IRCN V: An AI-Enhanced Crowdsourcing Platform for ChatGPT Development

Hong Linh Nguyen, Van Tam Mai, Anh Viet Nguyen, Duy Khanh Nguyen, and Ngoc An Nguyen Ho  
FPT University Campus Da Nang, Da Nang, Vietnam

*Abstract*—This article discusses the widespread use of artificial intelligence (AI) in various industries and applications, including the development of conversational AI, such as ChatGPT. While chatbots can be a powerful tool for providing information, they are limited by the accuracy of the data they are given. There are potential drawbacks to using crowdsourcing for AI research. This research project aims to create a community to enhance the accuracy and capabilities of AI systems through collaboration and knowledge sharing. This article presents an overview of related work, problem formulation, methods, data set, evaluation, and results. The new and better aspect of this article is the proposal for a collaborative community to improve the accuracy and capabilities of AI systems. The article suggests that collaboration and knowledge sharing can drive innovation in AI research. In conclusion, this article highlights the importance of ensuring the accuracy of AI technology and leveraging community intelligence to enhance its capabilities.

Keywords—AI chatbot, crowdsourcing, chatGPT

# Introduction

AI usage is becoming increasingly prevalent today as technology has advanced significantly. AI is used in various industries, including healthcare, finance, education, retail, etc. In healthcare, AI is used for diagnostics, treatment recommendations, and personalized medicine. AI was applied to fraud detection, risk management, and portfolio optimization in finance. Besides, AI is being used to personalize learning experiences for students in education. Especially, AI is used in retail for supply chain optimization, customer service, and recommendation engines. Furthermore, AI is integrated into various products and services, such as virtual assistants, smart home devices, and autonomous vehicles. The current status of AI usage indicates that the technology is rapidly advancing and will continue to play a significant role in various industries and aspects of our daily lives.

Using the Generative Pre-trained Transformer (GPT) architecture, OpenAI developed the substantial language model known as ChatGPT. It is intended to comprehend natural language inputs and conversationally respond to them. ChatGPT can produce logical and pertinent answers to various questions and topics because it was trained on a large corpus of text data [1]. For technique, AI has been applied to create a ChatGPT application. The field of natural language processing (NLP) uses the deep neural network design known as GPT. One of the most well-known NLP models today is GPT, which was trained on enormous amounts of data to automatically predict and generate text from short to long. The largest and most recent version of GPT, GPT-4, is regarded as one of the most sophisticated NLP models available today. [1]. Chatbots like ChatGPT are limited in their ability to perform basic math, solve logical questions, and provide accurate information [7]. Despite being an artificial intelligence, ChatGPT has provided inaccurate information pretty. CNET used an AI tool to write an article on loans that contained incorrect information. The editorial board failed to verify the accuracy of the content, causing problems for CNET. This situation warns us to avoid using chatbots and other AI tools in journalism without proper verification [4]. Chatbots are designed to provide accurate information, but they can only work within the limits of the data they are given. Outdated or incorrect information can lead to the distribution of misinformation and misdirection [5]. As chatbots become more popular, it is essential to ensure that they provide up-to-date and accurate information by reviewing their accuracy.

This research aims to create a community that can improve the accuracy of AI technology, share knowledge, and develop new directions for technological advancement. As a crowdsourcing platform, IRCN V leverages the community's collective intelligence and expertise. By using this crowdsourcing platform, we could enhance the capabilities of AI systems and drive innovation in the field. Besides, through collaboration and knowledge sharing, we hope to foster a continuous learning and growth culture that will benefit the community and the broader field of AI research. However, crowdsourcing has several potential drawbacks that should be considered. One issue is the quality of work. Since the crowd is often not as tightly controlled as internal employees, the quality of work may not meet requirements. Another challenge is project management. Managing and allocating work to a large group of people can become complex and challenging. Additionally, there is a risk of sensitive or confidential information being leaked when collecting data from the crowd. Hence, in the early stages of the research, we identified potential solutions and developed models for how to apply the proposed solutions to the study's feasibility. We also developed scenarios of how the chatbot could be used and worked to create a testable demo model in real-life scenarios.

The remainder of this article is constructed as follows. The next section is an overview of related work, including the problem approach and the solution provided by the application. Section 3 details a description of the IRCN V platform, including architecture and functionalities. The experimental results were provided in section 4. Finally, we conclude with potential future directions in section 5.

# LITERATURE REVIEW

## Crowdsourcing platforms

In 2006, in the article "The Rise of Crowdsourcing" by Jeff Howe, published in Wired [3], crowdsourcing appeared for the first time, a contraction of crowd and outsourcing. This quote from the article describes simply what crowdsourcing is and how it's made possible by technological advances. The meaning of the word crowdsourcing itself is a portmanteau of crowd and outsourcing [11]. From an online application perspective, online crowdsourcing platforms are increasingly being used to capture ideas from the crowd. Global companies are adopting crowdsourcing ideas to connect with and get user feedback. The success of a crowdsourcing platform largely depends on members and their motivation to participate. Motivation determines the quality and quantity of contributions [12].

Crowdsourcing has become a helpful tool for understanding audience preferences and anticipating needs [13]. It is essential for businesses that depend on innovating or enhancing their products, such as fashion brands, food manufacturers, or restaurants [14]. For example, Starbucks provides a simple fill form where people can suggest new ideas, enhance existing services, or request product deliveries. More complex, Amazon MTurk and CrowdFlower [15], which are crowdsourcing platforms, have several jobs such as data categorization, metadata tagging, character recognition, voice-to-text transformation, data entry, email harvesting, sentiment analysis, ad placement on videos or surveys.

## ChatGPT

When chatGPT initially debuted in November 2022, it piqued the imagination of everyone: engineers, social network users, students, and so on. ChatGPT allows humans to communicate with AI by entering messages, and it will solve the problem and respond to the user based on the OpenAI paradigm. However, it still has significant limits on the information that comprises the proper query. And the developer is still striving to improve GPT. ''In the future, ChatGPT capabilities will be a helpful tool for enterprises in domains such as customer assistance, online learning, and market research.''[6].

We discovered the most faults in the GPT's calculations. Specifically, GPT appears to offer an inaccurate conclusion in current study inquiries on the mathematical abilities of chatGPT employing existing data. The GPT's mathematical abilities are restricted, falling "significantly short of an average math graduate"[7]. As a result, we discovered that GPT could grasp what the issue was saying but could not provide precise responses.

Human-suggested learning outcomes outperform those suggested by GPT, according to surveys and references to actual scientific articles: "assess the level of achievement in the first learning of ChatGPT by comparing the effectiveness of its suggestions with those given by the tutor who is an author with 77 participants on two algebraic subject areas, Elementary Algebra, and Intermediate Algebra" [8]. We discovered that chatGPT is only to a certain extent effective. According to the survey, GPT data could be more credible. GPT gives people knowledge that includes inaccessible data and that it learns from humans in the process. In reality, when utilizing chatGPT to compose essays in research domains, the chance of plagiarism is quite significant. chatGPT takes information from numerous sections in the most currently studied articles and offers the author the data; the information does not follow the author's study but contains comparable information. It would be even worse if the author did not write the essay himself [9].

ChatGPT's performance in numerous domains remains tough owing to its closed nature and regular upgrading due to continuous learning from user feedback. In the worst-case scenario, users give GPT erroneous information, causing it to lose its source of knowledge [10]. Since then, we have been doing research to strengthen the credibility of the data when supplying users who are not overly reliant on GPT but instead communicate with other users.

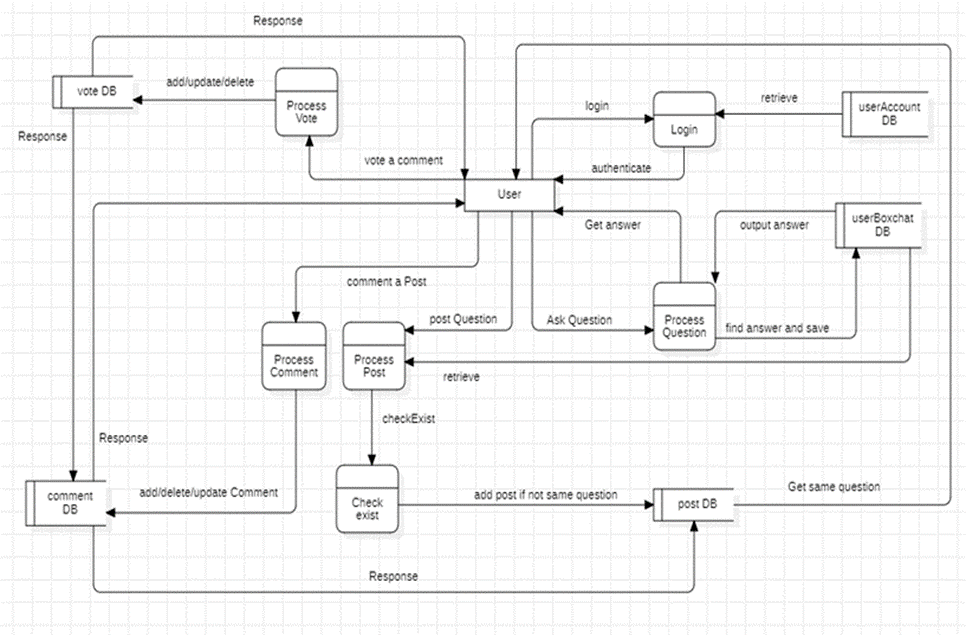
# IRCN V – CROWDSOURCING PLATFORM FOR COLLECTING FEEDBACK FROM EXPERT

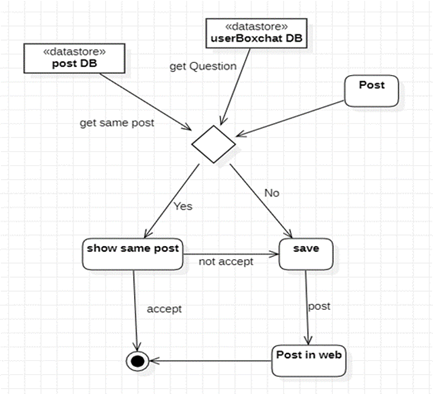
## The overview of IRCN V

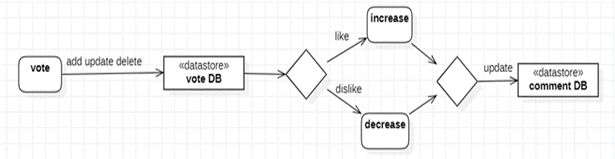
Get key API of ChatGPT is an effective approach to improve website user interaction. libraries and documentation are supported by OpenAI, so we were able to efficiently integrate ChatGPT's API into the JavaScript code. Any requests from user can be responded a way quickly. We were able to create a natural language processing (NLP) in our application.

Although it was successful in getting the openAI api, it Limited about user experience (UX) such as when saving the website data, it was reloaded, annoying to users as well as the website professionalism. so I used Asynchronous JavaScript and XML (Ajax). Ajax allows web pages to load necessary data and updates without necessarily reloading the entire website. This helps speed up page loads and improves user experience.

We deploy IRCN V as a web-based application that uses the Java language. Our platform relied on the MVC model a very popular and essential model for Java web applications. The functionalities of this system include essential features that create a reliable online environment and provide benefits for users. Firstly, the Sign-up and Login function allows users to create an account and log in to the system to access and use some features that will affect the community of application users. The Ask and Answer function allows users to query the chatbot to find answers to their questions using natural language processing (NLP) technology, saving time and increasing efficiency in information search. The Post question function allows users to post articles and search for related posts on their issues that the chatbot's answers have not satisfied, creating a space for discussion for knowledgeable people in that field to help. The system also applies the feature of suggesting related posts to minimize duplicate information. The Comment function allows members can act with their comments. Comments are sorted by default based on the votes they receive to help users find the best quality opinions. The system also applies rating and time limits to prevent abuse to ensure community quality. Finally, the Vote function allows users to evaluate and vote for posts or comments that they consider valuable or useful. This function ensures that posts or comments are evaluated based on their quality and actual value. Overall, the functionalities create a reliable online platform for users to exchange and discuss an issue in the field they are researching with others. At the same time, it improves the chatbot's responses to those with similar or related questions. The architecture of data flow in IRCN V is detailed in Figure 1.

**





## Functionalities

This section details functions in the IRCN V crowdsourcing platform that consist of user management (sign up and login), ask and answer, post question, and vote. All functions are described as follows.

*Sign up and log in*

* Sign up: Users can create an account on a website or application using the registration function. Before a person may register for an account on this application, they must enter the code sent to them by email. This code confirms their Gmail ownership and prevents the creation of fictitious accounts. The user's data, including name, address, email, and other pertinent information, will be saved on the server after correctly entering the code. The server stores user passwords in the database using argon2 memory password hashing for optimal security. By maintaining the confidentiality of saved passwords, this technique offers the best defense against prospective password-cracking attempts. Overall, this registration process provides a safe and reliable way for users to create an account on a website or app, protect their personal information and ensure that only authorized individuals only have access to the system.
* Login: The user enters their credentials into the system during the login procedure, and the system then contacts the server to verify the data. This verification process must ensure that only authorized individuals can visit the community page and communicate with the chatbot. The platform's capabilities, such as participating in forum conversations, receiving tailored suggestions, and utilizing the chatbot's services in real-time, are accessible to the user after successful verification. The system securely maintains user information in its database, making it possible to retrieve data as needed quickly. Users may easily travel between various platform regions and have a seamless experience by enabling simultaneous access to the community page and chatbot.

*Ask and answer*

A user will enter a question or request into the chat interface when they connect with the chatbot. Natural language processing (NLP) is used in the chatbot's programming to assess user input and select a suitable answer. The user's question and the chatbot's response will be saved in the database once the chatbot has determined the proper response. To ensure that the database can be searched using just one attribute, we use the forward slash and ampersand symbols /&...&/ This allows us to save the user's query and make it simple to query and display it to users. In addition to saving user queries, we also need to store bot replies. To accomplish this, we use the asterisk and forward slash symbols /\*...\*/

EXAMPLE: The attribute stored in the database, for instance, will be:

/&abc1&//\*xyz 1\*//&abc2&//\*xyz 2\*/

When the user logs in, the user will see the database displayed on the screen.

abc1

xyz 1

abc2

xyz 2

*Post question*

You can always post your query on a forum if you think the chatbot's response is inappropriate or want a more in-depth discussion. This forum is a website where users can post their thoughts, ask and answer questions, and discuss various subjects. This enables you to interact with other people who share your hobbies or have knowledge in your field of interest. When you post a question on the forum, the server automatically verifies if it matches other existing posts in its database by suggesting posts with similar content. This helps ensure that your question is unique and unanswered elsewhere. If related posts are available, the system can recommend them so you can review previous discussions and potentially find answers to your questions without waiting for a response, minimizing data redundancy. However, if there aren't any suggestions that match what you're looking for, or if you'd like to have a new conversation about your topic, you can wait for a response from other users. This can be a great way to get diverse perspectives and opinions from individuals with experience or expertise in your area of interest. Overall, forums are an excellent resource for anyone looking to engage in meaningful discussions, share their knowledge, or seek advice on a particular topic. The whole diagram of the post-question function is depicted in Figure 2.

*Comment*

The platform provides a vibrant space for users to actively participate in conversations by allowing them to answer questions and participate in discussions through comments on each post. To ensure that users don't miss out on the most popular or relevant discussions, the platform automatically displays the comments that received the most votes at the top. In addition, commenters are given the flexibility to make changes to their comments as they see fit, with 'add,' 'edit,' and 'delete' features available. However, the platform also recognizes the importance of maintaining a respectful and safe user environment. Users can vote on other users' comments to prevent abuse and promote constructive interactions. Only those with high authority can vote. At the same time, the platform imposes a time limit of five minutes per comment, encouraging users to focus on the quality of the comment and limiting controversial comments.

*Vote*

The phrase "Votes are counted as upvotes (+1) and downvotes (-1); if a comment receives a majority of downvotes, the vote count will be negative" implies that users have the option to express their opinion on comments by either upvoting or downvoting them. An upvote adds +1 to the total vote count, while a downvote subtracts -1 from the total count. If a comment receives more downvotes than upvotes, its total vote count will be negative. Specific requirements must be met to ensure that only eligible individuals cast their votes. Apart from being a registered user, one must also attain a certain level of prestige, which could be a significant number of followers or votes for favorable comments. This helps prevent fake accounts from being created to boost the number of votes or debunking other remarks artificially. By implementing these measures, the system aims to maintain a fair and transparent platform where users can express their opinions without fear of manipulation or exploitation. To have a more straightforward observation, Figure 3 illustrates the diagram of the vote function in IRCN V

# EXPERIMENTAL RESULTS

* The research results show that using data from community question-and-answer websites to build AI chatbot models can help improve their accuracy. During implementation, we used the ChatGPT API to provide answers to users' questions. Based on this model, we experimented and collected data from conversations between users and the chatbot to create questions and turn them into posts on the question-and-answer website when users request them.
* Our question-and-answer community provides users with many options when posting their questions, including searching for related posts on the community to ensure that their questions are answered quickly and accurately. We also used a method of selecting and listing relevant posts when users posted their questions. From there, users can evaluate whether the posts are useful to them and choose an appropriate approach to solve their problems.
* After testing the model on a Q&A website, we were able to collect data from user interactions with the chat AI with their permission. This data was then used to improve our chat AI, specifically using a selection method to evaluate related posts and identify common errors in chat AI responses. The results showed that this method helped improve the accuracy of the chat AI and provide users with a better experience.
* To evaluate the effectiveness of this method, we conducted some tests and evaluated the results. Specifically, we used a dataset from some public Q&A websites to train our Q&A and chat AI models. We then used the questions and answers from this dataset to create posts on our Q&A community website to create a sample dataset to attract users to participate in the community.
* After implementing the selection method and listing related posts for users, we have collected a large amount of data about questions and answers posted by users on our website. The data we collected includes both from ourselves and from friends and acquaintances who have used Chat GPT to ask questions but have not found answers. We collected these questions and posted them on community websites such as Quora, Stack Overflow, Reddit to find answers and then added that data to our database.
* To evaluate the effectiveness of this method, we conducted a survey with users and the results showed that most users highly appreciated the usefulness of the posts on the website, as well as the ability of the chat AI to answer questions. We also conducted tests on the accuracy of the chat AI after using the new data, and the results showed that the accuracy of the chat AI had been significantly improved. However, some users still have some opinions about this model, such as when encountering a new issue that the chat AI cannot answer satisfactorily, users have to wait for a sufficient period of time for the community to answer that question. That is also an issue that cannot be avoided when we built this model.

In summary, the method of selecting and listing relevant articles for users has helped us collect a large amount of data from the user community and improve the accuracy of our chat AI. However, this model still has some disadvantages that need to be improved. Therefore, we are researching and developing other methods to improve the quality of data and the accuracy of our chat AI.

Some challenges encountered during the project implementation include:

* Establishing project credibility with users: The reliability of the project has not been fully established among users, which may affect their trust in the output provided by chatGPT.
* Building a large enough community for verification: The community size may not be large enough to effectively verify the accuracy of the project's results, which could impact the robustness of the findings.
* Issues with creating credibility for community accounts: There have been difficulties in establishing credibility for user accounts within the community, which may affect the validation process and reliability of the feedback received.

# CONCLUSIONS

In conclusion, this study has demonstrated the significant potential of combining crowdsourcing with AI chatGPT in scientific research. Integrating these two powerful tools has allowed us to achieve more accurate and comprehensive results while reducing the time and resources required for data collection and analysis. The scientific implications of this research are far-reaching, as it opens up new avenues for exploring complex problems and generating innovative solutions. In terms of practical applications, the results of this study can be used to improve the efficiency and effectiveness of various industries and sectors. Looking to the future, we hope this research will inspire further exploration into the potential of combining crowdsourcing with AI chatGPT.

We look forward to seeing the exciting developments emerging in this field.

##### References

1. Lecler, Augustin, Loïc Duron, and Philippe Soyer. "Revolutionizing radiology with GPT-based models: Current applications, future possibilities, and limitations of ChatGPT." Diagnostic and Interventional Imaging (2023).
2. Lenart-Gansiniec, Regina, Wojciech Czakon, Łukasz Sułkowski, and Jasna Pocek. "Understanding crowdsourcing in science." Review of Managerial Science (2022): 1-34.
3. Howe, Jeff. "The rise of crowdsourcing." Wired magazine 14, no. 6 (2006): 1-4.
4. CNET's Article-Writing AI Is Already Publishing Very Dumb Errors, January 17 by Jon Christian
5. Borji, Ali. "A categorical archive of chatgpt failures." arXiv preprint arXiv:2302.03494 (2023).
6. Haleem, Abid, Mohd Javaid, and Ravi Pratap Singh. "An era of ChatGPT as a significant futuristic support tool: A study on features, abilities, and challenges." BenchCouncil transactions on benchmarks, standards, and evaluations 2, no. 4 (2022): 100089.
7. Frieder, Simon, Luca Pinchetti, Ryan-Rhys Griffiths, Tommaso Salvatori, Thomas Lukasiewicz, Philipp Christian Petersen, Alexis Chevalier, and Julius Berner. "Mathematical capabilities of chatgpt." arXiv preprint arXiv:2301.13867 (2023).
8. Pardos, Zachary A., and Shreya Bhandari. "Learning gain differences between ChatGPT and human tutor generated algebra hints." arXiv preprint arXiv:2302.06871 (2023).
9. Dahmen, Jari, M. Kayaalp, Matthieu Ollivier, Ayoosh Pareek, Michael T. Hirschmann, Jon Karlsson, and Philipp W. Winkler. "Artificial intelligence bot ChatGPT in medical research: the potential game changer as a double-edged sword." Knee Surgery, Sports Traumatology, Arthroscopy (2023): 1-3.
10. Aiyappa, Rachith, Jisun An, Haewoon Kwak, and Yong-Yeol Ahn. "Can we trust the evaluation on ChatGPT?." arXiv preprint arXiv:2303.12767 (2023)
11. Schenk, Eric, and Claude Guittard. "Crowdsourcing: What can be Outsourced to the Crowd, and Why?." (2009).
12. Janzik, Lars. "Contribution and participation in innovation communities: A classification of incentives and motives." International Journal of Innovation and Technology Management 7, no. 03 (2010): 247-262.
13. Gadiraju, Ujwal, Ricardo Kawase, Stefan Dietze, and Gianluca Demartini. "Understanding malicious behavior in crowdsourcing platforms: The case of online surveys." In Proceedings of the 33rd annual ACM conference on human factors in computing systems, pp. 1631-1640. 2015.
14. Zheng, Haichao, Dahui Li, and Wenhua Hou. "Task design, motivation, and participation in crowdsourcing contests." International Journal of Electronic Commerce 15, no. 4 (2011): 57-88.
15. Difallah, Djellel Eddine, Michele Catasta, Gianluca Demartini, Panagiotis G. Ipeirotis, and Philippe Cudré-Mauroux. "The dynamics of micro-task crowdsourcing: The case of amazon mturk." In Proceedings of the 24th international conference on world wide web, pp. 238-247. 2015.