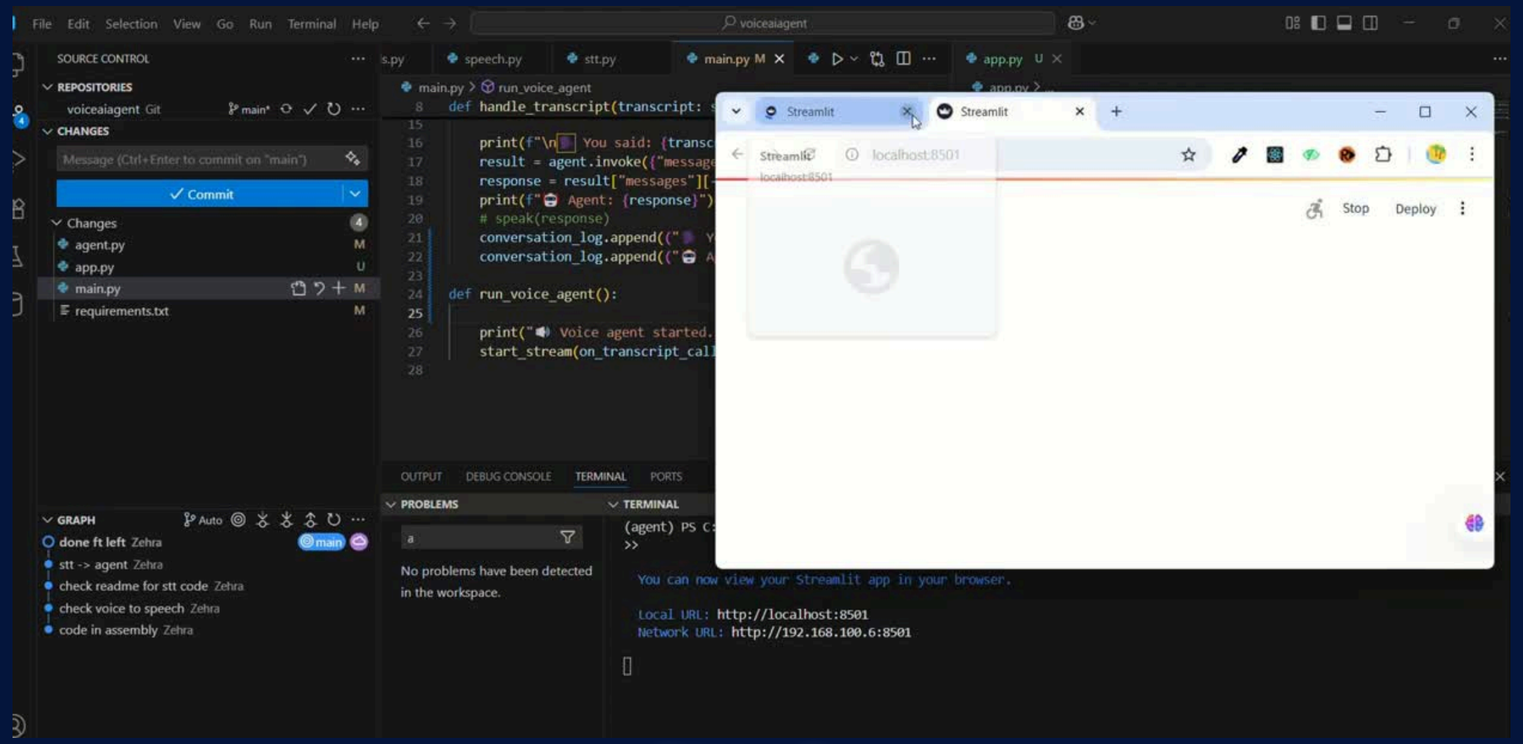
Automates web tasks like clicking, searching, and navigating pages

HAND FREE CONTROL

Designed for hands-free, accessible web control, useful even when input devices are unavailable

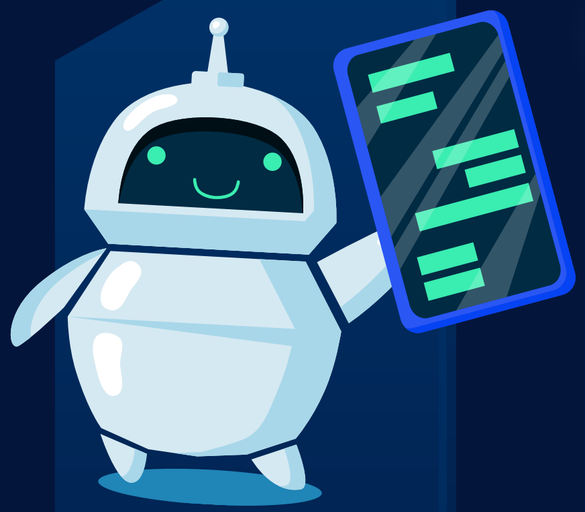


Demo



GitHubRepo:

https://github.com/ZehraArshad/Voice_Agent_Hackfest



Technologies used

FRONTEND

Streamlit

BACKEND

Selenium

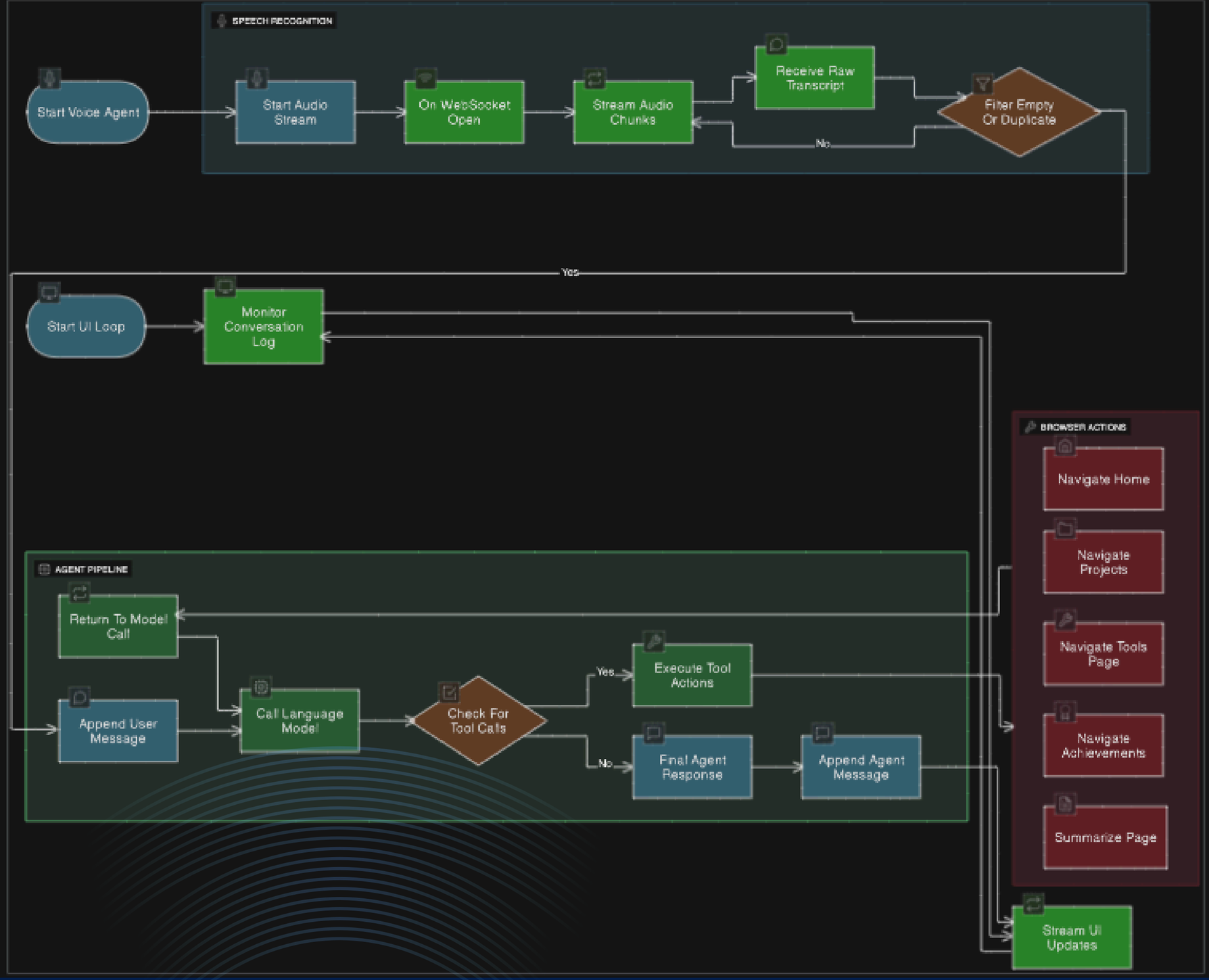
AI/LLM & DEPLOYMENT TOOLS

groq

AssemblyAI

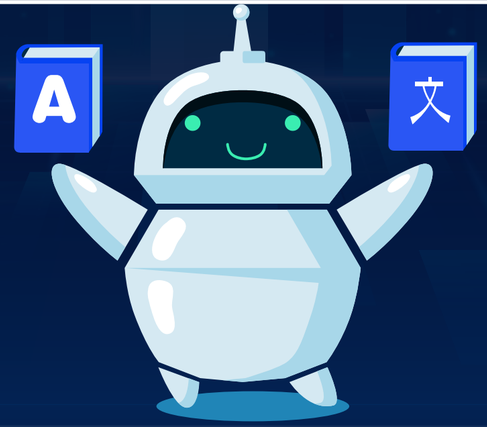
LangChain

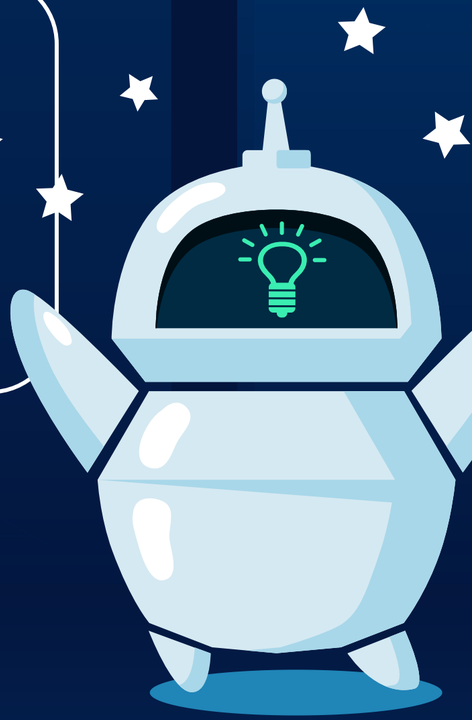
Workflow


SPEECH RECOGNITION
Start Voice Agent
Start Audio Stream
On WebSocket Open
Stream Audio Chunks
Receive Raw Transcript
Filter Empty Or Duplicate
No
Yes
Start UI Loop
Monitor Conversation Log
AGENT PIPELINE
Return To Model Call
Append User Message
Call Language Model
Check For Tool Calls
Yes
Execute Tool Actions
No
Final Agent Response
Append Agent Message
BROWSER ACTIONS
Navigate Home
Navigate Projects
Navigate Tools Page
Navigate Achievements
Summarize Page
Stream UI Updates

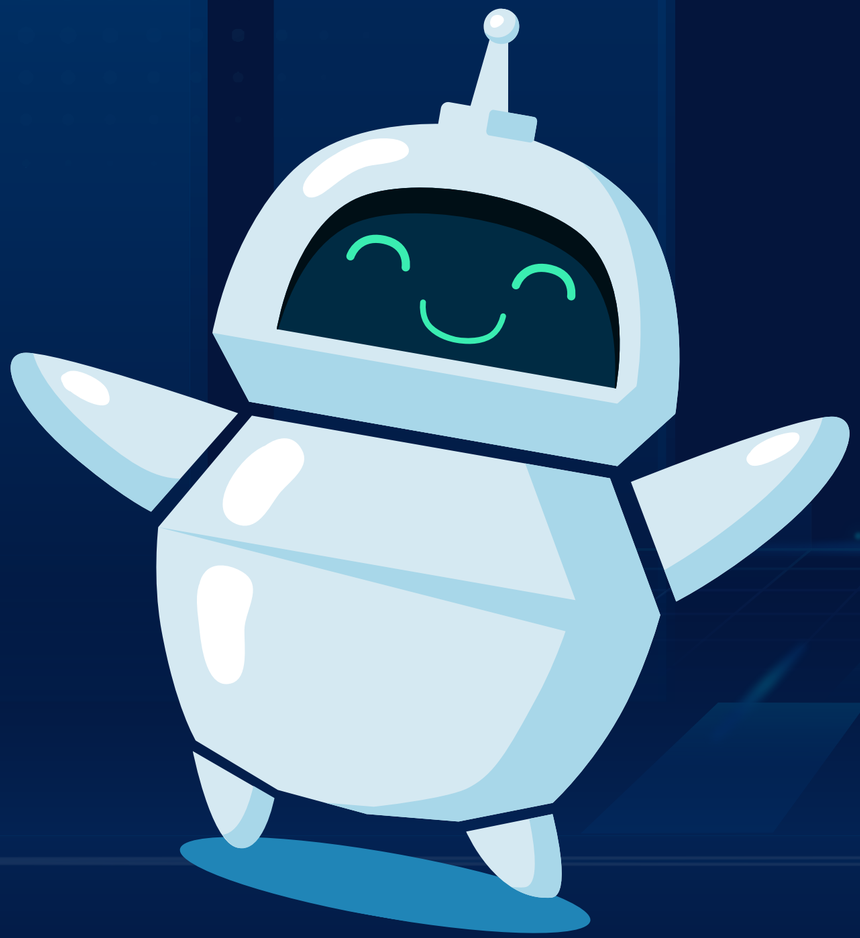
Conclusion:

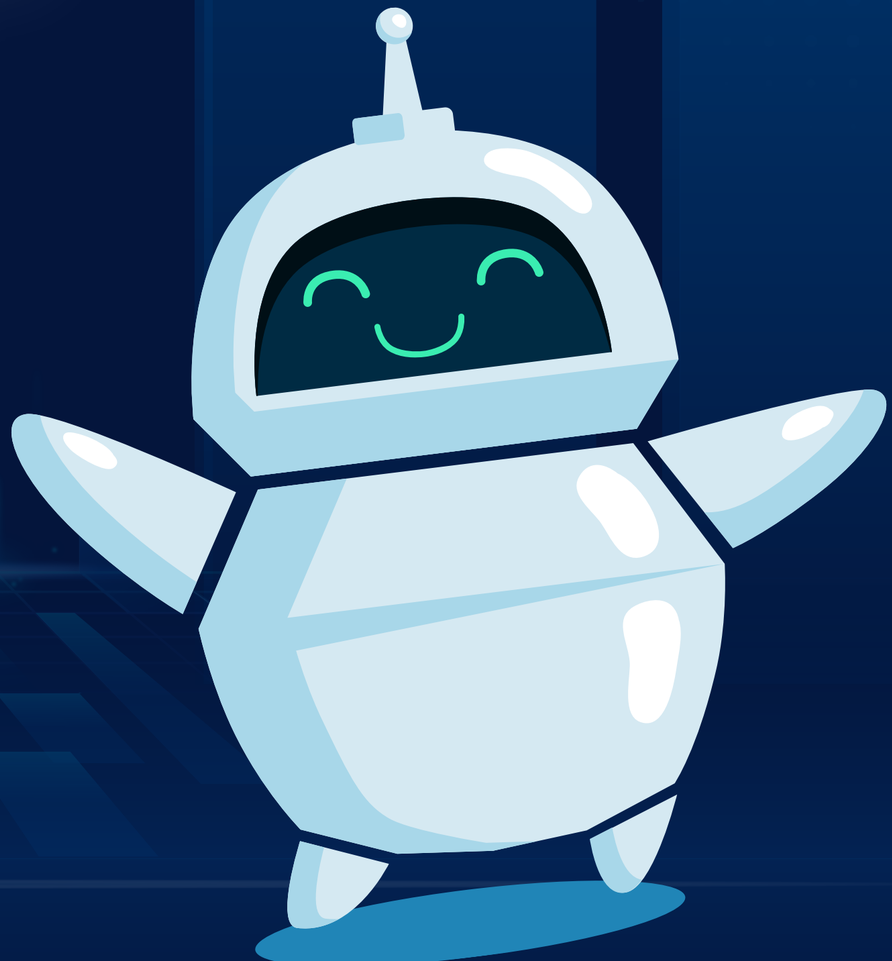
AutoBrowse revolutionizes the way users interact with the web by combining the power of voice control with intelligent automation. Its ability to understand natural language and execute browser commands hands-free not only enhances convenience but also promotes greater accessibility for all users.





Thank You!





Feel free to give feedback
@TeamAutonoMind