



IJRASET

International Journal For Research in
Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 12 **Issue:** IV **Month of publication:** April 2024

DOI: <https://doi.org/10.22214/ijraset.2024.61080>

www.ijraset.com

Call:  08813907089

E-mail ID: ijraset@gmail.com

HELIX: Autonomous AI Agent

Dr. Pravin Shinde¹, Himali Paradkar², Poojan Vig³, Sanchay Thalnerkar⁴, Vinay Jain⁵

¹Professor, ^{2,3,4,5}Student, Department of Artificial Intelligence and Data Science, Shah & Anchor Kutchhi Engineering College

Abstract: Artificial Intelligence has transformed the way we interact with technology, introducing us to agents that can think and make choices like humans. At the heart of this evolution is our project, 'HELIX'. Through 'HELIX', we've developed AI agents specialized in a variety of tasks: from digging deep into the web for research, streamlining email communication through automation, efficiently sending out emails in bulk, to strategically identifying and generating potential business leads. By weaving together cutting-edge machine learning algorithms and advanced language models, our system stands as a testament to the proficiency and versatility of AI. What makes 'HELIX' especially groundbreaking is its user-friendly approach, ensuring anyone, regardless of their technical background, can harness its power. As we embrace the dawn of this technological era, 'HELIX' not only showcases the potential of today's AI capabilities but also paves the way for future innovations, promising an expansive horizon for AI-driven solutions across myriad domains.

Keywords: Autonomous agents, Large Language Models, real-time data analysis, Automate repetitive tasks, reduce workload

I. INTRODUCTION

A. Background

In the realm of technological evolution, Artificial Intelligence (AI) stands as a beacon of innovation, transforming traditional paradigms across a multitude of domains. AI agents, in particular, have come to symbolize a new era in task automation and intelligent decision-making, catalysing enhanced efficiencies and optimized workflows across diverse applications. "HELIX," our innovative venture, is envisioned not merely as an AI agent but as a versatile and scalable platform, meticulously crafted to host and manage a diverse array of AI agents. Each of these agents is equipped with the capability to navigate and automate specific tasks through the integration of APIs, machine learning algorithms, and Language Models (LLMs). Designed with a foundation that intertwines robustness with adaptability, HELIX opens a plethora of possibilities, serving as a platform that not only fulfills the current automation needs but also stands prepared to embrace future developments and capabilities in AI-driven task automation^[1].

B. Objective

The primary target of this research is to develop and evaluate a AI autonomous agent that can reduce the workload of any person by automating the repetitive tasks. It aims to create a robust, scalable platform capable of hosting and managing a diverse array of AI agents, each specializing in specific task automation. This platform will feature a user-friendly interface to ensure accessibility for users of varying technical expertise.

Key to this project is the implementation of machine learning algorithms and Language Models (LLMs) to empower AI agents with adaptability and efficiency. Integration with various Application Programming Interfaces (APIs), including LLMs and Google API, will further enhance the agents' capabilities without compromising performance or security.

With scalability at its core, the platform will facilitate easy integration of new AI agents and allow for ongoing development and capability expansion. Ultimately, the project aims to significantly reduce manual effort and task execution time through intelligent automation, while prioritizing a user-centric design for enhanced operational efficiency and user satisfaction.

II. LITERATURE REVIEW

Numerous research efforts have explored autonomous agents using various technologies. Some notable examples include:

AlphaZero: An algorithm that independently learns optimal strategies for board games like chess, shogi, and Go, surpassing the performance of previous game-specific software.^[1]

BERT: BERT's introduction marked a paradigm shift in NLP, demonstrating the transformative impact of pre-trained models on a broad array of language understanding tasks.^[2]

[1] The paper titled "Mastering Chess and Shogi by Self-Play with a General Reinforcement Learning Algorithm (2017)" is a seminal work by DeepMind introduces AlphaZero, an algorithm that independently learns optimal strategies for board games like chess, shogi, and Go, surpassing the performance of previous game-specific software.

AlphaZero's success underscored the potency of generic reinforcement learning algorithms and marked a pivotal moment in the application of AI to game-playing, demonstrating superior performance without reliance on human-derived knowledge. This work represents a significant milestone in game AI, setting new standards for algorithmic efficiency and versatility.

[3] The paper "Attention is All You Need (2017)" proposes the Transformer architecture, relying entirely on attention mechanisms and parallel processing of input data, standing as the foundation for subsequent models like BERT and GPT. The Transformer model introduced a new era in machine learning, particularly in NLP, setting the stage for the development of highly influential models in subsequent years. Its impact is monumental, leading to a fundamental shift in the design of models for NLP and beyond.

[4] The paper "Proximal Policy Optimization Algorithms (2017)" authored by researchers at OpenAI, this paper introduces Proximal Policy Optimization (PPO), a reinforcement learning method achieving state-of-the-art performance across a diverse range of tasks, known for its stability and efficiency. PPO simplified the training of policy gradient methods, providing a reliable and efficient algorithm widely adopted in the deep RL community. This work holds a significant place in the advancement of deep reinforcement learning, promoting stability and efficiency in training.

III. METHODOLOGY

The methodology section of this research is to engineer a unified, intelligent platform capable of hosting and managing a myriad of AI agents, each adept at autonomously navigating through and executing specific tasks, thereby serving as a comprehensive solution for automating varied digital tasks across multiple domains. The specifics of the objectives are outlined as follows:

A. *Develop a Versatile AI Agent Platform*

Design and develop "HELIX" as a scalable and adaptable platform that can seamlessly integrate and manage various AI agents, each tailored to automate specific tasks.

Ensure that the platform is designed with a user-friendly interface to facilitate ease of use and management of the hosted AI agents.

B. *Intelligent Task Automation:*

Implement machine learning algorithms and Language Models (LLMs) to ensure that each AI agent is capable of executing tasks intelligently and can adapt and evolve to enhance its performance and capabilities over time.

Automate a diverse array of tasks, including but not limited to, web scraping, data analysis, email management, and lead generation, ensuring that each task is executed with enhanced efficiency and accuracy.

C. *API Integration for Enhanced Capabilities*

Integrate various APIs, including LLMs for content generation, Google API for email management, and other relevant APIs, ensuring that each AI agent can leverage external technologies to enhance its capabilities and functionalities.

Ensure that API integrations are seamless and do not compromise the performance or security of the AI agents.

D. *Facilitate Continual Development and Scalability:*

Ensure that "HELIX" is developed with a framework that allows for the easy integration of additional AI agents in the future, facilitating continual development and expansion of its capabilities.

Develop a system that can accommodate enhancements and modifications in the existing AI agents, ensuring that they can be refined and upgraded as per evolving requirements and technological advancements.

E. *Enhance Efficiency and Reduce Manual Effort:*

Through the intelligent automation of various tasks, significantly reduce the manual effort and time required to execute them, thereby enhancing overall operational efficiency.

Ensure that the automation does not compromise the accuracy or quality of task execution, maintaining or enhancing the standards of outcomes.

F. *Enable User-Centric Design and Operation:*

Ensure that users can easily interact with and manage the AI agents, facilitating a user-centric design that does not demand extensive technical expertise for operation and management.

Through the realization of these objectives, "HELIX" aims to stand as a beacon of intelligent automation, providing a unified solution that not only addresses the current spectrum of digital tasks but is also poised to adapt and evolve in the face of future developments and challenges in the digital landscape.

IV. SYSTEM ARCHITECTURE

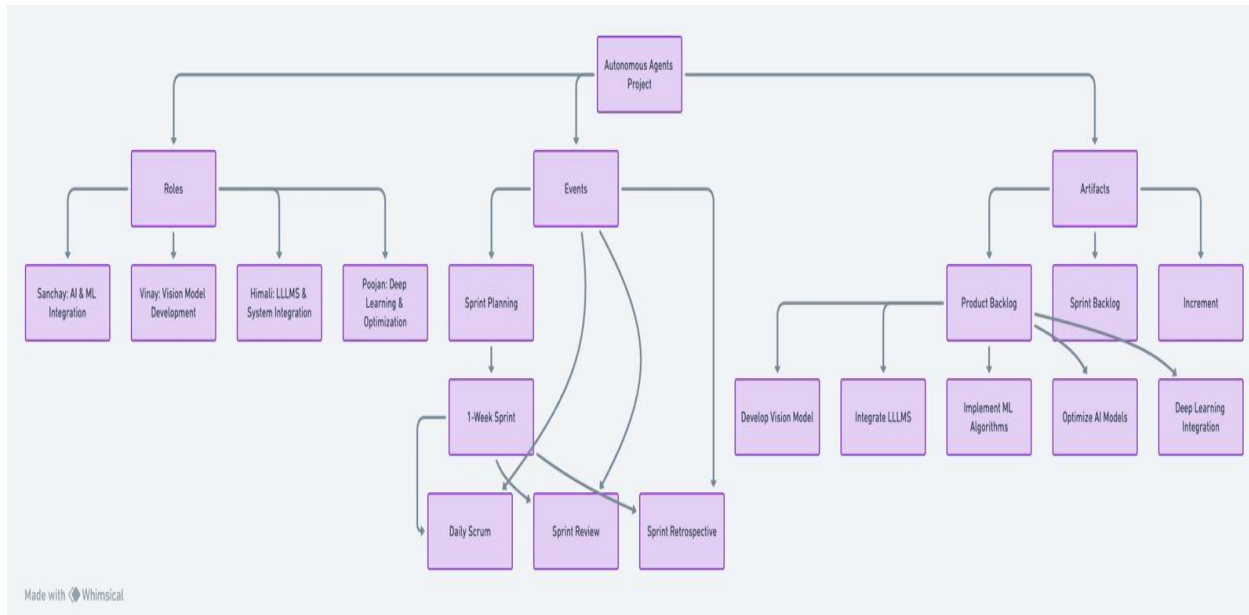


Fig 4.1: Framework

A flowchart representation of a Scrum-based project management approach for the Autonomous Agents Project. The chart breaks down roles, events, and artifacts, highlighting tasks like vision model development, ML integration, and deep learning optimization within the context of 1-week sprints.

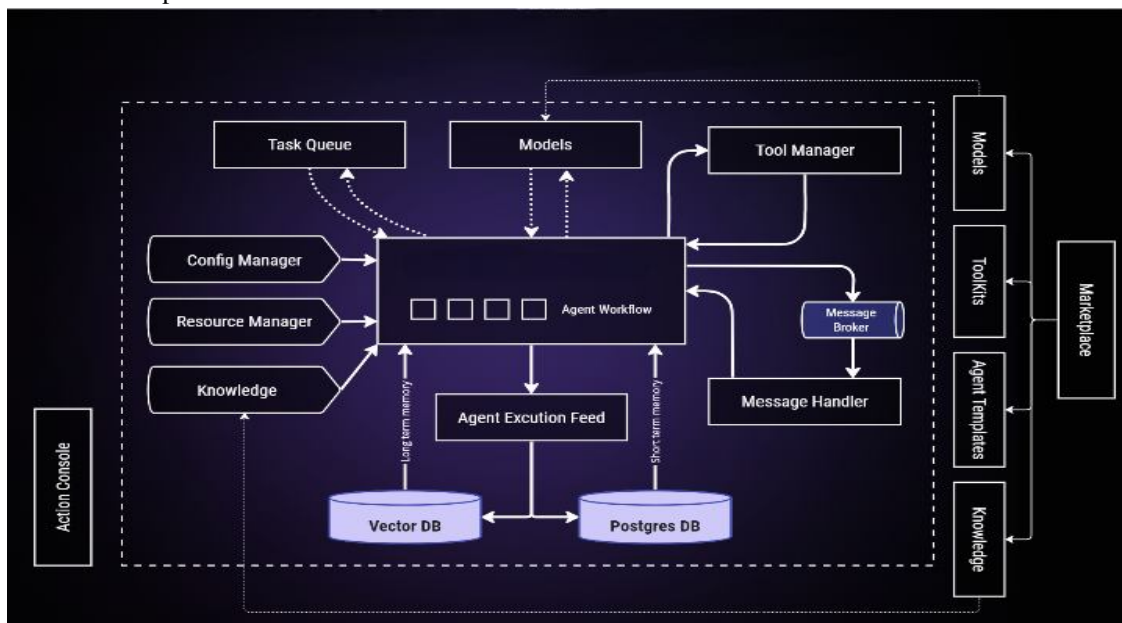


Fig 4.2: Agent workflow

A diagram showcasing an agent execution system, emphasizing agent workflows, task management, and interactions with multiple databases and modules, including configuration, resource management, and message handling.

V. RESULTS & DISCUSSIONS

A. Automate Lead generation process:

```
GPT: Hello! What would you like to browse today?
You: open vinay jains profile on linkedin search it on google

Sending ChatGPT request...

## PLAN ##
1. Go to google.com
2. Type 'Vinay Jain LinkedIn' into the search bar
3. Press Enter or click on the 'Search' button
4. Look for the search result that links to Vinay Jain's LinkedIn profile
5. Click on the link to open Vinay Jain's LinkedIn profile
## PLAN ##

Do you want to continue with this plan? (y/n): y
Sending ChatGPT request...
Going to https://www.google.com
Scraping page...
Sending ChatGPT request...
Typing "Vinay Jain LinkedIn" to q
Submitting form
Scraping page...
Sending ChatGPT request...
Clicking link "1300+ "Vinay Jain" profilesLinkedInhttps://www.linkedin.com > pub > dir > Vinay > Jain"
Scraping page...
Sending ChatGPT request...
Clicking link "Vinay Jain AVP - Risk Management at ICICI LOMBARD General Insurance Company Limited Mumbai ICICI Lombard General Insurance Company Limited, +4 more National Institute of Technology Raipur, +1 more"
Scraping page...
Sending ChatGPT request...
Current cost: 0.05 USD (17791 tokens)

GPT: Vinay Jain is an Associate Vice President - Risk Management at ICICI Lombard General Insurance Company Limited, based in Mumbai, Maharashtra, India. He has over 8 years of experience in the field of risk management and cybersecurity. Vinay has also worked at Reserve Bank Information Technology Pvt Ltd and Centre For Development of Advanced Computing. He holds a Bachelor of Technology degree in Computer Science & Engineering from National Institute of Technology Raipur.
```

Fig 5.1: Terminal window executing the command

Fig 5.1 shows representation of an agent's systematic approach to fulfill a user request through command-line interfaces (CLI). This process is initiated when the user expresses the need to find specific information online, in this instance, about "Vinay Jain" on LinkedIn.

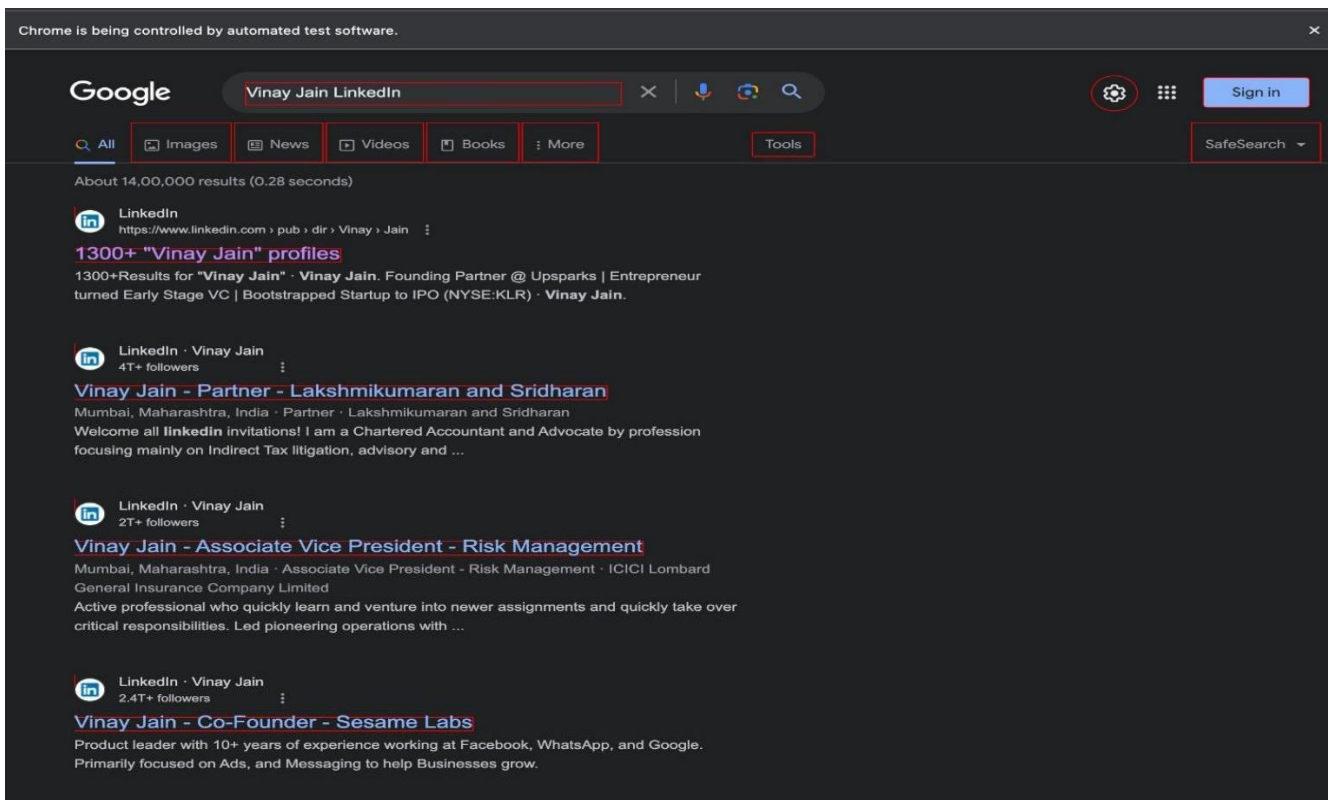


Fig 5.2: Automated browser

Fig 5.2 shows the screenshot of Google search results for "Vinay Jain LinkedIn", displayed through an automated browser controlled by Selenium software.

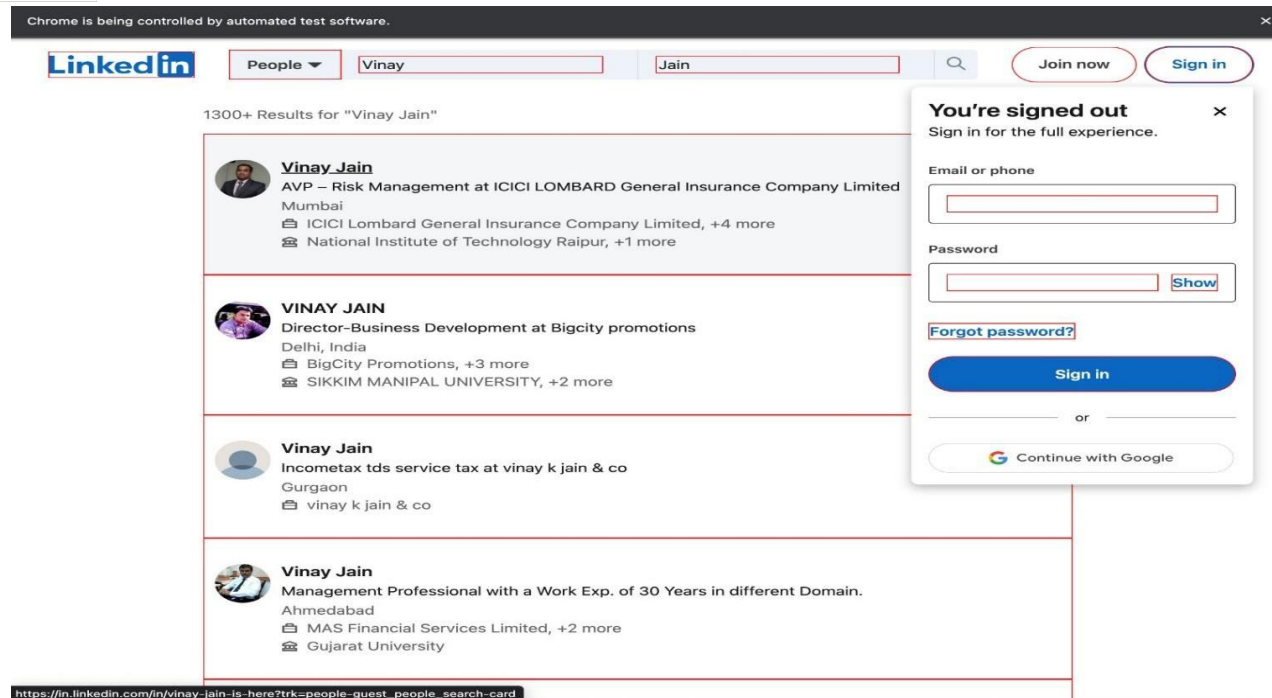


Fig 5.3: List of people with their name as “Vinay Jain”

Fig 5.3 shows the screenshot of LinkedIn search results showcasing various profiles with the name "Vinay Jain", displayed through an automated browser controlled by Selenium software.

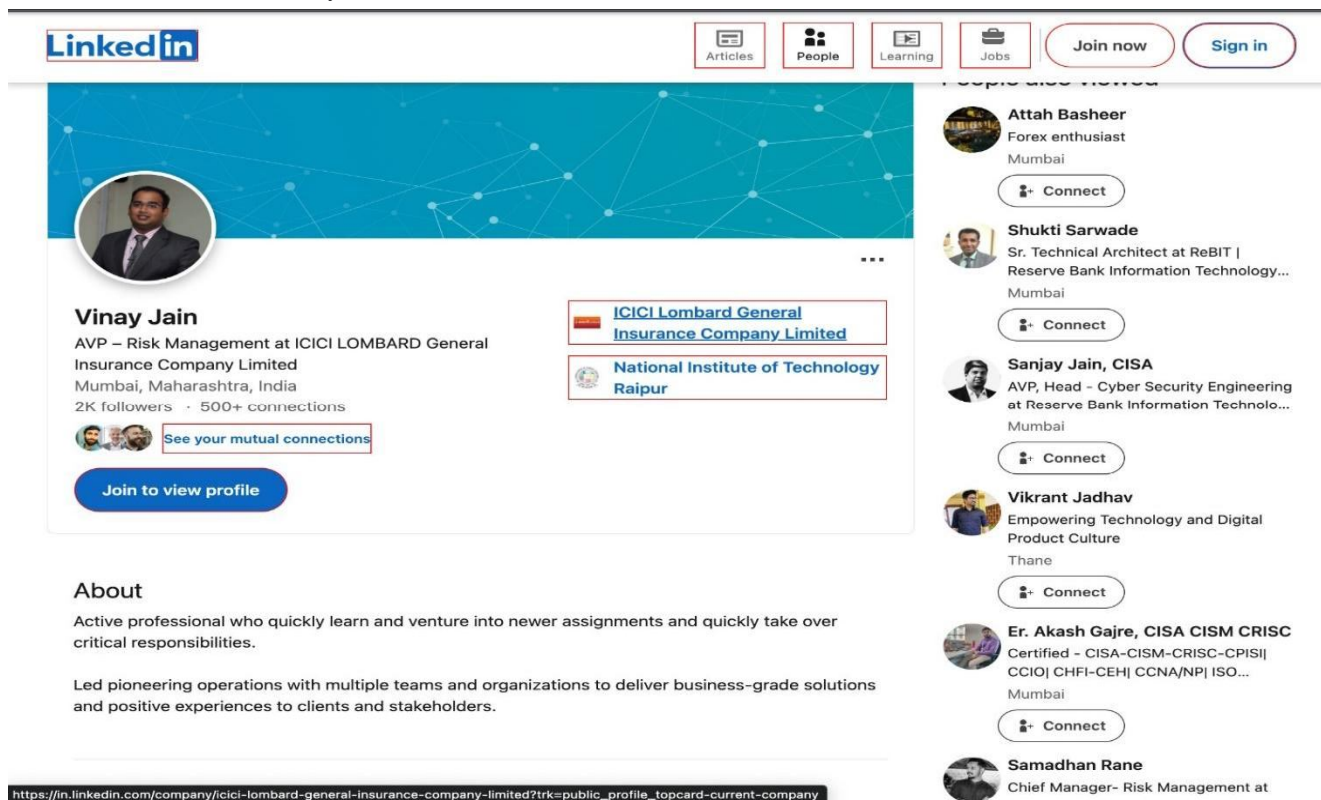


Fig 5.4: LinkedIn profile of a person named “Vinay Jain”

Fig 5.4 shows the linkedin profile of a person who appeared first on the seach results of linkedin page.

B. Automate Form filling process:

```
sanchaythalnerkar@Sanchays-MacBook-Air ~/Browser agent source env/bin/activate
source: no such file or directory: env/bin/activate
sanchaythalnerkar@Sanchays-MacBook-Air ~/Browser agent cd puppeteer-gpt
sanchaythalnerkar@Sanchays-MacBook-Air ~/Browser agent/puppeteer-gpt master node index.js --headless false
Using model: gpt-3.5-turbo-16k

GPT: Hello! What would you like to browse today?
You: read from personal.txt and fill https://www.jotform.com/form/241143010624439
```

Fig 5.5: Terminal window

Fig 5.5 shows the terminal window where the script is being executed and it had given the personal.txt file which contains basic details of the person and form link.

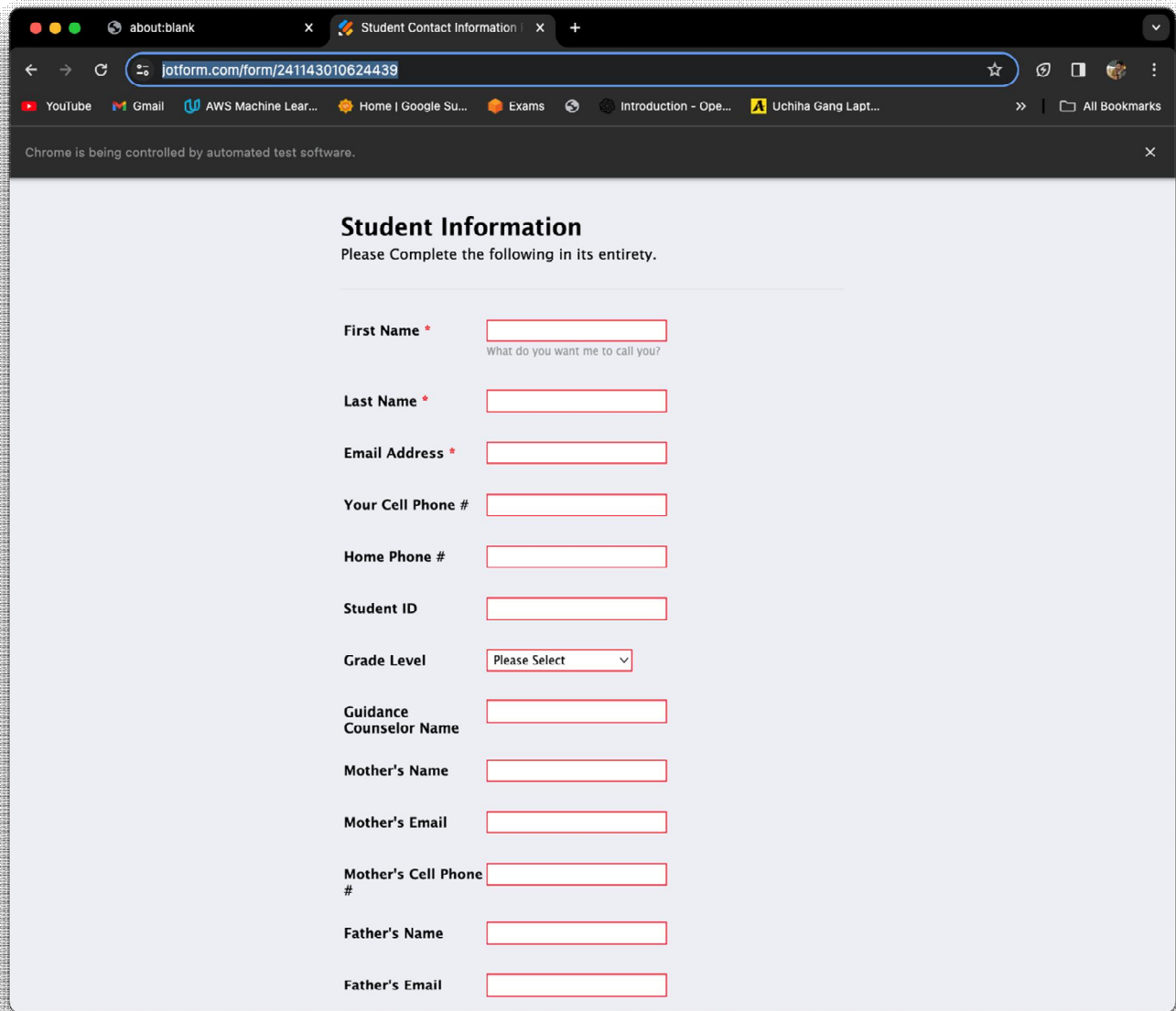
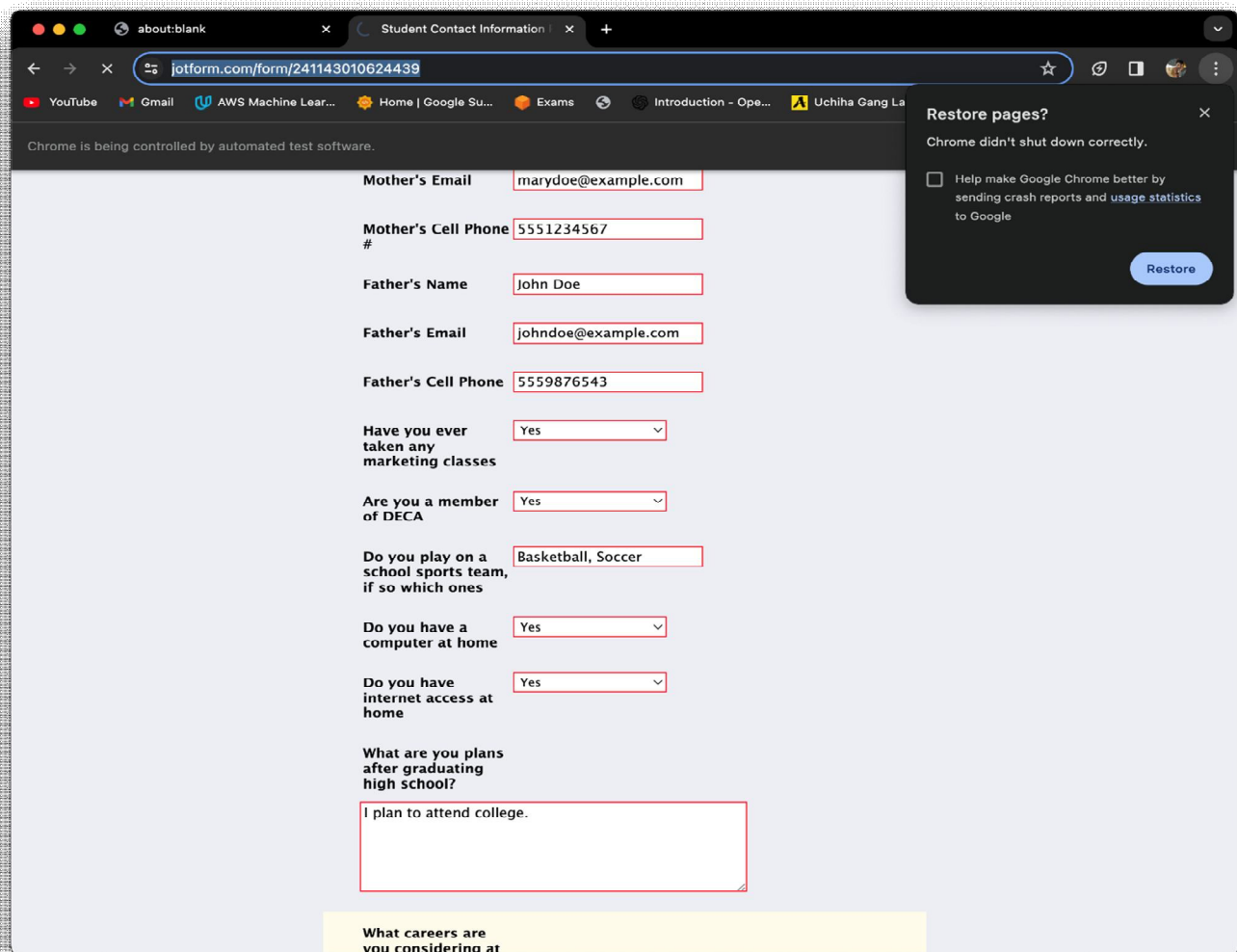


Fig 5.6: Form with fields detected in browser

Fig 5.6 shows the form opened in the browser and the agent have detected all the fields in the form that need to be filled marking them with red color.



The screenshot shows a web browser window with the URL `jotform.com/form/241143010624439`. The page title is "Student Contact Information". The form contains the following fields and values:

- Mother's Email: `marydoe@example.com`
- Mother's Cell Phone #: `5551234567`
- Father's Name: `John Doe`
- Father's Email: `johndoe@example.com`
- Father's Cell Phone: `5559876543`
- Have you ever taken any marketing classes: `Yes`
- Are you a member of DECA: `Yes`
- Do you play on a school sports team, if so which ones: `Basketball, Soccer`
- Do you have a computer at home: `Yes`
- Do you have internet access at home: `Yes`
- What are you plans after graduating high school?: `I plan to attend college.`

A "Restore pages?" dialog box is open in the top right corner, indicating that Chrome didn't shut down correctly. The dialog box contains the text "Restore pages?", "Chrome didn't shut down correctly.", and a checkbox for "Help make Google Chrome better by sending crash reports and usage statistics to Google". A "Restore" button is visible at the bottom right of the dialog box.

Fig 5.7: Data filled automatically in the form

Fig 5.7 shows the form where the data have been filled from the personal.txt file automatically without any human intervention.

VI. RESULTS SUMMARY

- 1) Measure Response Time: Response times for tasks like webpage loading and form submission were quantified using Python's time module, effectively measuring activity durations.
- 2) Calculate Throughput: Throughput was assessed by counting successful actions within specific time frames, such as forms submitted per minute or hour.
- 3) Monitor Resource Consumption: System resource usage, including CPU and memory, was tracked using Python's psutil library to identify potential inefficiencies.
- 4) Assess Reliability: Automation scripts were tested for consistency and stability by executing them repeatedly under consistent conditions.
- 5) Check Scalability: Stress-testing was performed by incrementally increasing load or operation frequency and measuring response times and resource usage.

- 6) Evaluate Error Rate: The error rate was quantified by logging and calculating the ratio of errors to total operations during script execution.
- 7) User Experience Evaluation: End-user experience was assessed by evaluating page loading speed and visual rendering using Selenium and potentially integrated tools like Lighthouse.

The comprehensive evaluation of the browser automation project revealed promising results across key performance indicators. Efficient response times, high throughput, and stable resource consumption demonstrated the system's robustness and scalability. These findings underscore the project's success in achieving its objectives of operational efficiency, scalability, and a positive user experience, positioning it as a reliable and adaptable solution for task automation.

VII. CONCLUSION & FUTURE DIRECTIONS

This paper represents a groundbreaking effort to address the challenges of digital task management and automation in today's rapidly evolving digital landscape. With its vision of a unified platform capable of hosting a diverse array of AI agents, HELIX has introduced a paradigm shift in how digital tasks are managed and executed. The robust technical infrastructure, strategic incorporation of cutting-edge technologies, and seamless API integrations have amplified the capabilities of the AI agents, enhancing efficiency, accuracy, and adaptability in task execution.

The user-centric design of HELIX underscores its commitment to accessibility and ease of use, democratizing the benefits of artificial intelligence and making sophisticated automation accessible to a broader audience. The platform's scalability, designed with a forward-thinking approach, ensures its relevance and adaptability to future technological advancements and evolving task requirements. The provision for easy integration of additional AI agents and capabilities further augments the platform's readiness for future expansions.

Future Directions:

Looking ahead, this paper is well-positioned to adapt and evolve in response to future developments and challenges in the digital domain. The platform's long-term vision and commitment to continual development and scalability lay a solid foundation for future advancements. Potential areas for future exploration include enhancing the platform's AI capabilities, expanding the range of automated tasks, and further refining the user experience to meet the evolving needs of users.

REFERENCES

- [1] Silver D, Hubert T, Schrittwieser J, Antonoglou I, Lai M, Guez A, Lanctot M, Sifre L, Kumaran D, Graepel T, Lillicrap T. Mastering chess and shogi by self-play with a general reinforcement learning algorithm. arXiv preprint arXiv:1712.01815. 2017 Dec 5.
- [2] Devlin J, Chang MW, Lee K, Toutanova K. Bert: Pre-training of deep bidirectional transformers for language understanding. arXiv preprint arXiv:1810.04805. 2018 Oct 11.
- [3] Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, Kaiser Ł, Polosukhin I. Attention is all you need. Advances in neural information processing systems. 2017.
- [4] Schulman J, Wolski F, Dhariwal P, Radford A, Klimov O. Proximal policy optimization algorithms. arXiv preprint arXiv:1707.06347. 2017 Jul 20.
- [5] M. Colledanchise, R. Parasuraman and P. Ögren, "Learning of Behavior Trees for Autonomous Agents," in IEEE Transactions on Games, vol. 11, no. 2, pp. 183-189, June 2019.
- [6] Barbara Schütt, Meng Zhang, Christian Steinhauser, Eric Sax, "Evolutionary Behavior Tree Generation for Dynamic Scenario Creation in Testing of Automated Driving Systems", 2023 7th International Conference on System Reliability and Safety (ICRSRS), pp.322-330, 2023.
- [7] Pattie Maes; Modeling Adaptive Autonomous Agents. Artif Life, 1993.
- [8] Silver, David, et al. "A general reinforcement learning algorithm that masters chess, shogi, and Go through self-play." Science 362.6419 (2018).
- [9] Schwartz, Sivan, Avi Yaeli, and Segev Shlomov. "Enhancing trust in LLM-based AI automation agents: New considerations and future challenges." arXiv preprint arXiv:2308.05391 (2023).
- [10] Abdulhakam AM Assidiq et al., "Real time lane detection for autonomous vehicles", 2008 International Conference on Computer and Communication Engineering, 2008.
- [11] Grimley, Michael J., and Brian D. Monroe. "Protecting the integrity of agents: an exploration into letting agents loose in an unpredictable world." XRDS: Crossroads, The ACM Magazine for Students 5.4 (1999): 10-17.
- [12] Dubiel, Mateusz, Sylvain Daronnat, and Luis A. Leiva. "Conversational Agents Trust Calibration: A User-Centred Perspective to Design." Proceedings of the 4th Conference on Conversational User Interfaces. 2022.
- [13] Liu B, Robertson E, Grigsby S, Mazumder S. Self-initiated open world learning for autonomous ai agents. arXiv preprint arXiv:2110.11385. 2021 Oct 21.
- [14] Zhou W, Jiang YE, Li L, Wu J, Wang T, Qiu S, Zhang J, Chen J, Wu R, Wang S, Zhu S. Agents: An open-source framework for autonomous language agents. arXiv preprint arXiv:2309.07870. 2023 Sep 14.
- [15] Zhou S, Xu FF, Zhu H, Zhou X, Lo R, Sridhar A, Cheng X, Bisk Y, Fried D, Alon U, Neubig G. Webarena: A realistic web environment for building autonomous agents. arXiv preprint arXiv:2307.13854. 2023 Jul 25.



10.22214/IJRASET



45.98



IMPACT FACTOR:
7.129



IMPACT FACTOR:
7.429



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089  (24*7 Support on Whatsapp)