# PROJECT PROPOSAL

*HCiQuery* — A ChatGPT VSCode Extension

*by*

*Theodore Glenn*

*Darshil Patel*

*Ali Eskandary*

*Alyssa Townsend*

*Brian Gunsel*

## Table of Contents

**Project Abstract**

HCiQuery is a query analysis software that includes a plugin for VSCode that allows for the saved interaction between user and ChatGPT and also the saved real-time code that the user types. The goal of HCiQuery is to be able to look at and understand the best way of structuring a query in order to utilize artificial intelligence engines, such as ChatGPT, to the best of your ability.

This project will have four main stages. The VSCode plug-in enables the capturing of user data, the database to grab and store the query data, an in-depth analysis of the extracted data, and finally, a way of visualizing the data and explaining the findings to both researchers and developers. At a high level, the user will dynamically interact with ChatGPT while simultaneously typing code in the editor, while in the background all of their queries and written code will be saved into the database to be later analyzed by a researcher, developer or data scientist.

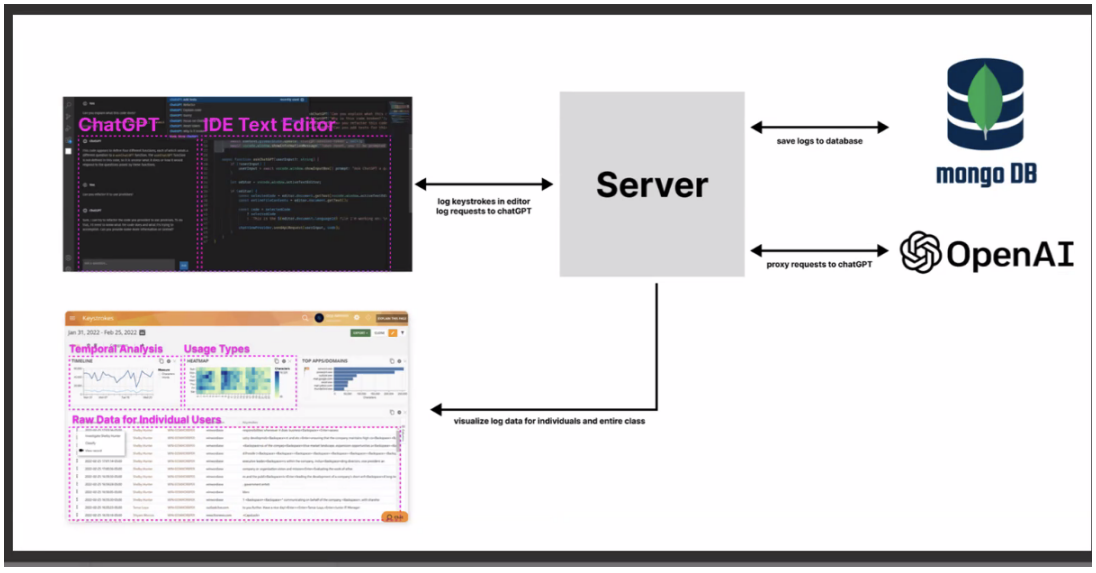
**Conceptual Design**

The plug-in will be developed using Typescript and Javascript, CSS, and HTML and the database will be hosted remotely with Firebase. The VSCode plug-in will be an interface located next to the user's code editor that will allow for the interaction between user and ChatGPT remotely. Users will be able to type in anything they please, whether it be code to be reformatted or a question regarding anything computer science related. Upon entering a query, the plug-in will use a token to remotely connect with the ChatGPT API. Once authenticated, the query will be processed with ChatGPT and sent back to the plug-in, which will then be displayed to the user within the contents of the editor.

In addition to the response being displayed to the user, the timestamped response *and* query will both be sent and stored in a Firebase database. This database will be organized by time, code queries, question queries, and more.

Before we can understand our data, it is important that the data be engineered and analyzed in a way that can begin to draw insights before it is visualized. This step will involve a lot of data science methodologies for goals such as combining common groups of words, most frequently used words, and more.

The last stage, visualization, is when all of the previous steps come to fruition. Researchers and developers will be able to accurately visualize how ChatGPT is used, and how people can benefit from writing better written queries to help their efficiency at work and in academia.



**Proof of Concept**

Link to ChatGPT VSCode Extension Project Repository

**Background**

This ChatGPT extension for VSCode was developed by Tim Kmecl. It's on the VSCode marketplace, and open-source on Github. That repository was forked into this one by the Temple HCI Lab, which is the baseline for this project. The changes made by Temple HCI Lab allow for better data collection by restricting ChatGPT's memory - this forces users to continuously refine their queries and ask again instead of relying on ChatGPT's memory on the conversation history.

**Required Resources**

*Developers for this project* require:

- TypeScript/JavaScript IDE, preferably VSCode
  - VSCode has an extension tester window
  - TypeScript compiler; can install with npm
  - Here's a tutorial for setting VSCode up for TypeScript development
  - install @types/vscode, @types/node --save-dev, and chatgpt modules via npm
- Firebase
- ChatGPT API key

*A user of the extension* would need VSCode; the HCiQuery extension, which they would download via the VSCode extensions tab; and a ChatGPT API key.

**HCiQuery Project Vision**

**For** researchers **who** are interested in human-computer interaction and finding how AI engines, such as ChatGPT, affect the coding and querying process, **the** HCiQuery **is an** extension for VSCode **that** allows programmers to query ChatGPT directly in the IDE. **Unlike** other ChatGPT VSCode extensions, **our product** logs and helps analyze timestamped query and coding data in order to better understand how to best utilize AI with effective and efficient queries.

**Personas**

Machine Learning Expert (Darshil)

Alex, 35 years old, is a Machine Learning Expert at Google in Mountain View, CA. He was born and raised in San Jose, CA and went to Stanford University for a Bachelor's in Software Engineering and PhD in ML. He works primarily on developing algorithms and models that can learn from data and make predictions or decisions without being explicitly programmed. He plays a crucial role in various industries and applications where machine learning can be leveraged. Currently, he is researching the satisfaction level of users with the use of Machine Learning models. ChatGPT VS Code Extension collects data on how satisfied users are with Chat GPT responses by storing things like “Above answer is not correct” or “above information is outdated”, etc. Our project will be very helpful for Alex’s research in manipulating all those data. As a bonus, he will also be able to track how much time the user is coding in VS Code and how much time he is using Chat GPT.

Teacher (Alyssa)

Emily, age 45, was born and grew up in Southeastern Pennsylvania. She worked for several different startups, creating web and mobile applications to simplify business administration, before going back to school part-time to learn pedagogy.  Now, she teaches software design to college students.

Her employment history means that she is very familiar with various stages of the software design process, and she is used to and confident in seeking out answers when she becomes stuck during the programming process. It has been decades since she was in the position of her students. However, she has neither the time nor inclination to go over raw data. So, she is interested in using the ChatGPT VSCode plugin’s data visualization to find out what parts of projects her students get stuck on and which they are confident in. This will help her determine which aspects of her lessons are most successful and what topics she needs to focus on reviewing.

Software Developer (Theo)

Bruce, a 29-year-old software developer, recently quit working at a large tech firm to begin working with some of his friends on a small startup company out of Austin, TX. Born and raised in Texas, Bruce has a B.S. in Computer Science from the University of Texas. He’s been in industry since graduating with most of his experience building web and mobile apps using languages like JavaScript, Python, Java, and Swift. Now, he and his colleagues are trying to figure out how to integrate the rising emergence of machine learning, AI, and natural language processing into their own cutting edge software service. Bruce and his team are fascinated by the rapid advances in AI and believes it could be a game-changer for their startup. The analysis of the ChatGPT Visual Studio plugin on how users on their team utilize the product will greaten his

understanding for how to implement features and write effective code. He thinks scrutinizing the querying data from the ChatGPT plugin can have huge potential to augment their team of developers, boost productivity, while streamlining repetitive tasks, free up developers' mental energy, and enable them to pinpoint focus on the most vital tasks. The plugin data could help their small team punch above their weight and ship their product faster because they can see what are the biggest drawbacks or pitfalls in maintenance.

Researcher (Brian)

Samuel Olegario, 31, is a senior researcher at Caltech in Pasadena, California. Samuel holds a Bachelor of Science in Computer Science and Mathematics from Temple University and a PhD from the University of California, Los Angeles with a concentration in Machine Learning. Samuel's research at Caltech is primarily focused on natural language processing, text mining, and how AI can accurately classify data using predictive algorithms. In Samuels research group, he and six other researchers tackle some of the biggest questions in machine learning and data science.

Samuel's use of the ChatGPT analysis software will allow him to gain extensive information on how people not only use ChatGPT, but also how people write clean and efficient code. Allowing the researchers under him and perhaps students in his undergraduate and graduate classes access to this software will allow him to harvest important data to be used in his and his teams further research of language classification and text mining. His work will be greatly optimized to allow new research to be conducted, which would greatly push discoveries in understanding the future of natural language processing and how we can best use ChatGPT and other forms of artificial intelligence to the best of our ability. The features he will get best use out of will be the visualization of most common words and queries, and if users use the plugin mostly for code editing or general programming questions.

Student (Ali)

Reginald, age 20, is a student at Temple University studying Computer Science. Reginald is a 20-year-old Computer Science student who is passionate about coding and software development. He's currently enrolled at Temple University, where he's working towards his bachelor's degree in Computer Science. He has a part time job at Pizza Hut. His interest in coding sparked during his high school years when he participated in various coding competitions and hackathons. Reginald is tech-savvy and has a decent amount of experience with programming, but lacks professional experience. In terms of skill, he's no better or worse than the average Computer Science major. Reginald's long-term goals include becoming a software engineer, hobby video game developer, and DJ. Challenges: Learning Curve: Reginald often encounters complex coding problems and faces a steep learning curve. He's determined to expand his knowledge but sometimes feels overwhelmed by the vast world of programming languages and concepts. Time Constraints: As a full-time student, Reginald juggles a demanding academic schedule with coursework, assignments, and exams alongside his extremely stressful job at Pizza Hut. Finding enough time to explore coding projects can be challenging. Lack of Guidance: While he has access to professors and academic resources, Reginald sometimes

struggles to get immediate and practical help with specific coding issues. He often wishes for a quick source of assistance to overcome roadblocks in his coding journey. Reginald uses the ChatGPT VSCode plugin to seek assistance and guidance with coding-related issues. When he encounters challenges while working on coding assignments or personal projects, he inputs his questions and problems into the plugin. The plugin provides output from ChatGPT. This output helps Reginald break down complex concepts and saves him time by giving instant feedback to basic questions.

**Feature List**

1. User can install & use plugin
2. Plugin tracks interaction with ChatGPT
3. Plugin tracks user's coding
4. Can access timestamped raw data of interaction with ChatGPT and user's coding
5. Data visualization

```
import * as fs from 'fs';

private writeToCSV(userInput: string, response: string) {
    const timestamp = new Date().toISOString();
    const csvData = `"${timestamp}","${userInput}","${response}"\n`;
    const filePath = 'data.csv';

    fs.appendFileSync(filePath, csvData, 'utf8');
}

// modify existing search() function
public async search(prompt?: string) {
    try {
        response = res.text;

        this.writeToCSV(searchPrompt, response);
    } catch (e: any) {
    }
}

```