

# Project Proposal

---



## Title Page

---

- **Title of Project:** Automated Note-Taking and Summarization Solution for University Lectures
- **Sponsoring Company:** Nanyang Technological University
- **Team Name:** Magnificent 7
- **Team Members:** Team Information
- **Date:** 14 February 2024

## Executive Summary

---

Nanyang Technological University (NTU) is addressing the substantial challenges presented by its traditional approach to lecture note-taking, which is both expensive and inefficient. The reliance on student note-takers has led to inconsistencies in note quality due to variations in accuracy, completeness, and the note-takers' understanding of the material. This system incurs significant financial costs and does not scale well to the volume of lectures at NTU.

To resolve these issues, NTU is proposing an innovative solution that leverages state-of-the-art speech-to-text and AI technologies to automate and improve the note-taking process. This comprehensive strategy includes:

### Automated Note-Taking

The introduction of advanced speech recognition technology to automatically convert lecture audio into precise transcripts, supporting a range of audio and video formats for both live and uploaded lectures. This aims to enhance the accuracy and reliability of lecture documentation while reducing manual labor and associated costs.

### Transcript Management Database

The development of a sophisticated database system for the effective storage, retrieval, and management of lecture transcripts. This system will feature an advanced search function, a user feedback mechanism for quality control, and a threshold policy for upload limits to ensure the maintenance of high-quality content.

### AI-Generated Summaries

The use of generative AI to create succinct summaries of lecture transcripts, customizable by difficulty level and length, and further enriched by the optional inclusion of lecture slides for context. This feature is designed to provide students with quick, tailored access to the core content of lectures, facilitating efficient revision and study.

### **Interactive Chatbots**

The implementation of AI-powered chatbots to simulate a personalized tutoring experience, enabling students to ask questions and receive accurate, contextually relevant responses based on the lecture materials. This initiative aims to enhance students' understanding and engagement with the material.

Despite the innovative aspects of this proposal, NTU recognizes the potential challenges and limitations inherent in such a technological undertaking, particularly the reliance on external APIs for transcription and summarization services, which may impact the quality of the outputs. Additionally, the manual input of metadata for the database could introduce human error and inefficiencies.

In conclusion, NTU's proposed enhancements to its academic support systems represent a significant advancement in educational technology, offering scalable, efficient, and user-friendly alternatives to conventional note-taking methods. By adopting these technological solutions, NTU aims to improve the educational experience for its students while also achieving operational efficiencies.

## **Table of Contents**

---

### **Contents**

---

#### **Title Page**

#### **Executive Summary**

#### **Table of Contents**

#### **Statement of Problem**

#### **Objectives**

Automated Note-Taking Process

Database Utilization for Transcript Management

Summary Generation with Generative AI

Interactive Chatbots for Enhanced Learning

Critical Design Issues, Constraints, and Limitations

#### **Technical Approach**

Front-End Development

Speech-to-Text Transcription

Summarization and AI Integration

Database Management

#### **Project Management**

Week 3: Introduction and Planning

Weeks 4-5: Sprint 1 - Initial Development

Weeks 6-7: Sprint 2 - Feature Development and Integration

Weeks 8-9: Sprint 3 - Advanced Features and Testing

Weeks 10-11: Sprint 4 - Finalization and Testing

Week 12: Documentation and Deployment

### **Deliverables**

Detailed Requirement Specifications

Physical Prototype

Detailed Description of Test Procedures

Computer Program Code, Flowchart, Documentation

System Components or Modules

Entire System

User-Friendly Instructions Including Training for Personnel

### **Budget**

API Costs

### **Communication and Coordination with Sponsor**

Communication Schedule and Responsibilities

Form of Communication

Information Distribution

### **Team Qualifications**

Nathanael Axel Wibisono (Project Manager, Lead Developer)

Sherwin Samson (QA Manager, QA Engineer)

Jiang Jiaxi (Back-end Developer)

Zhang Mengao (Release Engineer/Manager)

Yang Yida (Back-end Developer)

Liu Changsong (Front-end Developer)

## **Statement of Problem**

---

Nanyang Technological University (NTU) is facing significant challenges with its current approach to note-taking for academic lectures. The university employs student note-takers at a cost of \$18 per hour to document lectures across a wide array of subjects. With hundreds of modules being conducted on a weekly basis as part of approximately 140 courses, the financial burden on the university is substantial, amounting to tens of thousands of dollars each month. This traditional method of note-taking not only incurs a high financial cost but also lacks reliability and efficiency for several reasons:

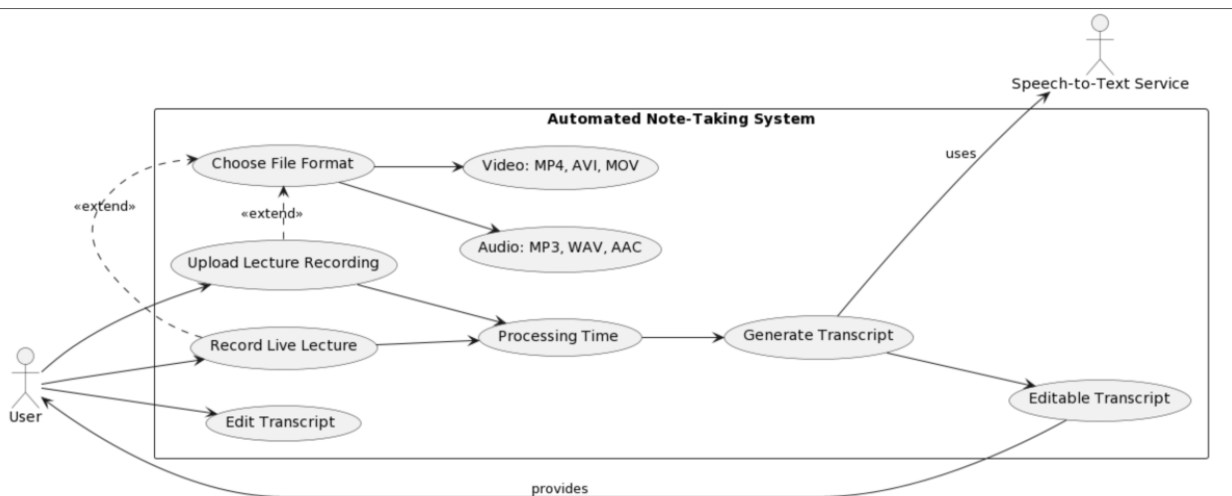
1. **Accuracy and Completeness:** There is no guarantee that the notes taken are accurate or complete. The quality of the notes can vary significantly depending on the note-taker's understanding of the lecture material, attentiveness, and note-taking skills. This variability can lead to inconsistent quality of notes, potentially affecting students' learning outcomes.

**2. Contextual Understanding:** The student note-takers, often undergoing the same course, may not possess the necessary depth of knowledge or context to capture the nuances and key points of the lectures accurately. This lack of expertise can result in notes that are superficial or miss critical information, making them less useful for study and revision purposes.

**3. Scalability and Efficiency:** The manual process of note-taking is not scalable to the volume of lectures conducted at NTU. It is labor-intensive and does not leverage technology to streamline the process, leading to delays in note availability and increased workload on student note-takers.

## Objectives

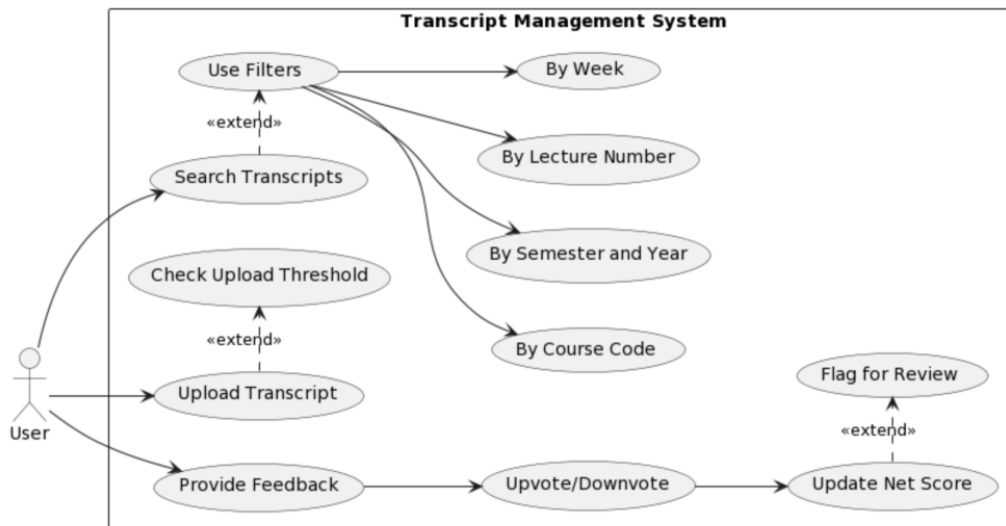
### Automated Note-Taking Process



Implement advanced speech-to-text technologies to automate the conversion of lecture audio into accurate transcripts. This automation will be facilitated through two primary methods: direct recording of live lectures and the option for users to upload existing lecture recordings. By leveraging cutting-edge speech recognition and processing capabilities, this objective aims to eliminate the need for manual note-taking and reduce associated costs, while ensuring comprehensive and accurate documentation of lecture content.

- **Facilitation Option:** Provide users with the versatility to submit content in either audio or video formats.
- **Wide Range of Compatible Audio Formats:** Transcripts should be able to be generated from three of the most popular audio formats: MP3, WAV, AAC
- **Wide Range of Compatible Video Formats:** Transcripts should be able to be generated from three of the most popular video formats: MP4, AVI, MOV
- **Processing Time:** Ensure that the time taken from the end of a lecture recording/upload to the availability of the transcript does not exceed 10 minutes for a 1-hour lecture.
- **Editable Transcript:** Enable users to interactively edit transcripts, ensuring they can accurately correct or refine key terms and content, enhancing the overall quality and reliability of the generated text.

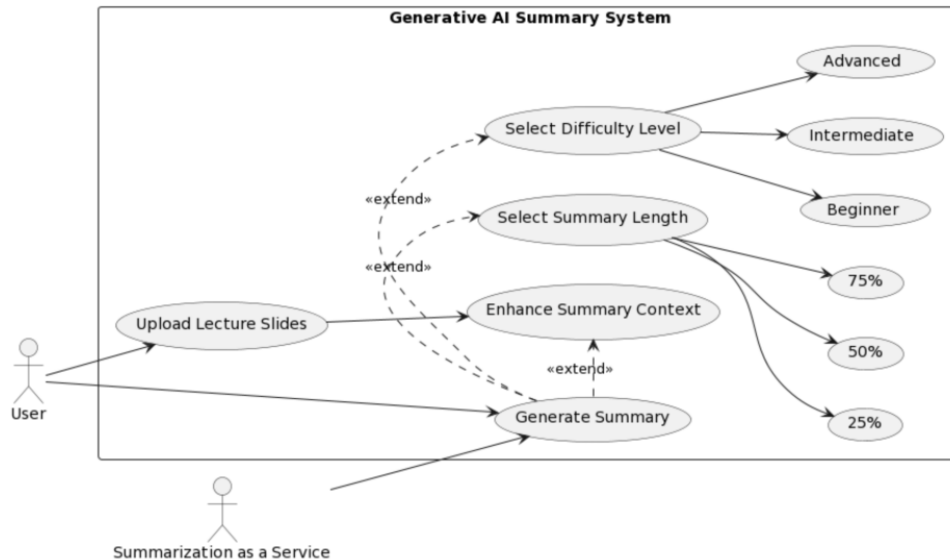
## Database Utilization for Transcript Management



Develop a robust database system to facilitate efficient storage, retrieval, and management of lecture transcripts. This system will allow users to search for specific transcripts based on multiple parameters such as course code, semester and year, lecture number, and week. Additionally, the database will incorporate a user feedback mechanism through an upvote/downvote system to assess the quality and accuracy of the generated transcripts. Transcripts with a net score (upvotes minus downvotes) below -100 will be flagged for review or removal, ensuring the maintenance of high-quality content within the database.

- **Upload Threshold Enforcement:** The system shall automatically check the existing count of available transcripts for a given course code, semester & year, lecture number, and week. If there are 3 or more transcripts with a net score greater than -100, the system will not permit additional uploads for that specific lecture
- **Search Functionality:** Enable precise search results within 2 seconds of query submission, using filters such as course code, semester & year, lecture number, and week.
- **Feedback Mechanism:** Implement an upvote/downvote system for each transcript, with real-time updates to the net score (upvotes minus downvotes). Automatically flag transcripts with a net score below -100 for review and hide it from selection.

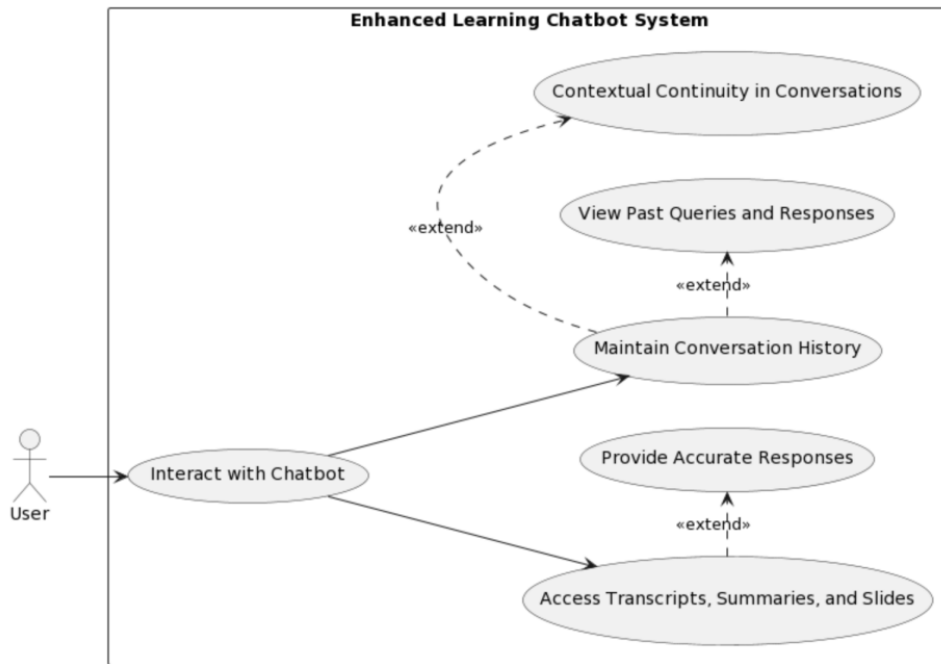
## Summary Generation with Generative AI



Utilize generative AI technologies to produce concise and coherent summaries of the generated transcripts. These summaries will be tailored based on varying levels of difficulty (beginner, intermediate, advanced) and desired length (25%, 50%, 75%), catering to different student needs and preferences. The system will also allow users to upload relevant lecture slides, which the AI can use as additional context to enhance the accuracy and relevance of the summaries. This feature aims to provide students with quick and customizable access to the essence of lecture content, facilitating efficient study and revision processes.

- **Difficulty Customisation:** The Generative AI system must be capable of adjusting the complexity of summaries according to three predefined difficulty levels: beginner, intermediate, and advanced in the choice of vocabulary, concept depth, and assumed prior knowledge
- **Length Customisation:** The system shall offer users the option to specify the length of the summary as a percentage of the original transcript's length, with selectable options at 25%, 50%, and 75%. The generated summary must adhere to the chosen length with a tolerance of  $\pm 5\%$  to accommodate natural language constraints.
- **Optional Slides Upload for Added Context:** Users shall have the option to upload relevant presentation slides alongside the lecture transcript. The Generative AI will analyze these slides to incorporate key points, terms, and concepts into the summary, providing a richer and more contextually accurate output. The system must ensure compatibility with common presentation formats such as PPT, PPTX, and PDF.

## Interactive Chatbots for Enhanced Learning



Implement interactive AI-powered chatbots that can engage with users in a conversational manner, allowing them to ask questions related to the lecture summaries. These chatbots will utilize the information from the transcripts, summaries, and any uploaded slides to provide accurate and helpful responses. This interactive component is designed to mimic a personalized tutoring experience, enabling students to clarify doubts, explore topics in depth, and reinforce their understanding of the lecture material.

- **Content Integration:** The chatbots will have integrated access to the transcripts, summaries, and uploaded slides. This enables them to draw upon the comprehensive content repository to provide accurate and contextually relevant responses to user inquiries.
- **Contextual Memory & Conversation History:** The chatbot shall maintain a memory of previous interactions with each user, allowing for continuity in conversations and enabling users to view their past queries and the chatbot's responses for seamless learning and reference.

## Critical Design Issues, Constraints, and Limitations

- **API Dependency for Transcription:** Reliance on Whisper API may limit transcription accuracy for low-quality audio or diverse accents, impacting the quality of generated transcripts.
- **API Dependency for Summarisation:** Summarization effectiveness using ChatGPT API could vary based on the complexity and specificity of lecture content, potentially affecting summary quality and relevance.
- **Manual Metadata Entry:** The requirement for database admins to manually input metadata such as course code, semester & year, lecture number, and week for search functionality may introduce human error and inefficiencies.

## Technical Approach

## Front-End Development

- The application's user interface will be built using **Next.js**, a React framework that offers server-side rendering and generating static websites, enhancing the application's performance and SEO.
- **Tailwind CSS** will be utilized for styling, providing a utility-first CSS framework for rapidly designing custom user interfaces without leaving the HTML.
- The entire front-end development will be undertaken with **TypeScript**, adding static type definitions to enhance code quality and predictability.

## Speech-to-Text Transcription

- For real-time transcription of lectures, the **Whisper API** will be integrated. This API is known for its exceptional accuracy in converting speech to text, even in challenging acoustic environments, making it ideal for capturing the nuances of live lectures.

## Summarization and AI Integration

- The application will leverage the **OpenAI GPT API** for generating summaries of the transcribed lectures. This advanced AI will analyze the text to highlight key points and concepts, providing students with concise summaries to aid in their revision.

## Database Management

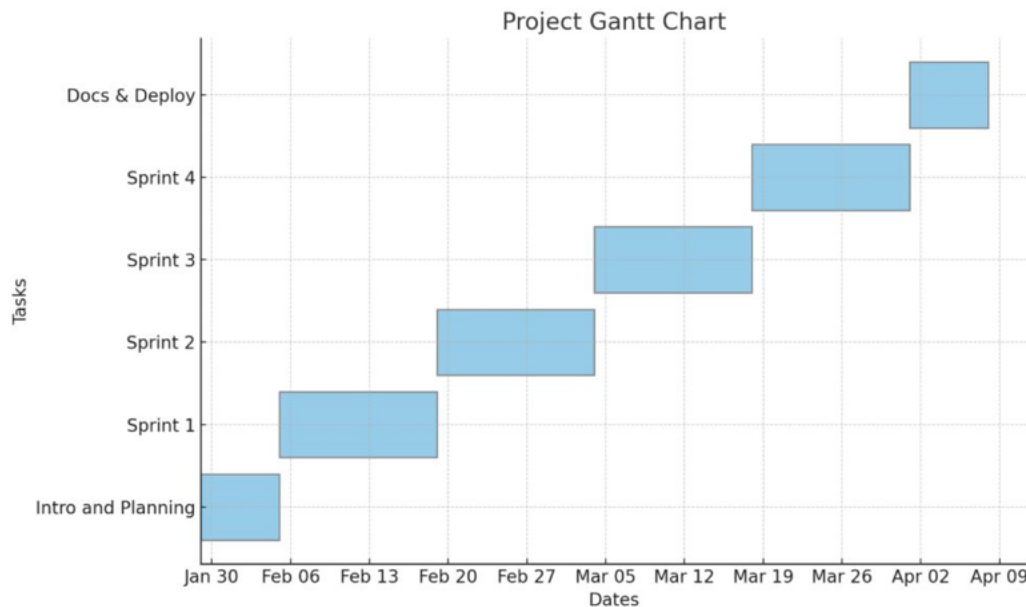
- To store and manage the data efficiently, **PostgreSQL** will be used as the database system. This choice is motivated by PostgreSQL's robustness, scalability, and support for complex queries and data types, which are essential for handling the diverse data types generated by the application.

## Project Management

---

The project will adopt an Agile project management approach, allowing for iterative development, frequent reassessment, and adaptation of plans. This approach will enable the team to respond to changes in project requirements and priorities effectively. The project will be divided into two-week sprints, with each sprint focusing on a specific set of features and tasks.





Gantt Chart illustrating the project timeline and milestones.

## Week 3: Introduction and Planning

- Introduction to project scope, objectives, and team members.
- Initial planning phase, including setting up development environments and defining sprint goals.

## Weeks 4-5: Sprint 1 - Initial Development

- **Front-End Team:** Start developing the basic UI components using Next.js, Tailwind CSS, and TypeScript.
- **Back-End Team:** Set up the initial server infrastructure and database schema in PostgreSQL.
- **AI Integration Team:** Research and prototype integration methods for the Whisper API and OpenAI GPT API.

## Weeks 6-7: Sprint 2 - Feature Development and Integration

- **Front-End Team:** Continue UI development and start integrating with back-end services.
- **Back-End Team:** Develop API endpoints for handling lecture recordings and user data.
- **AI Integration Team:** Implement Whisper API for real-time transcription and integrate initial summarization features using the OpenAI GPT API.

## Weeks 8-9: Sprint 3 - Advanced Features and Testing

- **Front-End Team:** Implement advanced UI features and ensure responsive design. Start preliminary user testing.
- **Back-End Team:** Optimize database interactions and ensure security measures are in place.
- **AI Integration Team:** Fine-tune AI summarization features and test for accuracy and reliability.

## Weeks 10-11: Sprint 4 - Finalization and Testing

- **All Teams:** Focus on bug fixing, performance optimization, and finalizing features.
- **QA Team:** Conduct extensive testing, including unit tests, integration tests, and user acceptance testing.

## Week 12: Documentation and Deployment

- **All Teams:** Prepare user manuals, technical documentation, and deployment guides.
- **Deployment Team:** Deploy the application to the production environment and perform final checks.

## Deliverables

---

The Deliverables Agreement outlines the specific items to be provided upon the project's completion. Each deliverable is detailed below with its respective delivery date and a thorough description. Each deliverable will undergo a review and approval process with the project sponsor to ensure compliance with the agreed-upon specifications and quality standards. This agreement is a commitment to delivering high-quality outputs in a timely manner.

### Detailed Requirement Specifications

- **Delivery Date:** Week 4
- **Description:** A document detailing all technical and functional requirements of the web application, including user roles, features, performance criteria, and system integration specifications.

### Physical Prototype

- **Delivery Date:** N/A
- **Description:** Not applicable for this web-based application project.

### Detailed Description of Test Procedures

- **Delivery Date:** Week 10
- **Description:** Documentation of testing strategies, methodologies, test cases, and scenarios for various testing phases of the web application.

### Computer Program Code, Flowchart, Documentation

- **Delivery Date:** Continuous delivery from Week 5 to Week 11
- **Description:** The complete source code of the web application, system architecture flowchart, and comprehensive documentation covering the codebase, APIs, and functionalities.

## System Components or Modules

- **Delivery Date:** Continuous delivery from Week 5 to Week 11
- **Description:** Delivery of system components including front-end UI, back-end server infrastructure, database schema, and AI integration modules, with individual documentation.

## Entire System

- **Delivery Date:** Week 12
- **Description:** The final, fully integrated and tested web application, ready for deployment, encompassing all components, modules, and functionalities as specified.

## User-Friendly Instructions Including Training for Personnel

- **Delivery Date:** Week 12
- **Description:** Comprehensive user manuals and tutorials for end-users and NTU personnel, including guides, FAQs, and training materials for the web application's effective use.

## Budget

---

The estimated budget for the initial design phase of the project is outlined below, covering major categories such as personnel, equipment, and supplies. This budget aims to provide a comprehensive overview of the anticipated costs associated with the development of the web application. It is important to note that this budget is an initial estimate and may be subject to changes based on project requirements, unforeseen challenges, and adjustments in the scope of work. Detailed financial planning and regular budget reviews will be essential to ensure the project remains within the allocated budget.

## API Costs

- **Whisper and OpenAI GPT APIs:** An estimated cost of \$20 for initial usage of the Whisper API for real-time transcription and the OpenAI GPT API for summarization. This cost may vary based on the scale of usage and specific pricing models of the APIs.

## Communication and Coordination with Sponsor

---

Effective communication and coordination with the project sponsor are crucial for the successful completion of the project. The following outlines the structured approach for interaction with the sponsor, Chen Yiwen, ensuring clarity, accountability, and responsiveness throughout the project lifecycle.

## Communication Schedule and Responsibilities

- Bi-weekly TA meetings will be held with Chen Yiwen to discuss project progress, address any concerns, and plan for upcoming phases.

- The project manager will be responsible for organizing these meetings, setting the agenda, and ensuring that all relevant team members are prepared to discuss their respective areas.
- Chen Yiwen, as the project sponsor, will provide guidance, feedback, and decisions on critical project matters during these meetings.

## Form of Communication

- The primary form of communication will be through virtual meetings, utilizing platforms such as Zoom or Microsoft Teams, to accommodate the schedules and locations of all participants.
- In addition to bi-weekly meetings, weekly updates will be provided via email, summarizing the progress made and highlighting any critical issues or decisions required.
- For urgent matters, direct communication may be initiated via phone calls or instant messaging platforms, with follow-up documentation provided via email.

## Information Distribution

- Meeting minutes, weekly update emails, and any other relevant documentation will be shared with Chen Yiwen and all project team members to ensure everyone is informed and aligned.
- Documentation will be transmitted via Wiki for easy access and reference.

## Team Qualifications

---

### Nathanael Axel Wibisono (Project Manager, Lead Developer)

Nathanael Axel Wibisono is a distinguished student from Nanyang Technological University Singapore, pursuing a BEng in Computer Science with a second major in Business, projected to graduate in May 2024 on Honours with Highest Distinction. His academic excellence is evident from recognition through multiple Dean's List awards and prestigious scholarships such as the Dr. Goh Keng Swee Scholarship and the ASEAN Scholarship. He is currently a quantitative software developer in a high-frequency trading firm, specialising in high-performance computer networks in C++. Previously, he was a software engineer in Goldman Sachs, F5, Anacle Systems Limited, and Deloitte.

His final year project, mentored by Ethereum founder Vitalik Buterin, focuses on developing an interoperable energy trading infrastructure on blockchain platforms. Nathanael has also showcased his technical prowess in competitions, securing first place in the Goldman Sachs Intern Hackathon 2023 with a project on generative AI and achieving first place in Singapore in market-making competition held by Optiver.



Nathanael Axel Wibisono. [LinkedIn Profile \(https://www.linkedin.com/in/nathanaxel/\)](https://www.linkedin.com/in/nathanaxel/)

## Sherwin Samson (QA Manager, QA Engineer)

Sherwin Samson is a final year student from Nanyang Technological University Singapore, pursuing a BEng in Computer Science with a second major in Business. Primarily specialised in Artificial Intelligence and Data Science. With a passion for all things AI: Computer Vision, NLP, Signal Processing.

Email: sherwin002@e.ntu.edu.sg

Phone Number: 91711551



Sherwin Samson. [LinkedIn Profile \(https://www.linkedin.com/in/sherwinsamson/\)](https://www.linkedin.com/in/sherwinsamson/)

## Jiang Jiaxi (Back-end Developer)

Jiang Jiaxi is a final year student from Nanyang Technological University Singapore, pursuing a BEng in Computer Science, projected to graduate in May 2024 on Honours with Highest Distinction. Primarily specialised in Artificial Intelligence and Data Science. With a passion for creating seamless user experiences and a strong background in HTML, CSS, and JavaScript, she's dedicated to crafting intuitive interfaces that enhance user engagement and satisfaction.

Email: jiano117@e.ntu.edu.sg

## Zhang Mengao (Release Engineer/Manager)

Zhang Mengao is a final year student at Nanyang Technological University Singapore, pursuing a BEng in Computer Science, specializing in AI and Data Science. Expected to graduate in May 2024 with Honours and the highest distinction, Zhang has consistently achieved academic excellence, supported by the NTU Science and Engineering Scholarship.

His passion lies in exploring state-of-the-art AI technologies, with a particular interest in Natural Language Processing and Audio Processing. Currently a Research Assistant at Temasek Laboratories@NTU, Zhang contributes to research projects related to Speech transcript post-processing. Previously, he interned at Huawei International, gaining practical experience in AI research. His work has resulted in a published paper at Interspeech 2023, showcasing his innovative approach to speaker verification systems.

Email: ZH0024AO@e.ntu.edu.sg



Jiang Jiaxi. [LinkedIn Profile \(https://www.linkedin.com/in/jiangjiaxi/\)](https://www.linkedin.com/in/jiangjiaxi/)

## Yang Yida (Back-end Developer)

Yang Yida is a fourth-year Computer Science undergraduate student at NTU, Singapore. He has a robust foundation in full-stack software development, through internship in diverse industries. Yida also has enthusiasm in research, particularly in computer vision and multimodal problems. Currently, he is focusing on 3D scene generation.



## Liu Changsong (Front-end Developer)

Liu Changsong is a final-year computer science undergraduate at NTU, specializing in artificial intelligence and data science. With a keen interest in computer vision (CV) and natural language processing (NLP), Changsong has contributed to research in these areas by publishing two papers on punctuation restoration, with one being honored with the Best Student Paper award. This not only highlights his ability to solve complex problems but also underscores his potential to make impactful contributions to the field of data science and beyond.



Zhang Meng'ao. LinkedIn Profile (<https://www.linkedin.com/in/mengao-zhang/>)



Yang Yida. LinkedIn Profile (<https://www.linkedin.com/in/yida-yang/>)



Liu Changsong. [LinkedIn Profile \(https://www.linkedin.com/in/changsong-liu/\)](https://www.linkedin.com/in/changsong-liu/)

---

Retrieved from "[http://155.69.100.30/s22324SCELMagnificent7/index.php?title=Project\\_Proposal&oldid=193](http://155.69.100.30/s22324SCELMagnificent7/index.php?title=Project_Proposal&oldid=193)"

---

This page was last edited on 4 April 2024, at 10:58.