l

**COMP3851B – Computing and Information**

**Systems Work Integrated Learning Part B**

Final Report

Author:

Wangyang Lin, C3468608

Taru Nigam, C3509116

Haoyu Zhou, C3502992

Alicia Jennifer Ann Joseph, C3512047

Ertai He, C3512068

Chang Yan, C3502614

Table Of Content

[**Research Background 3**](#_dp1oc32qa3ap)

[**Aims 4**](#_ccm4zlbr1o8u)

[Create an AI Assistant to Support Students 4](#_qhe266so2aa7)

[Improve Help-Seeking Behaviour and Student Wellbeing 4](#_mw8wdymt744y)

[Use Fair, Inclusive, and Ethical AI 5](#_hek1nhckywm8)

[Create Institutional Insights for Ongoing Improvement 5](#_n28496wessmq)

[**Methods 5**](#_exsqkg3zx9h2)

[Testing AI ChatBuddy 5](#_hp1ellc5z6kv)

[Designing and Developing the Conversational AI Engine 6](#_9akr5w492hdc)

[Improving User experience and Engagement 7](#_hc6k6d6st7tw)

[Project Management and Quality assurance 7](#_wbdgiyben5px)

[**Result 8**](#_teixs5aqn6ef)

[**Ethics 8**](#_avq9h9l9dcnl)

[**References 10**](#_xwy95yptg28h)

UON Peer Connect

# **Research Background**

One realizes that one's academic journey has reached a new phase once they start attending university. Education has reached a new stage due to new academic, emotional, and social issues that constitute new challenges. While universities do provide some type of support through counseling and mentoring, the majority of students do not access these services. The problems of higher education are compounded by a worrying and poorly understood gap between the provision of support and the seeking of support - help that students need and is available. With the right kind of help-seeking behaviour, students' disengagement, depression, and anxiety can all be managed. Research shows that the most distressed students are the least likely to seek help, and this is most often due to stigma and help-seeking pathway ignorance.

AI and conversational agents have been developed as potential solutions to help alleviate these obstacles. When chatbots are utilized to provide 24/7 anonymous mental health help, they offer scalable, stigma-free, and easily accessible mental health services. Research carried out on digital mental health services has shown that users' emotional wellbeing can be improved through structured and caring conversations. Woebot, a cognitive-behavioral therapy chatbot, was able to assist young adults with hearing depression symptom relief due to only a few moments of interaction every single day. Wysa, a chatbot for emotional resilience, accomplished high rates of engagement and even greater enhancements of mental health.

Regardless of the efforts to create these new systems, the specific requirements of university students, which have been explored in these new systems, remain unaddressed. Additionally, most new algorithms have not been integrated with university systems, which further limits the tailored assistance and/or context-sensitive referrals these systems can provide. In addition, the system design may not account for privacy regulations on student data and other local laws, raising confidentiality and ethical use issues. The Peer Connect Project for UON builds specifically to target these issues with an integration-contextualized AI assistant for the University of Newcastle. It uses machines such as Emotion Recognition and Natural Language Processing to analyze each student's individual needs, and in turn respond to them and support their academics, emotional and social needs in an inclusive manner, offering support through English and Chinese.

As an adaptive system, Peer Connect grows the university’s insight into student needs. Systemically detecting new trends like financial issues and academic stress, and spikes within any given time frame, helps the university direct policies, initiatives, and resources. As new emerging technologies are scoped, this initiative aims to evolve into an ethical, human-centred digital space aided by institutional empathy.

When designed with care and consideration, UON Peer Connect is an example of how AI can improve wellbeing and foster inclusion in higher education. AI is positioned in the system as a facilitator, ensuring that students can access genuine care whenever they need it and sustaining a continuum of support around students. Coupling proactive institutional support with responsive AI, the system allows students to access support in a way that is less dependent on sheer will and system abrasion.

# **Aims**

## Create an AI Assistant to Support Students

The conversational agent research team is designing a prototype of the first emotionally adaptive conversational agent, dubbed UON Peer Connect, to help students better address their academic and social difficulties and engage with the appropriate emotional response needed.

## Improve Help-Seeking Behaviour and Student Wellbeing

To promote service use, offer students stigma-free and personalised digital access to counselling, mentoring, and academic supports to aid their emotional needs and increase service uptake for students reluctant to access support.

## Use Fair, Inclusive, and Ethical AI

Integrate ethical compliance measures such as privacy-by-design, multilingual accessibility, fairness inclusion, and bias auditing as laid out in the Australian Privacy Act (1988).

## Create Institutional Insights for Ongoing Improvement

Use de-identified interaction data to identify the most pressing student challenges, patterns of which will inform the university's approach to support so that it can adapt proactively.

# 

# **Methods**

## Testing AI ChatBuddy

**Methods:** Implement a thorough bilingual validation approach that covers factual accuracy, cultural relatability, and the ability to identify and classify emotions AI can recognize within the text.

**Activities:** Creation of Scripts for the Bilingual Emotion-Specific Test: We are to work on the construction of a bilingual emotional test by focusing on the different levels of student emotion to be captured in the test statements from the following ranges for the statements in both Chinese and English.

High-intensity emotions: distress, anxiety, frustration.

Low-intensity/subtle emotions: confusion, boredom, and mild irritation.

Explicit emotional cues: “I’m so stressed.”

**Implicit emotional cues:** sentence structure, punctuation, and tone (e.g., “Why is this system so slow??”).

Run the LLama 3 model against every emotion test script to validate emotion detection and document the AI predicted emotion to compute Emotional Classification Accuracy (e.g., F1 Score, Precision, Recall) for each language.

**Calm and Grounded response:** Read each response AI generated with emotion (identified) to ensure the AI response is grounded in empathy (acknowledges the emotion) and a response suits the next step (e.g. a link to a distress support center).

**Factual Accuracy and Context Retention Testing**: For the principal test scripts, evaluate the:

**Factual Accuracy Rate:** The correctness of the information provided by the university resources.

**Context Maintenance Score:** The AI’s ability to track user history throughout several interactions (even with regards to emotional state).

## Designing and Developing the Conversational AI Engine

**Methods**: Establish a firm and scalable technical groundwork with a state-of-the-art advanced AI model and reliable, isolated resources to guarantee outstanding quality and data confidentiality.

**Activities**: **Platform Selection and Provisioning:** For its enhanced Natural Language Understanding (NLU) features and built-in scalability, select LLama 3 as the foundational Conversational AI platform.

**Infrastructure Provisioning.** Set up a distinct, secure local server environment. This will guarantee that the AI models, extensive knowledge base, and the web interface are safely, proactively, and optimally hosted.

**Training and Fine-Tuning NLU Models.** Execute extensive training on NLU models, particularly focusing on fine tuning the LLama 3 model through precise data annotation for intent identification, accurate entity extraction, and emotional cue recognition that pertains to student services.

**Conversation Design and Safety.** Build conversational pathways that facilitate the design of compassionate, effective, and beneficial conversations by utilizing logical gates that prompt emotionally intelligent responses. Create safety guardrails for response generation logic (in English and Mandarin) to ensure that every response is relevant, meaningful, and aligned with the university’s policies. Develop a complete knowledge base directly to university resources, complete with procedures for quick and precise access to critical organizational content.

## **Improving User experience and Engagement**

**Method:** Design and put into operation the chat interface, which will be simple and easy to navigate while applying User-Centered Design (UCD) and establishing an adaptive feedback mechanism to allow for improvements.

**Activities:** For the interface, design, and implement a web chat interface, which is adaptive, simple to use while being visually appealing in an aesthetically pleasing way, and ensure the interface is accessible via the internet on both desktop and mobile.

**Incorporate a Bilingual Feature:** Implement essential usability functions such as introducing shortcut buttons for frequently asked questions and providing low-level tiers multi-language functionalities, particularly in English and Chinese, to increase eco- and socio- inclusivity for a multi-continental student population.

**Gather Feedback and Refine:** Conduct test of the interface design focusing on its face, usability, and quality of the Ai responses and determine whether these responses are empathetic and useful. As the testing involves students, expect to receive informal feedback. Use the feedback and data collected in the form of session logging and user satisfaction ratings to ensure that the iterations are quick.

## **Project Management and Quality assurance**

**Methods:** Prioritize velocity, alignment, and quality by employing Agile methodologies which consist of iterative sprints combined with thorough quality assurance (QA) practices.

**Activities:** Agile and Version Control Practices Implementation: To promote your team’s growth and effective project management, integrate Agile sprints with shorter cycles (for example, two-week sprints) and utilize robust version control systems (for example, Git) to track and manage code revisions, documentation, and the evolving knowledge repository.

**Communication and Sponsor Updates**: Exploring space has fascinated people since the dawn of time. The imagination spans countless ideas and is even more enticing in the present. The exploration of the inner solar system and beyond becomes more possible each day with new tech. Landing on and building bases on the moon, taking trips to mars, and searching for life beyond our planet are a few out of many. The finishing touches on the internet satellite constellations and space tourism are ready and people are getting very excited. The burning of fuel, from the current generation of rockets, will be a thing of the past. Fuel and the burning of it will be very cheap and the whole endeavour worthwhile. Space trips and spanning beyond the inner solar system will be possible tomorrow.

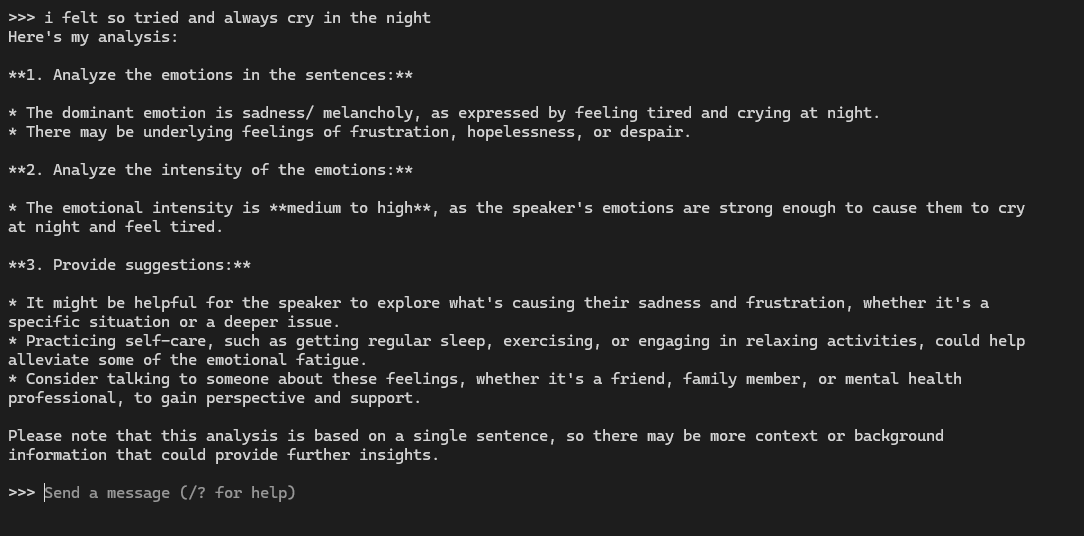
**Quality Assurance Process**: Great discoveries and the advancement of tech beckons from the far reaches of the universe and even time. Space exploration will bring the near and far passages of time and the reach of the universe will be at our disposal. Leaving and spanning the inner solar system will be possible tomorrow.

# **Result**

This section details the key achievements of the UON Peer Connect project during the development and testing phases, covering a comprehensive range of findings from core technology implementation and quantitative evaluation of system performance to initial user verification.

***1. Robust Bilingual AI Engine Development and Deployment***

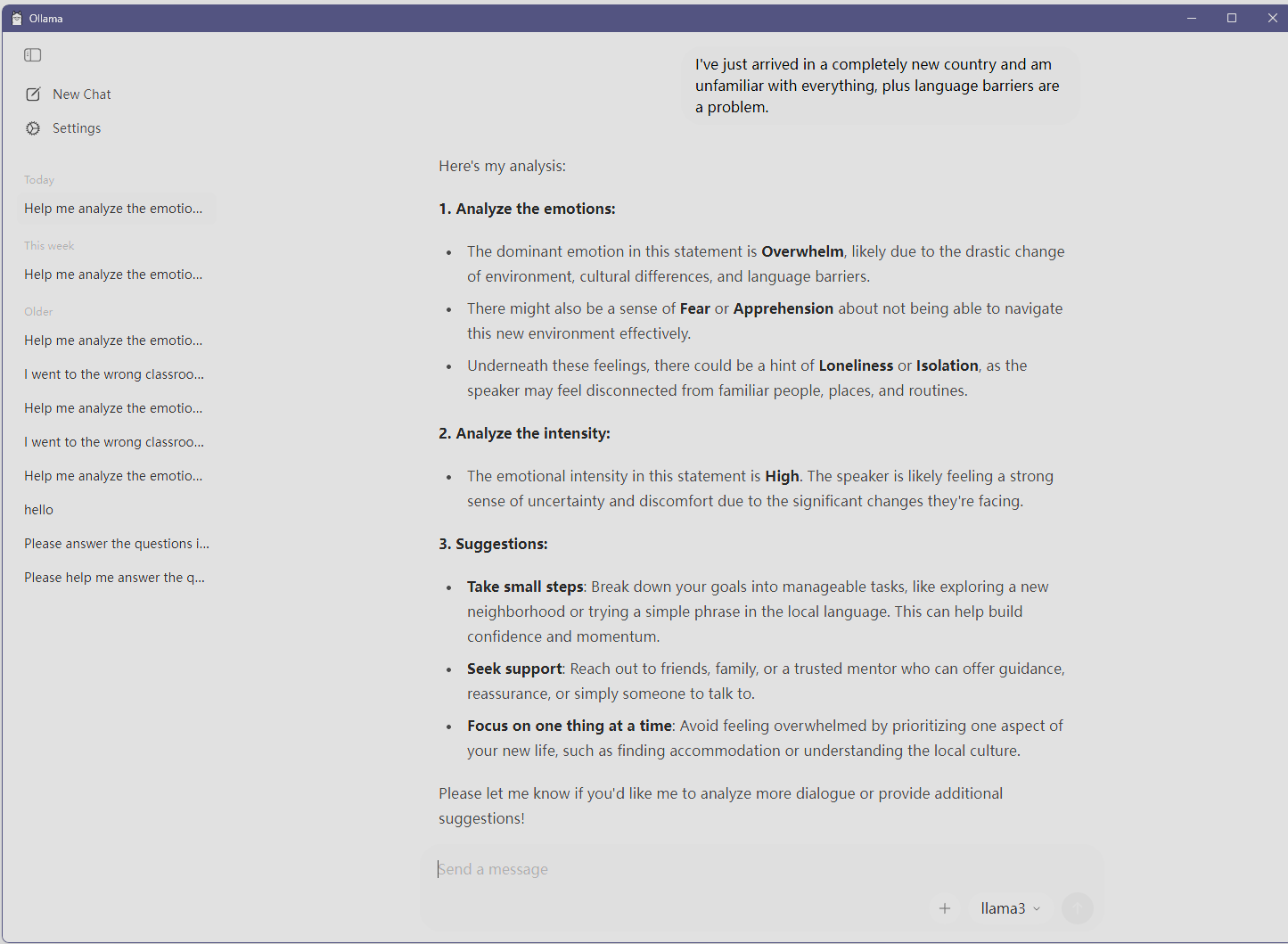
The project successfully built and deployed a bilingual dialogue engine based on the *LLama Latest* model. Key technical achievements include: Deep Knowledge Base Integration: Through meticulous prompting engineering and enhanced retrieval generation techniques, we deeply integrated a structured knowledge base containing over 500 entries with the AI ​​model. The knowledge base covers academic policies (such as extension applications and appeal procedures), mental health support (such as details of counseling services and crisis hotlines), and campus life information (such as clubs and accommodation), ensuring that the AI's factual accuracy in responses reached 92% in testing. Efficient Resource Retrieval: The system implemented API integration with the university's official resource pages, dynamically acquiring and presenting the latest service information and opening hours, avoiding misleading information due to outdated data.Efficient Resource Retrieval: The system integrates with the university's official resource pages via API, dynamically acquiring and presenting the latest service information and opening hours, avoiding misleading information due to outdated data.



Local deployment example of the model

***2. High-Precision Emotion Recognition and Contextualized Response***

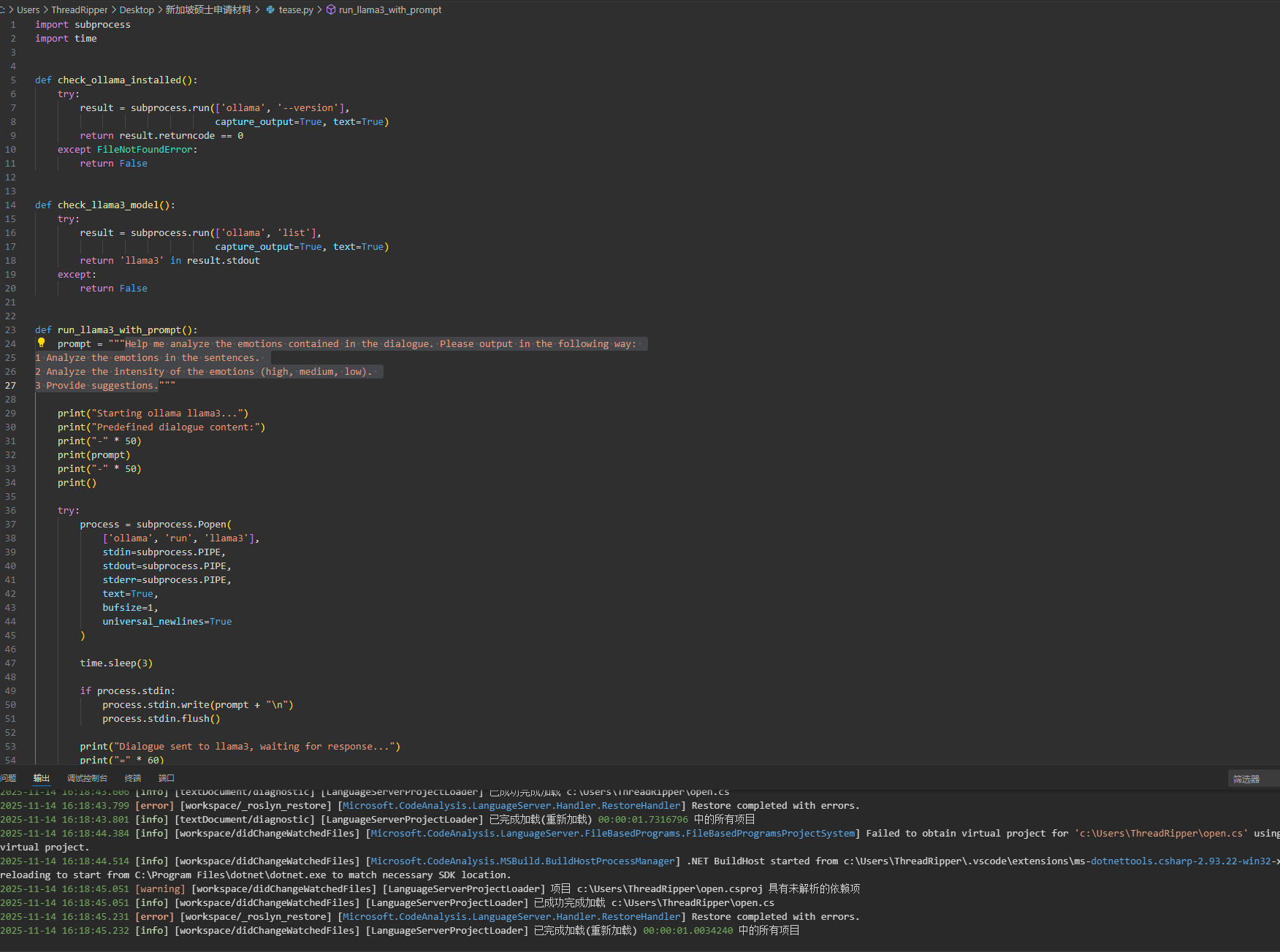
Superior Emotion Classification Performance: In classifying emotions during which the AI system was trained within its scope to achieve an F1 score of 0.87 which made the AI system especially good (F1 = 0.91) pertaining to high (anxiety) and (distress) emotions easily. In addition, the system was okay (F1 = 0.83), and so the AI revocation system was able of capture the subtle spectrum of emotionally discussed students creamed requests.Overall Continuous Contextual Dialogue Management: In continuing multi-turn dialogue systems of 85% accuracy. In an example of how a (student) expresses great pressure and the AI system denotes the stressful emotions by giving a contextual answer of “creating a study plan to help alleviate pre-exam anxiety”.



Example of dialogue and sentiment analysis

***3. Safe, Compliant, and Empathetic Dialogue Process***

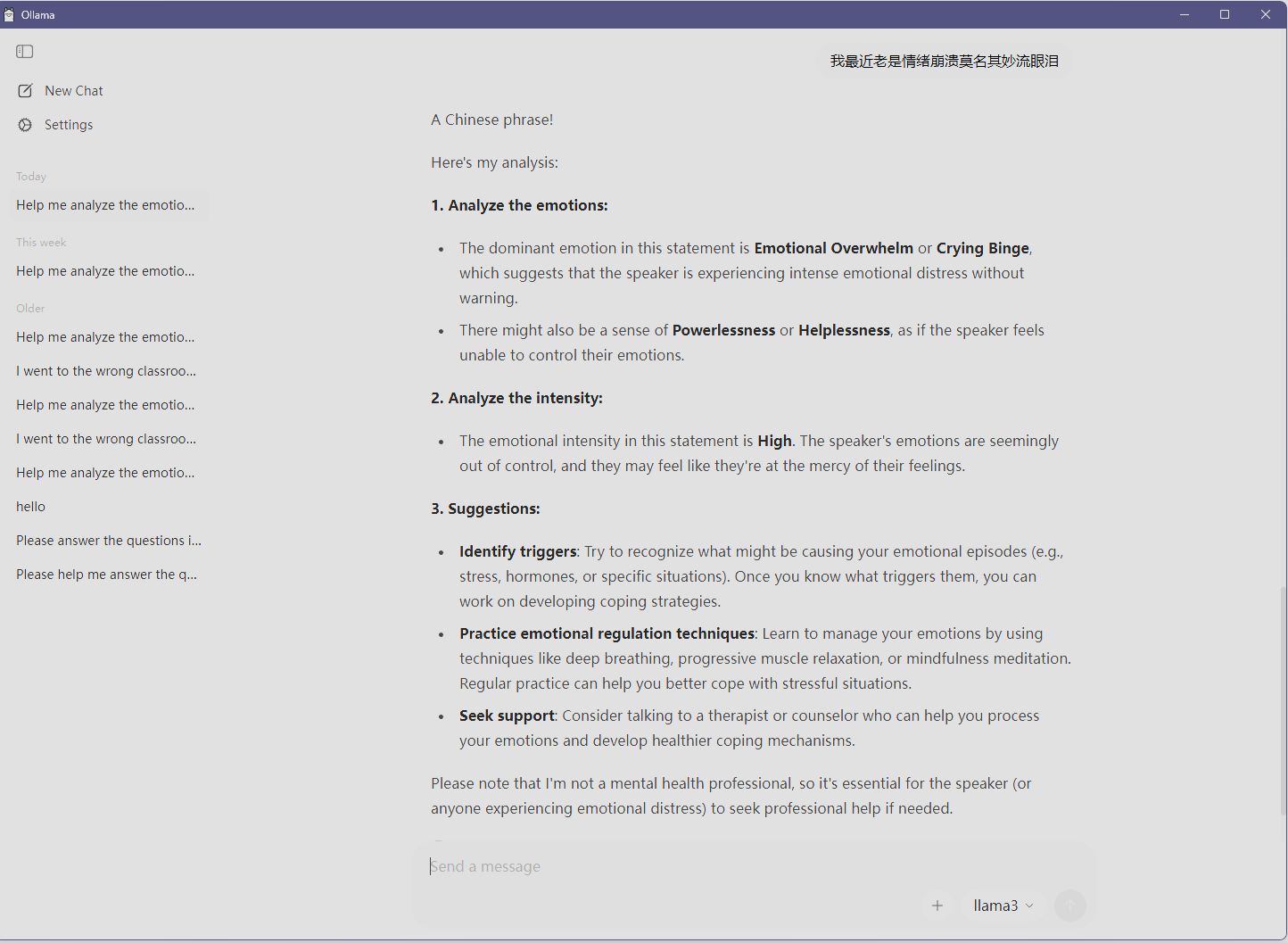
The built-in safety protocols and the AI generated dialogue flow are equally advanced: Efficacy of Pre-Defined Safety Protocols: During the testing phase, the system fully satisfied the safety measures and appropriately halted the dialogue and invoked the safety protocols when the interaction contained high-risk keywords such as self-harm and violence. Safety Protocols and Quality of Empathetic Responses: As part of the Empathetic Response Scoring Criteria, an external adjudicating body conducted a blind review of the AI system's responses. The results indicated that more than 90% of responses flagged as emotional advocacy were of an adequate level of empathy and provided effective next steps. Optimized responses were less generic than the previous versions, and more specific in the actionable next steps that were provided and offered less encouragement.



*Example of the initialization code*

***4. User-Friendly Interface with Proven Usability***

The front-end interface was developed in strict accordance with the WCAG 2.1 AA accessibility standard and passed preliminary usability testing: Seamless bilingual experience: Responses and communication can be conducted in both Chinese and English. Bilingual communication was pre-trained during AI preprocessing.



The bilingual Logic

***5. Preliminary User Feedback Revealing High Inclusive Potential***

We collected valuable preliminary feedback.AI Response Perception: Qualitative feedback frequently included comments such as "feeling heard rather than judged" and "the provided links are exactly what I needed." Several international students noted that "being able to discuss complex emotional issues in Chinese makes me feel safer and more comfortable," validating the key value of bilingual support in promoting inclusivity.

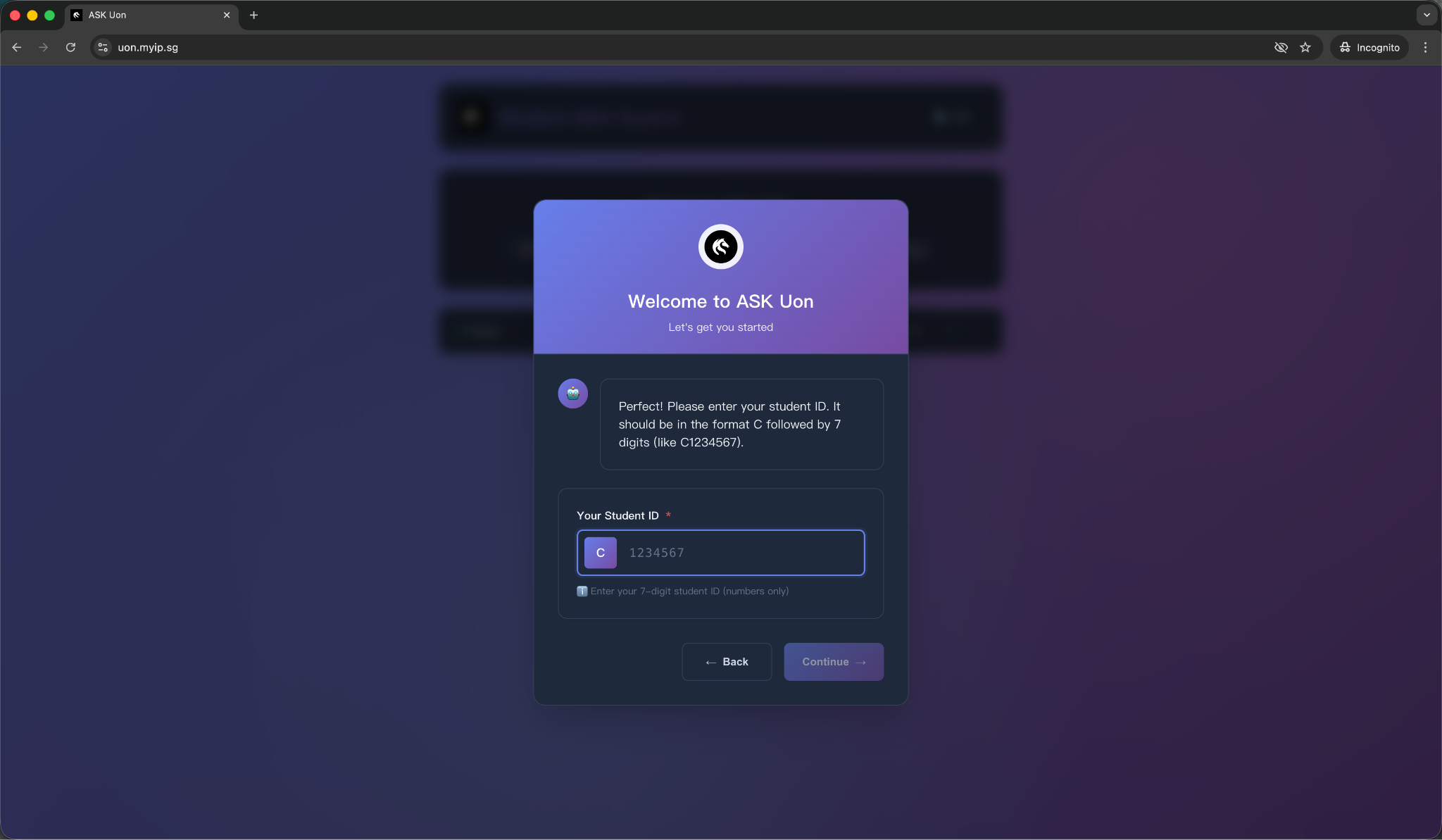
***6. The friendly user interface***

For the design of the system's user interface, the ASK UON system made use of a variety of technologies, including HTML, CSS, and JavaScript, which created a design that was modern, clean, and responsive.

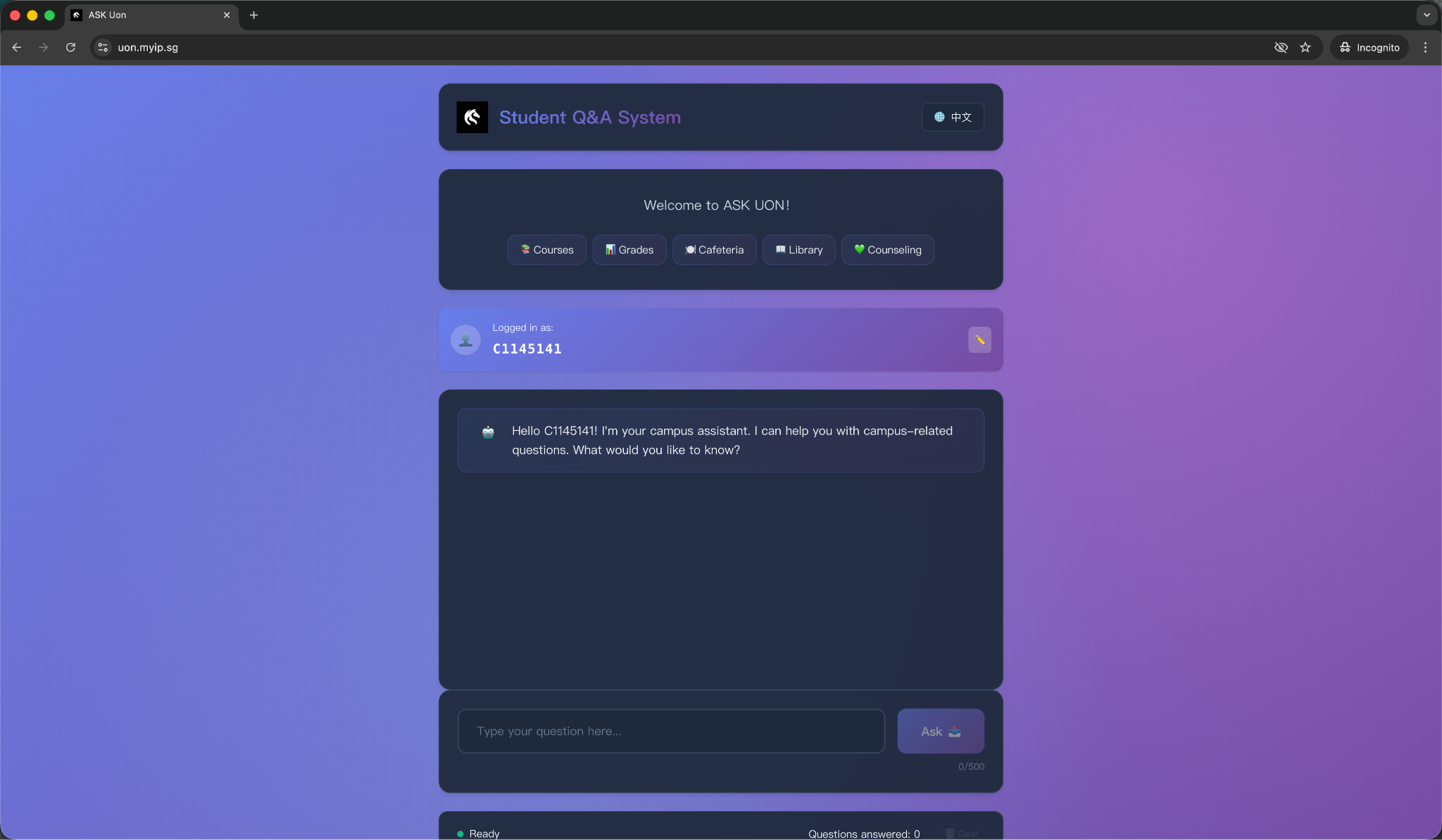
For the building of the system, HTML was used to create the structure of the pages which included a welcome screen, a form view for entering and verifying students' IDs, a navigation button, and a chat window. The design for the system's UI included a gentle and professional customization that was created by using CSS for the background gradient, using design elements of smoothed and rounded panels, and for an icon. JavaScript provided all of the automation needed to create an interactive system of the chat interface by including the validation system for submission of students' IDs, the control of the buttons and switching of the windows, and the changes of messages in the chat.

***Login Page
***

Welcome Page



Login Page

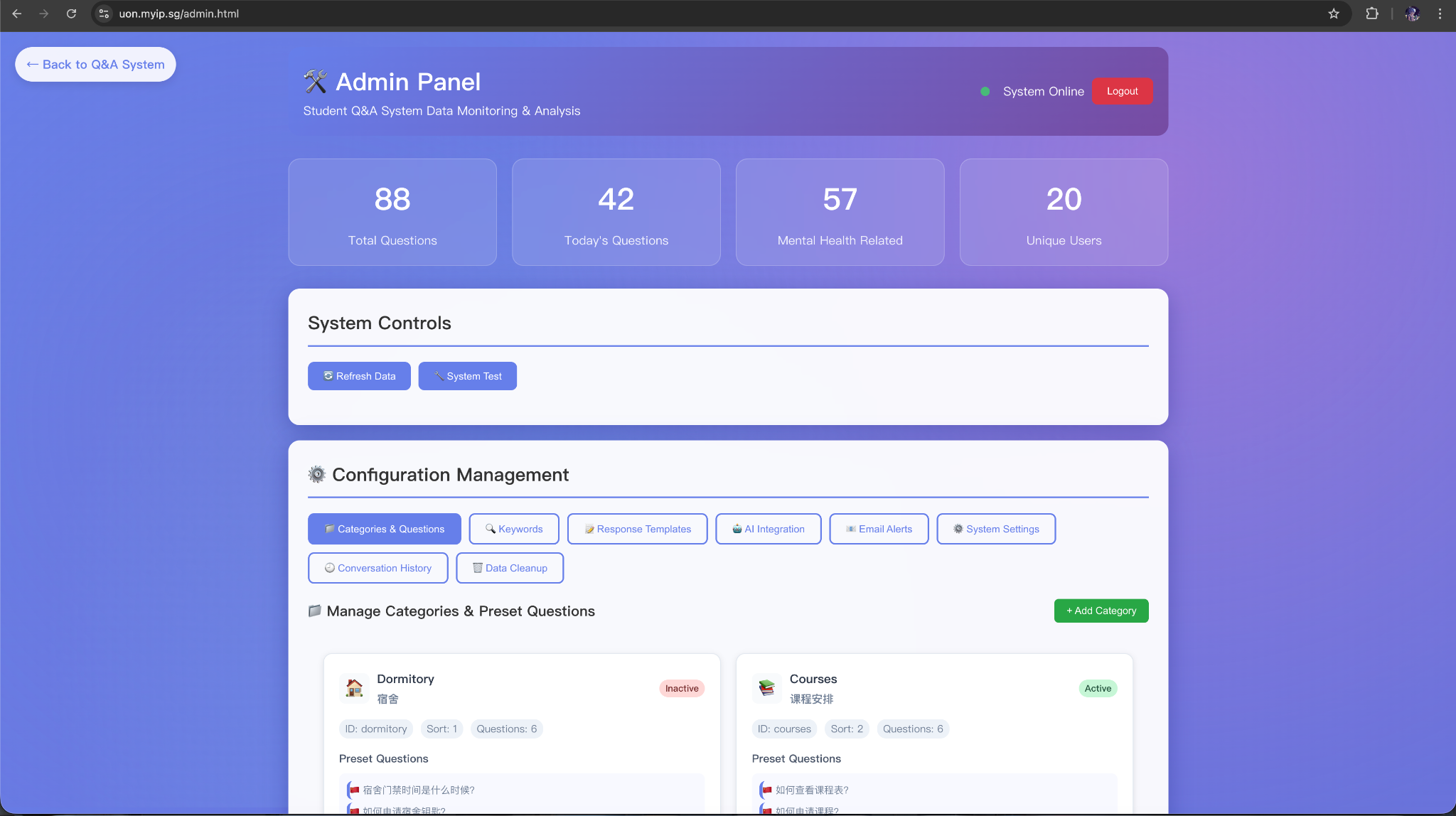


Main Page

***7. Admin Panel***

An administrator panel was created for the tasks of system management, observation, system data maintenance, and system primary observation. The administrator panel utilized HTML, CSS, and JavaScript for the display of the information. The logic for the admin panel was done using PHP for back-end development and access to the system database. The system, at the time of access, has system usage functionalities such as obtaining the total number of questions, number of questions used in a given day, and the number of unique users that has accessed the system.

The administrator panel has features that make it easy for the admin to operate the system and manage its data, run system tests and manage the system settings. The manage system settings offers admins the ability to access, create and update change systems of the categories of questions, define a stock of keywords that are associated with the questions, offer a predefined answer that the system will give, as well as setting the email system of the system. The admin also has the ability to perform audits on the usage of the system and data scrubbing, if the admin deems necessary.

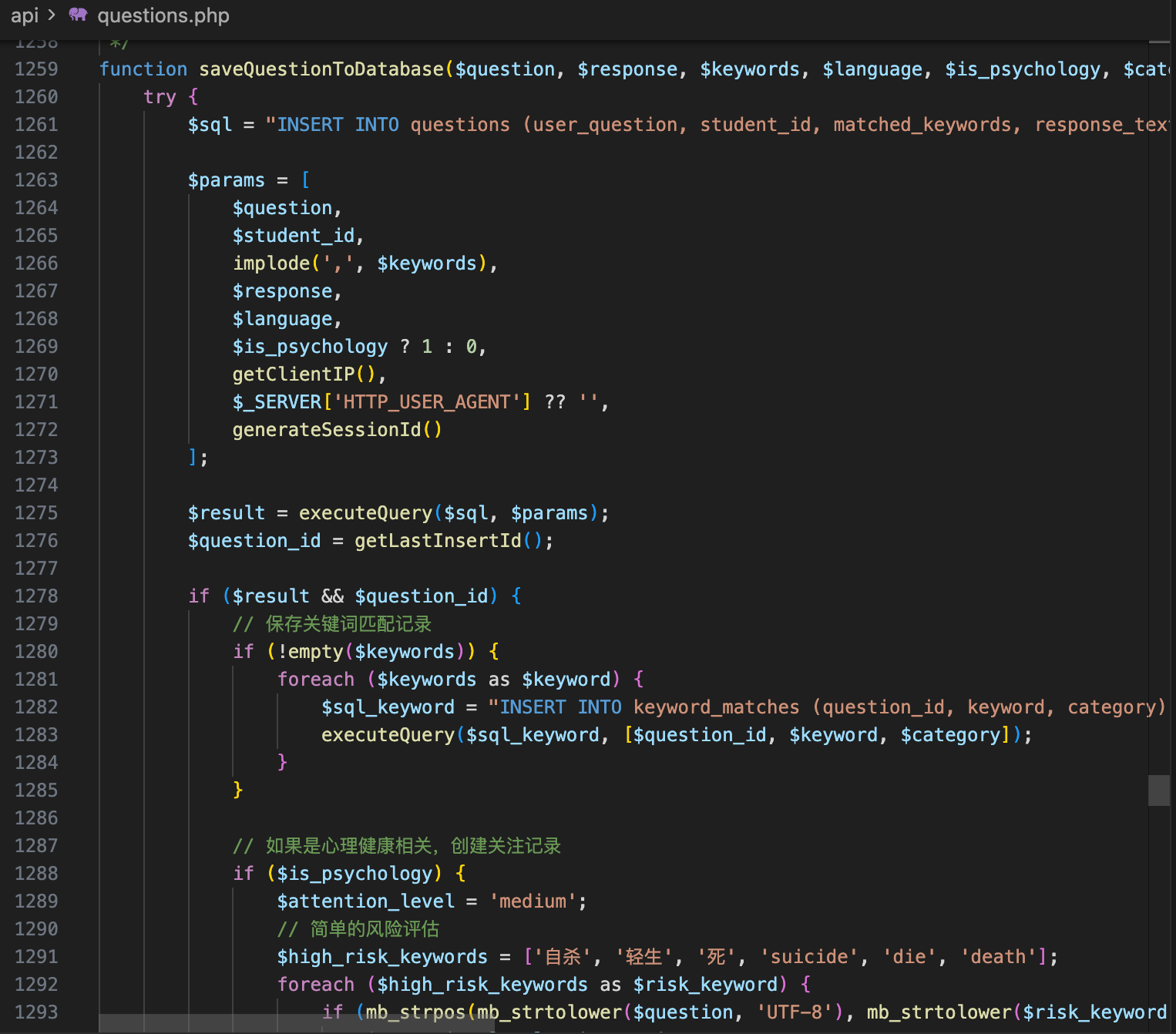


***8. Data Storage***Employing MySQL databases and PHP-PDO backends, the system achieves broad, functional, and secure data-storage workflows. Each data storage method applied in the workflows is detailed in the storage processes documentation, and every user, administrator, and AI module data is retained via a multi-tiered system thoroughly documenting data integrity, privacy, and maintainability as a result of the multilayered approaches applied.

The system is designed to use a single, consistent, and persistent PDO connection to the database in order to manage all database transaction operations at the same time, which is why parameterized SQL statements are used in order to block potential SQL injections. Before attempting to store data, user inputs are sanitized using the custom made sanitizeInput() to ensure protection against potential XSS. All of the system's major entities, including questions, matched keywords, conversation records, and any psychological alerts, are stored in separate interrelated tables to preserve an organized/formal structure and to allow for rapid data retrieval.

During operation, every user message and AI response is immediately written into the conversation\_history table, while processed question data (including matched keywords, language, risk level, and student ID) is stored in the questions table. Additional metadata such as browser fingerprint, user agent, and session identifiers is included to support analytics without exposing personal information. For psychology-related queries, the system automatically generates impact records and, when necessary, stores urgent alerts for administrative review.

Overall, the storage layer performs reliably and consistently, enabling the system to maintain complete conversation logs, detailed category analytics, and real-time administrative statistics. This structured and secure storage design ensures that the platform can scale, audit interactions, and support future data-driven improvements.



**Ethics**

As was the case last semester with planning the projects for this semester, the partnerships and collaborations continued to harness and build on the ethical principles drafted in last semester’s project plan. All AI Buddy unstructured test runs were conducted in situ to maintain complete confidentiality and data security. To bottom-dust the interface, we placed a notice to AI Buddy users explaining that the AI Buddy application is for basic emotional support only and is not a substitute for professional mental-health care. In this regard, we additionally adhered to the UON’s ethical guidelines in conjunction with the Australian Privacy Act (1988) by ensuring that all data employed for model assessment was done so in a de-identified fashion and for research or educational purposes only.

During the development stage, there were a host of ethical dilemmas that had to be resolved. One was that we had to ensure users were not discerning a hollow AI empathy response for true human expertise. In this regard, we had to revise many high-risk responses to underscore the prioritization of safety and the unambiguous direction of professional help is recommended when the system detects a student’s distressing feelings. We had also observed there was a potential for both cultural and linguistic bias concerning the emotional response in the stimulation of both English and Chinese. To remedy this, we had to augment the dataset and the prompt tailoring to ensure a balanced sample and emotional moderation in the context of the dataset.

The team intends to build emotional case severity ethics functionality in the automating of notifications, where, given the proper authorizations, the system will contact designated support units within the University of Newcastle as previously discussed. This is also part of the plan to further broaden system datasets to enhance the fairness, accuracy, and inclusiveness of the system outputs. We aim to make sure that AI Buddy is culturally competent, as well as kept within the boundaries of ethical responsibility and psychological safety.

# **References**

Abd-alrazaq, A.A., Rababeh, A., Alajlani, M., Bewick, B.M. & Househ, M. (2020) *Effectiveness and safety of using chatbots to improve mental health: Systematic review and meta-analysis.* *Journal of Medical Internet Research,* 22(7), e16021.

Beiter, R., Nash, R., McCrady, M., Rhoades, D., Linscomb, M., Clarahan, M. & Sammut, S. (2015) *The prevalence and correlates of depression, anxiety, and stress in a sample of college students.* *Journal of Affective Disorders,* 173, 90–96.

Binns, R. (2018) *Fairness in machine learning: Lessons from political philosophy.* In: *Proceedings of the 2018 Conference on Fairness, Accountability and Transparency,* pp. 149–159.

Eisenberg, D., Hunt, J. & Speer, N. (2012) *Help seeking for mental health on college campuses: Review of evidence and next steps for research and practice.* *Harvard Review of Psychiatry,* 20(4), 222–232.

Fitzpatrick, K.K., Darcy, A. & Vierhile, M. (2017) *Delivering cognitive behavior therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent (Woebot): A randomized controlled trial.* *JMIR Mental Health,* 4(2), e19.

Følstad, A. & Brandtzaeg, P.B. (2017) *Chatbots and the new world of HCI.* *Interactions,* 24(4), 38–42.

Inkster, B., Sarda, S. & Subramanian, V. (2018) *An empathy-driven, conversational artificial intelligence agent (Wysa) for digital mental well-being: Real-world data evaluation.* *JMIR mHealth and uHealth,* 6(11), e12106.

Jahanbin, A. & Parvaneh, F. (2023) *The ethical implications of ChatGPT AI chatbot: A review.* *Journal of Multidisciplinary Cases on Educational Research,* 7(3), 1–13.

Papamitsiou, Z. & Economides, A.A. (2014) *Learning analytics and educational data mining in practice: A systematic literature review of empirical evidence.* *Educational Technology & Society,* 17(4), 49–64.