COOPERATIVE EDUCATION REPORT:

Rapid AI: An AI Application for Thematic Analysis and Enhanced Qualitative Research

WRITTEN BY:

Ram Chandra Giri

S.ID. 6405140006

Date: 9 November, 2024

This report is submitted in partial fulfillment of the requirements for Cooperative Education,
Faculty of Information Technology, Academic Semester 1/2024,
Siam University.

Title: Rapid AI: An AI Application for Thematic Analysis and Enhanced Qualitative Research

Internship Period: 9 July 2024 to 8 November 2024

Written by: Ram Chandra Giri

Department: Bachelor of Science Program in Information Technology

Academic Advisor: Dr. Parham Porouhan, Ph.D.

We have approved this cooperative report as a partial fulfillment of the Cooperative Education Program Semester 1/2024.

Oral Presentation Committees

.....

(Dr. Parham Porouhan, Ph.D.)

Academic Advisor

.....

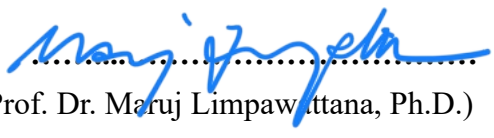
(Dewi Ratnawulan)

Job Supervisor

.....

(Arj. Michael Slater)

Cooperative Committee



.....

(Asst. Prof. Dr. Maruj Limpawattana, Ph.D.)

Assistant President and Director
of Cooperative Education

DECLARATION

I, the undersigned, declare that this report is a result of my research assignment carried out in the year 2024. It has not been previously submitted to any other university or any other examination(s).

Signature

Ram Chandra Giri

Title: Rapid AI: An AI Application for Thematic Analysis and Enhanced Qualitative Research

Credits: 5

By: Ram Chandra Giri

Advisor: Dr. Parham Porouhan, Ph.D.

Degree: Bachelor of Science Program in Information Technology (Intl. IT)

Major: Information Technology

Faculty: Information Technology

Semester/Academic year: 1/2024

Abstract

In today's digital age, the application of Artificial Intelligence (AI) has become increasingly crucial in enhancing the efficiency and effectiveness of communication within educational institutions. This project focuses on the development and deployment of UniBot, a generative AI-powered chatbot designed specifically for Thai universities. UniBot aims to revolutionize the admissions process and engagement with prospective students by providing instant, personalized support and streamlining communication tasks. Traditional communication methods often fail to meet the demands of modern university environments, leading to delays and inefficiencies. UniBot addresses these challenges by automating responses to common inquiries, improving information accessibility, and reducing the workload on university staff. The project also explores the technical and organizational challenges associated with implementing AI solutions in educational settings. By leveraging Natural Language Processing (NLP) and Retrieval-Augmented Generation (RAG), UniBot is designed to handle a wide range of queries with high contextual relevance and accuracy. This project outlines the key steps in developing and deploying UniBot, including data collection, NLP functionality optimization, and system deployment within a university department. Ultimately, UniBot aims to enhance communication efficiency, empower prospective students, and contribute to the digital transformation of Thai higher education.

Keywords: RAG, qualitative research, AI, thematic analysis, Rapid AI, natural language processing

ACKNOWLEDGEMENT

I would like to express my sincere gratitude to my supervisor, Dewi Ratnawulan, for her invaluable guidance, encouragement, and support throughout my internship. I am also deeply appreciative of the entire team at Rapid Asia Co., Ltd., whose collaborative spirit and openness fostered an enriching environment for both learning and professional development.

My heartfelt thanks also extend to my academic advisor, Dr. Parham Porouhan, as well as to Asst. Prof. Dr. Maruj Limpawattana, Assistant President and Director of Cooperative Education, for their mentorship and invaluable support. I am grateful to the Cooperative Education Committee at Siam University for providing this valuable opportunity for experiential learning and growth.

Lastly, I extend my deepest appreciation to my family, whose unwavering support and encouragement have been a constant source of strength throughout my academic and professional journey.

Name : Ram Chandra Giri

S. ID. : 6405140006

TABLE OF CONTENTS

ABBREVIATION TABLE

Abbreviation	Full Form
AI	Artificial Intelligence
API	Application Programming Interface
BMRS	Business and Marketing Research Solutions
CAWI	Computer-Assisted Web Interviewing
CAPI	Computer-Assisted Personal Interviewing
ESOMAR	European Society for Opinion and Marketing Research
KAP	Knowledge, Attitudes, and Practices
LLM	Large Language Model
M&E	Monitoring and Evaluation
NaN	Not a Number
NGO	Non-Governmental Organization
NLP	Natural Language Processing
RAG	Retrieval-Augmented Generation
TF-IDF	Term Frequency-Inverse Document Frequency
TMRS	Thailand Marketing Research Society
UK	United Kingdom
UI	User Interface
UX	User Experience
QA	Quality Assurance
SWOT	Strengths, Weaknesses, Opportunities, and Threats
UN	United Nations
USA	United States of America
WHO	World Health Organization

CHAPTER 1 : INTRODUCTION

1.1 Company Profile

Founded in February 2010 and headquartered in Bangkok, Thailand, Rapid Asia Co., Ltd. is a prominent data research and consulting company with expertise in Monitoring and Evaluation (M&E) services for the social development sector. The strategic location of its Bangkok headquarters enables the company to closely collaborate with regional clients while effectively managing a diverse portfolio of global projects. To further support regional outreach, Rapid Asia also operates BMRS Asia, an office located in Cambodia. The organization's approach to social research and impact evaluation established methodologies designed to generate actionable insights. These include baseline and endline studies, Knowledge, Attitudes, and Practices (KAP) surveys, impact assessments, and program evaluations. Each project is structured to support data-driven decision-making, empowering clients to address complex social issues and drive meaningful change in their respective sectors. With extensive experience spanning over 40 countries across Asia, Africa, Eastern Europe, and South and Central America, the company has addressed a wide range of thematic areas, including migration, agriculture, disaster relief, education, human rights, human trafficking, climate change, microfinance, health, and food security. The organization's ability to work across such diverse domains underscores its commitment to adaptability and specialized expertise in high-impact social research. As a member of ESOMAR, a global organization for market, opinion, and social research, the company is committed to maintaining the highest standards of data quality and ethical integrity. These standards are rigorously applied to all projects, including collaborations with trusted subcontractors, to ensure that each engagement reflects both professionalism and respect for the communities and issues it addresses. This commitment to quality and ethical responsibility positions Rapid Asia as a trusted partner in providing evidence-based insights to support policy change and sustainable development on an international scale.

1.1.1 Mission

To empower social development organizations and policymakers with high-quality, data-driven insights that drive sustainable change. Through rigorous research and

evaluation, with aim to provide evidence-based solutions that address complex social issues across diverse communities and thematic areas.

1.1.2 Vision

To be a global leader in social research and impact evaluation, recognized for our commitment to ethical practices, innovative methodologies, and meaningful contributions to positive social change. Rapid Asia envision a world where data-informed decisions lead to measurable improvements in social equity, resilience, and human well-being.

1.1.3 Values and Guiding Philosophy

This approach to social research and evaluation is built on a foundation of core values that guide every aspect of the work. With a commitment to meaningful change, Rapid Asia focus remains on simplifying complex challenges, engaging deeply with clients, fostering innovation, upholding integrity, and cultivating a team driven by passion and dedication. These principles shape the delivery of impactful insights that empower partners to address pressing social issues effectively and ethically.

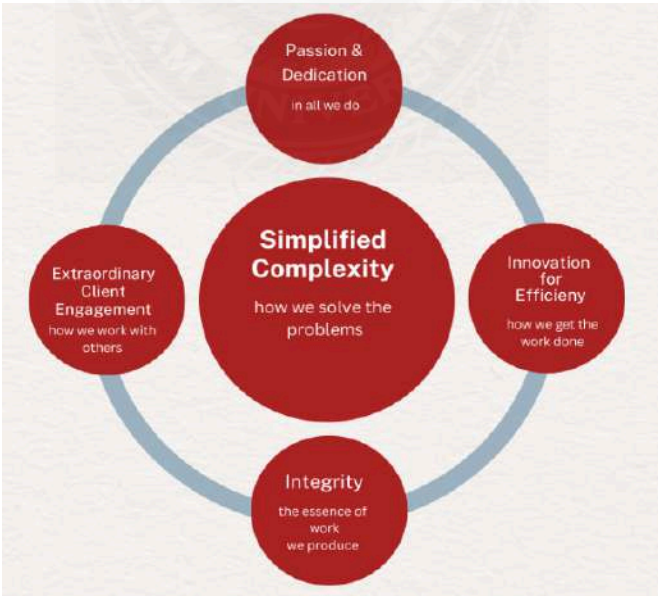


Fig 1. Guiding Philosophy Diagram

- **Simplified Complexity**: Embracing simplicity in addressing complex challenges, the belief is that straightforward solutions are the most effectively understood, remembered, and implemented, ensuring lasting impact.
- **Exceptional Client Engagement**: Client and stakeholder engagement are central to each project. This approach explores a deep understanding of project challenges and needs, leading to collaborative, customized solutions.
- **Innovation for Efficiency**: With a focus on timely and efficient outcomes, there is a continuous pursuit of innovation and the integration of new technologies to enhance operational efficiency and streamline research processes.
- **Integrity**: Integrity is fundamental to the work, upheld through best practice models, rigorous methodologies, and ethical standards that produce reliable, high-quality results.
- **Passion and Dedication**: Guided by a shared commitment to meaningful social impact, the team remains driven to provide insights that make a difference. This dedication is evident in every project, with team members fully invested in delivering valuable outcomes.

1.1.4 Services

Rapid Asia offers a comprehensive range of services designed to address diverse social research and evaluation needs. These include Enumerator Training to ensure reliable data collection, KAP Score Modeling for assessing knowledge and behavior change, and Baseline Studies to establish benchmarks for program evaluations. Advanced methodologies like Impact Assessments, Program Evaluation, and Policy Analysis support strategic decision-making and program improvement. Additionally, innovative tools such as Real-Time Data, Qualitative Work, and the m-panel®application enable timely insights and continuous monitoring. The organization also provides Global CAPI Surveys for robust data collection and tailored solutions to meet unique client requirements, ensuring actionable, high-quality outcomes.

1.1.5 Company's logo

RAPID ASIA
Guiding Sustainable Change

1.2 Organizational Structure

This organizational chart presents the structure of Rapid Asia, illustrating how each division and role supports the company's work in social research. Every department, from Project Management to Data Analytics and IT, plays a vital role in delivering valuable insights and high-quality service to clients.

1.2.1 Diagram of the Organizational Structure

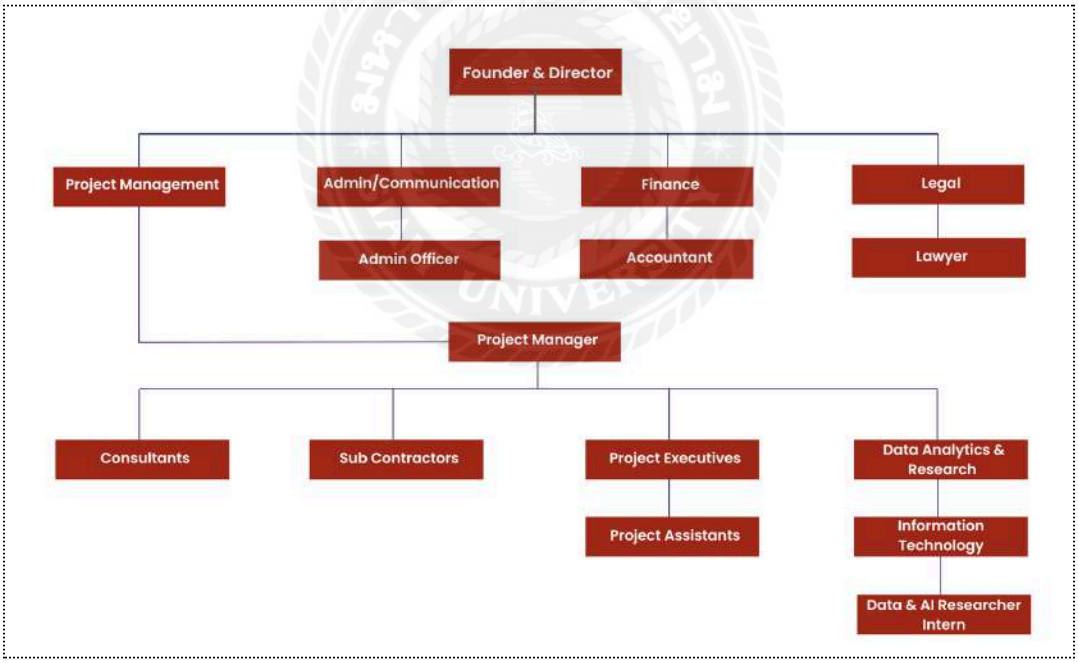


Fig 2. Organizational Structure Diagram

The *Founder & Director* leads Rapid Asia, setting its vision and strategy while ensuring that each division aligns with organizational goals. This role involves high-level decision-making and fostering key client and stakeholder relationships. The *Project Management* division oversees

projects from start to finish, coordinating across departments to meet milestones, budgets, and client expectations. They work closely with clients to integrate their needs and address any project challenges, ensuring smooth execution and client satisfaction.

The *Admin/Communication* division, led by an Admin Officer, manages internal communication, documentation, and logistical support, ensuring smooth day-to-day operations throughout the organization. The *Finance* division, managed by an Accountant, oversees budgeting, cash flow, and compliance, working with Project Management to ensure financial stability for each project and the organization as a whole.

The *Legal* division, staffed by a Lawyer, ensures compliance with regulations, handles any legal matters, and protects the organization's interests, particularly across diverse regulatory contexts.

The *Project Manager* coordinates divisions to bring projects to completion, ensuring alignment with organizational goals and maintaining quality standards throughout each project.

Consultants bring specialized expertise to specific projects, enhancing company's ability to provide tailored solutions to meet unique client needs. *Subcontractors* are hired for specific tasks like field data collection and technical support, allowing the organization to scale operations and manage large or dispersed projects effectively.

Project Executives handle project workflows, tracking progress and managing client expectations to ensure deliverables meet standards. *Project Assistants* support project teams with administrative tasks like record-keeping and scheduling, enabling senior staff to focus on core responsibilities. The *Data Analytics & Research* division generates insights for social research projects, using quantitative and qualitative techniques to support the company's evidence-based approach.The *Information Technology* division manages the technical infrastructure for data collection and analysis, supporting secure and efficient data handling for research projects.

Finally, the *Data & AI Researcher Intern* role (which I held) is part of the Information Technology division, focused on using AI and machine learning for data analysis. This role includes developing AI tools to streamline data processing, extract themes, and derive insights from large datasets, advancing the company's capabilities in handling complex social research data.

1.2.2 Job Position

During my internship at Rapid Asia Co., Ltd., I worked as a Data and AI Researcher Intern within the Technology Department. This department focuses on advancing data analysis and implementing tech solutions to enhance the organization's research and consulting capabilities. It plays a key role in managing essential data functions and supporting a range of projects across the company, ensuring that technology and data analytics are integrated into each project's success. Within the Technology Department, I was part of the Data Research & Analysis Team, a group dedicated to leveraging data research for informed, evidence-based decisions. This team manages all stages of data collection, processing, and analysis, essential for the organization's mission to deliver actionable insights for social development. My primary assignment involved developing and deploying qualitative data solutions technology. Through high-quality data support, our team ensures that research findings are backed by thorough analysis, covering everything from survey data and desk reviews to extracting final key findings. This approach helps deliver reliable, impactful recommendations to stakeholders.

1.3 Intention and Motivation for Choosing Rapid Asia

My decision to choose Rapid Asia Co., Ltd. for my cooperative education was driven by its impactful work in the field of social research, particularly in monitoring and evaluation (M&E) for social development initiatives. Rapid Asia's reputation as a research and consulting company that provides data-driven insights for policy change and social improvement aligned with my personal and professional aspirations to use technology for meaningful change. Additionally, its international scope provided an ideal platform for me to gain exposure to diverse projects and real-world challenges in data processing for social research. I was motivated to contribute my skills in data and AI to a company that values innovation and actively pursues methods to improve research accuracy and efficiency. Working on developing *Rapid AI,* an artificial intelligence application for qualitative data analysis allowed me to apply my technical knowledge in a setting where it could drive real-world impact, making Rapid Asia an ideal choice for my cooperative study.

1.4 Strategic Analysis of Rapid Asia

A strategic analysis of Rapid Asia reveals key insights into its competitive positioning, strengths, and areas for growth within the social research and consultancy sector. The *SWOT framework* below offers a comprehensive view of company on how it can navigate both opportunities and potential challenges as it continues to expand its impact:

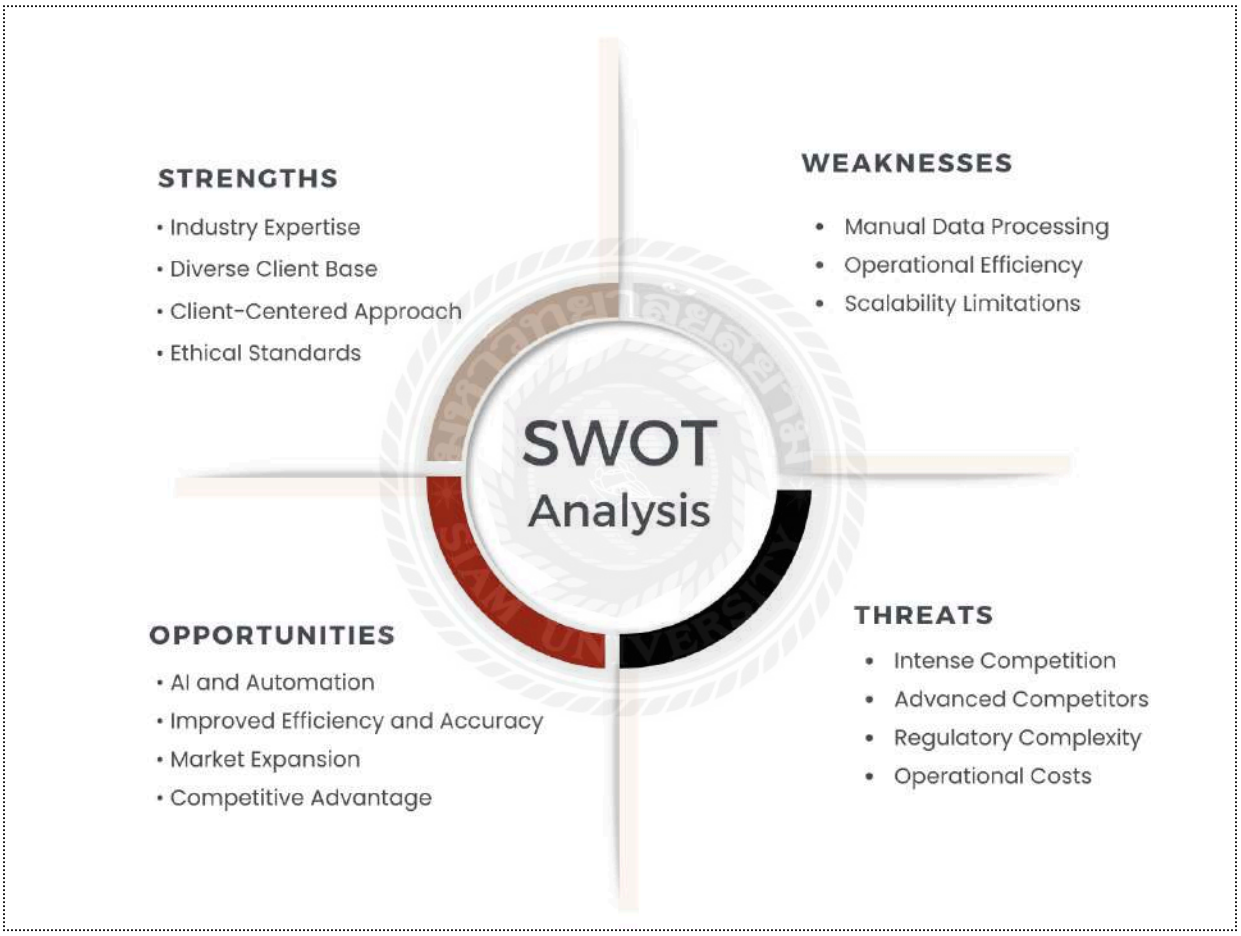


Fig 3. SWOT Analysis of Rapid Asia

Strengths: As a leader in Monitoring and Evaluation (M&E), known for its deep expertise and extensive experience across multiple regions, including Asia, Africa, and Latin America, the company has built a strong, diverse client base, encompassing government agencies, non-governmental organizations (NGOs), and international development bodies. Its success is attributed to a client-centered approach, which involves close collaboration to tailor research solutions that precisely meet project objectives. This commitment to customization enables the

company to deliver data-driven insights that are not only relevant but actionable, helping clients make informed, evidence-based decisions. Additionally, its emphasis on ethical practices and adherence to ESOMAR standards enhance its credibility and reputation in the industry, fostering long-term client relationships and partnerships.

Weaknesses: While Rapid Asia has demonstrated strong performance, it faces limitations due to a significant reliance on manual data processing, particularly within qualitative research. The manual handling of data, though meticulous, can introduce risks of human error and inconsistencies, especially in large-scale projects requiring extensive data interpretation. This reliance also impacts operational efficiency, as manual processes are often time-consuming and costly which can delay project timelines. Furthermore, a lack of advanced automated tools places constraints on the company's ability to scale its qualitative analysis efforts, potentially limiting its capacity to handle an increasing volume of projects or expand into new service areas. This gap in technological infrastructure suggests an area where investment in automation and AI could yield substantial benefits.

Opportunities: The adoption of AI and automation presents a significant opportunity to enhance its data processing workflows. By integrating technologies like machine learning and natural language processing (NLP) tools-such as the in-house Rapid AI solution-the company can streamline the analysis of qualitative data, improving both the speed and consistency of insights generated. This technological enhancement would not only increase operational efficiency but also improve the scalability of the company's services, enabling it to take on larger, more complex projects without compromising quality. Moreover, the use of AI could provide Rapid Asia with a competitive edge in emerging markets, where timely, data-driven decisions are critical in responding to social issues. The expansion into these markets with enhanced capabilities could diversify the company's client base and revenue streams, while reinforcing its position as an innovative leader in social research.

Threats: The social research and consultancy market, especially for M&E services, is highly competitive, with numerous players competing for similar government and NGO contracts. Competitors who have already embraced AI and automated solutions may have an advantage in

terms of faster delivery times and reduced costs, potentially attracting clients seeking efficiency. Additionally, Rapid Asia must navigate the complexities of varied regulatory frameworks and ethical standards across the regions it serves. These regulatory requirements, which can vary significantly from one country to another, present challenges in maintaining standardized processes and ensuring compliance. Adhering to these diverse standards is essential for sustaining the company's reputation but may also increase operational costs and administrative workload, particularly as the company expands into new territories.

In this way, Rapid Asia's established strengths in M&E, combined with its client-focused approach, provide a solid foundation for future growth. By addressing the current reliance on manual processes, particularly in qualitative research, and investing in advanced AI tools, the company can enhance its operational efficiency, increase the scale of its offerings, and improve the quality of its insights. Embracing technology will not only allow the company to overcome its current limitations but will also enable it to remain competitive in an industry that is increasingly leaning toward digital transformation. Furthermore, proactive adaptation to regulatory challenges and a commitment to ethical research practices will support sustained credibility and trust, allowing the organization to expand its impact in the evolving landscape of social research and consultancy.

1.5 Objectives of the Cooperative Study

The primary objective of this cooperative study was to design and implement *Rapid AI*, an innovative AI-powered tool intended to support and streamline qualitative research at Rapid Asia Co., Ltd. Recognizing the time-intensive nature and potential for inconsistencies in traditional qualitative data analysis, this project was conceived to address these challenges by automating theme, key findings and supportive quotes identification and extraction. By building this application, the study aimed to contribute meaningfully to the company's commitment to delivering high-quality, timely insights for data-driven decision-making in the social research sector. This study held particular significance as it represented an opportunity to apply cutting-edge AI technologies within a real-world context, bridging the gap between theoretical knowledge and practical application. The objectives were carefully designed to reflect this, with specific aims that included:

- **Automating Qualitative Data Analysis**: The study sought to create a tool that could automatically identify themes and key findings within large qualitative datasets. This automation not only reduces the manual effort and time required for data processing but also allows researchers to focus on higher-level analysis and interpretation, enhancing the overall research quality.
- **Enhancing Data Consistency and Accuracy**: By automating key parts of the qualitative analysis process, *Rapid AI* was intended to improve data consistency and accuracy by reducing human error, a common challenge in manual analysis. This enhancement in reliability directly impacts the quality of insights and ensures that clients receive dependable, data-driven recommendations.
- **Accelerating Report Turnaround**: A critical objective was to speed up the reporting process for qualitative data, allowing the research team to deliver results more quickly and align with clients' timelines and expectations. This efficiency not only benefits project management but also strengthens client relationships by consistently meeting deadlines with high-quality output.
- **Applying AI Knowledge in a Real-World Setting**: Personally, this project represented a unique opportunity for me to gain hands-on experience in developing and deploying an AI application within the field of qualitative research. The ability to test and refine AI concepts in a practical environment allowed for deeper learning and skill development, which I view as essential to my career in data and AI engineering.
- **Supporting Strategic Goals**: The study aimed to contribute to the company's broader goals of innovation and operational efficiency. By implementing an AI-driven solution, the project sought to position the company as a leader in technological advancement within the M&E sector, enabling it to tackle larger and more complex projects in the future. In addition, by applying my theoretical knowledge in data analytics, web design and content management into practical not only helped me to learn confidently but also enhance the company's goal.

This cooperative study not only provided a valuable opportunity to develop technical and practical skills in AI but also contributed to Rapid Asia's mission to provide efficient, evidence-based solutions for social research. My interest in this project derived from a desire to

apply AI in meaningful ways that drive social impact, making this study both professionally enriching and personally fulfilling.

CHAPTER 2 : CO-OP STUDY ACTIVITIES

2.1 Job Description

As a Data and AI Researcher Intern, my primary responsibility was to develop and implement *Rapid AI*, an AI-based web application designed to support the company's qualitative data analysis needs. The objective of the application was to automate the extraction of themes, key findings, supporting quotes and summarizing the report from large qualitative datasets, significantly reducing the manual work required in thematic analysis. This role required technical expertise in AI programming specially natural language processing (NLP), Large Language Models, data preprocessing, and the ability to work collaboratively with the research team to ensure that the application met the practical needs of the company's ongoing projects. My role involved direct interaction with both the project management team, finance and top leadership to gather requirements, test the application, and make iterative improvements. This hands-on experience allowed me to apply my academic knowledge in a real-world setting, contributing a functional tool to the company's operations while enhancing my skills in data science and AI. In addition, my role involved designing websites, building prototypes, data analysis and data research and so on.

2.2 Job Responsibilities

In my role as a Data and AI Researcher Intern at Rapid Asia, I was responsible for a variety of tasks that directly contributed to the development of *Rapid AI* and supported other research and operational activities. My responsibilities were divided into primary and secondary categories:

Primary Responsibilities

- **Development of Rapid AI**: Designed and built this application from scratch, focusing on creating efficient algorithms for theme extraction and key findings identification using

natural language processing (NLP) techniques. This included integrating machine learning models to enhance data analysis accuracy and reliability.

- **Website Design**: Developed a website prototype for BMRS Asia, a regional company for Rapid Asia that is located in Cambodia with an interactive, user-friendly design. This involved structuring the site for clear navigation, aligning content effectively, and ensuring the interface met the research team's operational needs.
- **Survey Coding**: Created survey questionnaires based on specific client requirements, ensuring compatibility with CAWI (Computer-Assisted Web Interviewing) and CAPI (Computer-Assisted Personal Interviewing) systems, as per standards used by organizations like the World Bank.
- **Data Preparation and Preprocessing**: Managed large datasets by cleaning, organizing, and structuring data to optimize its usability for AI models. This was a crucial step to guarantee data quality and ensure model outputs were accurate and reliable.
- **Collaboration with Data and Project Management Teams**: Engaged with both the data and project management teams to align application's features with their needs. This included regular meetings to understand requirements and gather feedback, refining application's functionalities to support qualitative analysis workflows.
- **Testing and Optimization**: Conducted extensive testing to ensure accuracy, reliability, and performance. Evaluated the application's outputs, identified areas for improvement, and made iterative adjustments to meet the research team's standards for high-quality analysis.
- **Documentation and User Guide Development**: Created comprehensive documentation and user guides such as 'User Manual', a toolkit to help the research team understand and effectively use *Rapid AI*. The documentation included technical details, troubleshooting steps, and best practices, enabling smooth adoption and usage within the team.

Secondary Responsibilities

- **Social Media Management**: Assisted in managing the company's social media accounts, focusing on sharing updates and insights related to data analytics and research findings.

This included creating and scheduling posts, monitoring engagement, and responding to user interactions.

- **Data Analytics**: Supported the research team by conducting data analysis on survey results and research data, providing insights that contributed to project conclusions and client presentations. This often involved quantitative analysis and visualizing key metrics for easier interpretation.
- **Content Creation**: Developed content for both internal and external communications, including research summaries, blog posts, and informational materials. These efforts supported Rapid Asia's goal of sharing insights and maintaining transparency with stakeholders.

2.3 Job Process Diagram

As an intern, my role followed a structured 9-step workflow for developing and deploying AI applications (see Fig.4). The process began with Problem Identification, followed by Requirement Analysis, Data Collection, and Model Selection, forming the foundation for effective solutions. It concluded with App Development, Feature Testing, Deployment, and Maintenance, ensuring reliable and impactful outcomes.

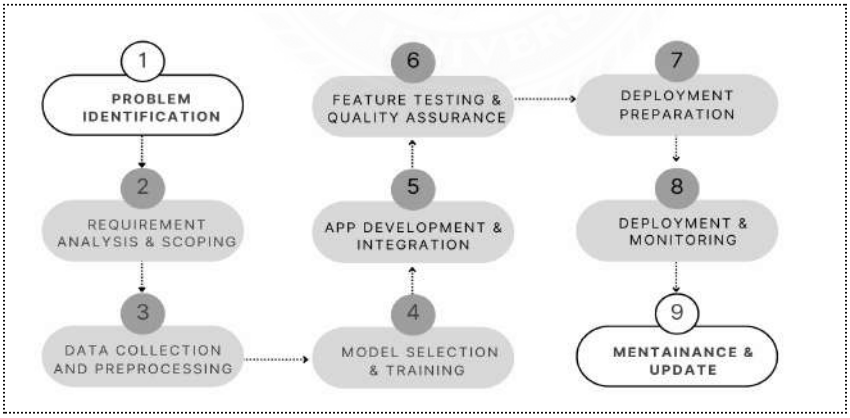


Fig 4. Job Process Diagram

2.3.1 Problem Identification

The initial step of the project is to clearly define the core problem and understand the underlying need for an AI-driven solution. This involves identifying specific challenges or

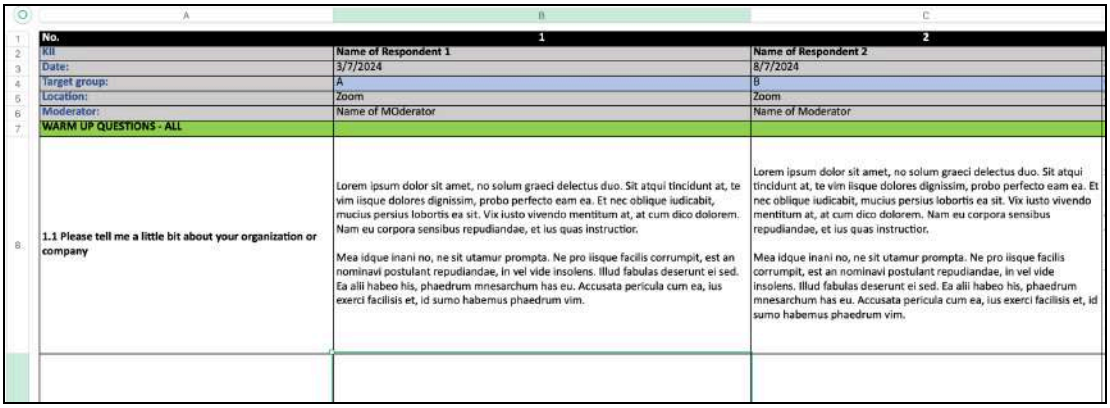
gaps within Rapid Asia's existing research workflows, such as time-intensive manual processing and inconsistencies in qualitative data analysis. The aim is to pinpoint areas where AI and automation could create value, such as reducing manual efforts, enhancing data consistency, or improving reporting speed. By thoroughly understanding the problem, this stage establishes a foundation for the entire project, resulting in a well-defined problem statement and a justification for why the AI solution is essential for the company's goals.

2.3.2 Requirement Analysis & Scoping

Following problem identification, this phase involves a comprehensive analysis of project requirements. It starts with gathering input from project teams, to understand their goals, expectations, and any specific requirements for the AI tool. Through meetings and discussions, the project's essential features-such as theme extraction, triangulation, and report generation-are identified. Additionally, data requirements are assessed, taking into account the format, volume, and quality of available data. A detailed project scope document is then created, outlining all required features, functionalities, and any limitations.

2.3.3 Data Collection & Preprocessing

Data is the backbone of any AI project, and this phase is dedicated to gathering and preparing the data required for effective analysis. Relevant datasets are collected, including any necessary training data for Natural Language Processing (NLP) models. Once gathered, the data undergoes a thorough cleaning process to handle issues like missing values (NaN), duplicates, and irrelevant information. The data is also structured into segments (e.g., metadata, section heading, questions, responses) to facilitate analysis. For user-uploaded data, real-time preprocessing capabilities are implemented within the app to ensure seamless analysis. The result of this phase is a preprocessed and labeled dataset that is ready for model training and analysis.



No.	1	2
KII	Name of Respondent 1	Name of Respondent 2
Date:	3/7/2024	8/7/2024
Target group:	A	B
Location:	Zoom	Zoom
Moderator:	Name of MOderator	Name of Moderator
WARM UP QUESTIONS - ALL		
1.1 Please tell me a little bit about your organization or company	Lorem ipsum dolor sit amet, no solum graeci delectus duo. Sit atqui tincidunt at, te vim iisque dolores dignissim, probo perfecto eam ea. Et nec oblique iudicabit, mucius persius lobortis ea sit. Vix iusto vivendo mentitum at, at cum dico dolorem. Nam eu corpora sensibus repudiandae, et ius quas instructior. Mea idque inani no, ne sit utamur prompta. Ne pro iisque facilis corrumpit, est an nominavi postulant repudiandae, in vel vide insolens. Illud fabulas deserunt ei sed. Ea alii habeo his, phaedrum mnesarchum has eu. Accusata pericula cum ea, ius exerci facilisis et, id sumo habemus phaedrum vim.	Lorem ipsum dolor sit amet, no solum graeci delectus duo. Sit atqui tincidunt at, te vim iisque dolores dignissim, probo perfecto eam ea. Et nec oblique iudicabit, mucius persius lobortis ea sit. Vix iusto vivendo mentitum at, at cum dico dolorem. Nam eu corpora sensibus repudiandae, et ius quas instructior. Mea idque inani no, ne sit utamur prompta. Ne pro iisque facilis corrumpit, est an nominavi postulant repudiandae, in vel vide insolens. Illud fabulas deserunt ei sed. Ea alii habeo his, phaedrum mnesarchum has eu. Accusata pericula cum ea, ius exerci facilisis et, id sumo habemus phaedrum vim.

Fig 5. Preview of the Sample Survey Dataset in Microsoft Excel Format.

2.3.4 Model Selection & Training

With clean, structured data in place, the next phase is to select and train suitable AI models for tasks such as theme extraction, key findings identification, and quote recognition. Depending on the project requirements, NLP models-such as OpenAI's API, fine-tuned BERT, or GPT-are evaluated and chosen for their strengths in handling qualitative data. To optimize data extraction, prompt templates for OpenAI's 'gpt-4o-mini' model calls are developed. Training the models involves using sample datasets *(See. above Fig.5)* and may incorporate techniques like transfer learning to improve performance. Multiple levels of testing is conducted to evaluate the accuracy and reliability of the models in identifying themes and quotes. Model parameters *(See. Below Fig. 6)* are adjusted as needed to balance accuracy and processing speed. By the end of this phase, trained models or refined prompt templates are ready for integration into the app.

```
try:
    # Call the OpenAI ChatCompletion API with required parameters
    response = client.chat.completions.create(
        model="gpt-4o-mini",  # Use the correct model name
        messages=messages,
        max_tokens=1000,
        temperature=0.4,
    )
    return response.choices[0].message.content.strip()

except Exception as e:
    return f"⚠ Error generating narrative: {str(e)}"
```

Fig 6. Initialized Large Language Model

2.3.5 App Development & Integration

This phase is focused on creating the application's user interface and backend functionalities while integrating the trained AI models. The Streamlit framework is set up to serve as the app's user interface, providing a simple, interactive environment for users. Key UI elements-such as file upload, report type selection, and progress indicators-are implemented to make the app intuitive and easy to use. On the backend, AI models are integrated to perform theme extraction, triangulation, and report generation. Additional features, including user authentication, are developed to enable sign-in, logout, and password reset functionalities. OpenAI API integration is optimized to ensure efficient parallel processing and cost-effectiveness, creating a smooth experience for end-users. At the end of this phase, the app has a fully functional user interface and backend, ready for testing.

Fig 7. User Interface of Rapid AI (Login Page)

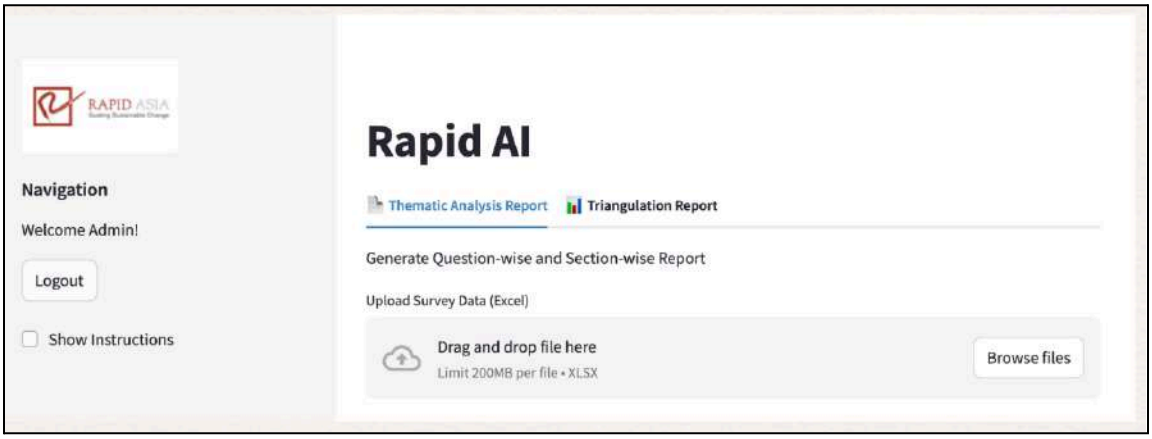


Fig 8. User Interface of Rapid AI (After Successfully Logged in)

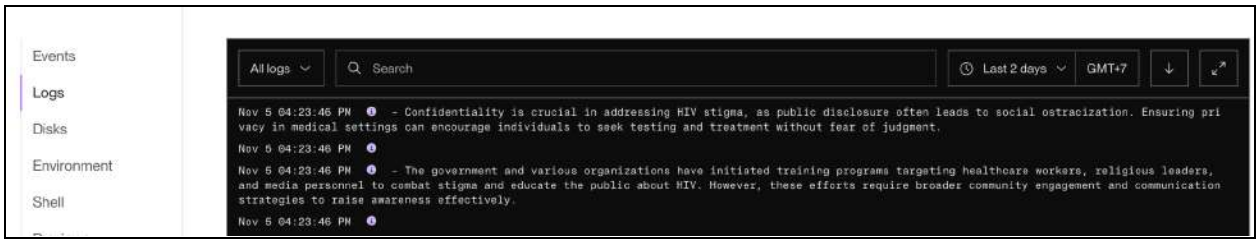


Fig 9. Application events and logs file in backend server (On cloud Platform)

i) Infrastructure of the Application:

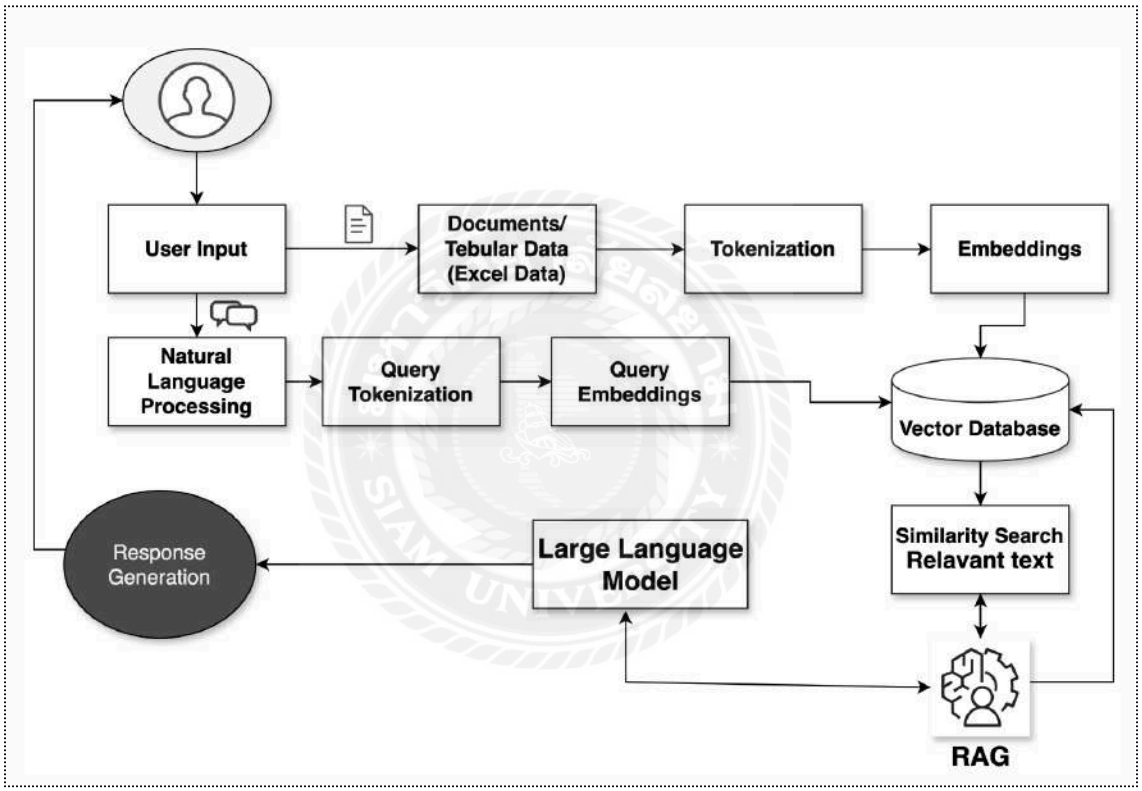


Fig 10. Rapid AI workflow

In implementing the workflow, I developed a sophisticated and integrated pipeline that effectively combines multiple AI components to process and analyze data efficiently. This system represents a significant advancement in handling complex user inputs while maintaining high accuracy and response quality.

- **System Architecture and Workflow**

The pipeline begins with a dual-stream processing approach, initiating from the user input interface. When a user interacts with the system, their input triggers two parallel processing paths: document handling and natural language processing. This divided approach ensures comprehensive data processing while maintaining operational efficiency. In the document processing stream, the system handles various forms of structured data, particularly focusing on Excel files and tabular information. This data undergoes a systematic tokenization process, transforming raw content into processable units. These tokens are then converted into embeddings - numerical vector representations that capture the semantic essence of the content [2]. This transformation enables efficient content analysis and retrieval in later stages of the pipeline.

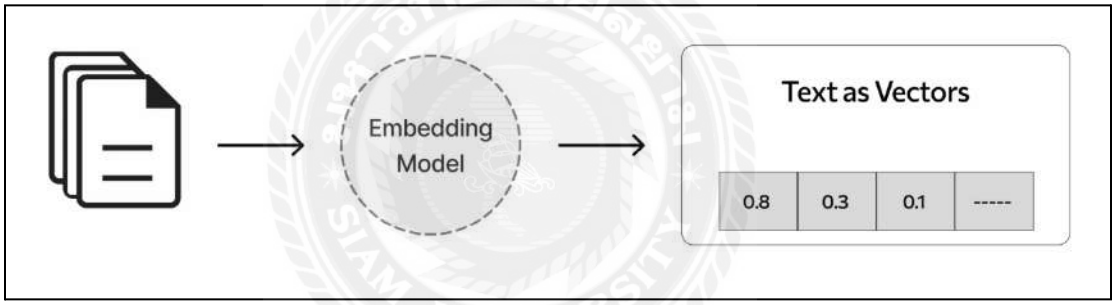


Fig 11. Systematic tokenization process

Simultaneously, the natural language processing stream manages user clicking behavior based queries through sophisticated linguistic analysis. The system processes these queries through multiple stages, including initial NLP analysis, query tokenization, and the generation of query embeddings (See. Fig 11.). This careful processing ensures that user intentions are accurately captured and translated into a format that can be effectively matched with relevant information. The Data Triangulation Process Flow diagram *(See below, Fig. 12)* illustrates a streamlined approach to combining insights from multiple qualitative data sources using AI-driven natural language processing (NLP) [4,5]. This mixed method workflow involves several key steps:

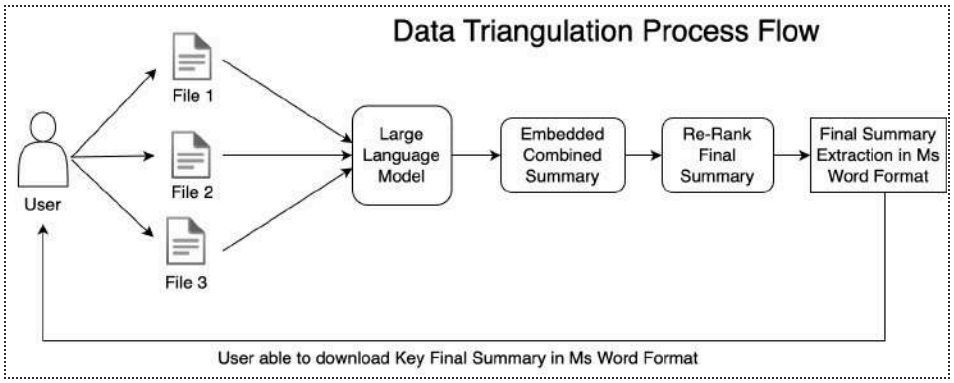


Fig 12. DataTriangulation Process

- **File Upload by User**: The process begins with the user uploading three separate documents *(File 1, File 2, and File 3)*. These files typically contain section-wise data, such as themes, key findings, and quotes, from qualitative research.
- **Large Language Model (LLM) Processing**: The uploaded files are processed through a Large Language Model (such as OpenAI's GPT). The model leverages advanced NLP [1] capabilities to read and understand the contents of each file. This step involves analyzing themes, extracting key findings, and identifying meaningful quotes that support the insights in each document [3].
- **Embedded Combined Summary**: The model then generates an Embedded Combined Summary, which combines key insights from all three documents into a single, cohesive summary[7]. This embedded summary includes themes and findings without redundancy, ensuring that the core message from each file is preserved and merged effectively.
- **Re-Ranking of the Final Summary**: After generating the combined summary, a re-rank step organizes the content based on relevance and significance. This step ensures that the most important themes and findings are prioritized, making the final output more coherent and impactful.
- **Final Summary Extraction in MS Word Format**: The re-ranked summary is then formatted and exported as a downloadable Microsoft Word document. This final summary provides the user with a concise, structured, and professional report of the key insights across all three uploaded files.
- **Download**: The user is now able to download the Key Final Summary in MS Word format, ready for further use, sharing, or reporting purposes.

ii) Innovation in Qualitative Data Analysis: My work on *Rapid AI* represented a step forward in the use of AI for qualitative research at Rapid Asia. By integrating AI into qualitative analysis, it demonstrated the potential for future technological advancements within the company. This project served as a foundation for future innovations, showing how AI tools can enhance the speed, accuracy, and scope of qualitative data analysis. My contributions thus laid the groundwork for ongoing exploration of AI applications in research, helping Rapid Asia stay at the forefront of innovation in the M&E sector.

iii) Core Processing and Integration: At the heart of the pipeline lies the vector database, a crucial component that serves as a central repository for all vector representations. This database efficiently manages both document and query embeddings, facilitating rapid information retrieval and matching. The similarity search system works in conjunction with the vector database to identify and retrieve relevant text based on user queries. A key innovation in this applied pipeline is the implementation of RAG (Retrieval-Augmented Generation) technology. This component bridges the gap between information retrieval and response generation, ensuring that responses are not only accurate but also contextually appropriate. The RAG system works seamlessly with the large language model to enhance the quality and relevance of generated responses.

iv) Response Generation and Output Refinement: The final stage of this pipeline involves refined response generation through the large language model. This model processes the retrieved information and contextual data to produce coherent, relevant answers to user queries. The response generation system ensures that outputs are well-structured and aligned with user expectations in the microsoft word document.

v) Operational Excellence and System Benefits: This integrated approach offers numerous advantages in terms of operational efficiency. The parallel processing capabilities significantly reduce response times, while the multiple validation points throughout the pipeline ensure consistent output quality. The modular architecture allows for easy scaling and maintenance, making the system both robust and adaptable to changing requirements.

The pipeline's design particularly excels in managing complex queries that require both document analysis and natural language understanding. By combining these capabilities, the system can provide more comprehensive and accurate responses than traditional single-stream processing approaches.

vi) Technical Implementation Considerations: Throughout the implementation, data flow and system integration were carefully managed. The pipeline efficiently routed information between components, ensuring synchronization and optimal resource use. Quality control measures were embedded at each stage, with continuous monitoring of performance and validation of outputs.

vii) Future-Forward Design: The system was designed to be flexible and future-ready, with modular components that make updates and enhancements easy. This setup allows the pipeline to grow alongside new technologies and adapt to changing user needs, creating a solid foundation for future development. This implementation has resulted in a powerful and reliable system that can handle complex data processing while maintaining high standards of accuracy and quality. By combining advanced AI capabilities with efficient data management, this pipeline represents a meaningful step forward in automating information processing and response generation. Overall, this pipeline demonstrates the potential for practical, sophisticated AI applications in real-world settings. Its success provides valuable insights that will guide future improvements and set the stage for further advancements in AI-powered data processing.

2.3.6 Feature Testing & Quality Assurance

Quality assurance(QA) is crucial to ensuring that the application functions correctly and meets industry standards. Each feature, such as report generation (Thematic Analysis Report and Triangulation Report), file upload, and authentication, is tested individually through unit testing to confirm that it performs as expected. Integration testing follows, where different components are verified to work seamlessly together. Load testing is conducted to evaluate the app’s stability and responsiveness when processing large datasets, simulating real-world usage scenarios. Edge cases, such as handling empty cells, multiple files, or unexpected file formats,

are tested to ensure robustness. Feedback from pilot users are collected to make necessary usability and performance adjustments. The outcome is a QA report detailing test results and any areas for further improvement.

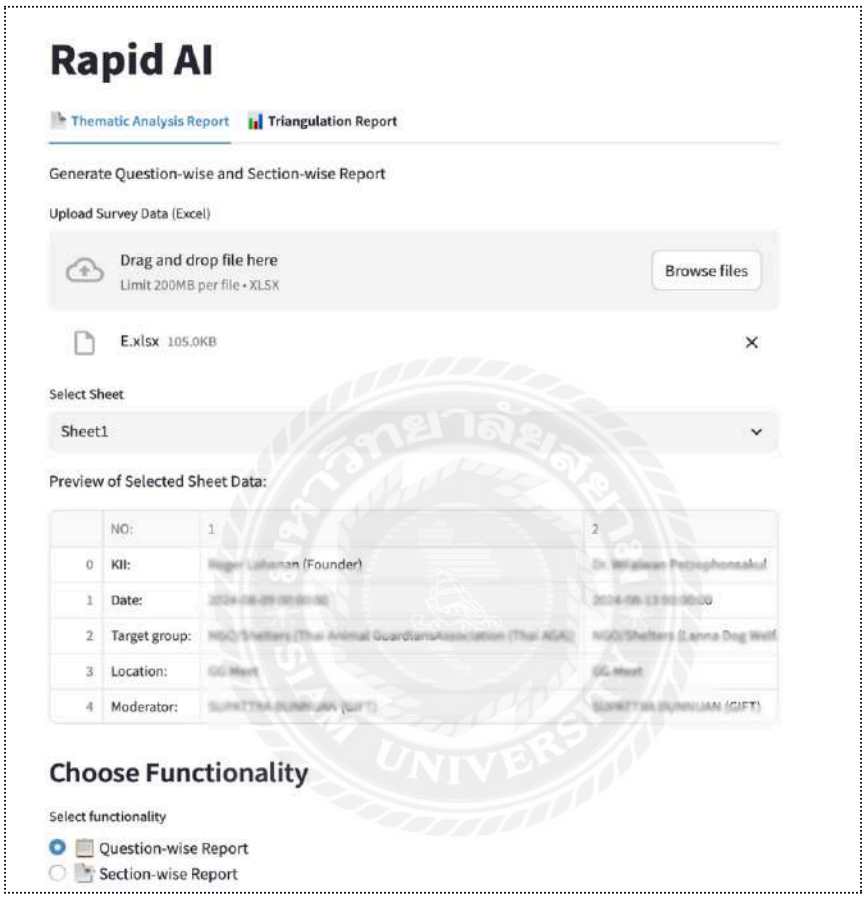


Fig 13. Feature Testing Functionality

2.3.7 Deployment Preparation

In preparation for production deployment, this phase focuses on finalizing configurations and optimizing the codebase. Essential configuration files, such as requirements.txt, .gitignore, and secrets.toml, are created to manage dependencies and secure credentials. A deployment environment is set up on the Render cloud platform for the infrastructure of the application. Security measures are implemented, especially for handling user data and authentication, to comply with data protection standards. A staging server is used to test the app in a controlled environment, verifying its readiness for production. Code optimizations are made to handle

anticipated loads and API call limits, resulting in a production-ready codebase and deployment plan.

Fig 14. Essential configuration files

2.3.8 Deployment & Monitoring

Once the app is ready, it is deployed to a production server on Render where it becomes accessible to users. Monitoring tools are set up to track key metrics, such as user activity, API usage, and error logs, providing insights into the app's performance in real time.

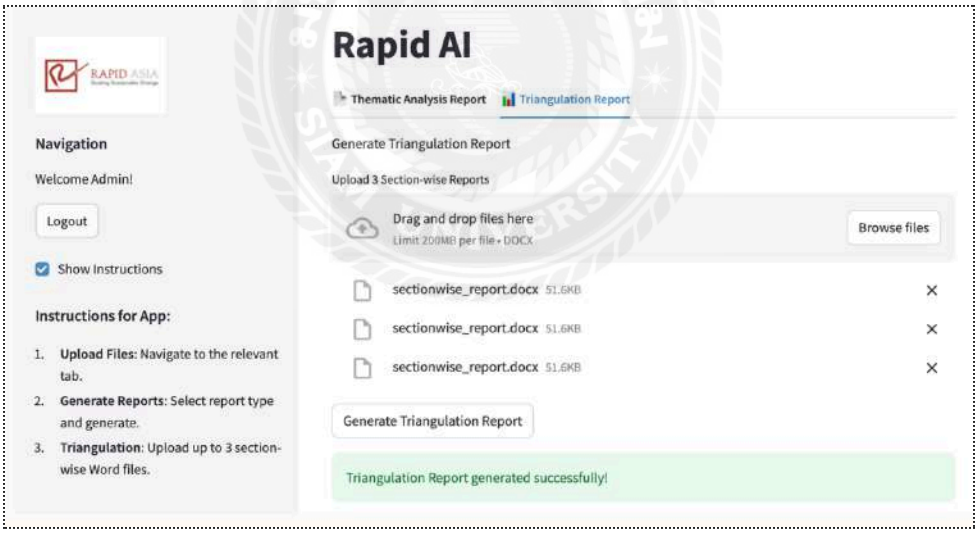


Fig 15. Triangulation Feature Testing and Monitoring

Response times, uptime, and user interactions are closely observed to ensure smooth operation. Alerts are configured to notify the team of critical issues, such as high API usage or potential server outages. Continuous feedback from users is encouraged, enabling ongoing improvements. This phase ensures that the live application meets performance expectations and offers a reliable experience for end-users.

2.3.9 Maintenance & Updates

The final phase involves the ongoing maintenance and enhancement of the application to keep it functioning optimally. Regular reviews of models and APIs are conducted to maintain performance, with updates applied as necessary. Based on user feedback, new features or improvements may be added to enhance usability and functionality. The data uploaded by users will not be stored in the app's database, it will be cleared up automatically once users complete the task. So that there will be no chance of data loss. Compatibility with external API changes is monitored to prevent disruptions. Additionally, user support and troubleshooting assistance is provided as needed. This phase ensures that the app remains up-to-date, responsive to user needs, and capable of delivering consistent, high-quality functionality over time.

2.4 Contributions as a Co-op Student

As a co-operative Education student, my contributions were integral to advancing the company's qualitative data analysis capabilities. Tasked with developing *Rapid AI*, an AI-driven tool, I focused on creating solutions that would streamline the analysis of qualitative data, enhancing both operational efficiency and data reliability. My work was directly aligned with Rapid Asia's commitment to providing data-driven, evidence-based insights. The following key contributions reflect the impact of my efforts:

- **Enhanced Operational Efficiency**: One of the primary benefits of *Rapid AI* was its ability to automate complex data analysis tasks, particularly the extraction of themes, key findings and supportive quotes from large datasets. Previously, these tasks required significant manual effort, consuming time and resources. By automating these processes, it enabled the research team to complete analyses more quickly and consistently (see Figures below). This increased efficiency allowed the company to shorten project turnaround times, meeting client deadlines more reliably and freeing up resources for other critical tasks.

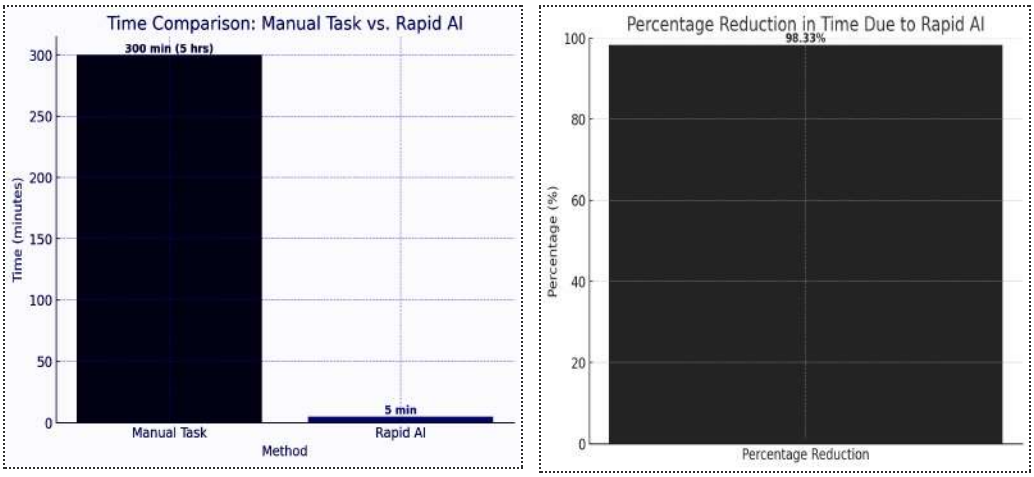


Fig 16. Time Reduction in Percentage and Time Comparison

- **Improved Data Quality and Reliability**: Manual data analysis can introduce variability and human error, which may affect the accuracy of insights. It was designed to minimize these risks by automating thematic analysis, thereby standardizing the process and enhancing the precision of data interpretation. The increased accuracy of thematic insights also strengthened the validity of the findings, enhancing the credibility of Rapid Asia's work.

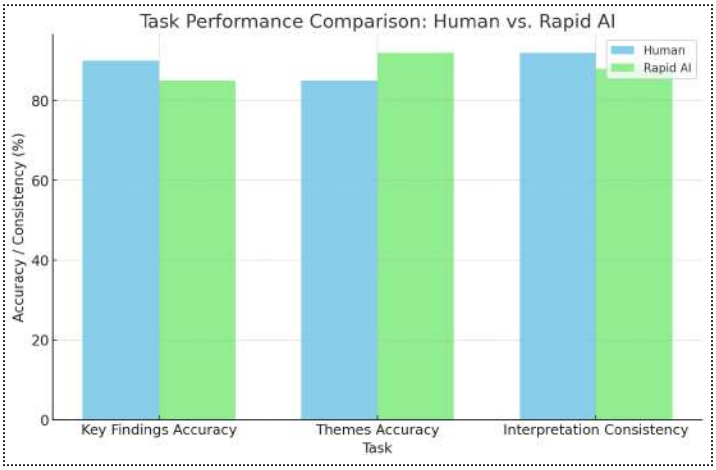


Fig 17. Task Performance Comparison

- **Scalability of Research Processes**: Another significant contribution of *this application* was its scalability, allowing the company to handle larger datasets *(up to 200MB, see Fig.*

18) and more complex projects with ease. By providing a robust solution for managing and analyzing high volumes of qualitative data, this application positioned Rapid Asia to take on larger, data-intensive projects that would have been challenging with manual analysis methods. This scalability not only supported the company's growth but also improved its competitiveness in the consulting and M&E market.

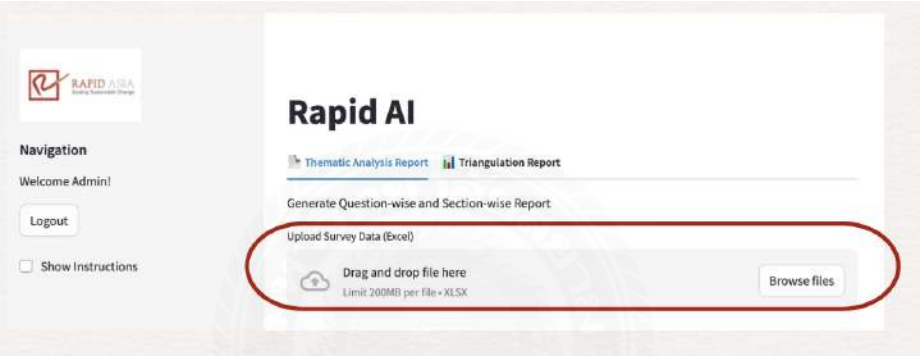


Fig 18. larger dataset handling feature

- **Technical Documentation and Knowledge Transfer**: Beyond the technical development, I contributed to the project's sustainability by creating comprehensive documentation and user guides. This documentation provided the research team with clear instructions on how to use *Rapid AI* and troubleshoot common issues, making it possible for the team to use the tool independently. By ensuring that the documentation was both thorough and accessible, I facilitated a smooth transition to using the tool and empowered team members to maximize its capabilities without requiring extensive

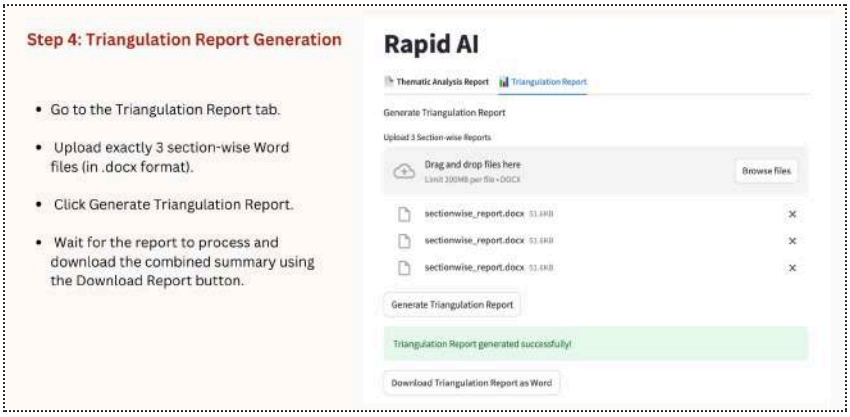


Fig 19. Task Performance Comparison

support. Through these contributions, I not only gained valuable experience in applying AI solutions to real-world challenges but also made a lasting impact on Rapid Asia's qualitative research capabilities. The development of this application not only enhanced the operational efficiency, scalability, and data quality within the organization but also set a precedent for future technological advancements in the field.

CHAPTER 3 : LEARNING PROCESS

3.1 Problem Statements

During my cooperative study at Rapid Asia Co., Ltd., a primary challenge was the time-intensive and inconsistent nature of manual qualitative data processing. The research team frequently managed large qualitative datasets, comprising open-ended responses from interviews, focus group discussion and surveys. Manually processing this data to identify themes and key findings was labor-intensive, often likely to human error, and introduced variability in the quality of analysis. This reliance on manual methods constrained the team's efficiency, especially for large-scale projects, and created delays in delivering insights to clients in a timely manner. Additionally, ensuring the accuracy and consistency of themes and key findings across different datasets was another significant issue. Given the inherent variability in qualitative responses, the process was highly subjective and could lead to inconsistencies that impacted the reliability of client reports. These challenges underscored a crucial need for a standardized, automated tool that could streamline the qualitative analysis process, minimize subjective biases, and improve processing speed, consistency, and data quality across projects.

3.2 Solution Development

To tackle these above mentioned challenges, I developed *Rapid AI*, an AI-powered application tailored for Rapid Asia's qualitative research needs. This tool was specifically designed to automate the extraction of themes and key findings from large, text-heavy datasets, leveraging advanced natural language processing (NLP) techniques. The development of this application followed a structured approach involving research, design, testing, and training, all centered on producing a user-friendly and efficient solution. The solution development process unfolded through several key stages:

- **Research and Planning**: My work started with an in-depth exploration of Natural Language Processing (NLP) methods and AI models to identify the most effective solutions for handling tasks such as theme extraction, sentiment analysis, and data summarization. This involved conducting a comprehensive literature review to understand existing techniques and analyzing successful applications in similar domains. Through this process, I identified key algorithms, including K-Means for clustering, topic modeling, Sentence-BERT for semantic similarity, and OpenAI GPT models for summarization and key finding generation.

  These algorithms were chosen for their proven ability to process large volumes of qualitative data and detect meaningful patterns within text, ensuring that the tool could meet high standards of accuracy and relevance. This foundational research not only informed my decisions on model selection but also helped me prioritize processing efficiency alongside accuracy, creating a robust framework for developing AI solutions tailored to Rapid AI's qualitative analysis needs.

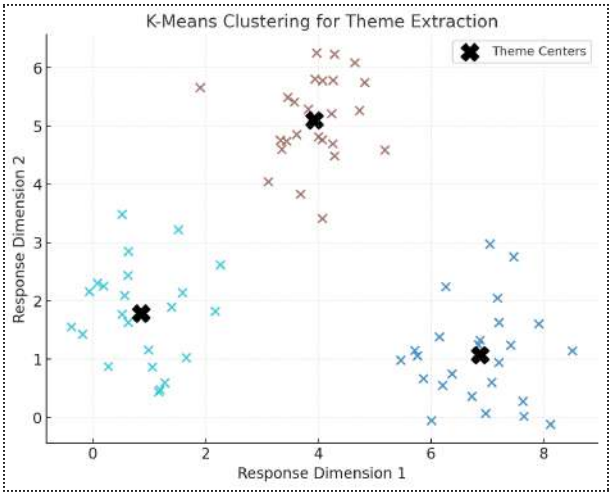


Fig 20. Thematic Analysis using K-Means Clustering Algorithm

- **Application Design and Development**: *Rapid AI's* architecture was built to streamline multiple aspects of qualitative data processing. The application was designed to automate preprocessing steps, such as data cleaning and structuring, and to support precise theme identification and key finding extraction. Leveraging OpenAI's API, I integrated advanced NLP techniques to enhance accuracy and scalability while ensuring that the

interface remained intuitive and accessible for end-users. The emphasis on accuracy and user-friendliness allowed the tool to seamlessly integrate into the research workflow, reducing manual labor while preserving data consistency.

- **Iterative Testing and Optimization**: Throughout development, I adopted an iterative approach to testing and refining its performance. Initial tests focused on the accuracy of theme and key finding extraction, while subsequent rounds of optimization fine-tuned the model to align with Rapid Asia's qualitative analysis standards. Feedback from the research team played a crucial role in refining the tool's outputs, and adjustments were made to enhance relevance and precision. This cycle of testing and improvement allowed application to meet the company's quality expectations and align with its analytical requirements.

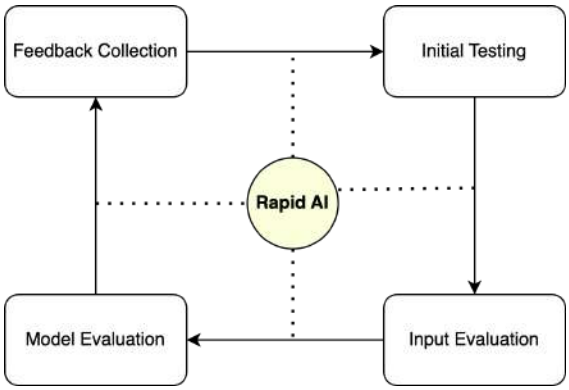


Fig 21. Iterative Testing and Optimization Workflow

- **Documentation and Training**: To ensure successful adoption, I created comprehensive documentation and a user guide, detailing the tool's functionalities, setup procedures, and troubleshooting tips. I also conducted training sessions to familiarize the research team, explaining its capabilities and answering any questions. This approach enabled a smooth transition, empowering team members to use *Rapid AI* independently and effectively in their daily workflow.

Through these stages, I addressed the core challenges identified, providing Rapid Asia with a reliable, automated process for qualitative analysis that substantially reduced processing time and improved consistency in insights.

3.3 Learning Outcomes

The co-op experience at Rapid Asia offered invaluable lessons on applying AI to real-world research challenges. The process of developing this application deepened my understanding of AI's role in enhancing qualitative research and sharpened my skills in technical problem-solving and collaborative development. Key learning outcomes included:

- **Practical Applications of AI in Qualitative Research**: This project provided hands-on experience in applying AI to solve qualitative analysis issues, from automating theme extraction to ensuring data consistency. Working directly on these applications demonstrated how AI can drive operational efficiency, streamline workflows, and deliver accurate insights in a field that typically relies on manual methods.
- **Project Management and Time Coordination**: Building *Rapid AI* required careful planning, coordination, and time management, as each phase-from requirement analysis to deployment-needed to align with project timelines. I learned to break down complex tasks into manageable steps, prioritize features based on feedback, and balance development speed with the need for accuracy and reliability. This experience improved my organizational abilities and my capability to manage end-to-end project development.
- **Cross-Functional Collaboration**: Close collaboration with the research team taught me the importance of a user-centered approach in AI development. Regular feedback sessions and progress meetings with the entire team allowed me to refine features and prioritize functionalities that directly addressed the team's pain points, underscoring the value of iterative, feedback-driven development. This process highlighted the importance of understanding user needs and aligning technical solutions with operational workflows.
- **Adaptability and Problem-Solving Skills**: Developing this application involved overcoming several challenges, from adjusting model parameters to optimize performance to handling unexpected issues in data processing. These situations required flexibility and an adaptive approach, allowing me to explore alternative solutions and modify my approach to ensure alignment with project goals. This experience improved my confidence and capability in troubleshooting AI applications and addressing technical constraints effectively.

3.4 Skills and Knowledge Acquired

During my cooperative experience at Rapid Asia, developing and optimizing *Rapid AI*, a super experience helped me build technical, analytical, and project management skills, crucial for AI-driven qualitative analysis. Here is a detailed breakdown:

Technical Skills:

- **Natural Language Processing (NLP)**: I implemented advanced NLP techniques to perform theme extraction, sentiment analysis, and similarity-based quote retrieval. Using libraries like NLTK and SpaCy, I developed algorithms that identified themes by clustering responses based on cosine similarity scores. For instance, responses sharing a similarity score of over 70% were grouped, allowing concise theme identification and reducing redundancy in qualitative insights. For similarity-based analysis, I applied *TF-IDF vectorization* and *cosine similarity* measures, which improved the accuracy of theme extraction by highlighting recurring concepts across different responses.

```
# Compute similarity between key findings and quotes
combined_texts = [key_findings] + [q[0] for q in quotes_with_sources]
vectorizer = TfidfVectorizer().fit_transform(combined_texts)
similarity_matrix = cosine_similarity(vectorizer[0:1], vectorizer[1:])
```

Fig 22. Similarity matrix applied for similar key findings and quotes

- **Python Programming**: My work on *Rapid AI* solidified my Python proficiency, particularly in handling large datasets with efficiency. I incorporated multi-threaded processing using concurrent.futures to optimize the app's performance. This was essential for processing and summarizing text in real-time, even with large datasets. The concurrent approach reduced processing time significantly by distributing tasks across threads, a critical feature given the volume of qualitative data. For example, splitting data into smaller batches allowed simultaneous processing, enhancing the app's responsiveness for users.
- **OpenAI API Integration**: A major part of my work involved integrating the OpenAI API to generate concise, human-like summaries and identify key insights from

unstructured text. I crafted customized prompts for OpenAI's language models to achieve highly tailored outputs. This integration required careful management of API tokens and rate limits, so I implemented a retry mechanism to handle rate limits gracefully. Additionally, I fine-tuned the model's temperature and max token settings to balance response creativity with conciseness, crucial for generating actionable insights.

```
# Call the OpenAI API
response = client.chat.completions.create(
    model="gpt-4o-mini",
    messages=[{"role": "system", "content": prompt}],
    max_tokens=max_tokens,
    temperature=temperature
)
```

Fig 23. OpenAI API , as a large model call

- **Automated Document Generation and Formatting**: To make *Rapid AI*'s output accessible to non-technical users, I automated the generation of well-structured microsoft word reports using python-docx. This involved designing functions that format extracted themes, findings, and quotes into justified paragraphs and bulleted lists. The final reports contained a professional layout, with clear section headings and alignment. I also included user-centric features, such as automatically detecting section headings and identifying quotes from metadata tags, making the output consistent and easy to review.

Professional Skills:

- **Project Management and Workflow Optimization**: Managing *Rapid AI*'s development required efficient project management. I was responsible for planning feature releases, coordinating tasks, and ensuring the app met its functionality and performance benchmarks. To handle user authentication, data validation, and session management, I structured modular workflows to streamline processes, which improved both the app's security and user experience.
- **Technical Documentation and User-Centric Design**: I developed detailed user guides and in-app instructions to make an application accessible for both technical and non-technical users*(see Fig.15)*. This involved writing clear documentation for tasks such

as report generation, API integrations, and error handling, along with implementing tooltips and contextual guides within the app interface.

Problem-Solving and Analytical Abilities:

- **Debugging and Troubleshooting**: Encountering real-world challenges with API rate limits and data processing bottlenecks allowed me to develop robust troubleshooting techniques. For instance, to avoid rate limits, I implemented a retry function with exponential backoff. Additionally, I used extensive logging to track the app's performance, especially for long-running tasks, which enabled me to identify and resolve issues promptly. A retry mechanism example used in the project:

```
except openai.error.RateLimitError as e:
    logging.warning(f"Rate limit exceeded: {e}")
    st.error("⚠ The service is currently busy. Please try again later.")
    return None

except openai.error.OpenAIError as e:
    logging.error(f"OpenAI API Error: {e}")
    st.error("⚠ An error occurred with the OpenAI API. Please try again.")
    return None
```

Fig 24. Error handling method

- **Analytical and Critical Thinking**: Building *Rapid AI* required balancing processing speed and accuracy in the extracted insights. I analyzed user feedback and monitored metrics like response time and model accuracy to optimize the app. These insights guided adjustments in model parameters and data processing techniques, helping the app deliver concise and relevant information consistently.

CHAPTER 4: CONCLUSION

4.1 Co-op Highlights

My internship at Rapid Asia Co., Ltd. has been a deeply rewarding experience, providing me with opportunities to grow both technically and professionally. One of the most significant achievements during my time here was the successful development and deployment of *Rapid AI,*

an AI-driven web application designed to automate the thematic analysis of qualitative data. This tool has transformed the way data is processed at the company by significantly reducing the time required for manual analysis while improving consistency and minimizing errors.

The journey of creating Rapid AI involved not only developing advanced NLP capabilities but also ensuring the tool was practical and aligned with the team's needs. Collaborating closely with the Research Department, I gained valuable insights into the challenges of social research, such as dealing with unstructured data and capturing nuanced insights. This collaboration allowed me to tailor Rapid AI's features, making it a reliable solution for extracting meaningful themes and findings from large datasets.

What stood out to me most was seeing the immediate impact of my work. Rapid AI has enhanced the team's ability to process data efficiently, giving them more time to focus on interpreting results and delivering actionable insights to clients. This alignment between technology and purpose reinforced the importance of designing tools that are both innovative and user-focused.

Beyond technical achievements, this internship also gave me the chance to apply my academic knowledge to real-world problems. From optimizing models to handling data workflows, I was able to translate theory into practice in meaningful ways. Additionally, working in a professional setting taught me the importance of effective communication, cross-functional collaboration, and maintaining attention to detail in every stage of development.

Overall, my time at Rapid Asia has been invaluable, not just for the skills I've developed but also for the perspective I've gained on how technology can be used to solve complex challenges in qualitative research. This experience has solidified my passion for AI-driven innovation and prepared me to contribute to projects that bridge technology and impact in meaningful ways.

4.2 Self-Evaluation

Throughout my time at Rapid Asia, I have made significant progress in both my technical skills and professional abilities. I am proud of my contributions to *Rapid AI*, particularly in the areas of NLP implementation, model testing, and iterative improvements based on user feedback. These experiences have strengthened my skills in Python, NLP, and AI-driven application development. Moreover, this internship provided a unique opportunity to collaborate across departments,

teaching me the importance of clear communication and user-centered design in creating effective solutions. Reflecting on my overall performance, I believe I successfully met the objectives set for my cooperative study, delivering a functional, efficient tool that supports Rapid Asia's research process. While I encountered challenges, particularly in optimizing the model's accuracy and speed, I was able to adapt and resolve issues through persistence and problem-solving. This experience has prepared me well for future roles in data science and AI, and I am confident that I will be able to leverage the skills gained during this internship in upcoming projects.

4.3 Limitations of the Study

While *Rapid AI* has proven effective in automating parts of the qualitative analysis process, there are some limitations to its current implementation:

- **Limited Scope of Theme Extraction**: The AI models used were effective in identifying primary themes but had limitations in recognizing nuanced contextual key findings or context-specific insights. This may require additional fine tuning or more advanced models to handle complex datasets.
- **Processing Time for Large Datasets**: While the application reduces the time required for qualitative analysis, it still experiences delays when processing extremely large datasets. Optimizing the model to handle such volumes more efficiently could improve performance.
- **Reliance on Initial Data Quality**: The accuracy of *Rapid AI* largely depends on the quality and consistency of the input data. Inconsistent or poorly formatted datasets can lead to errors in theme extraction, highlighting the importance of standardized data preprocessing.
- **Limited Adaptability for Varied Research Topics**: Since *Rapid AI* was tailored for specific research objectives, it may require adjustments to adapt to projects with vastly different thematic requirements or areas of focus.

4.4 Recommendations for Rapid Asia

Based on my time at Rapid Asia, I've identified a few key areas where the company could further innovate and enhance its efficiency in qualitative research. Here are my top recommendations:

- **Invest in Smarter Model Optimization and Advanced NLP Techniques:** To keep *Rapid AI* at its best, Rapid Asia could look into optimizing existing NLP models for faster processing and try out newer, advanced models to handle larger datasets. Improving model efficiency will help Rapid AI tackle bigger projects with ease. Additionally, adding more advanced contextual analysis capabilities would allow the tool to capture nuanced themes, making it even more valuable for complex qualitative studies.
- **Establish Standard Data Preprocessing Protocols:** Having a standardized approach to data cleaning and preparation across projects would help Rapid AI deliver consistent, reliable results. By setting up clear protocols for how data should be structured and preprocessed, Rapid Asia can ensure it is working with the best possible data-leading to more accurate and trustworthy insights.
- **Expand Rapid AI's Abilities for Context and Sentiment Analysis:** Adding functions like contextual and sentiment analysis to it would open up new ways to analyze qualitative data. With these enhancements, the application could detect tone, emotions, and even subtle themes in responses, providing richer, more meaningful insights that go beyond basic theme extraction. This would be especially valuable in studies that require deep social and emotional insights.
- **Offer Cross-Departmental AI and Data Tool Training:** As Rapid Asia brings in more AI and data tools, offering training across departments could help everyone make the most of these technologies. This kind of skill-building not only empowers staff to use AI more effectively but also encourages fresh, creative applications of these tools. It could lead to greater collaboration and more innovative research solutions across the organization.
- **Invest in Scalable, Cloud-Based Data Infrastructure:** To support the growing demands of data-heavy projects, Rapid Asia could benefit from a scalable, cloud-based infrastructure. Cloud solutions would provide the flexibility to adjust resources as

needed, ensuring application runs smoothly even with large datasets. Plus, secure and efficient data storage would further support the team as the company takes on larger projects.

- **Create a Central Knowledge Repository and Use Data Visualization for Client Reports:** Developing a central hub for research insights, project data, and documentation would make it easy for team members to access and leverage past findings. It would also reduce redundant efforts and promote knowledge-sharing across the team. Additionally, incorporating data visualization into client reports-like interactive dashboards and clear charts-would make reports more engaging and allow clients to quickly understand key takeaways, enhancing communication and client satisfaction.

These recommendations aim to help Rapid Asia continue to lead in innovative, efficient research by focusing on smarter data handling, cross-team collaboration, and powerful new features in *Rapid AI*. By investing in these areas, Rapid Asia can ensure its tools and processes are future-ready, giving the company a strong edge in the field of qualitative research.

REFERENCES

Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., … & Amodei, (2020). *Language models are few-shot learners.* In *Advances in Neural Information Processing Systems*, 33, 1877-1901.

Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). *Efficient estimation of word representations in vector space.* In *Proceedings of the International Conference on Learning Representations (ICLR)*.

O'Cathain, A., Murphy, E., & Nicholl, J. (2010). *Three techniques for integrating data in mixed methods studies.* BMJ, 341, c4587. doi:10.1136/bmj.c4587.

OpenAI. (n.d.). *OpenAI API documentation.* https://platform.openai.com/docs

Saldana, J., & Miles, M. B. (2014). *Qualitative data analysis: A methods sourcebook* (3rd ed.). SAGE Publications.

Streamlit. (2024). Streamlit API reference documentation. Streamlit, Inc.

Yin, R. K. (2018). Case study research and applications: Design and methods (6th ed.). SAGE Publications.