Dissertation Title: Enhancing Educational Quality Assessment in Thailand: Leveraging NLP and Machine Learning for Effective Benchmarking and Decision-Making

Programme title: Data Analytics

Name: Punyisa Phumiphol

Year: 2024

CONTENTS

List of Tables

List of Figures

Statement of compliance with academic ethics and the avoidance of plagiarism

I honestly declare that this dissertation is entirely my own work and none of its part has been copied from printed or electronic sources, translated from foreign sources and reproduced from essays of other researchers or students. Wherever I have been based on ideas or other people's texts I clearly declare it through the good use of references following academic ethics.

(In the case that is proved that part of the essay does not constitute an original work, but a copy of an already published essay or from another source, the student will be expelled permanently from the postgraduate program).

Name and Surname (Capital letters):

PUNYISA PHUMIPHOL

Date: 07/10/2024

ABSTRACT

Benchmarking plays a crucial role in External Quality Assurance (EQA), enabling educational institutions to measure their performance against defined standards and identify areas for improvement. In the context of Thailand's EQA, conducted by ONESQA, benchmarking is often hindered by the challenge of extracting useful insights from vast amounts of unstructured data.

This study attempted to address this issue by creating an NLP-based data extraction pipeline adapted to EQA benchmarking requirements through the mix-methods research, which combined quantitative data analysis with qualitative regards from stakeholders.

Stakeholder surveys found that approximately 87.5% of respondents identified the need for automated NLP techniques for transforming unstructured data into actionable insights, implying that the response emphasizes the practical relevance of creating technology to expedite and improve the benchmarking process in education. Moreover, the survey identified crucial benchmarking variables—such as student outcomes, instructional efficiency, and financial management—that are critical for assessing the performance of schools that should be used to analyze hidden patterns of the educational quality enhancing appropriate EQA benchmark settings. Leveraging these findings, the NLP pipeline was built using regular expression, pattern matching, and NER to capture the desired text from complicated documents, thereafter utilizing IF-IDF to vectorize and analyze meaningful insights with high accuracy, reaching a 98.33% match with annotated datasets and an F1 score of 1.0, which enabled the system to extract data effectively while also supporting advanced analytics and visualizations such as scatter plots and heat maps, which revealed hidden performance patterns for both regulatory and collaborative benchmarks.

The ability to turn unstructured data into clear, evidence-based insights enhances ONESQA's capacity to provide tailored recommendations for school improvement, making the benchmarking process more transparent and objective. While the pipeline demonstrated effectiveness, challenges remain in handling ambiguous language and expanding its scope to other EQA forms. Future research should focus on refining these aspects to further support data-driven decision-making and continuous improvement in Thai education.

INTRODUCTION

In Thailand's dynamic educational landscape, school quality assessment and evaluation are crucial for creating optimal learning environments. The Office for National Education Standards and Quality Assessment (ONESQA) plays a pivotal role in this effort, particularly through its external quality assessments. By conducting rigorous evaluations of educational institutions, ONESQA holds them accountable to stakeholders and ensures transparency in educational practices as instrumental in benchmarking, comparing educational institutions against established standards and best practices.

However, the process of benchmarking and selecting best practices presents significant challenges due to the diverse school environments, evolving indicators, and vast amounts of qualitative data (ONESQA, 2021). Despite the importance of benchmarking to identify areas for improvement, the sheer volume of schools requiring quality assessments can be time-consuming, particularly when dealing with qualitative data. Furthermore, traditional analysis and visualization methods often fall short in revealing weaknesses and areas for improvement, even though most schools meet all standards. In additional, while some educational institutions perceive the recommendations in external quality assessment reports as lacking clarity and concreteness (ONESQA, 2020). This issue drive ONESQA to explore different aspects and indicators for benchmarking, aiming to enhance the quality assessment process. Additionally, requesting more information from schools is not feasible due to their existing workloads and the potential violation of EQA regulations (ONESQA, 2024). The reliance on unstructured data pose significant challenges for information retrieval and analysis. Consequently, manual data extraction from school annual reports is fraught with challenges, including the time-consuming susceptibility to errors, potential human bias, and the dynamic nature of assessment criteria (ONESQA, 2022). Given these constraints, there is a pressing need to refine the quality assessment procedures to address these challenges effectively.

In response to these challenges, this dissertation project aims to leverage natural language processing (NLP) techniques to extract data from school documents and evaluate performance against quality standards and indicators. Additionally, NLP holds promise in extracting insights to inform the creation of meaningful benchmarks, thereby encouraging continuous improvement and development among schools.

To ensures the benchmarks are relevant and meaningful, reflecting the true context of each school. Creating clustering models facilitate the grouping schools based on various indicators. In this context, schools can be grouped into clusters such as high-performing, average, and needs improvement. ONESQA can provide more tailored recommendations allowing for more targeted improvement strategies and more precise benchmarking by comparing schools within the same cluster rather than across all schools. Finally, demonstrating the insights from the data and clustering models accessible, data visualization plays a crucial role to assist in highlighting key findings that allows stakeholders to explore the data and see how different schools compare based on various indicators and making data-driven decisions.

In summary, this dissertation endeavors to bridge the gap between traditional assessment methods and emerging technological solutions, with the goal of enhancing the effectiveness and efficiency of external quality assessment. Through the integration of NLP techniques, and develop robust analytical models that can handle the complexities of unstructured data and diverse school environments.

Research Questions

- What are the needs for automated data extraction and insights on the nature of educational data for setting EQA benchmarking and KPI analysis?
- How can NLP techniques be applied to effectively extract unstructured data and key information from EQA artifacts in Thai language to accommodate evolving assessment criteria to reduce evaluation time?
- How to demonstrate the utility of extracted data for meaningful analysis that can integrated for setting EQA benchmarking , and visualization suitable for stakeholders expectation?

Aims and Objectives

This dissertation aims to develop a robust automated data extraction pipeline and to illustrate its practical application in enhancing EQA benchmarking through informed analysis and visualization. The specific objectives are:

- To develop an NLP-based automated data extraction pipeline for EQA artifacts in Thai
- To demonstrate the extracted data utility through cluster analysis and visualization for EQA benchmarking

Methodology

1) Research Design

This study adopts a mixed-methods approach that aims to promote the EQA benchmarking process using NLP by integrating exploratory and descriptive research methods. Initially, an exploratory incrementally encompasses surveying ONESQA employees to determine the needs for establishing an NLP data extraction pipeline and gathering stakeholders' expectations of analysis and visualization for EQA benchmarking, which is consistent to Research Objective 1. Following the exploratory stage, the descriptive design was used to analyze the retrieved data, indicating its usefulness in benchmarking through clustering and visualization approaches, accordingly supporting Research Objective 2.

2) Data Collection Methods

The collection of data involves surveys and EQA artifacts, surveys collect numerical data on perceptions alongside qualitative information, serving to lead the development of the extraction process. EQA artifacts are processed with NLP to extract and evaluate quantitative and qualitative data, which includes text preparation, section identification, and entity extraction, guarantees a methodical approach to transforming raw data into suitable insights.

3) Analysis Techniques

Data analysis is separated into two phases including survey data is evaluated using descriptive statistics and content analysis, and EQA artifacts are processed using NLP pipeline, the performance of the pipeline was validated using precision, recall, and correlation coefficients against a manually annotated dataset.

4) Measures for Validity, Reliability, and Ethical Considerations

This study maintains validity and reliability by controlling confounding factors, employing representative samples, and obtaining pilot testing to assure consistent results. Data privacy through anonymization, secure storage, including the protection of sensitive information in EQA artifacts pose ethical issues. Survey participants were provided informed consent that are transparent concerning data usage.

Synopsis of Chapters

Chapter 1, Literature Review I, emphasizes the vital role of QA in sustaining and improving educational standards across worldwide, with an emphasis on Thailand's difficulties that drives for innovative approaches, such as NLP, to increase the efficiency and accuracy of QA operations.

Chapter 2, Literature Review II, investigates the use of NLP methods that are critical for establishing an automated data extraction pipeline for EQA artifacts in Thailand by focusing on selecting appropriate NLP techniques and proving the relevance of retrieved data with analysis and visualization. These observation need to ensure that the procedures are consistent with the research objectives and facilitate correct data analysis.

Chapter 3, Methodology, describes a mixed-methods approach which combines qualitative insights from surveys with quantitative analysis of EQA artifacts, providing the groundwork for explaining how to develop an automated data extraction pipeline align with the stakeholders guideline that can potentially revolutionize EQA benchmarking processes.

Chapter 4 presents the survey insights, the efficiency of an automated data extraction framework and its application in improving EQA benchmarking through analysis and visualization. The last section of chapter concludes with a discussion of the results, study limitations, and suggestions for future research.

The final chapter, Conclusion Remarks, summarizes the study's key findings, emphasizing the effective creation and implementation of the automated data extraction pipeline in EQA benchmarking. It concludes with practical recommendations for stakeholders on how to utilize automated data extraction for analysis and visualization to enhance EQA procedures and benchmarking strategies.

CHAPTER 1 LITERATURE REVIEW I: CHALLENGES AND NEEDS

This chapter commences with an exploration of Quality Assurance (QA) and its pivotal role in educational environments by examining the global challenges inherent in Educational Quality Assessment (EQA), with a focused analysis on the obstacles encountered by Thailand's ONESQA. Emphasizing the need for innovative solutions, it advocates for streamlining EQA processes to enhance their effectiveness in evaluating educational quality. Lastly, the chapter explores the application of Natural Language Processing (NLP) techniques investigating how NLP can overcome language-specific barriers to extract and analyze educational data.

1.1 Importance of Quality Assurance in Educational Performance Improvement

Quality Assurance (QA) in education is a systematic approach to ensuring that educational institutions consistently meet and maintain predetermined standards of quality. This procedure serves as a crucial mechanism for identifying and enhancing the quality of education across educational institutes worldwide.

Building upon this framework of ensuring consistent quality in education, QA establishes clear and consistent standards and benchmarks for educational institutions. These standards encompass various aspects. These benchmarks enable institutions to measure performance and pursue continuous improvement (Baartman *et al.*, 2007, 2011; Lucander & Christenson, 2020; Marciniak, 2018). Regular audits, evaluations, and assessments enhance accountability, ensuring transparency in operations. This transparency builds trust among students, parents, employers, and policymakers.

In Europe, the European Association for Quality Assurance in Higher Education (ENQA) plays a pivotal role in promoting cooperation among QA agencies. National bodies including the UK's Quality Assurance Agency (QAA), Finland's Finnish Education Evaluation Centre (FINEEC), and Germany's German Accreditation Council (GAC) tailor their QA processes to fit their educational contexts while aligning with European standards. Similarly, in Asia, organizations including India's National Assessment and Accreditation Council (NAAC), Malaysia's Malaysian Qualifications Agency (MQA), and Japan's National Institution for Academic Degrees and

Quality Enhancement of Higher Education (NIAD-QE), exemplify the region's commitment to maintaining high educational standards (Gapsalamov, et al., 2020).

Recognizing the global importance of QA, Thailand has developed its own robust QA system. The Office for National Education Standards and Quality Assessment (ONESQA) is pivotal in enhancing educational quality in Thailand. Through rigorous external assessments, ONESQA ensures accountability, and provides actionable insights for continuous improvement (ONESQA, 2020). This process improves educational quality and ensures transparency and builds trust among stakeholders.

In conclusion, the systematic implementation of QA processes is indispensable for ensuring educational excellence and fostering continuous improvement across institutions worldwide. The adoption of QA standards and benchmarks across various regions, including Europe, Asia, and Thailand, demonstrates a global commitment to educational excellence. Building on the established importance of QA, the next topic delves into the specific challenges faced in educational quality assessment, providing a nuanced understanding of the obstacles and potential solutions.

1.2 Challenges in Educational Quality Assessment

1.2.1 Global Challenges in Educational Quality Assessment

While EQA serves as a cornerstone for evaluating educational quality, it also presents various challenges that must be addressed. This section delves into these significant challenges and the need for new solutions to enhance EQA procedures.

1) Streamlined Data Collection and Analysis

Data collection and analysis provide a substantial difficulty in educational quality assessment. Traditional techniques of data gathering and analysis in educational institutions are labor-intensive and prone to inaccuracy. Educators and administrators have to dedicate a substantial amount of time to these approaches. Previous study by Jo (2023) indicates that standard approaches such as descriptive statistics, and correlation analysis are frequently insufficient for revealing underlying patterns and producing meaningful insights for educational improvements (Jo, 2023).

Similarly, correlation analysis detects links between variables nonetheless measures linear relationships, potentially overlooking more complicated (Jo, 2023, p. 28). Furthermore, information does not indicate a link between variables which is critical for making informed decisions. Traditional approaches cannot efficiently manage vast amounts of data, rendering them ineffective as educational systems grow and accumulate more information. These inefficiencies restrict the ability to perform comprehensive and efficient quality assurance, limiting the institution's ability to reach sensible decisions and implement essential improvements (Schellekens et al., 2023, p.2). As a result, there is a need for streamlined data collection and analysis to enhance the efficiency and accuracy of quality assurance processes.

2**) Stakeholder Engagement**

Throughout furtherance of the issues of administrative load and data analyzing, stakeholder participation remains a major issue in educational quality assessment. The literature possesses several examples of handles developed to assure assessment quality (Baartman et al., 2007, 2011; Lucander & Christenson, 2020; Marciniak, 2018). Baartman et al. (2007, 2011) created and validated a self-evaluation procedure for assessing the quality of assessment programs in competence-based education by evaluating program-level quality criteria. While Lucander and Christersson (2020) expanded on this technique, employing comparable criteria to create a process for quality assurance of assessment (PQAA) throughout whole educational programs. In addition, educational outlines and assessment plans were developed to visualize the assessments in the program's multiple phases. As a result, engaging stakeholders through mentioned methods, can lead to a better understanding of educational quality and promote continuous improvement. Effective engagement requires a more holistic and integrated approach into everyday educational practices.

In summary, Educational Quality Assessment (EQA) confronts two major issues worldwide including administrative constraints. Traditional data collection methods are prone to errors and inadequate for gaining meaningful insights. Stakeholder engagement is frequently limited owing to inadequate incorporation into daily processes.

1.2.2 Challenges in the Context of Thailand's ONESQA

Building on the global challenges identified, Thailand's ONESQA faces unique issues, particularly in the clarity and applicability of feedback provided to educational institutions.

1) Lacking Clarity and Concreteness Insights

The essence of an educational quality assessment lies not merely in the conclusions drawn from the data but in the actionable steps taken to improve the quality of schools (Chankseliani et al., 2021). A critical issue in enhancing schools in Thailand is the creation of a more effective and interactive learning environment for students, helping educational institutions adapt to a rapidly changing world by developing fundamental skills and catering to individual needs. Thus, measuring and fostering desirable learning environments in schools is an urgent protocol (Munna & Kalam, 2021).

ONESQA serves as a compass, providing direction and suggestions for the continuous improvement of educational institutions in Thailand. However, some institutions perceive the recommendations in external quality assessment reports as lacking clarity and concreteness (ONESQA, 2020). Effective feedback must reflect the true state of teaching and learning practices, be tailored to the context of each school, and be practical. The study by Sirin Tangpornpaiboon (2022) provides critical insights into the effectiveness of quality assurance mechanisms, such as ONESQA's evaluations, suggests that these mechanisms often fail to lead to substantive improvements due to a disconnect between assessment results and actionable feedback. However, in many cases, the assessment results do not translate into meaningful changes in teaching practices or improvements in student performance. This failure often arises from a disconnect between the data collected from assessments and the actionable feedback provided to educators (Tangpornpaiboon S., 2022, p. 72). According to this issue, ONESQA is aware of this issue and is actively seeking to address it by attempting to obtain specific data points from each school to provide tailored analysis. However, the practical challenges remain significant due to the reliance on unstructured data. Manual data extraction from school annual reports is time-consuming, prone to errors, and subject to human bias (ONESQA, 2021).

To overcome these challenges, there is a growing need to develop advanced techniques for automated data extraction, would enable more efficient and accurate analysis, leading to actionable feedback and tailored improvements for each school.

2) Data Collection and Analysis

Effective data collection and analysis are critical for educational quality assessments. However, Thai schools face significant challenges due to their existing workloads (ONESQA, 2024). This issue parallels a global challenge highlighted by Schellekens (2023), which underscores the significant administrative burdens faced in managing quality assurance processes (Schellekens et al., 2023, p. 3).

In the context of Thailand's EQA, the situation is further complicated by restrictions on directly requesting specific information from schools, which can lead to potential violations of EQA regulations. ONESQA mandates the collection of paperwork reported by schools to their authority for initial analysis to ensure compliance with quality standards. Additionally, ONESQA collects further data through observations and interviews conducted in the field (ONESQA, 2024). This process involves handling extensive qualitative data, which is challenging to analyze within a short time frame, ONESQA faces significant challenges in data collection (ONESQA, 2021).

Unlike Estonia, which has implemented a comprehensive public education database facilitating streamlined data collection and analysis. The Estonian Education Information System plays a crucial role in collecting and analyzing data on various educational parameters (Harju-Luukkainen et al., 2022, p. 202). The autonomy granted to schools within this system allows them to implement customized improvement plans based on these data insights. Another innovative approach to overcoming data collection and analysis issues in EQA is highlighted by Schellekens (2023). This study finds that a digital application can enhances the accuracy of quality assurance processes by providing real-time data analysis and reporting (Schellekens et al., 2023).

The innovations in educational quality assessments have demonstrated significant benefits in enhancing transparency and efficiency. However, ONESQA cannot adopt these innovations due to regulatory constraints (ONESQA, 2024). Specifically, ONESQA is not permitted to request specific information directly from schools, particularly through the development of a data collection platform that would require schools to input data. These obligations fall under internal quality assurance, which is managed by the schools or their authorities. As an external quality assurance body, ONESQA is prohibited from imposing additional administrative burdens on schools.

Overall, This presents a significant challenge for ONESQA in developing efficient data collection and analysis. To address these constraints, there is a clear need for advanced techniques capable of automating the extraction of unstructured data from existing school's report files. Natural Language Processing (NLP) techniques, in particular, could be highly effective in this context, offering a solution that minimizes the administrative load on schools while enhancing the accuracy and efficiency of data collection and analysis.

3) Benchmarking to Identify Areas for Improvement in EQA

Ensuring that benchmarks are relevant and meaningful is crucial for ONESQA in Thailand. Effective analysis of EQA insights is key to achieving this goal. Over the past 20 years, ONESQA data indicates that 80% of schools have met all standards at the highest level (ONESQA, 2020). While this high compliance rate suggests overall success, it may obscure underlying issues such as inequalities or specific areas needing improvement.

Despite impressive compliance, research from the OECD highlights significant disparities in educational resources and infrastructure between urban and rural areas (Vandeweyer M., et al., 2020). These disparities necessitate more efficient data analysis to improve educational outcomes. Effective benchmarking can reveal these hidden disparities, guiding resource allocation to the areas most in need. In the context of EQA, benchmarking involves comparing a school's performance against established standards or best practices to identify areas needing improvement. Effective benchmarking should address uniform high scores by identifying specific areas for continuous improvement, ensuring that even top-performing schools receive detailed analysis and tailored recommendations for ongoing development and data-driven decisions (Kayyali, 2023).

Advanced benchmarking techniques, however, provide deeper insights by considering multiple factors, revealing correlations between these factors and student performance. This approach is particularly important for EQA in Thailand, where almost schools have consistently met all standards at the highest level. While these schools excel in student outcomes, teaching approaches, and management processes, the current evaluation system by ONESQA merely confirms their excellence without providing a nuanced analysis due to traditional analysis and benchmarking methods often rely on simple metric (ONESQA, 2020).

By incorporating a more granular benchmarking process, ONESQA can help already excellent schools identify and focus on specific aspects where they can enhance performance, fostering a culture of continuous improvement. Benchmarking is essential for identifying areas for improvement in educational quality, but the process presents significant challenges for ONESQA due to diverse school environments and evolving indicators (ONESQA, 2021).

In summary, traditional data collection, analysis and visualization methods often fail to reveal weaknesses, even though most schools meet all standards. To address these challenges, a more sophisticated approach is required. This involves the adoption of advanced techniques, such as clustering algorithms, which can enhance benchmarking by identifying patterns and grouping similar schools based on performance metrics. This enables a more nuanced analysis of areas needing improvement, allowing ONESQA to provide more targeted and actionable insights, facilitating continuous improvement and equitable resource allocation.

1.3 NLP Techniques in Extracting and Transforming Unstructured Data

This sections will explore the intricacies of NLP techniques in addressing the challenges of data extraction to understand how to utilize and select the technique that could solve the challenges from education artifacts for streamline data collection.

1.3.1 Numeric Data Extraction

Numeric data extraction from unstructured text presents unique challenges, including variability in formats, units, and representations. A common approach is the use of regular expressions and rule-based methods to identify numerical patterns and extract relevant metrics (Kasliwal, 2018). Considering the aforementioned challenges, it is pertinent to explore NLP techniques starting with numeric data extraction, emphasizing the role of regular expressions and rule-based methods.

1) Regular Expressions and Rule-Based Methods

Regular expressions and rule-based methods serve as foundational tools in NLP, facilitating the extraction of numeric data from unstructured text. Regular expressions, defined as sequences of characters that delineate search patterns, empower NLP systems to pinpoint specific substrings within textual data, that encompass diverse numerical formats, integers, decimals, percentages, and currency symbols (Kasliwal, 2018). Conversely, rule-based methods entail the formulation of explicit rules the structure and format of numeric data embedded within text. Moreover, rule-based approaches often integrate contextual analysis alongside pattern matching, enriching the accuracy of numeric data extraction by considering surrounding text or linguistic cues. Consequently, both regular expressions and rule-based methods consistently deliver high precision and recall rates in numeric data extraction endeavors, owing to their deterministic nature (Kasliwal, 2018).

Hence, the selection of regular expressions and rule-based methods for numeric data extraction in educational materials is grounded in their proven effectiveness, adaptability, and computational efficiency, promising robust outcomes.

2) Term Frequency-Inverse Document Frequency (TF-IDF)

TF-IDF analysis involves measuring the significance of terms within individual documents and across the entire document collection. The term frequency (TF) component quantifies the occurrence of a term within a document, while the inverse document frequency (IDF) assesses the rarity of the term across the document collection. The TF-IDF-based summarization process involves tokenization of the text, calculation of TF-IDF scores for each term, scoring of sentences based on the TF-IDF scores of the terms they contain, and selection of top-ranked sentences to form the summary (Kasliwal, 2018, p. 99). This approach enables the extraction of essential numeric information from documents while maintaining the integrity and context of the original text.

In the context of extracting condensed structured numeric data from educational materials, TF-IDF is more appropriate choices in terms of being well-suited for identifying and extracting key numeric terms or phrases based on their importance within the document corpus.

1.3.2 Categorical Data Extraction

Named Entity Recognition (NER) techniques play a pivotal role in NLP by identifying and classifying specific entities within text, including names of people, organizations, and other predefined categories, thereby enhancing comprehension of text semantics and context. These techniques harness machine learning models (CRF, BERT) trained on annotated datasets to accurately recognize and categorize entities. These models adeptly navigate the contextual nuances of categorical data within text, distinguishing between entities with similar surface forms besides distinct semantic meanings to ensure precise categorization (Kroll M., et al., 2022). Previous research has empirically validated the effectiveness of NER techniques in diverse contexts conducted by Wu H. (2020) demonstrated the successful extraction of product categories from e-commerce reviews using CRF-based NER models. Similarly, Magoc T. (2023) showcased the enhanced performance of NER for social media texts through fine-tuning pre-trained language models (BERT). Additionally, NER seamlessly integrates with other NLP tasks, such as text classification and sentiment analysis, facilitating comprehensive data analysis and decision-making processes within the NLP pipeline (Wu H., et al., 2020; Magoc T., et al., 2023).

These empirical use cases underscore the robustness and applicability of NER techniques in real-world scenarios, thus affirming their selection as a key technique for extracting categorical data within the dissertation.

1.3.3 Narrative Data Extraction

With categorical data extraction methods outlined, our focus shifts to narrative data extraction. Abstraction-based summarization techniques, particularly those leveraging deep learning models, present a sophisticated methodology for extracting narrative data. This approach encompasses a comprehensive process, beginning with preprocessing tasks such as tokenization and sentence segmentation to prepare the narrative text. The utilization of a Sequence-to-Sequence (Seq2Seq) architecture, comprising an encoder and decoder, facilitates the encoding of the input narrative text into a fixed-length vector representation, effectively capturing its semantic meaning (Kasliwal, 2018).

Empirical studies by Khamphakdee & Seresangtakul (2023) underscore the effectiveness of deep learning-based Seq2Seq models with attention mechanisms in text summarization tasks, yielding state-of-the-art performance and capturing key ideas and concepts from the original text. Considering the nature of narrative data, characterized by descriptive and explanatory text with diverse linguistic expressions, abstraction-based summarization techniques emerge as a suitable approach for extracting relevant information. Leveraging deep learning models, particularly Seq2Seq architectures with attention mechanisms, enhances the effectiveness of abstraction-based summarization, particularly in extracting narrative data from educational materials, further advocating for their adoption in research and practice (Khamphakdee & Seresangtakul, 2023).

To sum it up, these methods can generate summaries that encapsulate the essence of the original text while preserving important contextual information and coherence, thereby mitigating the risk of information loss associated with extraction-based approaches.

1.4 Research Gaps

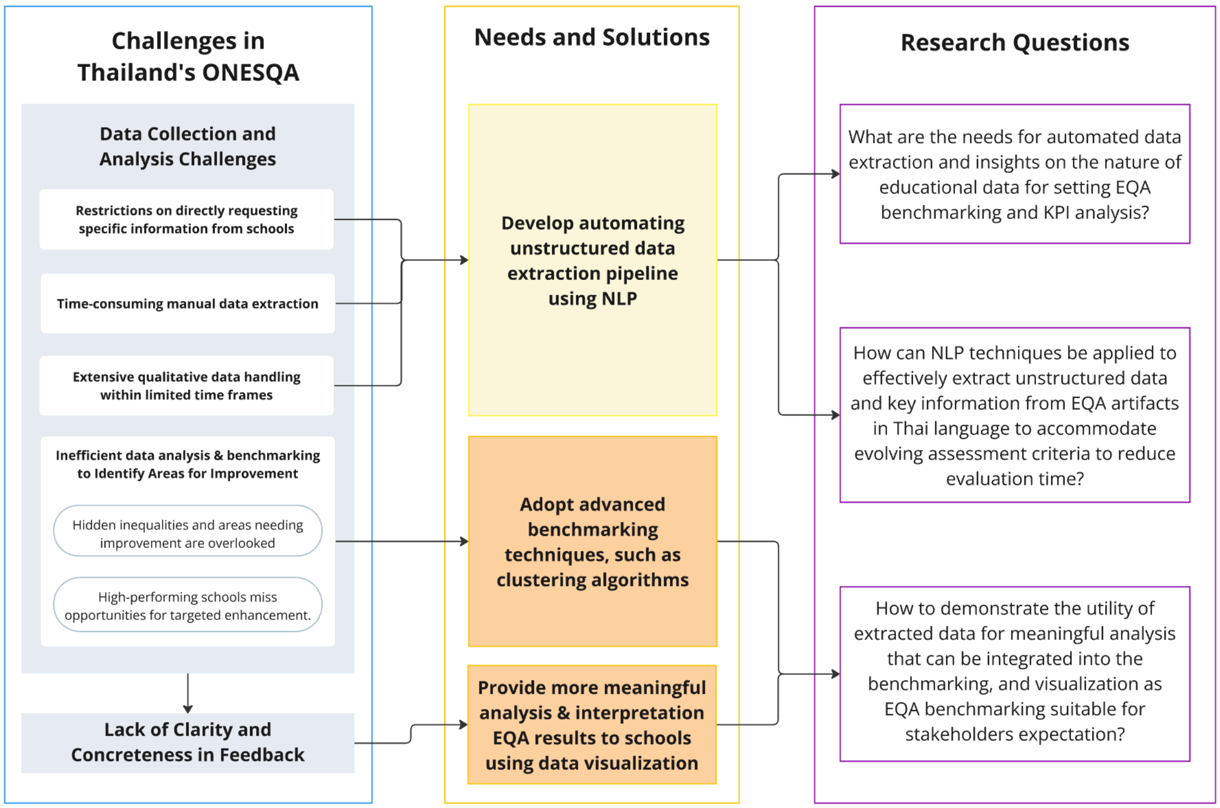


Figure 1: Research Gap

Despite the extensive application of NLP in various domains (Kasliwal, 2018; Wu H., et al., 2020; Magoc T., et al., 2023), its use in Thai education remains under-researched. The challenges unique to the Thai language (ONESQA, 2020; 2021; 2024), such as complex script and tonal nature, have not been sufficiently addressed in existing studies. Additionally, there is a lack of robust methods for automating the extraction and analysis of educational artifacts in Thai (Khamphakdee & Seresangtakul, 2023). This research aims to fill these gaps by exploring and developing NLP techniques that can efficiently process Thai educational data, thereby enhancing the accuracy and efficiency of EQA processes.

Summary

This chapter discussed the critical role of QA in maintaining and enhancing educational standards globally, with a particular focus on the challenges faced in Thailand. It advocates for innovative solutions using NLP to improve the efficiency and accuracy of QA processes. Moreover, the last section demonstrated principles utilizing NLP techniques for extracting and transforming unstructured data. These can provide robust frameworks for extracting and transforming unstructured educational data, paving the way for more efficient transformation and interpretation raw data into valuable insights. With a solid understanding of general NLP methodologies, the next chapter will examine use cases of NLP, particularly in the areas of text extraction and entity extraction, and explores how these techniques can be applied to educational data and solutions associated with NLP for Thai text in educational content.

CHAPTER 2 LITERATURE REVIEW II: USE CASES OF NLP

The previous chapter addressed the challenges and needs for developing data collection and analysis systems, emphasizing the importance of leveraging NLP to automate data extraction for enhancing EQA benchmarking. This chapter focuses on specific use cases of NLP, particularly in the areas of text extraction and entity extraction, and explores how these techniques can be applied to educational data in the Thai language.

2.1 Text Preprocessing and Vectorization

Text preprocessing and vectorization are fundamental steps in NLP that convert raw text into a structured format suitable for analysis. The methodologies involved include Tokenization, which divides text into smaller units; Stop Word Removal, which eliminates common words lacking significant meaning; Stemming, which simplifies words to their base or root forms; and Term Frequency-Inverse Document Frequency (TF-IDF), a statistical method for assessing the importance of a word within a document relative to a corpus.

The practical implementation of these methodologies is seen in the work of Zaki et al. (2022), who applied these techniques to align Course Learning Outcomes (CLOs) with Program Learning Outcomes (PLOs). By employing tokenization, stop word removal, and stemming, they prepared the text data for analysis, and used TF-IDF for vectorization to calculate cosine similarity between CLOs and PLOs, facilitating automated mapping (Zaki N., et. al., 2022). Likewise, Sirsat and Chavan (2016) utilized these preprocessing steps to extract and process information from news web pages, showcasing the effectiveness of regular expressions (regex) for pattern matching and data extraction (Sirsat, S. & Chavan, V., 2016).

To conclude, the aforementioned examples underscore the importance of these preprocessing methods, paving the way for their application in our research, an automated data collection pipeline, these preprocessing techniques can be used to clean and tokenize the text data. Applying TF-IDF will facilitate the vectorization of this text, making it possible to analyze and cluster schools based on performance metrics extracted from the PDFs. This approach ensures that the data is consistently formatted, leading to more accurate and meaningful analysis.

2.2 Pattern Matching and Entity Extraction

Pattern matching and entity extraction are critical for identifying specific pieces of information within large text corpora. The right methods needs to consider the complexity of identifying pattern of text and the nature of data. Techniques such as regular expressions (regex) and advanced algorithms that integrates regex with specific methods for complexity data are essential.

Regex (regular expressions) are sequences of characters that define search patterns, useful for text extraction and manipulation.. Sirsat and Chavan (2016) demonstrated the effectiveness of regex for pattern matching and text extraction from news web pages (Sirsat, S. & Chavan, V., 2016). Chen et al. (2023) introduced the Smore algorithm, which combines regex with semantic reasoning to handle complex data extraction tasks (Chen, Q et al, 2023).

Entity recognition involves identifying and classifying key information (entities) in text. Jofche et al. (2023) used BERT and BioBERT for entity recognition in their PharmKE platform, showcasing their capability to handle complex language tasks and improve information retrieval accuracy (Jofche, N. et. al., 2023).

To conclude, the aforementioned examples underscore the importance of these preprocessing methods, paving the way for their application in our research for the tasks of extracting specific patterns and entities from PDFs, we can implement regex to enhance pattern matching accuracy. Utilizing BERT models for entity recognition will further improve the extraction process, enabling precise identification of key information from school performance metrics in PDFs.

2.3 Information Extraction Techniques

The process of information extraction involves retrieving significant data from unstructured sources, including text documents and images. This task employs various techniques like Optical Character Recognition (OCR), Genetic Algorithms (GA), regex, and semantic reasoning. OCR is used for converting different formats of documents, such as scanned images or PDFs, into searchable and editable text. Genetic Algorithms, inspired by evolutionary principles, are used to optimize the performance of OCR and other extraction methods. Regex, or regular expressions, are utilized to define patterns for text extraction. Semantic reasoning helps in interpreting the meaning and context of the text, thereby improving the precision of data extraction.

In practical applications, Malashin et al. (2024) utilized OCR in conjunction with Genetic Algorithms and Neural Networks to extract text and critical information from document images. They employed PyTesseract and easyOCR for text recognition and used Genetic Algorithms to optimize OCR settings (Malashin, I., et al, 2024). Hansen et al. (2019) focused on extracting and classifying unstructured data from PDFs using OCR combined with deep learning techniques like Faster R-CNN, which helps in segmenting and detecting document elements (Hansen, M., et al, 2019). Chen et al. (2023) developed the Smore algorithm, which combines regex with semantic reasoning, to handle complex data extraction tasks, enhancing both the accuracy and efficiency of the information retrieval process (Chen, Q., et al, 2023).

Overall, the case studies highlight the effectiveness of information extraction techniques. To automate the extraction of text from PDFs, utilizing OCR for converting scanned documents into machine-readable text could be adopted. Regex can be used to identify and extract specific patterns in the text, while semantic reasoning will aid in understanding the broader context. These methods will constitute a robust data collection pipeline, ensuring the precise extraction of school performance metrics for detailed cluster analysis.

2.4 Feature Extraction

Feature extraction is a process that involves transforming raw data into a set of characteristics for further analysis. A widely used technique for this purpose is Term Frequency-Inverse Document Frequency (TF-IDF). TF-IDF is a statistical method that assesses the significance of a word in a document relative to a collection of documents. This technique is useful for identifying the most relevant words and features in text data.

In practical applications, Zaki et al. (2022) utilized TF-IDF to vectorize text data in their research on mapping Course Learning Outcomes (CLOs) to Program Learning Outcomes (PLOs). This method enabled them to effectively assess the importance of different terms in their documents (Zaki N., et. al., 2022). Similarly, Meesad (2021) used TF-IDF for feature extraction in the context of fake news detection, demonstrating its versatility in various NLP tasks (Meesad, P. 2021).

As evidenced by these implementations, the role of these techniques is crucial, and we aim to incorporate them into our thesis. For the purpose of feature extraction from text data in PDFs, TF-IDF will be used to identify and quantify the most significant terms. This will enable us to convert the text data into a structured format, which is essential for subsequent analysis, including clustering and classification of school performance metrics.

2.5 Model Training and Classification

Model training and classification are essential components in the field of machine learning, involving the training of algorithms to categorize data into predefined classes. Various techniques are utilized for these purposes, including Support Vector Machines (SVM), k-Nearest Neighbors (k-NN), and more advanced models like BERT, XLNet, NeuralCoref, and Semantic Role Labeling (SRL). SVM and k-NN are traditional machine learning algorithms known for their effectiveness in classification tasks. BERT and XLNet are transformer models that are highly effective in a range of NLP tasks, such as text classification and comprehension. NeuralCoref addresses co-reference resolution, which is crucial for understanding the context when different words refer to the same entity. SRL helps in identifying the roles played by entities in sentences.

For instance, Jofche et al. (2023) utilized these advanced models in the PharmKE platform for extracting and analyzing pharmaceutical data. Their work demonstrated the effectiveness of BERT, XLNet, NeuralCoref, and SRL in handling complex linguistic tasks and improving the accuracy of information retrieval (Jofche, N. et. al., 2023). Additionally, studies by Chen et al. (2023) and Meesad (2021) highlighted the versatility and robustness of these models across various classification tasks in NLP applications (Chen, Q., et al, 2023; Meesad, P. 2021).

In summary, these case studies underline the potential utility of these techniques in the context of our thesis. For clustering schools based on performance metrics, we could training classification models using SVM or k-NN to categorize the extracted data into distinct performance categories. Furthermore, employing advanced models like BERT and XLNet can enhance text understanding and improve classification accuracy. This combined approach will provide a comprehensive framework for analyzing school performance data extracted from PDFs.

2.6 Handling Specific Language Challenges

Certain NLP tasks require specialized techniques to handle language-specific challenges, such as segmentation and sequence processing. Sparse Distributed Representations (SDRs) are effective for segmenting text in languages without explicit word boundaries, especially Thai. Long Short-Term Memory (LSTM) networks, a type of recurrent neural network (RNN), are adept at handling sequences of data, retaining context over long dependencies.

A practical application of these methods can be seen in the work of Soisoonthorn et al. (2023), who applied SDR-based techniques for Thai word segmentation. Their study demonstrated significant improvements in accuracy for languages like Thai, which pose unique segmentation challenges due to the absence of spaces between words (Soisoonthorn, T. et al., 2023). Similarly, Meesad (2021) used LSTM networks for sequence handling in the context of fake news detection, highlighting the model's capability to retain context and enhance classification performance (Meesad, P. 2021).

In addition to these techniques, recent research by Phatthiyaphaibun et al. (2023) has expanded the toolkit for addressing the unique challenges of the Thai language. They utilized a combination of Conditional Random Fields (CRFs), LSTM networks, and BERT (Bidirectional Encoder Representations from Transformers) to tackle issues like tokenization and part-of-speech tagging. CRFs are particularly useful for sequence labeling, a critical task when explicit word boundaries are absent. LSTM networks aid in capturing long-term context, essential for understanding the full meaning in complex sentences. BERT, with its extensive pre-training on large text corpora, provides robust capabilities for a range of NLP tasks (Phatthiyaphaibun et al., 2023).

However, the development of NLP solutions for Thai is significant challenges due to the lack of annotated datasets. Furthermore, the complex structure of the Thai language, with its diverse expressions and intricate grammar rules, adds another layer of difficulty. To address these challenges, Phatthiyaphaibun et al. (2023) suggest that future efforts should focus on creating larger and more diverse datasets and refining algorithms to better handle Thai's specific linguistic features (Phatthiyaphaibun et al., 2023). Integrating these insights into broader NLP applications, particularly in the context of educational data analysis, allows for more accurate text extraction and data processing. By understanding and adapting to the specific needs of languages like Thai, NLP practitioners can improve the quality of text classification, clustering, and other critical tasks, thus enhancing the overall performance and utility of NLP systems in multilingual contexts.

To conclude, Figure 2 illustrates the overview of structured framework for information extraction (IE), emphasizing the application of various NLP techniques and methodologies, as synthesized from the literature review, and is divided into six distinct stages. Each stage of the framework showcases a range of NLP techniques that reflect the strategic choices made by researchers to overcome specific challenges in data extraction. Additionally, this framework underscores the importance of carefully selecting NLP techniques, as researchers emphasize the need to consider data characteristics, the quality of annotated datasets, computational performance, and alignment with the objectives of extraction at each stage. The structured approach presented in this framework ensures that each task is executed with precision, leading to accurate and efficient information extraction.

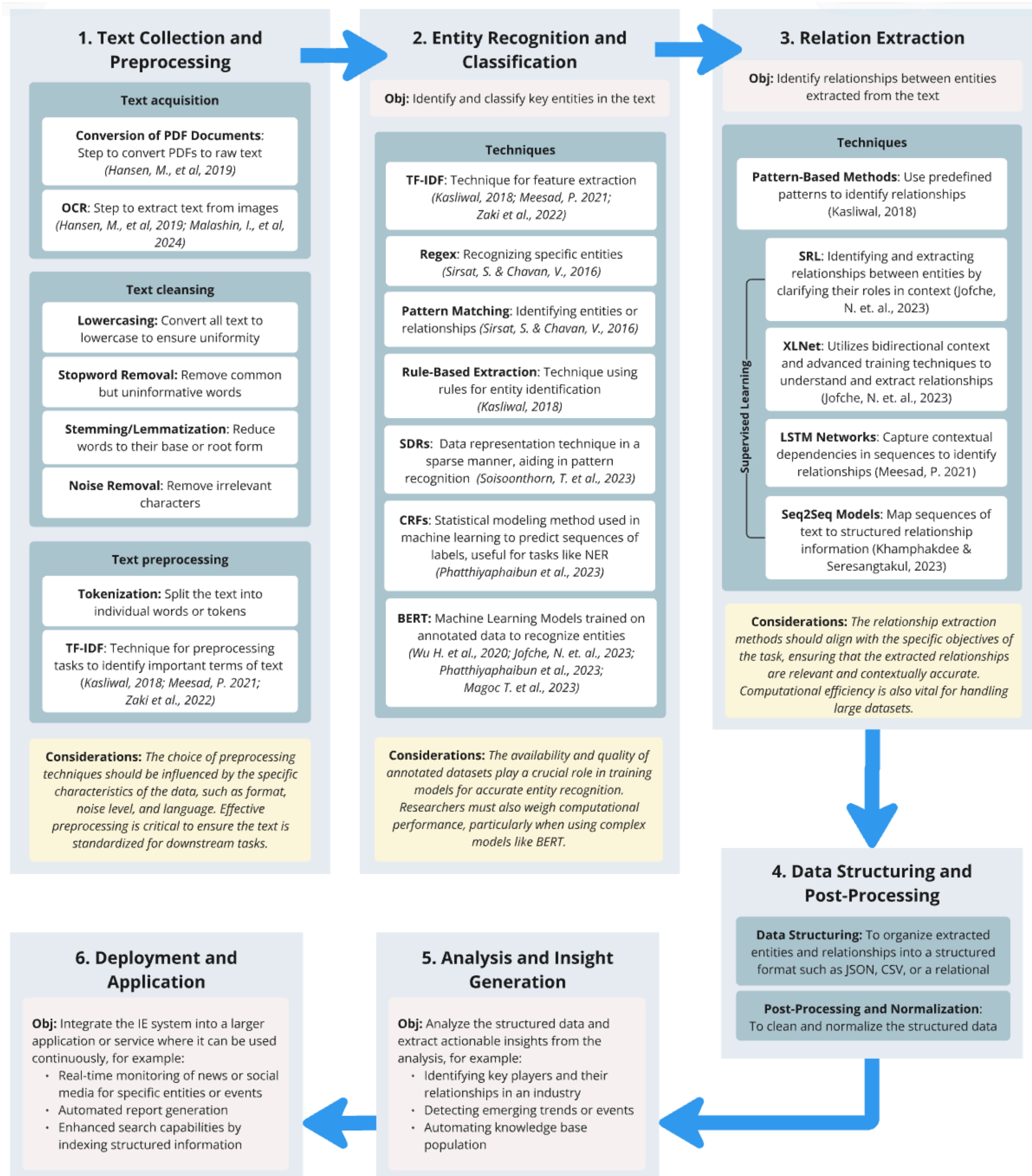


Figure 2: Summary of NLP Techniques for Data Extraction Use Cases

2.7 Research Gaps

In reviewing the existing literature, it is evident that while NLP techniques have been explored in various educational contexts, there is a significant gap in research focused on the practical deployment of these technologies for EQA. Current research in NLP has focused primarily on fundamental techniques such as text preprocessing and vectorization (Sirsat, S. & Chavan, V., 2016; Zaki N., et. al., 2022), which are essential for analyzing educational data. However, there is a significant gap in studies that advance beyond these basics to develop comprehensive NLP-based automated data extraction pipelines tailored specifically to Educational Quality Assurance (EQA) in the Thai language.

This gap is particularly evident in the application of advanced techniques like the combination of Regex, BERT, and OCR for pattern matching and entity extraction tasks (Hansen, M., et al, 2019; Chen, Q., et al, 2023; Jofche, N. et. al., 2023; Malashin, I., et al, 2024), especially when dealing with complex and unstructured Thai educational documents, such as handwritten notes or scanned reports. These challenges are further compounded by the unique linguistic features of the Thai language (Soisoonthorn, T. et al., 2023; Phatthiyaphaibun et al., 2023), such as the lack of explicit word boundaries, which are insufficiently addressed by traditional NLP approaches.

This research aims to fill these gaps by addressing the intricacies of processing unstructured Thai educational data (Research Question 2), and developing a comprehensive automated data extraction pipeline (Research Question 1). Furthermore, the study will illustrate the utility of extracted data using advance analysis and visualization techniques to integrate extracted data into EQA benchmarking and present insights that align with stakeholders’ expectations (Research Question 3).

Summary

To sum up this chapter, the integration of NLP techniques and frameworks provides a comprehensive methodology for automating text extraction from PDF files and conducting meaningful analysis. The preprocessing techniques play a crucial role in standardizing the text data. Pattern matching and entity extraction methods are employed to retrieve specific content from the text. Feature extraction methods then convert this text into structured data, which is necessary for further analysis. Additionally, specialized handling techniques are used to tackle language-specific challenges that may arise. These insights were synthesized into a structured framework that integrates these techniques into a cohesive methodology for information extraction. The framework organizes the process into distinct stages and guides the selection of appropriate techniques based on specific objectives and tasks, ensuring accurate and efficient extraction of relevant.

CHAPTER 3 METHODOLOGY

The primary aim of this study is to develop an automated benchmarking framework using NLP techniques for extracting relevant data from External Quality Assurance (EQA) artifacts and to cluster this data for insightful analysis. This chapter discusses the research design, data collection methods, data analysis procedures, and ethical considerations.

3.1 Research Design

Based on the research questions and objectives, a combination of exploratory and descriptive research designs is most suitable for this study. Initially, exploratory research will be conducted to investigate the need for automated data extraction through an online survey of Onesqa leaders and officers. This phase will also involve developing an NLP-based data extraction pipeline by reviewing literature to identify existing NLP techniques and their applications in similar contexts. The goal is to develop initial models by selecting NLP techniques appropriate for specific tasks related to Thai EQA artifacts. These steps align with Research Objective 1 and Research Questions 1 and 2.

Following the prototyping phase, the second phase will employ a descriptive research design. This phase will demonstrate the value of the data extracted through NLP and surveys by thoroughly analyzing it and presenting it visually, aligning with Research Objective 2. The analysis will focus on identifying trends and insights, addressing the challenge of unclear and abstract insights, and benchmarking to identify areas for improvement in EQA, specifically where most schools achieve the highest standards of educational quality, in line with Research Question 3.

The final step involves creating visual representations to make the findings clear and understandable for stakeholders, aligning with Research Question 4. Employing these research designs will enable a thorough exploration and description of the application of NLP techniques in extracting valuable data from Thai EQA artifacts, ultimately aiding in educational benchmarking and improvement.

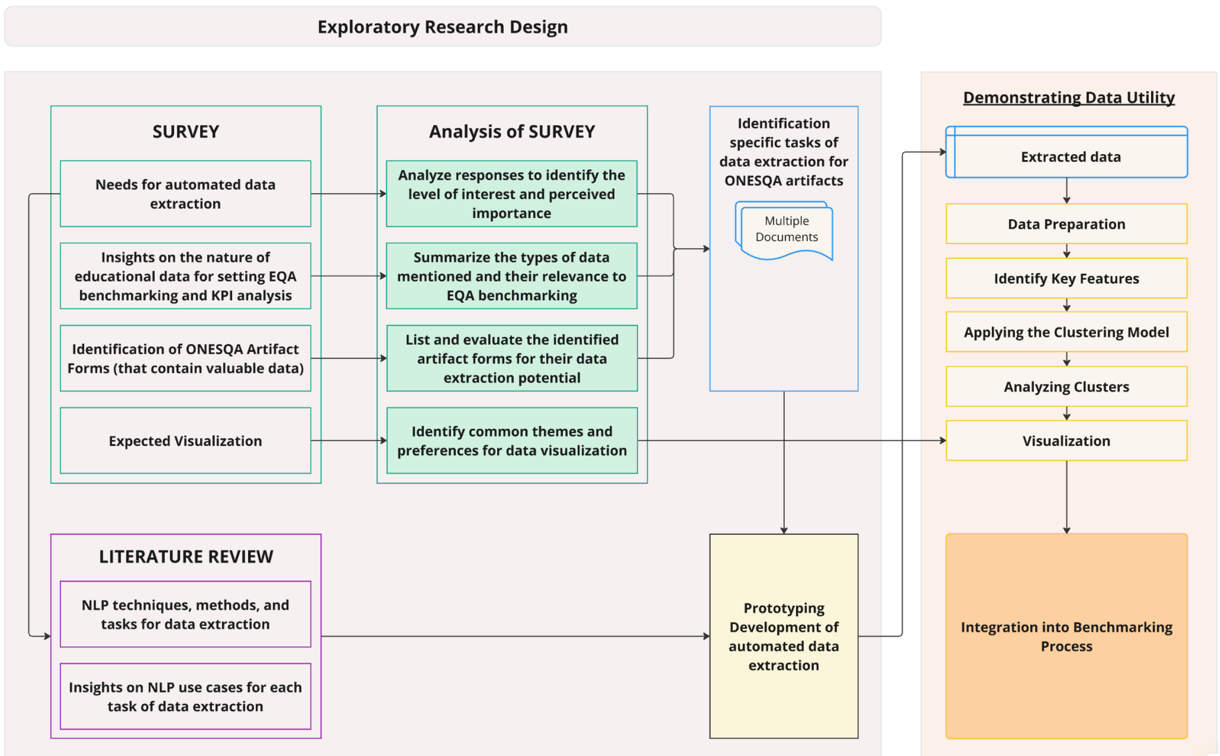


Figure 3: Research Design

3.2 Research Methodology

1) Description

This study employs a mixed-methods approach to address the complexities of enhancing EQA benchmarking through automated data extraction using Natural Language Processing (NLP), combining qualitative and quantitative research to provide a comprehensive understanding of the data. Data will be collected from two main sources including surveys of ONESQA leaders and officers and EQA artifacts. The Surveys will capture both numerical data on the perceptions of ONESQA leaders and qualitative feedback on their experiences and expectations. Furthermore, EQA artifacts will provide both the quantitative benchmarks and qualitative descriptions of educational quality. This combination allows for a comprehensive analysis of the quality assurance processes, enabling both detailed textual analysis and broader quantitative insights.

2) Rationale

By combining both qualitative and quantitative methods, offers the flexibility to adapt the research design as new insights emerge. Initial qualitative findings from surveys assist inform the subsequent quantitative analysis of EQA artifacts. Additionally, the ultimate goal of this study is to enhance EQA benchmarking processes. A mixed-methods approach can ensures that both the statistical accuracy of quantitative data and the rich, contextual understanding of qualitative data contribute to a holistic view of educational quality. This comprehensive perspective is essential for developing effective and meaningful benchmarking tools and strategies.

3.3 Data Collection Methods

This study collects data from two primary sources including surveys and EQA artifacts. Surveys administered to ONESQA internal and external stakeholders aim to identify the needs for automated data extraction, understand the nature of educational data for EQA benchmarking and KPI analysis, identify valuable data within ONESQA Artifact Forms, and determine expected visualizations. Insights from these surveys will guide the development of a prototype automated data extraction pipeline by pinpointing key documents for analysis. Subsequently, EQA artifacts, including reports and evaluations, will be processed using NLP techniques to extract both qualitative and quantitative data. Detailed information on target population, sampling methods, and analysis procedures is provided in Table 1.

Table 1: Data Collection

Aspect	Survey	EQA Artifacts
Objectives	- Investigate the need for automated data extraction - Gather insights on the nature of educational data for EQA benchmarking and KPI analysis - Identify valuable data within ONESQA Artifact Forms - Understand expected visualizations	- Provide detailed qualitative descriptions of educational quality - Offer numerical benchmarks for KPIs and EQA benchmarking
Target Population	ONESQA stakeholders (schools, school authorities, government, students, parents, assessor and ONESQA staffs)	Document involved in the EQA process
Sample	Selected ONESQA internal and external stakeholders (the samples size will show in Table 2 below)	Selected documents, reports, and evaluations that contains valuable variables for EQA benchmarking from different schools that was identified by ONESQA staffs (the samples size will show in Table 3 below)
Sampling Method	Purposive sampling (internal and external stakeholders can provide valuable insights and practically and efficiently without needing to survey every stakeholder directly)	Stratified random sampling (ensuring representation from various context of schools by focusing on school type, size, location, and authority)
Data Collection Instruments	Structured questionnaires (combining closed-ended and open-ended questions)	Document review checklist (ensuring all relevant aspects are captured)
Types of Data Collected	- Quantitative: Numerical measures of perceptions and experiences - Qualitative: Open-ended responses on needs, insights, and preferences	- Quantitative: Numerical data related to educational quality - Qualitative: Descriptive and evaluative text on educational quality

Aspect	Survey	EQA Artifacts
Data Processing and Analysis	- Analyze responses to identify interest and perceived importance - Summarize types of data and relevance to EQA benchmarking - List and evaluate artifact forms - Identify common themes and preferences for data visualization	- Preprocess, clean, and prepare the textual data for analysis - Apply NLP techniques (tokenization, entity recognition, rule-based extraction) - Extract numerical data for benchmarking and KPI analysis

3.4 Sample size

The calculated sample size for the survey respondents, considering the finite population of 100, is 80 that to achieve a 95% confidence level with a 5% margin of error. The detailed allocation respondent by department and role are illustrated in Table 2.

Table 2: Sample Size of Survey Respondents

Type	Section/ Department	Role of respondents	Population (N of Internal stakeholders= 48, N of External stakeholders= N/A)	Sample Size (N = 78; Internal stakeholders= 48, and External stakeholders= 30)
Internal Stakeholders	Assessment & Certification Bureau	Leader	6	5
		Officer	28	22
	Development & Promotion Bureau	Leader	4	3
		Officer	6	5
	Policy & Strategy Bureau	Leader	5	4
		Officer	5	4
External Stakeholders	Department of Local Administration (DLA), Ministry of Interior	Responsible for quality monitoring and management for local administration schools	N/A	5

Type	Section/ Department	Role of respondents	Population (N of Internal stakeholders= 48, N of External stakeholders= N/A)	Sample Size (N = 78; Internal stakeholders= 48, and External stakeholders= 30)
	Office of the Basic Education Commission, Ministry of Education	Responsible for quality monitoring and management for basic schools	N/A	5
	Office of the Higher Education Commission, Ministry of Education	Responsible for quality monitoring and management for university demonstration schools	N/A	5
	Office of the Private Education Commission, Ministry of Education	Responsible for quality monitoring and management for private schools	N/A	5
	Bangkok Metropolitan Administration (BMA)	Responsible for quality monitoring and management for schools in Bangkok	N/A	5
	Department of Local Education Development, Ministry of Culture	Responsible for quality monitoring and management for Buddhist monastic schools	N/A	5

The calculated sample size of EQA artifacts is **133** that to achieve a 95% confidence level with a 5% margin of error, randomly selecting EQA artifacts from 133 schools from the total population of 202. The detailed allocation of samples by size and type of school are illustrated in Table 3.

Table 3: Sample Size of EQA Artifacts

Type of school (divided by governing bodies)	School Size	Population (N=202)	Sample Size (N=133)
Local Administration	S (1-300)	19	13

Type of school (divided by governing bodies)	School Size	Population (N=202)	Sample Size (N=133)
	M (301-1000)	7	5
	L (1001-2000)	0	0
	XL (2001 up)	0	0
Basic School	S (1-300)	70	47
	M (301-1000)	28	19
	L (1001-2000)	7	5
	XL (2001 up)	6	4
Private School	S (1-300)	26	18
	M (301-1000)	20	14
	L (1001-2000)	13	9
	XL (2001 up)	6	4

3.5 NLP-Based Automated Data Extraction Pipeline Framework

The framework is divided into specific tasks, including text preparation, section identification, entity extraction, and accurate validation, which enhances modularity and simplifies management and debugging. Each module employs techniques aligned with its specific tasks, such as leveraging pre-trained models and custom patterns for entity recognition, allowing the framework to effectively handle complex language patterns and specific terminologies. This structured approach aims to establish a clear process from raw data to actionable insights, ensuring systematic handling of each stage of data processing. The detailed structure and workflow are illustrated in Figure 2.

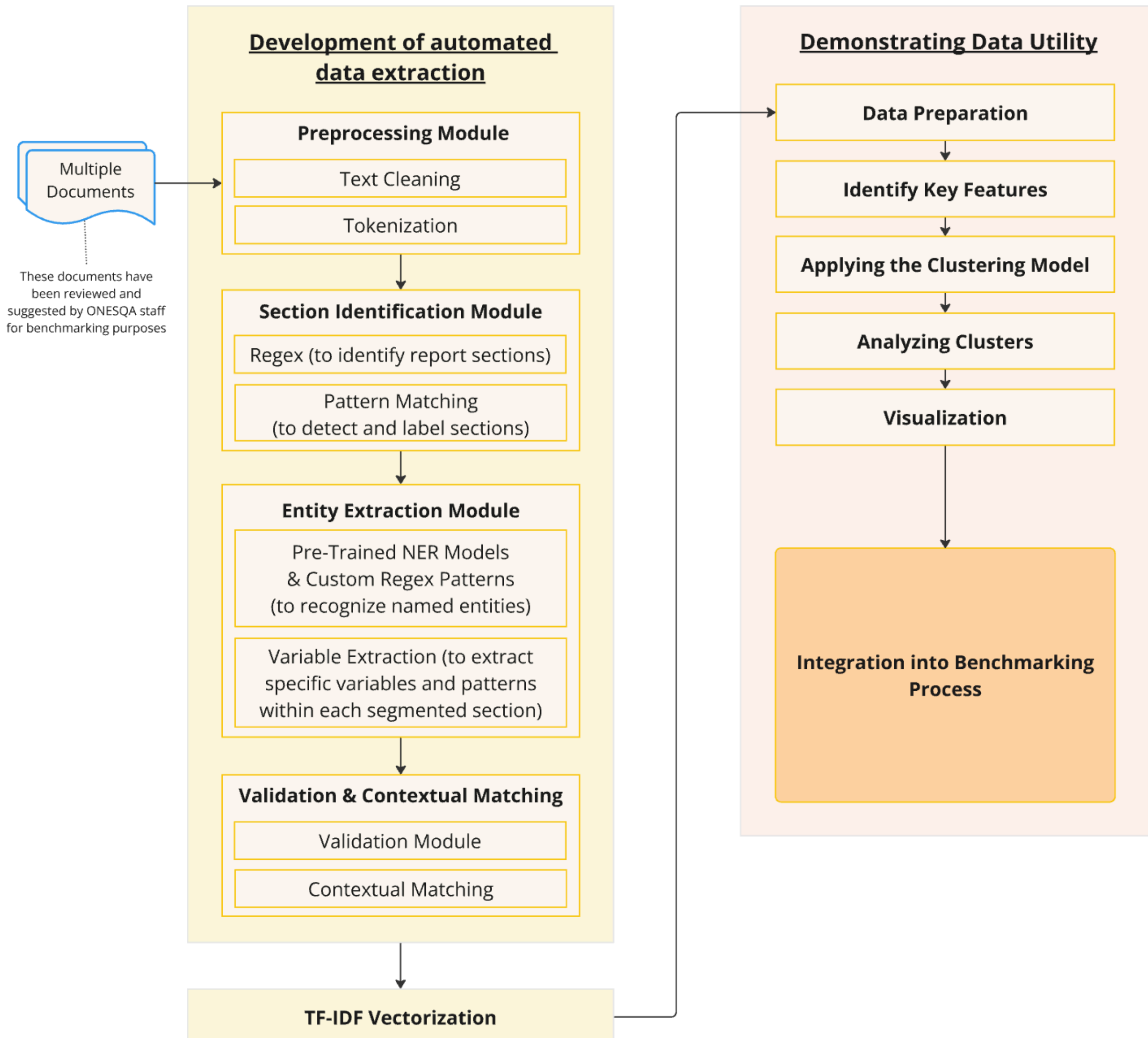


Figure 4: NLP-Based Automated Data Extraction Pipeline Framework

3.6 Data Analysis

Data for this study is sourced from two primary areas: surveys and EQA artifacts. The analysis is conducted in two phases, applying the following approaches.

1) Survey Analysis: Analyze using descriptive statistics and content analysis.

2) EQA Artifact Analysis: Develop and test an automated data extraction pipeline using NLP techniques. The extracted data will then be analyzed using statistical methods to compare with a manually annotated dataset. Additionally, demonstrate the utility of the extracted data through clustering analysis and visualization. The precision and recall metrics were used to measure the accuracy and completeness of the extracted variables. Precision indicates the proportion of correct positive predictions, recall measures the proportion of actual positives correctly identified, and the F1 score balances both precision and recall as a single performance metric.

3.7 Validity, Reliability, and Ethical Considerations

This study upholds internal and external validity through validated survey instruments, thorough NLP model validation, and representative sampling. Pilot testing and model reliability are employed to ensure consistent survey measures and stable data extraction. Ethical considerations are prioritized, including informed consent, confidentiality, and data security. Measures are also taken to reduce bias and ensure fairness in survey design and NLP models, as detailed in Table 4.

Table 4: Validity, Reliability, and Ethical Considerations

Aspect	Survey	EQA Artifacts
1. Validity		
1.1 Internal Validity	Controlling confounding variables in surveys such as work experience of responses by include questions in the survey to gauge the experience level	Controlling for bias in NLP extraction such as misidentifying section due to boilerplate language in the EQA reports by refining the NLP model to ignore common and non-informative text patterns

Aspect	Survey	EQA Artifacts
1.2 External Validity	Generalizing survey findings by collecting representative sample of ONESQA internal and external stakeholders that includes all departments that can use for cross-check the findings with a different group of responses	- Generalizing NLP model performance by using diverse set of EQA artifacts, including reports from various types of schools (as detailed in Table 2) and different document formats to train the model - Validating the NLP models by comparing the automated extraction results against manually annotated datasets.
2. Reliability	Using pilot testing with a small group (3 leaders, and 7 officers) to assist identify any inconsistencies and modify it before full deployment.	Using model reliability to checks whether the NLP model produces stable results every time it processes the same data.
3. Ethical Considerations		
3.1 Data Privacy	- Inform respondent anonymity and confidentiality. - Securely store and manage data collected from participants.	- Anonymize school and personal data in reports. - Protect sensitive information in EQA documents.
3.2 Informed Consent	- Obtain explicit consent from participants before collecting survey data. - Clearly explain the purpose, use, and storage of their data.	- Submit an official request to ONESQA for access to existing EQA Artifacts, detailing the research purpose, and intended use. - Ensure schools understand how their data will be used in research.
3.3 Fairness & Bias	- Design the survey to be inclusive and representative. - Avoid leading questions or biased phrasing.	- Ensure the sampling of EQA artifacts is representative of all school types. - Avoid selective analysis that could introduce bias.
3.4 Transparency	- Inform participants about the survey's goals, how their data will be used, and any risks involved.	- Clearly communicate how EQA data will be used in research. - Provide access to findings based on the extracted data in a way that is understandable to stakeholders.

Summary

This chapter draws the foundation for demonstrating how an automated data extraction pipeline can potentially revolutionize EQA benchmarking practices. By proposing the structured research design, employs an exploratory design to understand stakeholder needs, leading to the creation of a prototype NLP-based data extraction tool, addressing research objective 1. Following this, a descriptive design is used to demonstrate the practical application and utility of the extracted data within the EQA context, addressing research objective 2. These methodologies set the stage for Chapter 4, where the detailed analysis and findings will be presented and discussed.

CHAPTER 4 FINDINGS, ANALYSIS, AND DISCUSSION

Chapter 4 begins with presenting the analysis of survey collected from ONESQA staff, these insights guide the development of the NLP data extraction pipeline, addressing research objective 1. Following this, the practical application of the extracted data by demonstrating its use in various advanced analysis techniques enhancing the setting of appropriate EQA benchmarking standards is shown for serving the research objective 2. The final section addresses study discussion, limitations and proposes future directions, including refining automated techniques and exploring broader applications within the EQA context.

4.1 Survey Analysis and Key Insights

The content validity of the questionnaire was assessed using the Index of Item-Objective Congruence (IOC) tool, this evaluation was conducted by three experts in EQA. A survey was subsequently administered, targeting key stakeholders involved in EQA within ONESQA. The survey received responses from 79 participants (101% response rate), including external stakeholders from six relevant organizations (30 respondents) and internal stakeholders, comprising leaders (14 respondents) and officers (35 respondents). The questionnaire was designed to identify current challenges across three key areas including (1) needs for automated data extraction, (2) identification of artifacts, and key variables for EQA benchmarking, and (3) preferences in analytics and visualization suitable for EQA benchmarking. A mixed-methods approach was employed, integrating closed-ended questions and open-ended responses to capture qualitative insights.

4.1.1 Needs for automated data extraction

1) Analysis of Needs for Automated Data Extraction

The survey assessed the needs for automated data extraction by dividing the analysis into 4 key areas including necessity, importance, utility, and feasibility of implementing an NLP pipeline for automated text extraction within ONESQA.

2) Findings of Needs for Automated Data Extraction

Survey results indicated that 64.8% to 88.8% of respondents, including both internal and external stakeholders, recognized the overall need for an NLP data pipeline to extract sufficient variables for comparative analysis. Specifically, 87.5% to 88.8% of respondents expressed a need for such a tool, while 73.3% to 74.7% highlighted the importance of tools that assist in extracting unstructured data to generate valuable insights. The perceived utility of the NLP pipeline was rated moderately, ranging from 64.8% to 67.9%. Additionally, a majority of respondents (71.1% to 74.7%) believed that implementing the NLP pipeline would be feasible as part of the process for establishing appropriate EQA benchmarks.

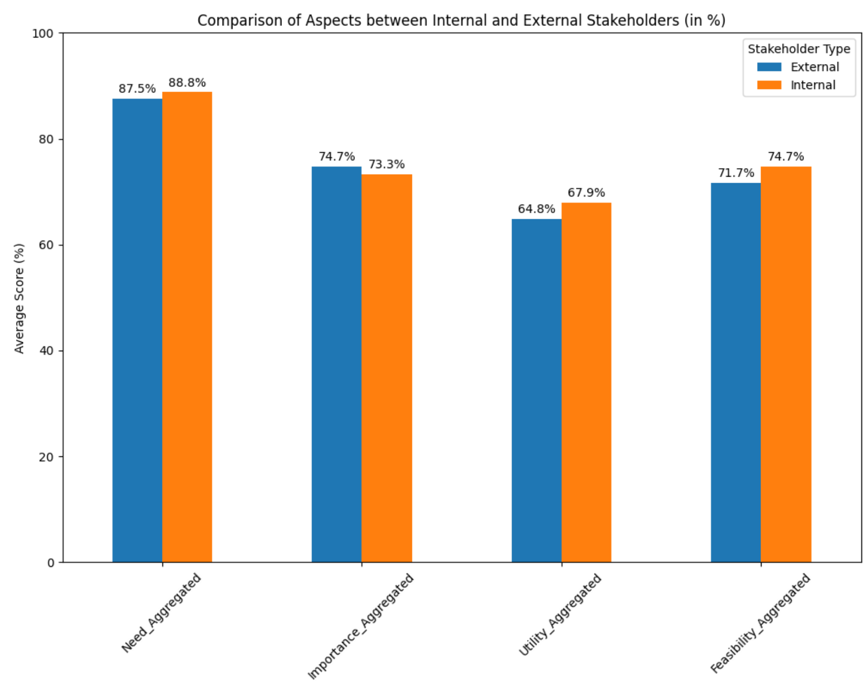


Figure 5: Comparison of needs aspects between internal and external stakeholders

Further analysis in Figure 6 reveals that differences in internal stakeholder positions influence their perceived needs. Notably, officers rated the need, utility, and feasibility higher than external stakeholders, with differences ranging from 1.2% to 6%. Despite variations in positions between internal and external stakeholders, as well as between leaders and officers, the results consistently showed a shared concern for the utility of the NLP pipeline, highlighting its potential benefits in achieving seamless EQA benchmarking.

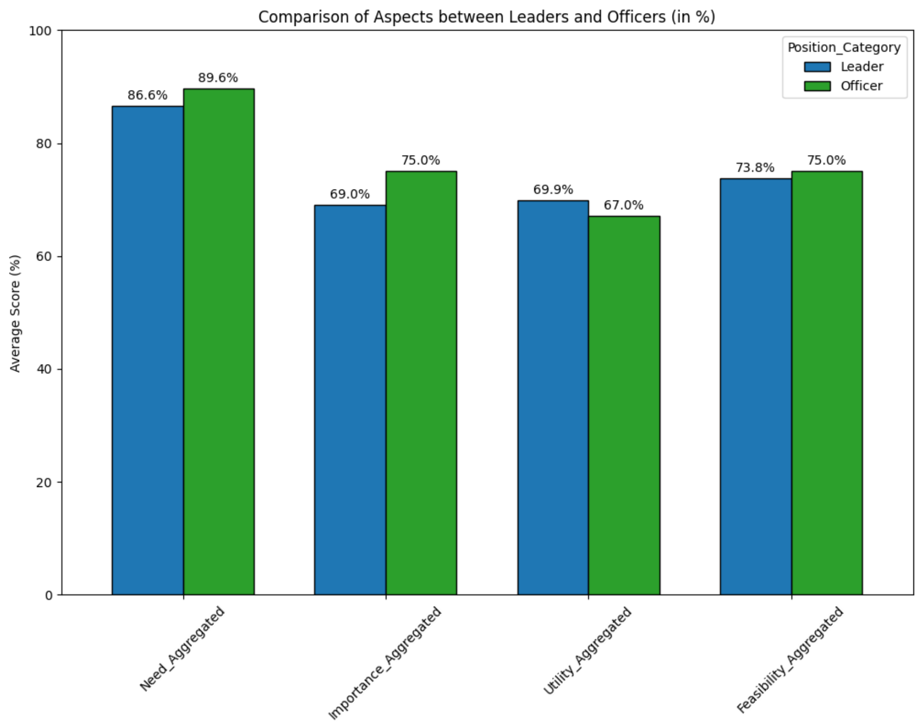


Figure 6: Comparison of needs aspects between leaders and officers

4.1.2 Identification of Artifacts, and Key Variables for EQA Benchmarking

1) Analysis of Identification of Artifacts, and Key Variables for EQA Benchmarking

The survey identified critical variables necessary for effective EQA benchmarking. Key Performance Indicators (KPIs) such as student-teacher ratios, graduation rates, and standardized test scores were frequently highlighted by respondents. Notably, 68% of participants indicated that ONESQA did not collect this data in any formal manner. Instead, assessors recorded these metrics in personal notes, using them to justify the ratings or scores for each quality standard rather than inputting the raw data into the prescribed forms, as detailed in Table 5.

2) Findings of Identification of Artifacts, and Key Variables for EQA Benchmarking

In this context, the ratings or scores of indicators for each standard are recognized as critical variables for effective EQA benchmarking. Specifically, Financial Management Performance

(FMP), Teaching and Assessment Effectiveness (TAE), and Student Outcomes (SO) at each level were identified as essential for clustering schools based on their performance.

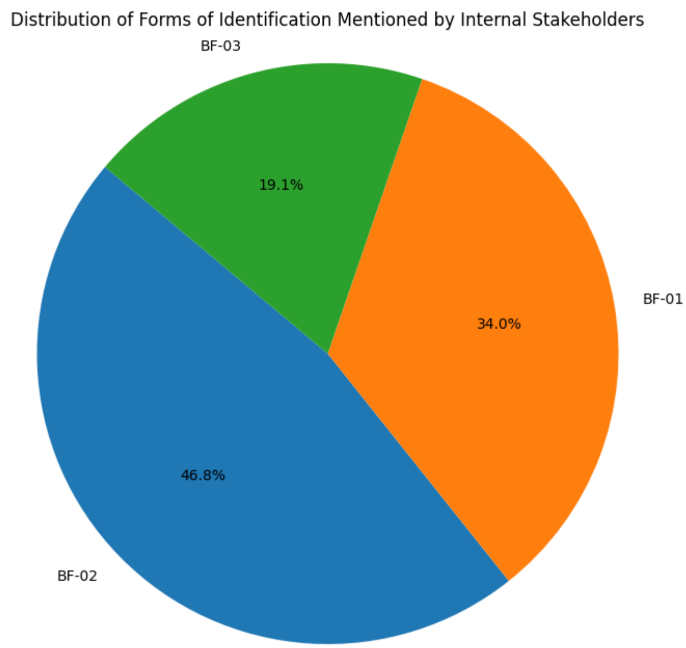


Figure 7: Distribution of EQA Artifacts Forms

Additionally, 46% of participants noted that all key variables for the benchmarking process are contained within the BF-02 forms. These insights were instrumental in selecting the variables used in developing the NLP pipeline, ensuring that it aligns with user expectations and supports effective EQA benchmarking.

Table 5: List of Key Variables for EQA Benchmarking

List of Key Variables for EQA Benchmarking	Percentages
1. Key Variables represent the student outcomes	
- Student Outcome Indicators (SO Indicators Scores)	60.84%
- Graduation Rates (Not Collected by Onesqa)	2.04%
- Student Satisfaction (Not Collected by Onesqa)	2.04%
- Best Practices of SO	35.08%

List of Key Variables for EQA Benchmarking	Percentages
2. Key Variables represent the teaching approach efficiency	
- Teaching approach efficiency Indicators (TAE Indicators Scores)	56.01%
- Teachers with qualification (Not Collected by Onesqa)	12.26%
- Best Practices of TAE	31.73%
3. Key Variables represent the facilitating and management performances	
- Facilitating & management performances Indicators (FMP Indicators Scores)	54.88%
- Stakeholders Satisfaction (Not Collected by Onesqa)	6.81%
- Best Practices of FMP	38.31%

4.1.3 Stakeholder Expectations for Analytics and Visualization in EQA Benchmarking

1) Analysis of Stakeholder-Recommended Methods and Visualizations

This section reviews stakeholders' expectations regarding the most suitable analysis methods and visualization themes for EQA benchmarking, taking into account differences in work experience among respondents. The analysis focuses on comparing the suggested methods and visualizations for each specific benchmarking purpose.

2**) Findings of Stakeholder-Recommended Methods and Visualizations**

Figure 8 presents the analysis methods and visualizations recommended by both internal and external stakeholders. The findings indicate that the choice of analysis method and visualization type should be tailored to the specific benchmarking objective. For example, when assessing the impact of teaching and school management effectiveness on student quality, 63.3% of respondents favored comparative analysis, while 73.4% recommended a scatterplot chart with annotations as the most appropriate visualization.

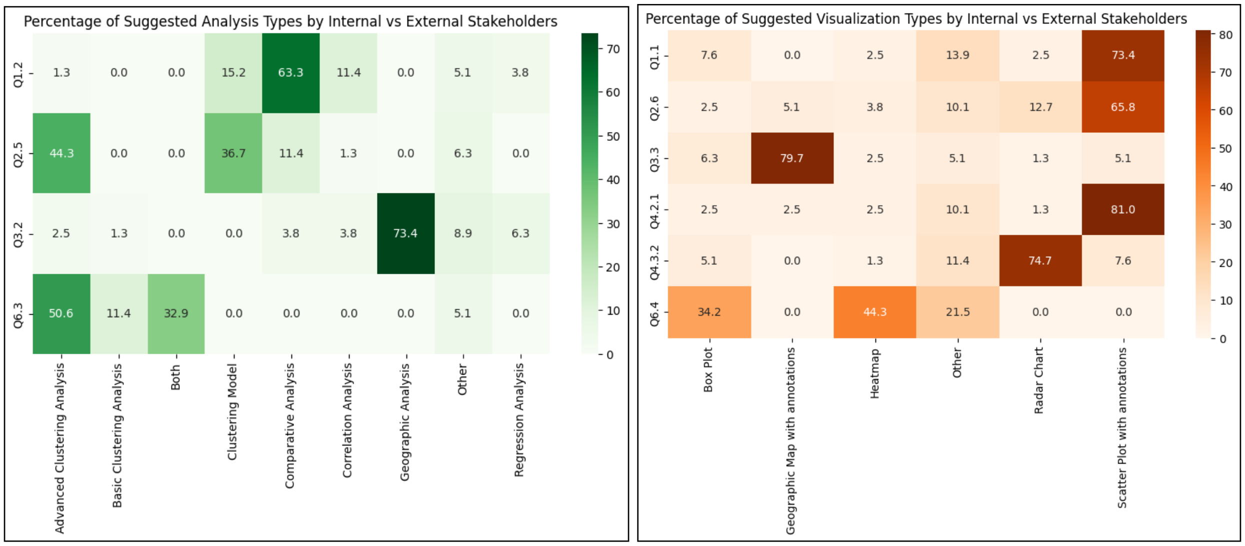


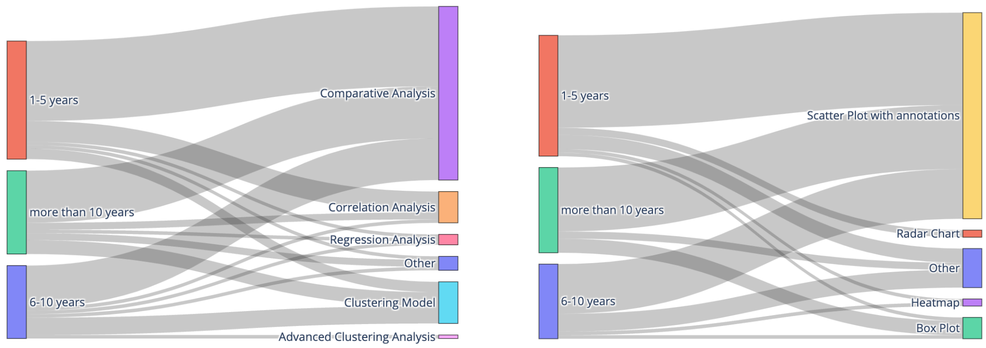
Figure 8: Percentage of suggested analysis methods and visualizations

For grouping schools by quality performance characteristics, 44.3% of respondents identified advanced clustering analysis, with 65.8% supporting the use of a scatterplot chart with annotations for visualization. Additionally, when it came to identifying high-performing schools and showcasing their best practices through collaborative benchmarking, 73.4% of respondents recommended geographic analysis, and 79.7% endorsed the use of geographic maps with annotations.

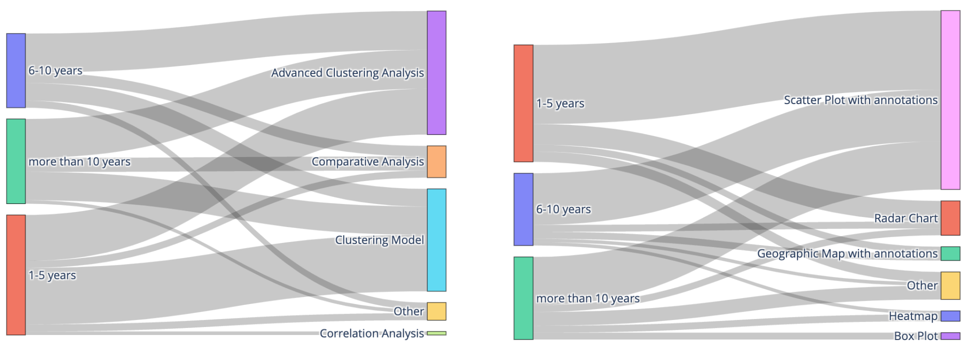
Regarding regulatory benchmarking, 81% of respondents suggested scatterplot charts with annotations as the most suitable visualization, followed closely by radar charts at 74.7%, both of which effectively represent school performance rankings and provide a clear scoreboard for each school. Furthermore, for comparing school clusters based on performance characteristics, 50.6% of respondents favored advanced clustering analysis, with heat maps (44.3%) and box plots (34.2%) cited as the most appropriate visualization techniques.

Despite variations in work experience, respondents across all levels demonstrated a consistent preference for the recommended analysis methods and visualization themes, as detailed Figure 9.

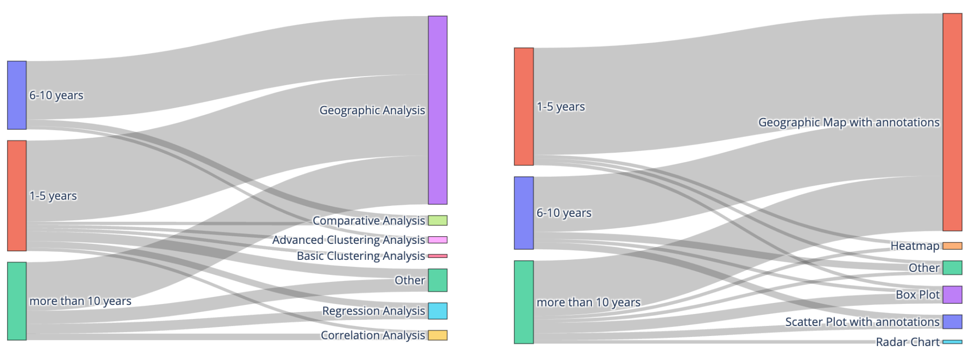
Analytical techniques and visualization themes for assessing the impact of teaching and school management on student quality



Analytical techniques and visualization themes for grouping schools by quality performance characteristics



Analytical techniques and visualization themes for identifying high-performing schools and showcasing their best practices



Analytical techniques and visualization themes for comparing school clusters based on performance characteristics within each

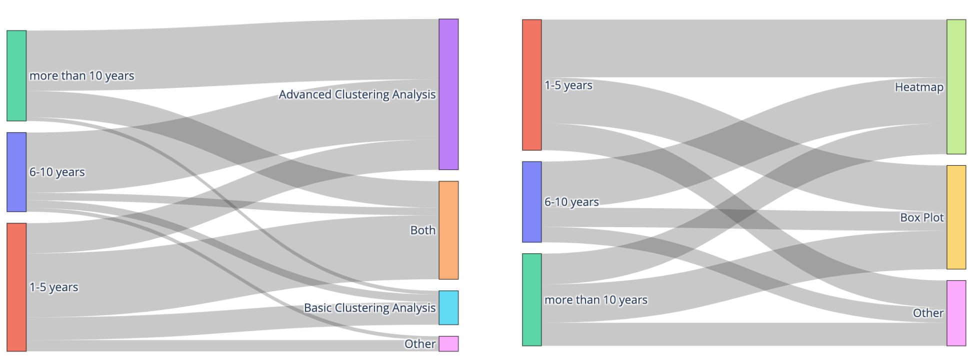


Figure 9: Survey-Based Analytical Techniques and Visualization Themes by Purpose

4.2 Development and Performance of the NLP Extraction Pipeline

This subsection describes the development of the NLP extraction pipeline based on survey findings. The pipeline was designed to accomplish two main tasks, extracting indicator ratings and identifying text relevant to best practices for each indicator. The development process focused on extracting ratings and identifying relevant text, as suggested by the respondents.

4.2.1 NLP Extraction Pipeline Development Process Based on Survey Findings

1) Analysis of "BF-02" document for NLP Extraction Pipeline Development

The “BF-02” document features a complex structure that required thorough analysis to design an efficient NLP-based extraction process. It includes multiple structured sections, consistently covering both Early Childhood Education (ECE) and Basic Education (BE). Each section presents indicators with corresponding ratings (numerical scores) and best practice descriptions (textual content).

Key challenges encountered during the analysis included **repetitive headings** that identical headings, such as “Standard 1” and “Indicator 1.1,” appeared repeatedly, necessitating differentiation in the code to avoid confusion between sections. **Thai Numbering,** the use of Thai numerals instead of Arabic ones added complexity, requiring additional steps to ensure accurate extraction. **Data Positioning,** indicator ratings appeared both in the detailed sections and a summary table, raising the risk of duplicate data extraction. **Mixed Data Types,** the need to extract both numerical ratings and descriptive best practices demanded a tailored, multifaceted approach. **Document Size, t**he document’s length (exceeding 80 pages) increased processing time. Optimization strategies, such as indexing, were employed to accelerate data retrieval. Additionally, **inconsistent formatting,** although most sections followed a consistent format, minor layout deviations required fallback logic to maintain robustness in the extraction process. These challenges were addressed through careful code design and optimization, ensuring the pipeline could effectively extract the necessary data.

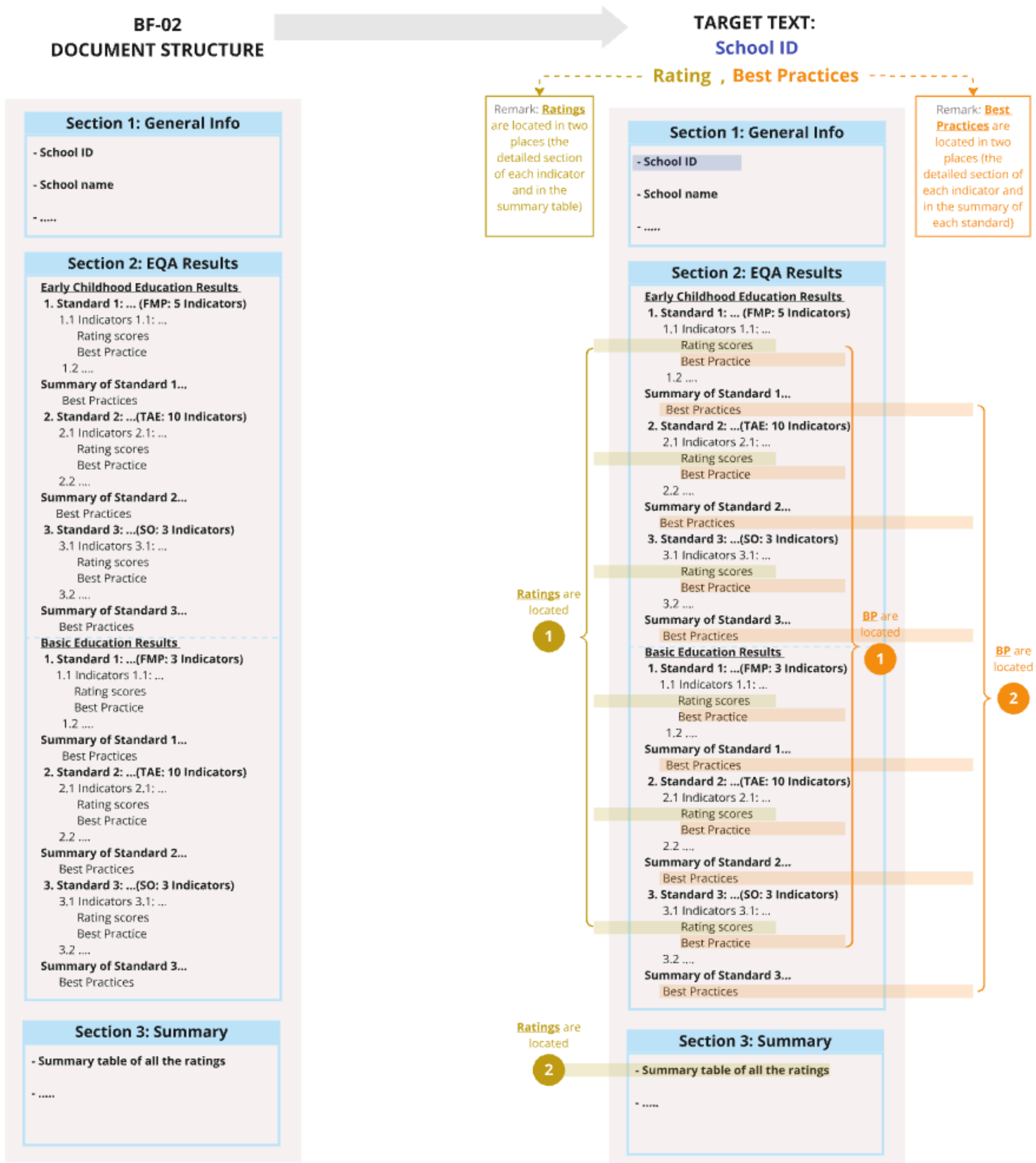


Figure 10: "BF-02" Document Structure

2) Findings of NLP Extraction Pipeline Development Process

To ensure the accuracy and efficiency of the data extraction, each task (numeric data extraction and text extraction) is handled appropriately aligned with the challenge mentioned. The process is split into two distinct code sets including rating extraction and best practices text extraction.

2.1) Part 1: Rating Extraction

This model focuses on extracting the numerical ratings for each indicator in both ECE and BE. This extraction targets these specific positions to accurately capture the numerical data associated with each indicator (Code 1: Rating of indicators Extraction, available at: https://colab.research.google.com/drive/).

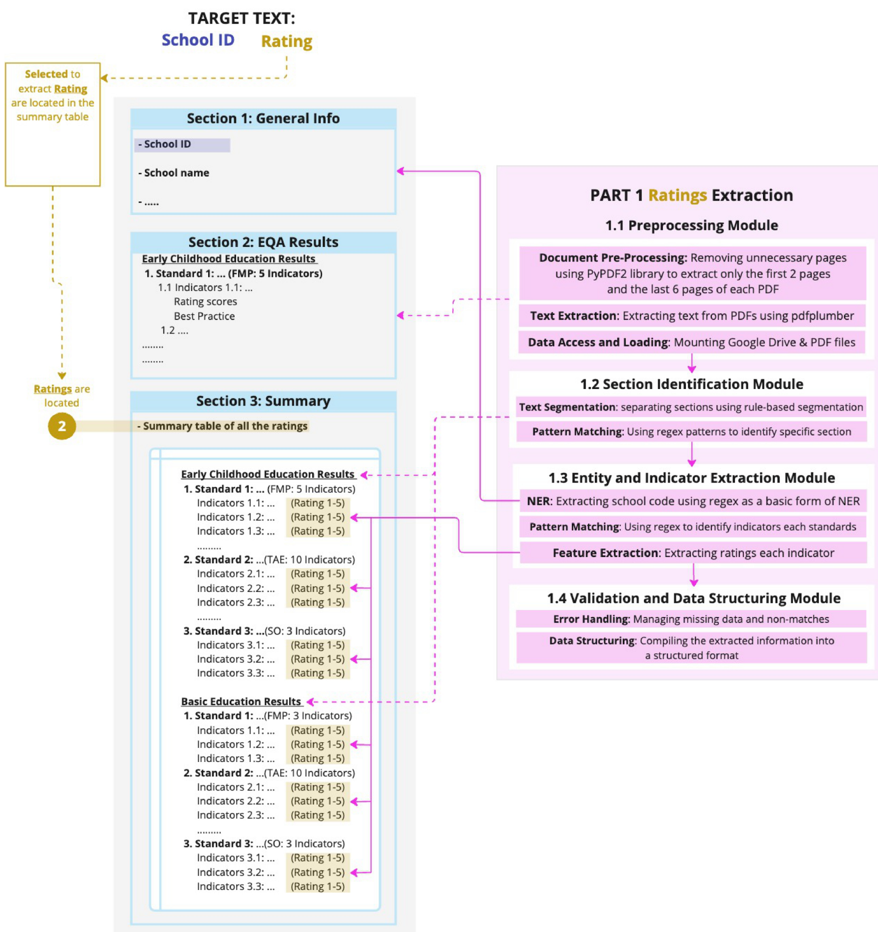


Figure 11: The Structure of Rating Extraction Development (Code 1)

2.1.1) Document Pre-Processing

Removing Unnecessary Pages Before extracting the target data, unnecessary pages were removed using a Python script. The code leveraged the PyPDF2 library to extract only the first 2 pages and the last 6 pages of each PDF, which contained the relevant information.

2.1.2) Relevance Tasks for Rating Extraction

The Rating Extraction process for the "BF-02" document focuses on extracting numerical ratings for the 18 indicators of ECE and the 16 indicators of BE. These ratings are located in two main sections of the document, the detailed sections of each indicator and the summary tables. The extraction process targets these sections specifically using rule-based methods, as the relevant tasks involved in the extraction below.

(1) **Preprocessing Module** prepares the document for extraction by performing fundamental text processing tasks. **Text Extraction from PDF**, the process_pdfs function utilizes the pdfplumber library to extract raw text from all pages of the PDF files in the specified directory. The extracted content is then passed through the text processing pipeline for further analysis. **Conversion of Thai Numerals to** Arabic, this conversion ensures consistency across the extracted data.

(2) **Section Identification Module** focuses on identifying and isolating relevant sections in the text based on structural patterns. **Text Segmentation,** the function extract_info uses regular expressions to segment the document into predefined sections. **Pattern Matching for Sections**, the function employs regular expressions to match specific patterns related to educational levels and standards, ensuring that the text corresponding to the indicators and ratings is accurately captured for each section.

(3) **Entity and Indicator Extraction Module** handles the extraction of critical information. **School Code Extraction**, the extract_info function includes a regex-based method to extract the school code, which is often written in Thai numerals. The code is then converted into Arabic numerals using the thai_to_arabic function for consistency. **Indicator and Rating Extraction,** the function extract_ratings_for_level is designed to capture the ratings for each indicator within a standard. Using regex to identify each "**ตัวชี้วัด**" (Indicator) by its number and then extract the associated rating.

(4) **Validation and Data Structuring Module** organizes and validates the extracted data. **Error Handling,** if no rating is found for an indicator, None is appended to the result, ensuring that missing data is handled gracefully without breaking the extraction process. **Data**

Structuring, the extracted ratings and school code are organized into a structured format using the results_to_dataframe function.

2.2) Part 2: Best Practices Text Extraction

This model is designed to extract the descriptive text under the "Best Practice" sections for each indicator. This part deals with extracting narrative text, requiring different NLP techniques to identify and isolate the relevant content ensuring that no important information was missed. (Code 2.1-2.2: Best Practices Extraction, available at: https://colab.research.google.com/drive/1rOwk7urCKeoLOpEg9lumhZzB6aOLCUwz?usp=sharing).

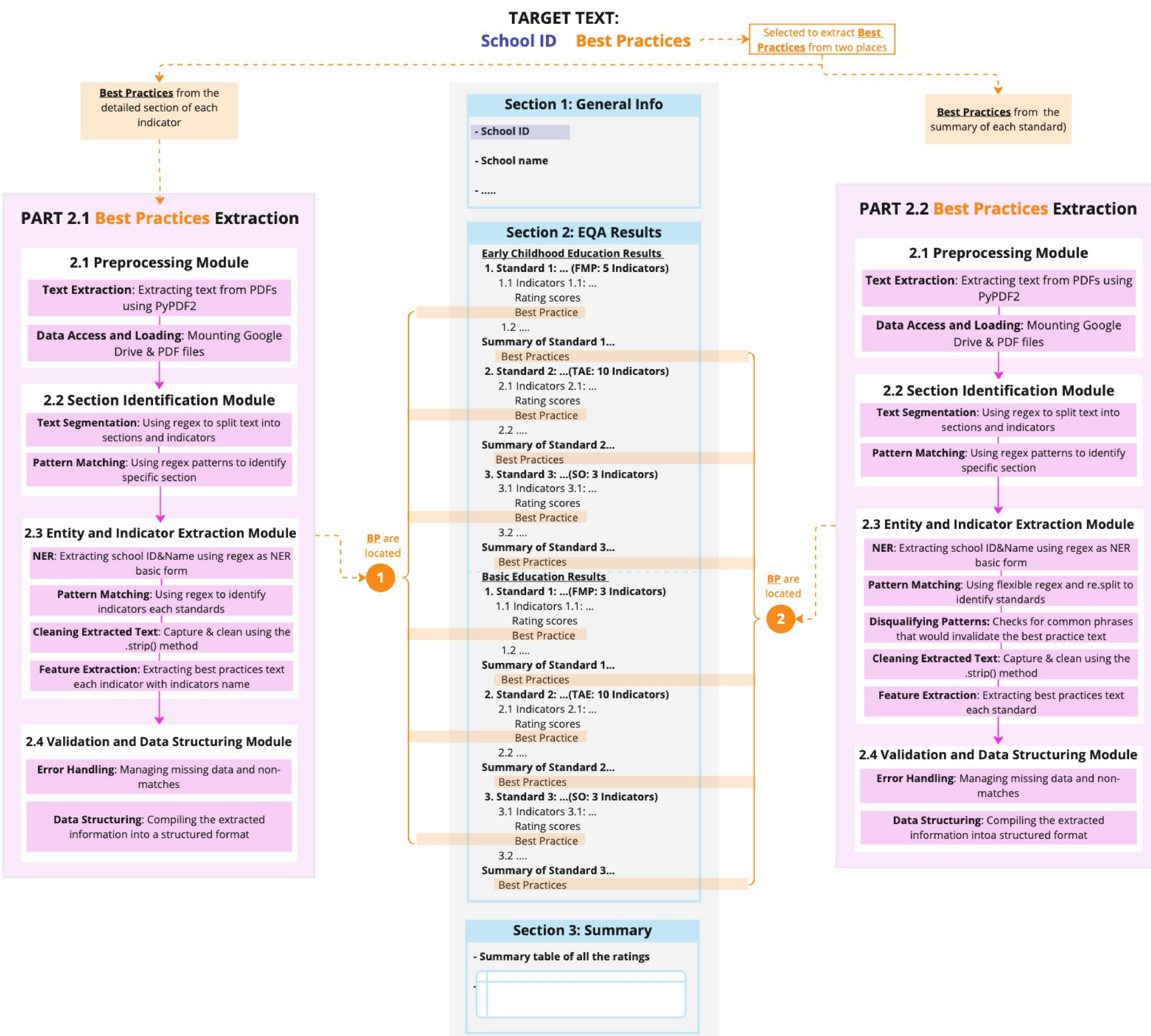


Figure 12: The Structure of Best Practices Extraction Development (Code 2.1-2.2)

(1) **Preprocessing Module** prepares the document for extraction by performing basic text processing including **Text Extraction (from PDF)**, the extract_text_from_pdf function reads the content of the document using the PyPDF2 library, converting the entire PDF content into a structured text format for further processing. **Basic Text Cleaning,** using the thai_to_arabic function ensures that the data is in a standardized format for easier extraction.

(2) **Section Identification Module** isolates the relevant parts of the document, separating content based on the structure. **Text Segmentation,** the segmentation is handled through rule-based methods that split the text where certain headings, such as "มาตรฐานที่" (Standard) and "ตัวชี้วัดที่" (Indicator). **Pattern Matching,** Regex are used to identify specific sections of the document. The structure of "BF-02" uses repeated headings and indicators for both academic levels, making it necessary to capture the position of ratings correctly.

(3) **Entity and Indicator Extraction Module** focusing on extracting specific pieces of data such as school codes, indicators, and their ratings. A basic form of NER is implemented using regex to extract the school ID, names, and indicator names. **Pattern Matching for Indicators**, the extract_desired_text function uses regex to capture each indicator's number and name. By identifying the sections labeled with "ตัวชี้วัดที่" (Indicator), the code isolates the content specific to each indicator. **Feature Extraction for Ratings**, within each indicator section, the code identifies and extracts the rating or score associated with that particular indicator.

(4) **Validation and Data Structuring Module** focuses on validating and organizing the extracted data into a structured format for easy analysis and review. **Error Handling**, the code is designed to manage any missing data or failed extractions. **Data Structuring**, the extracted data, including school ID, indicator number, and rating, is organized into a structured format (data frame).

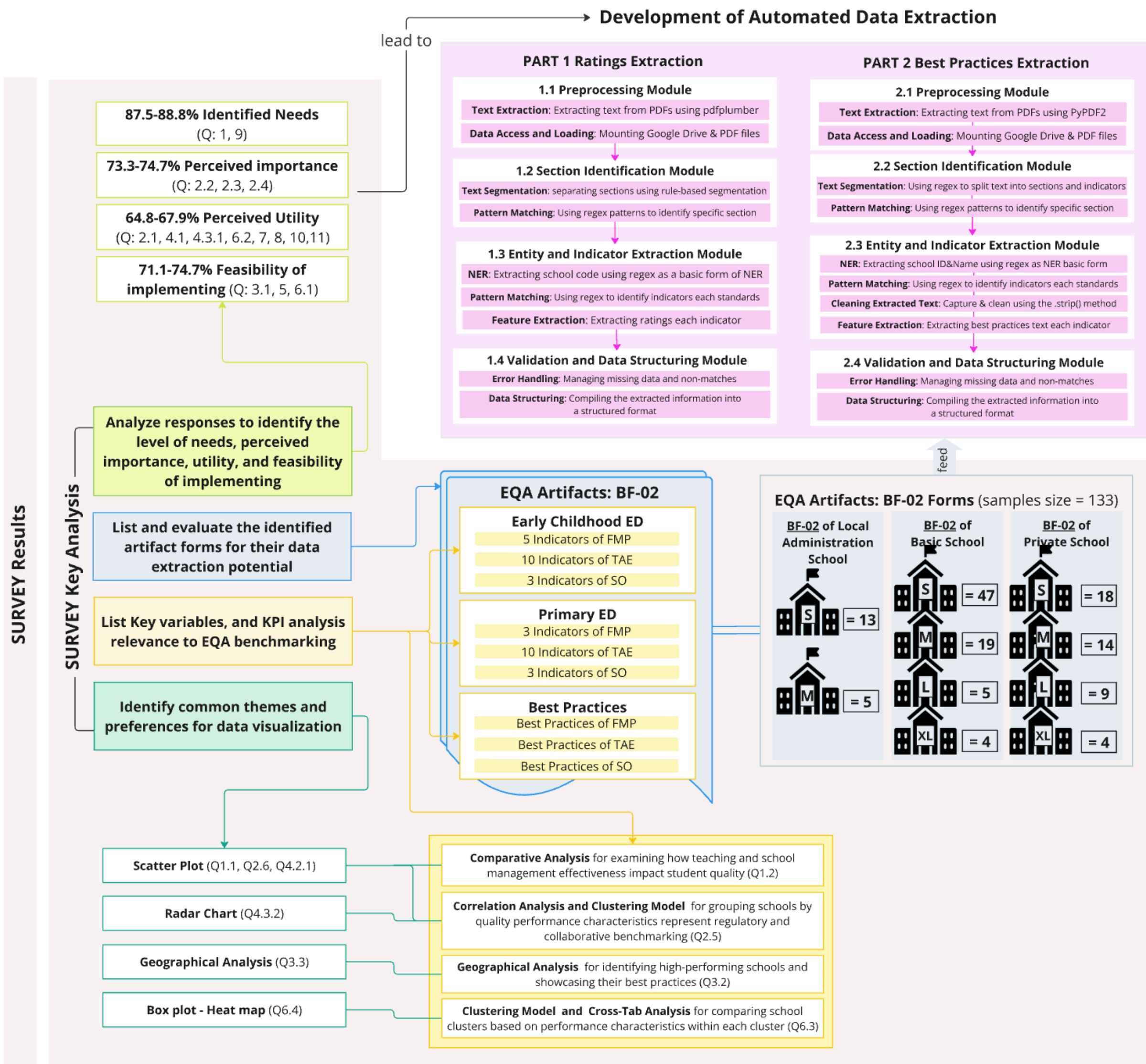


Figure 13: The results of NLP Extraction Pipeline Development

The development process provided critical insights into the effectiveness of the NLP pipeline. For the extraction of indicator ratings, the pipeline consistently captured numerical values with high accuracy. This success was largely due to its precise understanding of the document structure and the application of a refined extraction process, which included cleaning functions and rigorous debugging (Figures 10-11). Furthermore, the pipeline effectively identified and classified relevant text related to best practices from various sections of the document, offering a comprehensive overview of best practices associated with each indicator. By utilizing the NLP techniques implemented in code parts 2.1 and 2.2, the model was able to accurately extract relevant information, regardless of the variation in text placement, ensuring complete coverage of all relevant data.

4.2.2 Evaluation Metrics for NLP Extraction Performance

1) Analysis of NLP Extraction Performance

This section evaluates the performance of the NLP pipeline for two key tasks, extracting indicator ratings and identifying text related to best practices. Precision, recall, and F1 scores were used to measure the model's accuracy and completeness.

2) Findings of NLP Extraction Performance

2.1) Rating of Indicators Extraction Performance

The performance of the NLP extraction model in capturing indicator ratings demonstrates high efficiency. As shown in the metrics, the model achieved consistent accuracy across all indicators. Precision scores were perfect, indicating that the model reliably extracted the correct ratings without error. Similarly, recall values confirm that the pipeline successfully captured the full set of relevant ratings across indicators, leaving little to no data unaccounted for. The F1 Score, a balance of precision and recall (1.00), further verifies the model's overall reliability and robustness.

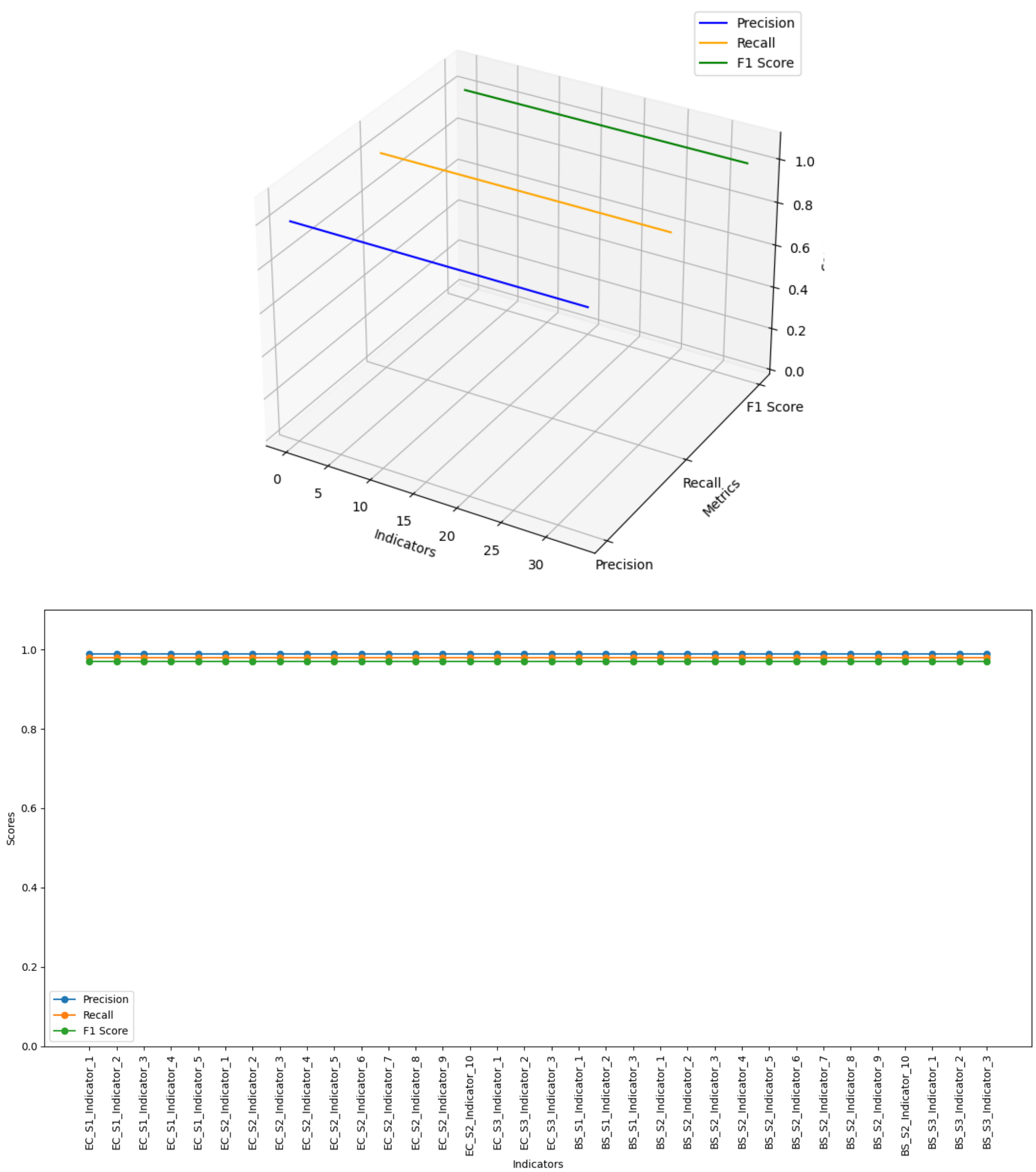


Figure 14: Rating of Indicators Extraction Performance

2.2) Best Practices Text Extraction Performance

The NLP extraction model's performance in identifying relevant best practices from the **BF-02** document was evaluated using standard metrics. These metrics assess both the accuracy of extracting best practices and the completeness of the information. The results, as visualized in Figure 15, show consistently high scores across all metrics, with precision nearly reaching 1.0, indicating a strong ability to identify relevant practices.

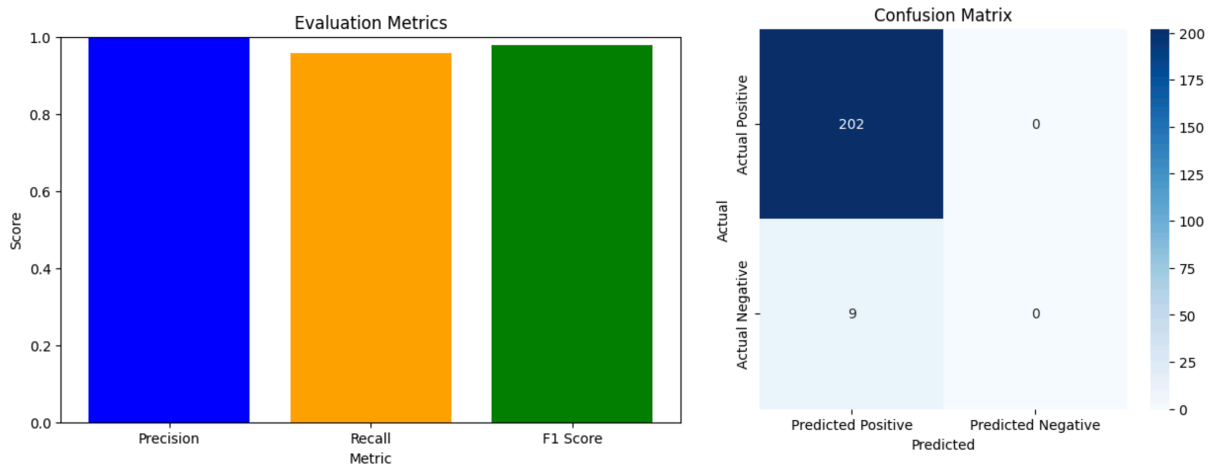


Figure 15: Best Practices Text Extraction Performance

The extraction process involved two sets of code, **2.1** and **2.2**, which captured a total of **240 rows** of relevant best practices from **43 schools** listed in the document. Code **2.1** identified **135 rows**, while code **2.2** captured **105 rows**. The outputs of these two codes were appended together, and **TF-IDF vectorization** was applied to manage duplicates in school IDs, as some schools were linked to multiple best practice projects. Due to the nature of the data, where schools could have multiple best practices, using only the "School ID" as an identifier would have led to duplicates. To address this, the column of indicator names was vectorized using TF-IDF, then merged with the "School ID" to create a unique identifier. This approach was applied to both the extracted dataset and the annotated dataset, ensuring accurate evaluation.

The **annotated dataset** consisted of **249 rows**, and the **extracted dataset** included **240 rows**. Using **cosine similarity** to compare the two datasets, **236 rows** in the extracted dataset achieved high-quality matches (similarity score of 1), representing **98.33%** accuracy. This shows that the model was able to capture almost all relevant text with exact matches to the annotated dataset.

The performance of the extraction model is highly accurate, with a **precision of 1.00**, indicating that all identified matches were correct. The **recall score of 0.96** shows that the model successfully captured the vast majority of relevant identifiers, although a few were missed. This balance between precision and recall is reflected in the **F1 score of 0.98**, demonstrating a strong overall performance. Specifically, **202 identifiers** from the annotated dataset were correctly matched with those in the extracted dataset, representing **81.21%** of the total annotated identifiers. However, **9 identifiers** were not found, accounting for **3.61%** as false negatives. Notably, no false positives were recorded, resulting in a **0% false positive rate**.

Overall, these results demonstrate the model's exceptional precision and recall, resulting in almost perfect alignment between the extracted and annotated datasets. The minimal discrepancies further underscore the model's robustness in identifying best practices across schools.

4.3 Application of Extracted Data for EQA Benchmarking

This section would to show the benefit of how the extracted data could be useful for uncovering hidden patterns of school performance and setting appropriate benchmarks for External Quality Assurance (EQA) by using analysis methods and visualization that suggested by survey respondents mentioned at survey results section (as page 51-53).

4.3.1 Demonstrating Advanced Analysis Techniques

The following sections outline how the data was analyzed and visualized to achieve this objective, highlighting its relevance in both **Regulatory Benchmarking** and **Collaborative Benchmarking** processes.

1) Section 1: Analysis and Visualization for Regulatory Benchmarking

To assess school performance and support EQA benchmarking, the study applied a two-step clustering process based on key performance indicators. This process first involved creating **quadrant-based clusters** using **Facilities & Management Performance (FMP)**, **Teaching Approaches Efficiency (TAE)**, and **Student Outcomes (SO)** as key factors, followed by refining the analysis through **K-means clustering**.

1.1) Quadrant-Based Clustering

After aggregating the scores for FMP, TAE, and SO, a **2D scatter plot** was generated, with FMP on the y-axis, TAE on the x-axis, and SO represented by the color of each dot. This visualization offered a comprehensive view of school performance across these three critical dimensions.

To create clusters, **median values** of FMP and TAE were used to divide the schools into distinct groups. This initial clustering helped identify groups of schools with similar performance characteristics, serving as a foundation for setting appropriate EQA benchmarks based on their strengths and weaknesses (Figure 16 and Table 6). These benchmarks provide insights into performance variation across schools and can inform targeted improvement strategies.

Table 6: School Performance Clustering Results Using Quadrant-Based

Cluster	Name	Quadrant	Teaching Efficiency	Facilities-Management Performance	Student Outcomes
Cluster 1	High Performers	Upper-Right (Q1)	Higher the median	Higher the median	Excellent (dark green dot)
Cluster 2	Teaching-Focused Achievers	Lower-Right (Q4)	Higher the median	Lower the median	Very good to Excellent (green and dark green dot)
Cluster 3	Teaching-Oriented Schools	Lower-Right (Q4)	Higher the median	Lower the median	Good to Adequate (light green and yellow dot)
Cluster 4	Well-Managed Potential	Upper-Left (Q2)	Lower the median	Higher the median	Very good to Excellent (green and dark green dot)
Cluster 5	Facility-Driven Schools	Upper-Left (Q2)	Lower the median	Higher the median	Good to Adequate (light green and yellow dot)
Cluster 6	Low Performers	Lower-Left (Q3)	Lower the median	Lower the median	Good to Adequate (light green and yellow dot)
Cluster 7	Mixed Performers	Scattered Across Quadrants	Mixed	Mixed	Mixed

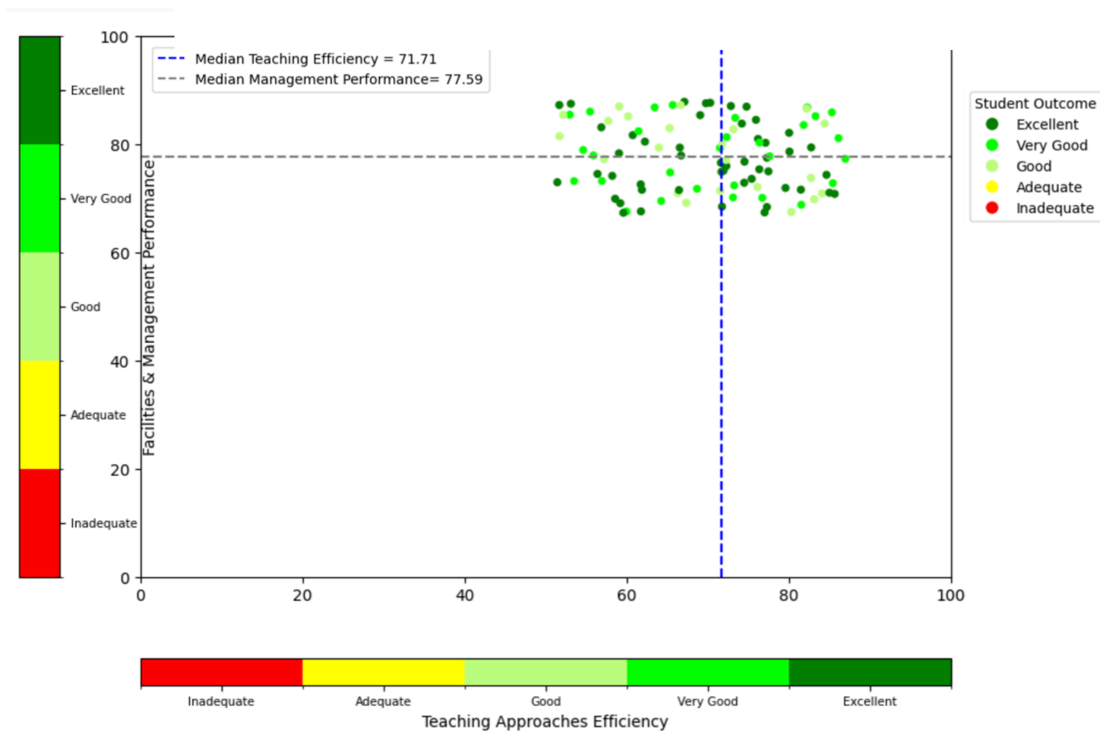


Figure 16: Scatter Plot of Comparative Analysis with Quadrant-Based Clustering

1.2) K-Means Clustering for Mixed Performers

One particular group, Cluster 7, showed varied performance patterns that could not be clearly interpreted through quadrant-based clustering. **K-Means clustering** was applied for comprehension the diversity in this cluster using performance variables such as SO, TAE, and FMP, sub-clusters were identified within Cluster 7, revealing three distinct performance patterns;

- **Yellow Sub-cluster**: High SO but average TAE and FMP.
- **Teal Sub-cluster**: High TAE and FMP but average SO.
- **Purple Sub-cluster**: Moderate performance across all indicators.

These sub-clusters provide further granularity in assessing school performance and suggest more tailored interventions. For instance, **high performers** could serve as examples for others, while **mixed performers** would benefit from customized strategies targeting specific areas of weakness (Figure 17 and table 7).

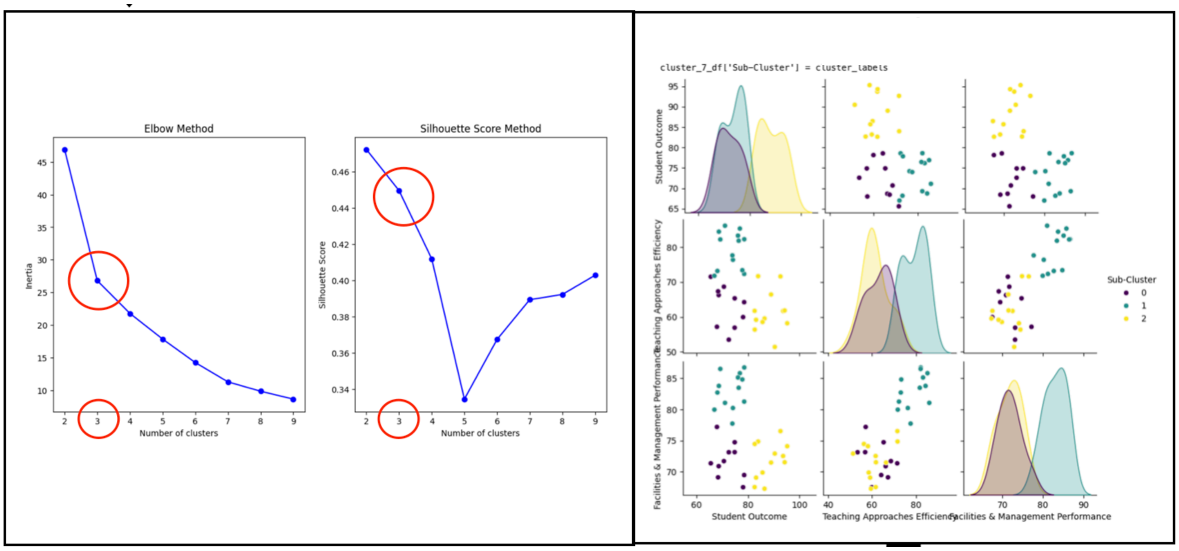


Figure 17: K-Means Clustering of Sub-Groups Within Cluster 7 by School Performances

Table 7: School Performance Clustering Results Using Quadrant-Based and K-Means Algorithms

Cluster	TAE	FMP	SO	Clustering Reason
Cluster 1: High Performers (Upper-Right: Q1)	Higher the median	Higher the median	Excellent (dark green dots)	Schools in this cluster excel in both teaching efficiency and facilities management, leading to excellent student outcomes.
Cluster 2: Teaching-Focused Achievers (Lower-Right: Q4)	Higher the median	Lower the median	Very good to Excellent (green and dark green dots)	These schools are grouped together because of their strong teaching efficiency, despite weaker facilities management, resulting in generally positive outcomes.
Cluster 3: Teaching-Oriented Schools (Lower-Right: Q4)	Higher the median	Lower the median	Good to Adequate (light green and yellow dot)	Similar to Cluster 2, but with slightly lower student outcomes. The clustering reflects their high teaching efficiency but lower facilities management performance.

Cluster	TAE	FMP	SO	Clustering Reason
Cluster 4: Well-Managed Potential (Upper-Left: Q2)	Lower the median	Higher the median	Very good to Excellent (green and dark green dots)	Schools in this cluster have strong facilities management but lower teaching efficiency. They still achieve good to excellent student outcomes.
Cluster 5: Facility-Driven Schools (Upper-Left: Q2)	Lower the median	Higher the median	Good to Adequate (light green and yellow dots)	Similar to Cluster 4 but with lower student outcomes. They are grouped based on their high facilities management and the potential for improvement in teaching efficiency.
Cluster 6: Low Performers (Lower-Left: Q3)	Lower the median	Lower the median	Good to Adequate (light green and yellow dots)	This cluster contains schools that struggle in both teaching efficiency and facilities management, leading to poor student outcomes.
Cluster 7: Outcome-Focused Achievers (Scattered Across Quadrants)	Close to the median both above and lower, near the line of median	Close to the median both above and lower, near the line of median	Very good to Excellent (green and dark green dots)	These schools have high student outcomes while performing averagely in other areas.
Cluster 8: Balanced Performers (Scattered Across Quadrants)	Higher the median	Higher the median	Close to the median (light green and green dots)	These schools excel in Facilities & Management and Teaching Approaches Efficiency but have average student outcomes, indicating a balanced focus on resources and teaching, with room for improvement in outcomes
Cluster 9: Consistent Moderates (Scattered Across Quadrants)	Close to the median both above and lower, near the line of median	Close to the median both above and lower, near the line of median	Close to the median (light green and green dots)	these schools have moderate performance across all variables, without any particular area standing out as either strong or weak.

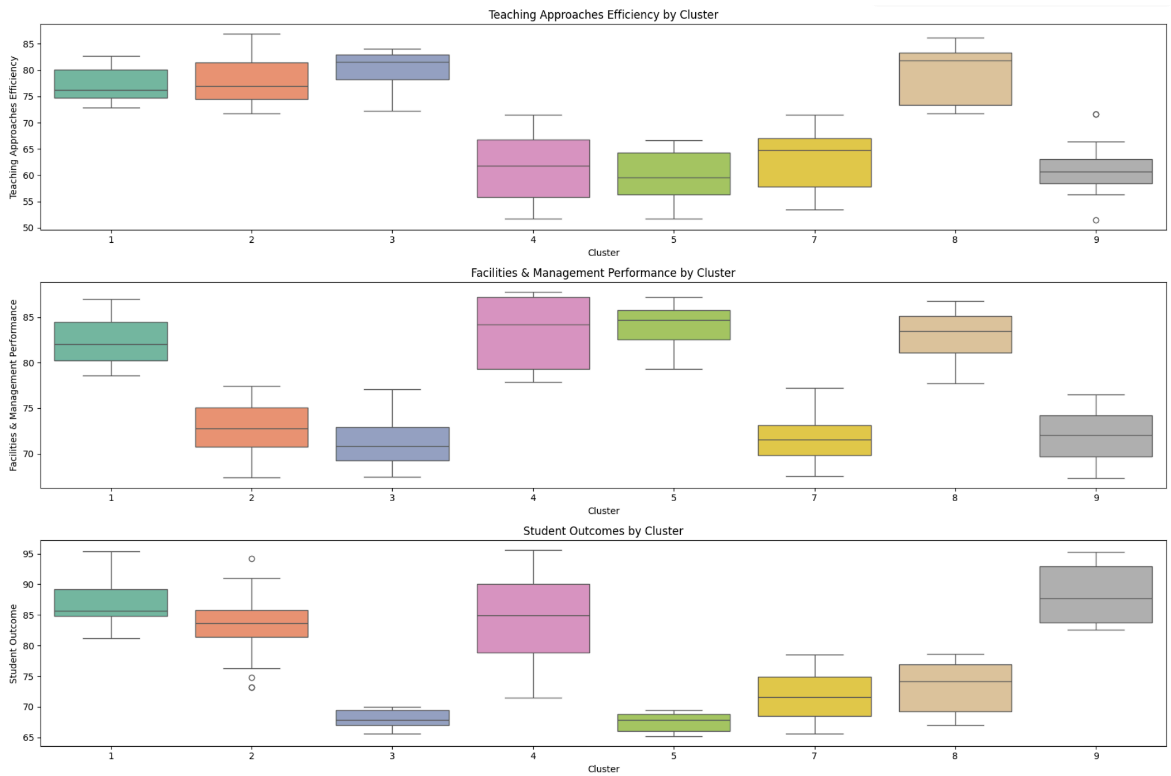


Figure 18: Box Plot of School Performance Characteristics by Cluster

By dividing schools into these clusters, the analysis offers a deeper understanding of how schools vary in performance across multiple dimensions, providing a foundation for **tailored benchmarking**. Figures 18 and 19 demonstrate how these clustering algorithms support EQA by identifying strengths and weaknesses across schools. Clustering enables the grouping of schools with similar performance characteristics, helping to pinpoint specific areas for improvement.

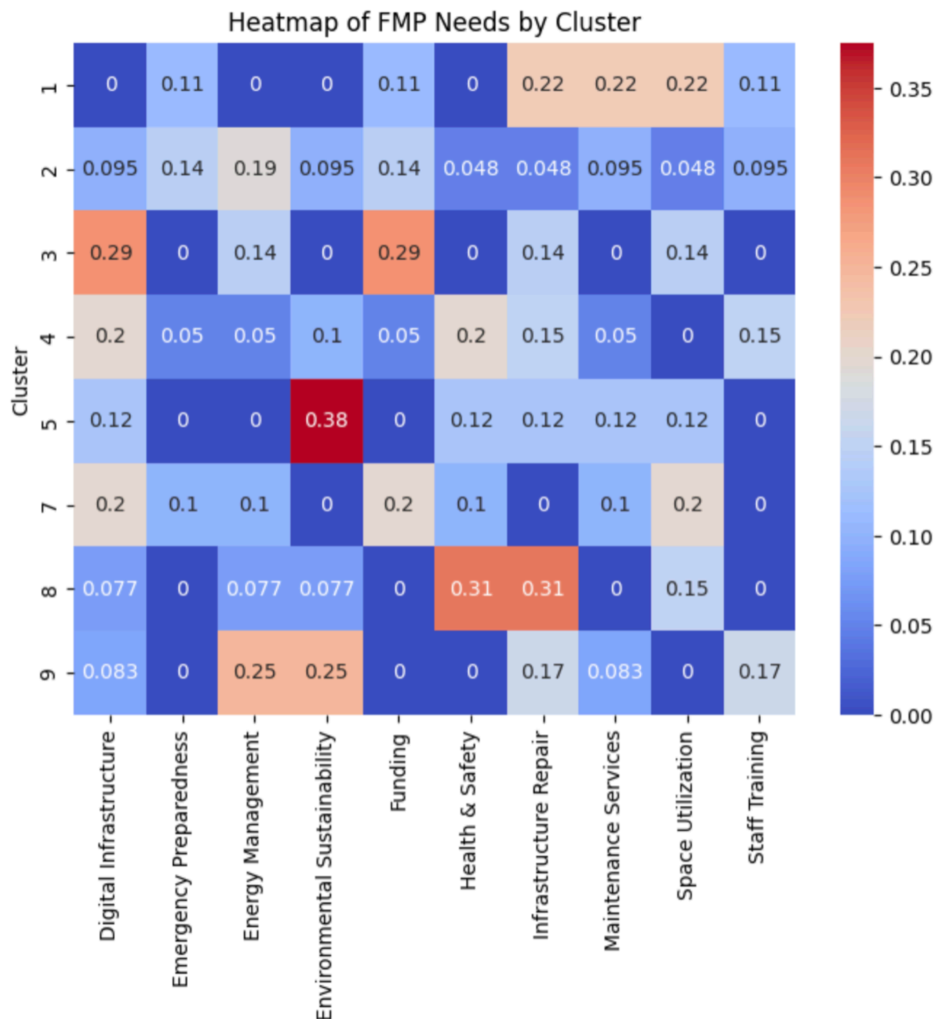


Figure 19: Heatmap of Performance Needs Across School Clusters

2) Section 2: Analysis and Visualization for Collaborative Benchmarking

In addition to regulatory benchmarking, the extracted data was used to demonstrate **Collaborative Benchmarking**, where schools are compared with their peers to foster the sharing of best practices aims to enhance overall educational quality by promoting collective improvement.

2.1) Best Practices and Geographic Analysis

To identify and honor high-performing schools, **geographic mapping** was employed. This method allows for the visualization of where high performers are located and highlights the spread of **best practices** across regions (Figures 20-21). Geographic analysis facilitates the recognition of top schools besides encourages other institutions to adopt successful strategies, contributing to continuous improvement in educational standards (available at: https://drive.google.com/file/d/1jNeKKLnGnaE7JCRAuVEzNwhU2g1FlTDk/view?usp=sharing).

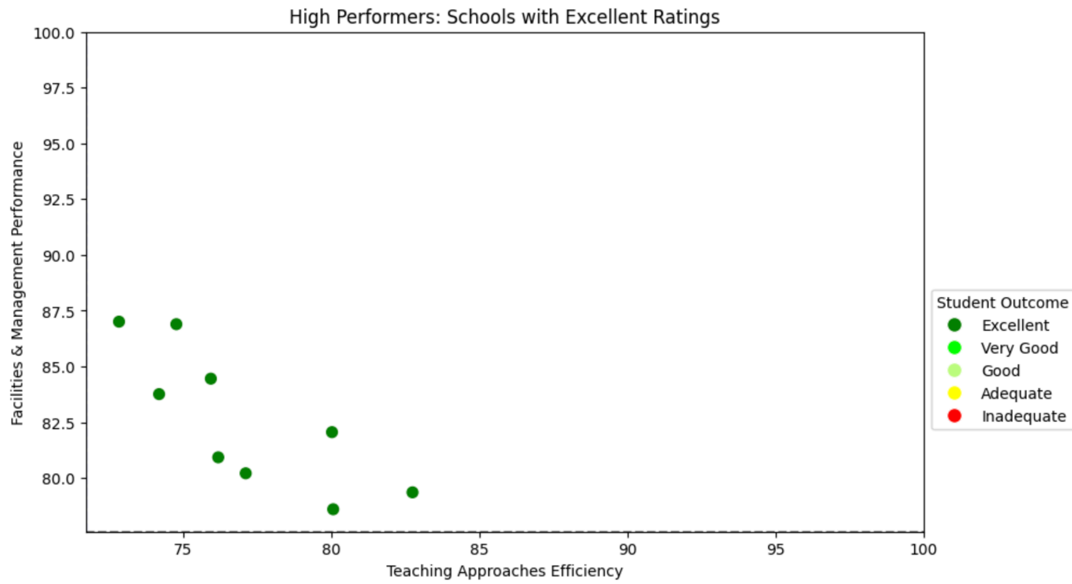


Figure 20: Scatter Plot of High-Performing Schools with Excellent Ratings

Figure 21: Geographic Map of High-Performing Schools with Best Practices for Adoption

2.2) Individual School Performance and Benchmarks

Using data visualizations such as **individual school performance scoreboards** (Figure 22), the study compared schools with established benchmarks, both overall and within their clusters. This comparative analysis helps schools understand how they rank in relation to the benchmarks and highlights areas where further improvements can be made. Moreover, this visualization empowers schools to set their own goals, whether they aim to meet average performance levels or continue excelling beyond their current benchmarks (Figure 23).

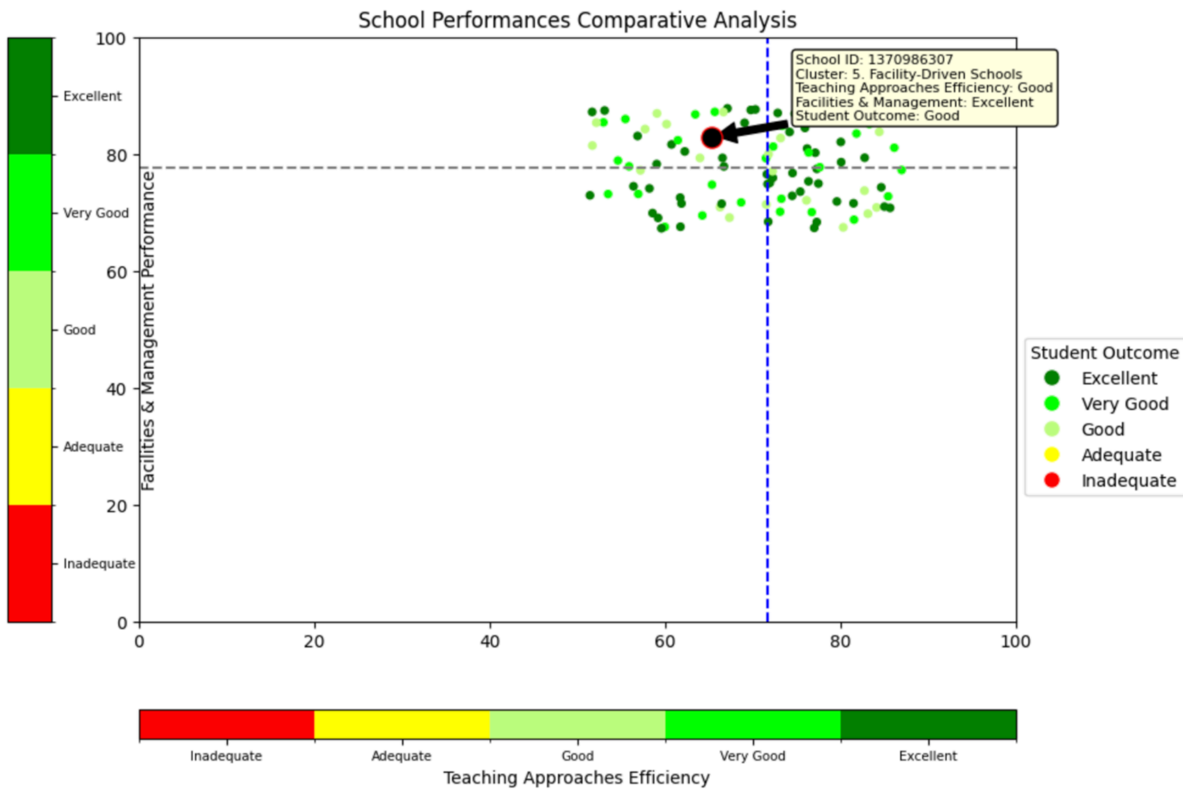


Figure 22: Scatter Plot of Individual School Performance Scores with Ranking Comparison

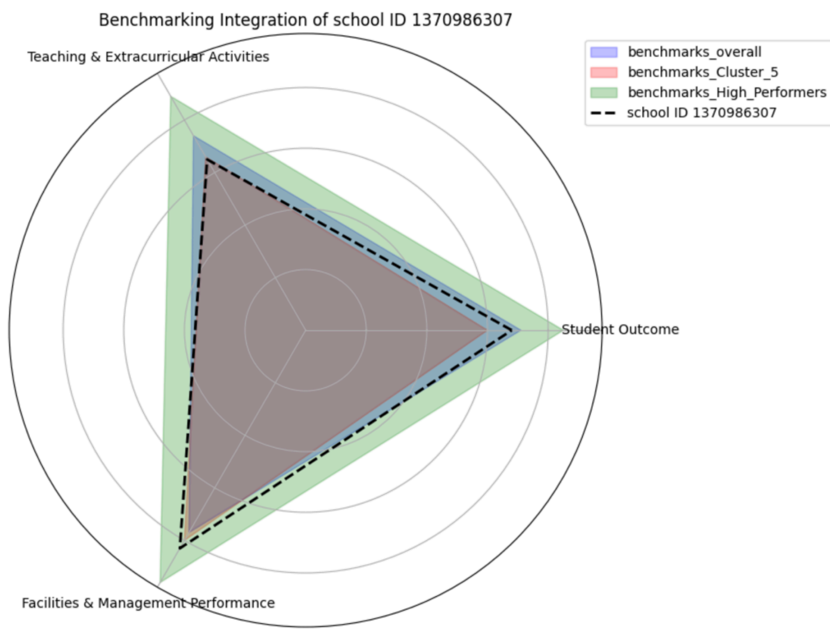


Figure 23: Radar Chart of Individual School Performance vs. Overall, Cluster, and High-Performer Benchmarks

4.3.2 Final Thoughts on the Utility of Extracted Data

The research objectives have been clearly met through a combination of data analysis, visualization, and feedback from ONESQA stakeholders. The exploratory survey was crucial in shaping the research direction, revealing that ONESQA stakeholders expected both advanced analytical approaches and clear visualizations. This led directly to the development of the NLP pipeline, ensuring that the extracted data would be relevant and useful for performance evaluation.

The extracted data enabled analysis and visualization of school performance, for instance, quadrant-based clustering identified distinct performance groups, providing a more detailed understanding of how schools perform across key indicators, directly supporting research objective 2. This method also allowed ONESQA to set specific performance benchmarks for each cluster (see Figures 16-23). These benchmarks align with ONESQA's goals of regulatory and collaborative benchmarking, helping to establish clearer, more realistic performance categories, and targets for school improvement. This approach makes it clear why schools received certain scores and ensures that decisions are based on actual performance data, enhancing transparency and robustness in the evaluation process.

4.4 Discussion

In the final section, we examine how advanced clustering techniques and visualizations can uncover hidden performance patterns, demonstrating their seamless integration into the EQA benchmarking process to support more effective and transparent evaluation standards.

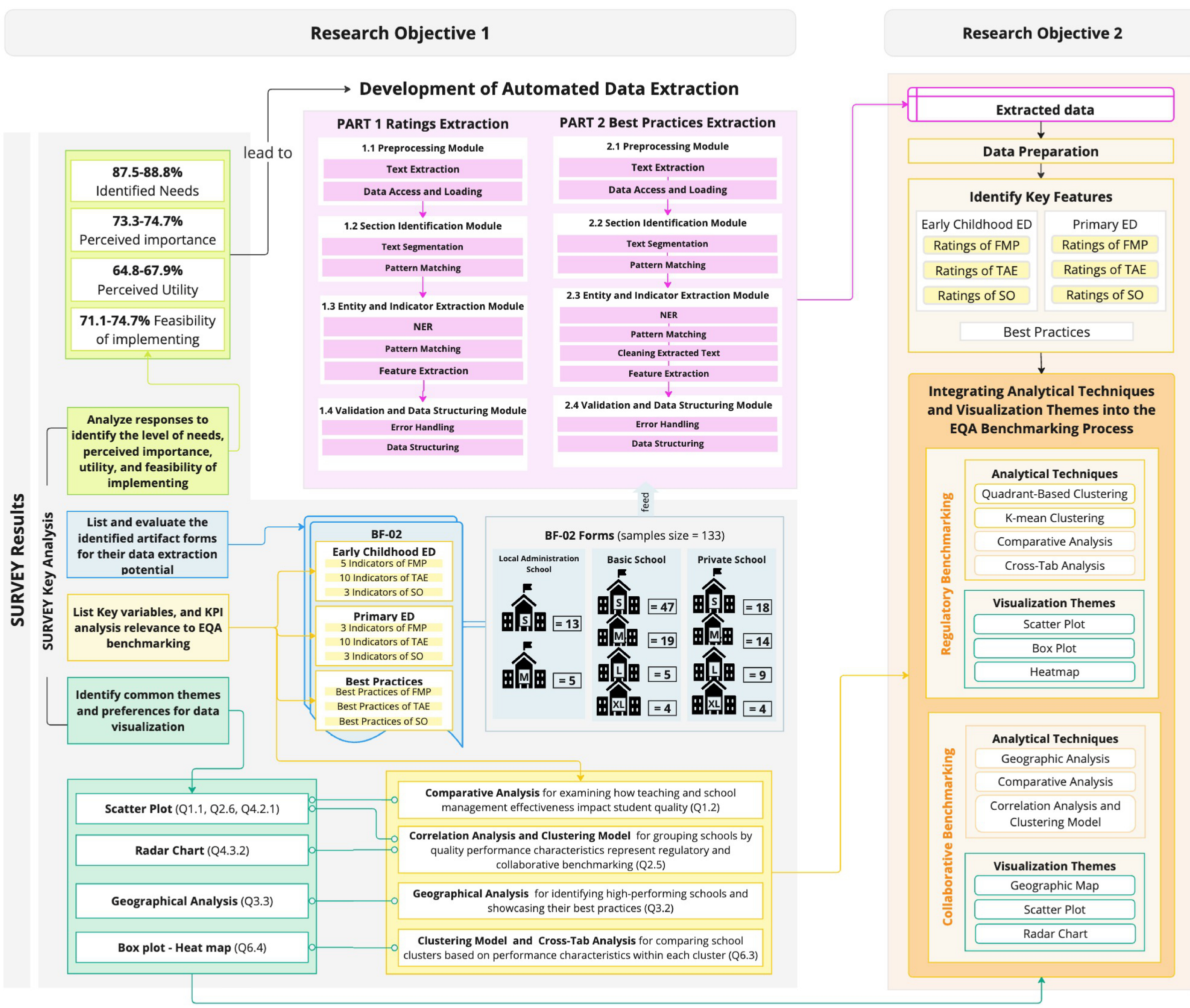


Figure 24: Research Results

4.4.1 Discussion on key insights of survey

1) Identifying the Need for NLP-Based Automated Data Extraction

A majority of survey respondents (87.5%-88.8%) recognized the importance of implementing an NLP-based pipeline for extracting unstructured data, highlighting its necessity for benchmarking purposes. Furthermore, 71.1%-74.7% of respondents believed that the NLP pipeline was feasible and would enhance the EQA benchmarking process. This observation is consistent with findings from Harju-Luukkainen et al. (2022) and Schellekens (2023), who highlighted the inefficiency of manual processes and the necessitate of streamline data collection. Studies like those by Sirsat and Chavan (2016) emphasize that while NLP offers significant advantages, interest and adoption remain limited in certain sectors. Similarly, Chen et al. (2023) and Jofche et al. (2023) noted that automating such tasks could significantly reduce human error and improve overall accuracy, aligning with the positive feedback for automation in this study.

2**) Preferences for Analytics and Visualization Techniques**

Stakeholders recommended advanced clustering analysis and scatterplot visualizations, with 73.4% favoring geographic maps for collaborative benchmarking. This supports the study's second objective of using visualization techniques to present extracted data in a way that enhances understanding and decision-making. Relate this to the study of Jo (2023), which emphasize the value of advanced analytics, combining ML and statistical methods to extract meaningful insights from large datasets and using visualizations to make data insights accessible in educational contexts.

4.4.2 Discussion on key insights of NLP Pipeline for Data Extraction development

The NLP-based extraction pipeline demonstrated exceptional performance, achieving a 98.33% accuracy rate in extracting both numerical ratings and descriptive best practices. This aligns with findings by Zaki et al. (2022), who emphasize the value of preprocessing and vectorization in enhancing data reliability, and Chen et al. (2023), who highlight the role of regex and text segmentation in improving extraction accuracy. However, despite its success with a small dataset of 133 samples, the model must be adapted to handle the larger datasets typical of the EQA process, often exceeding 1,000 files. To ensure accurate extraction on this scale, the algorithm requires enhancements in table recognition to capture ratings embedded in tables effectively.

A significant challenge emerged in the segmentation of Thai text, where the absence of explicit word boundaries reduced the pipeline's recall rate. This was evident in the 15.27% of identifiers (38 out of 249) that were neither correctly matched nor identified as missing. This challenge is consistent with the observations of Soisoonthorn et al., and Phatthiyaphaibun et al. (2023), who discuss language-specific barriers in NLP, particularly in Thai. Future improvements should focus on integrating language-specific models or dictionary-based approaches to better address these segmentation issues and enhance overall accuracy.

4.4.3 Discussion on Application of Extracted Data for EQA Benchmarking

The key findings demonstrate the high utility of extracted data for advanced analysis and visualization, making it a sufficient input for EQA benchmarking. By employing scatter plots, geographic maps, radar charts, and K-Means clustering, performance across critical areas can be thoroughly assessed. These tools are more effective than traditional descriptive statistics or correlation analysis, which often fail to reveal hidden patterns and non-linear relationships in the data (Jo, 2023).

The extracted data's integration into the benchmarking process is both effective and adaptable, fostering dynamic benchmarks that are context-sensitive. This supports the arguments by Baartman et al. (2007, 2011), Lucander & Christenson (2020), and Marciniak (2018), who emphasize that benchmarking, when done effectively, enables institutions to measure performance against standards and pursue continuous improvement. It supports both regulatory and collaborative benchmarking, providing a macro (systemic) and micro (individual school) understanding of strengths and weaknesses. This dual perspective aligns with Tangpornpaiboon's (2022) critique, as it allows ONESQA to deliver more personalized and actionable recommendations, bridging the gap between data collection and meaningful feedback.

In summary, the integration of advanced data extraction and visualization techniques into the EQA benchmarking process aligns with the evolving demands of educational quality assurance.

It enables deeper insights, supports continuous improvement, and provides a balanced approach to system-wide and institution-specific evaluation (ONESQA, 2020).

4.4.4 Discussion of the Benefits of Research Results for Stakeholders

This research provides clear value to multiple stakeholders. ONESQA, as the main party to benefit, will be able to establish realistic performance benchmarks for schools and deliver more tailored recommendations. This research encourages ONESQA to focus on both regulatory and collaborative benchmarking, which is central to the EQA agency's mission. By integrating analysis into the benchmarking process, the visualizations provide a clear and engaging way for schools and policymakers to understand their performance in relation to others, data-backed decisions make it easier for ONESQA to explain how schools received their ratings, which is crucial for building trust and consensus among stakeholders, thus enhancing overall engagement in the quality assurance process.

The schools themselves, especially those in the middle and lower clusters, will receive targeted interventions to enhance both teaching efficiency and student outcomes. Furthermore, high-performing schools can use the peer comparisons to maintain and further elevate their standards. This transparency encourages schools to actively engage in the benchmarking process, fostering a culture of continuous improvement.

Moreover, policymakers and government agencies will also benefit by gaining a clearer understanding of where resources and policy interventions should be focused to support continuous improvement across the education sector.

4.5 Study Limitations

Despite the research results show strengths, the study had several limitations. First, the survey sample excluded some stakeholder groups, affecting the generalizability of the results. Additionally, after applying advanced analysis and visualization, a follow-up survey may be needed to confirm whether the outcomes align with stakeholder expectations and effectively integrate into the EQA benchmarking. Second, the NLP pipeline, while effective, faced challenges in handling ambiguous or context-dependent language, leading to some inaccuracies in entity extraction. Additionally, the reliance on pre-existing literature to develop the pipeline meant that certain innovative features, such as deep learning models, were not fully explored.

4.6 Recommendations for Future Research

Future research should consider integrating advanced NLP techniques, such as transformer-based models, to enhance the accuracy of data extraction. Developing models that better recognize the document structure of all EQA forms would also be valuable. Additionally, adapting the NLP pipeline to extract information for ONESQA's back office could significantly reduce processing time, streamlining the overall EQA process. Implementing NER to accurately identify school names, particularly given their complexity in Thai, would further enhance the pipeline's functionality for advanced analysis in educational contexts.

CONCLUSION

The primary objective of the research was to create and deploy a Natural Language Processing (NLP) data extraction pipeline to improve external quality assurance (EQA) benchmarking processes at ONESQA. The major investigation component of this work was surveying key stakeholders to determine the specific requirements for an automated data extraction system, which expects the insights could directly lead to tailored development of suitable NLP data extraction along with ONESQA challenges. The survey results highlighted the necessity of an NLP-based approach along with directly informed the development process.

Survey data played a crucial role in shaping the research objectives. A significant 87.5% to 88.8% of respondents emphasized the importance of automated tools for extracting unstructured data to generate valuable insights and improve benchmarking processes. This primary research also identified crucial criteria required for effective EQA benchmarking, including student results, instructional approach efficiency, and financial management performance, which influenced the creation of an NLP pipeline capable of reliably extracting useful and adequate data from complicated texts, therefore filling a crucial gap in existing EQA processes and aligning the tool with real-world requirements.

The NLP pipeline's efficiency was proved by its excellent accuracy rates, which included a 98.33% match with annotated datasets and an F1 score of 1.0 for relevant text of best practices and numerical ratings extraction. Understanding and implementing stakeholder demands enabled the pipeline to effectively acquire vital data with minimum differences, hence supporting the objective of enhancing automated EQA methodologies.

Equally important was how the primary research informed the utility of extracted data through advanced analytics and visualization techniques. Stakeholders expressed specific expectations for the kinds of analysis and visualizations that would be most beneficial for EQA benchmarking. The study responded to these expectations by employing clustering methods and visualizations, scatter plots, heatmaps, geographic maps and etc, that was able to uncover hidden patterns in school performance and set context-sensitive benchmarks.

The utility of extracted data by analytics and visualization proposed in this study enhances stakeholder engagement, transparency, and accountability by making the EQA process more data-driven and objective. By automatically extracting data from schools' records as evidence-based reports, the pipeline reduces subjective judgment and fosters open participation in the evaluation process which encourages schools to engage in discussions for improvement, as the insights are based on their own performance data. In this case, teachers and administrators can have meaningful conversations about specific areas of improvement using objective evidence. This allows ONESQA to compare school performance, identify trends that might not be apparent from raw data, and visualize the results using tools using scatter plots, heat maps, radar charts and etc, which could simplify complex data, assisting stakeholders understand performance trends more clearly. Besides, for guiding benchmark setting and support schools on their improvement paths by concentrating on transparency and ongoing improvement, the proposed analytics in this study reveal which schools thrive and the reasons that drive sustainable progress in education.

Although the research provided significant insights, it faced several limitations. The vital point need to work in the future regarding the NLP pipeline; even though it performed well, it has to be strengthened in order to accurately extract data from new EQA forms and emphasized the summarization of text extracted rather than only identifying the desired text. The pipeline was built around a rule-based technique for identifying document portions and headers; however, this algorithm failed with documents that had dangling text. Future study ought to probe into more advanced approaches for dealing with such circumstances, which would result in more successful text capture and should focus on developing NLP model to interpret or generate key information rather that text capturing. Furthermore, the dangling text found in this study should be resolved by integrating NLP methods with machine learning algorithms and/or deep learning models to battle with unclear language and failed to thoroughly analyze new areas such as deep learning models.

To sum it up, this work enhances the area of educational quality assurance by demonstrating how an NLP-based data extraction pipeline can accelerate ahead of the previously time-consuming benchmarking procedure. The pipeline enables for faster and more efficient comparisons of institutions and educational results by automating the extraction and structuring of massive volumes of textual data, which builds a foundation for more effective and transparent EQA standards by using powerful analytics and visualization enhancing stakeholders engagement in terms of simplifying the harder analytics alignment with their perceptions. These technologies improve the ability to interpret complicated qualitative data, thereby bringing insights more accessible to stakeholders. Future study ought to aim at developing these approaches and examining their broader applicability to establish systems that allow for continuous improvement and support the continued growth of educational quality assurance standards.

REFERENCES

Baartman, L.K., Prins, F.J., Kirschner, P.A. and Van der Vleuten, C.P. 2007, “Determining the quality of assessment programs: a self-evaluation procedure”, *Studies in Educational Evaluation*, vol. 33, no. 3/4, pp. 258-281. Available at: https://doi.org/10.1016/j.stueduc.2007.07.004 [Accessed 4 April 2024].

Baartman, L.K., Prins, F.J., Kirschner, P.A. and Van der Vleuten, C.P. 2011, “Self-evaluation of assessment programs: a cross-case analysis”, *Evaluation and Program Planning*, vol. 34, no. 3, pp. 206-216. Available at: https://doi.org/10.1016/j.evalprogplan.2011.03.001 [Accessed 4 April 2024].

Chen, Q., Banerjee, A., Demiralp, Ç., Durrett, G. and Dillig, I. 2023, “Data extraction via semantic regular expression synthesis”, Proceedings of the ACM on Programming Languages, vol. 7, pp. 1848-1877. Available at: https://acm.org [Accessed 16 June 2024].

Deshmukh, A.A. 2021, Unsupervised Syntactic Structure Induction in Natural Language Processing, Master’s thesis, University of Waterloo. Available at: http://hdl.handle.net/10012/17347 [Accessed 29 July 2024].

Gapsalamov, A., Bochkareva, T., Vasilev, V., Akhmetshin, E. and Anisimova, T. 2020, “Comparative analysis of education quality and the level of competitiveness of leader countries under digitalization conditions”, *Journal of Social Studies Education Research*, vol. 11, no. 2, pp. 133-150. Available at: https://learntechlib.org [Accessed 7 July 2024].

Hansen, M., Pomp, A., Erki, K. & Meisen, T. 2019, "Data-Driven Recognition and Extraction of PDF Document Elements", *Technologies,* vol. 7, no. 3, pp. 65, DOI:10.3390/technologies7030065.

Jackson, N. 2002, “Growing knowledge about QAA subject benchmarking”, *Quality Assurance in Education*, vol. 10, no. 3, pp. 139-146, DOI:10.1108/09684880210435912.

Jo, A. 2023, Visualization And Analysis Of Student Data Using Machine Learning And Statistical Methods. Master's thesis, Tampere University, Master's Degree Programme in Information Technology. Available at: https://tuni.fi [Accessed 7 July 2024].

Jofche, N., Mishev, K., Stojanov, R., Jovanovik, M., Zdravevski, E. & Trajanov, D. 2023, "PharmKE: Knowledge Extraction Platform for Pharmaceutical Texts Using Transfer Learning", *Computers,* vol. 12, no. 1, pp. 17, DOI:10.3390/computers12010017.

Kasliwal, N., 2018, Natural Language Processing with Python Quick Start Guide : Going from a Python Developer to an Effective Natural Language Processing Engineer, Packt Publishing, Limited, Birmingham. Available from: ProQuest Ebook Central. [2 March 2024].

Kayyali, M., 2023, "The Relationship between Rankings and Academic Quality", *International Journal of Management, Sciences, Innovation, and Technology,* vol. 4, no. 3, pp. 1-11. Available at: https://www.researchgate.net/publication/371982309_The_Relationship_between_Rankings_and_Academic_Quality [Accessed 16 June 2024].

Khamphakdee, N. and Seresangtakul, P. 2023, "An Efficient Deep Learning for Thai Sentiment Analysis", *Data,* vol. 8, no. 5, pp. 90, DOI:10.3390/data8050090.

Kroll, M., Rettinger, A. and Gemulla, R. 2022, "Named Entity Extraction for Knowledge Graphs: A Literature Overview", *Journal of Web Semantics,* vol. 72, 100677. Available at: https://doi.org/10.1016/j.websem.2021.100677 [Accessed 16 June 2024].

Lucander, H. and Christersson, C. 2020, "Engagement for quality development in higher education: a process for quality assurance of assessment", *Quality in Higher Education*, Vol. 26 No. 2, pp. 135-155, DOI:10.1080/13538322.2020.1761008.

Magoc, T., Everson, R. and Harle, C.A. 2023, "Enhancing an enterprise data warehouse for research with data extracted using natural language processing", *Journal of Clinical and Translational Science,* vol. 7, no. 1, DOI:10.1017/cts.2023.575.

Malashin, I., Masich, I., Tynchenko, V., Gantimurov, A., Nelyub, V. & Borodulin, A. 2024, "Image Text Extraction and Natural Language Processing of Unstructured Data from Medical Reports", *Machine Learning and Knowledge Extraction,* vol. 6, no. 2, pp. 1361-1377, DOI:10.3390/make6020064.

Marciniak, R. 2018, “Quality assurance for online higher education programmes: design and validation of an integrative assessment model applicable to Spanish universities”, *International Review of Research in Open and Distributed Learning*, Vol. 19, No. 2, pp. 126-154, DOI:10.19173/irrodl.v19i2.3443.

Meesad, P. 2021, "Thai Fake News Detection Based on Information Retrieval, Natural Language Processing and Machine Learning", *SN Computer Science,* vol. 2, no. 6, pp. 425, DOI:10.1007/s42979-021-00775-6.

Nuankaew, P., Nasa-Ngium, P., Kunasit, T. & Nuankaew, W.S. 2023, "Implementation of Data Analytics and Machine Learning in Thailand Education Sector", *International Journal of Emerging Technologies in Learning (Online),* vol. 18, no. 5, pp. 175-191, DOI:10.3991/ijet.v18i05.36871.

Office for National Education Standards and Quality Assessment (ONESQA) 2021, “การศึกษารูปแบบการประเมินคุณภาพภายนอกแบบเสมือนจริง [Study on Virtual Models for External Quality Assessment]”, จุลสาร สมศ. [ONESQA Journal], vol. 22, no. 1, pp. 6-12. Available at: https://www.onesqa.or.th/upload/download/202204051424371.pdf [Accessed 4 April 2024].

Office for National Education Standards and Quality Assessment (ONESQA). 2020, การสังเคราะห์การประเมินคุณภาพภายนอก [Analysis of the Fourth Cycle of External Quality Assessment and Future Directions for External Quality Evaluation]. Internal report, ONESQA.

Office for National Education Standards and Quality Assessment (ONESQA) 2024, ONESQA Self-Evaluation Report. Internal report, ONESQA. Available at: https://www.onesqa.or.th/upload/download/202405091620022.pdf [Accessed 4 April 2024].

Pang, B. and Lee, L. 2008, "Opinion Mining and Sentiment Analysis." *Foundations and Trends in Information Retrieval*, vol. 2, No. 1–2, pp. 125–135, DOI:10.1561/1500000011.

Phatthiyaphaibun, W., Chaovavanich, K., Polpanumas, C., Suriyawongkul, A., Lowphansirikul, L., Chormai, P., Limkonchotiwat, P., Suntorntip, T. & Udomcharoenchaikit, C. 2023, *PyThaiNLP: Thai Natural Language Processing in Python*, Ithaca. Available at: https://arxiv.org/pdf/2312.04649.pdf [Accessed 7 July 2024].

Russell, J. and Markle, R. 2017, "Continuing a culture of evidence: assessment for improvement", *ETS Research Report Series*, Vol. 2017 No. 1, pp. 1-10, DOI:10.1002/ets2.12136.

Schellekens, L.H., van d.S., Vleuten, C.P.v.d., Prins, F.J., Wools, S. and Bok, H.G.J. 2023, "Developing a digital application for quality assurance of assessment programmes in higher education", *Quality Assurance in Education,* vol. 31, no. 2, pp. 346-366, DOI:10.1108/QAE-03-2022-0066.

Sirsat, S.R. and Chavan, V., 2016. Pattern matching for extraction of core contents from news web pages. In: *2016 Second International Conference on Web Research (ICWR)*. [Online] Available at: https://www.researchgate.net/publication/304456801_Pattern_matching_for_extraction_of_core_contents_from_news_web_pages [Accessed 7 July 2024].

Soisoonthorn, T., Unger, H. & Maliyaem, M., 2023, "Thai Word Segmentation with a Brain-Inspired Sparse Distributed Representations Learning Memory", *Computational Intelligence and Neuroscience : CIN,* vol. 2023, pp.1-10, Available at: https://doi.org/10.1155/2023/8592214 [Accessed 7 July 2024].

Tangpornpaiboon, S., 2021. Understanding Discrepancies in Trends Results of PISA, TIMSS, and O-NET and Implications for Education Policies in Thailand. PhD thesis, University College London

Udomcharoenchaikit, C., Boonkwan, P. & Vateekul, P., 2023, "Towards improving the robustness of sequential labeling models against typographical adversarial examples using triplet loss", *Natural Language Engineering,* vol. 29, no. 2, pp. 287-315, DOI:10.1017/S1351324921000486

Wu, H., Yamal, J.M., Yaseen, A. and Maroufy, V. 2020, Statistics and Machine Learning Methods for EHR Data: From Data Extraction to Data Analytics, CRC Press LLC, Milton. Available from: ProQuest Ebook Central. [8 April 2024].

Zaki, N., Turaev, S., Shuaib, K., Baby, A. and Mohamed, E. 2022, "Automating the Mapping of Course Learning Outcomes to Program Learning Outcomes using Natural Language Processing for Accurate Educational Program Evaluation". DOI: 10.21203/rs.3.rs-2196467/v1. Available at: https://www.researchgate.net/publication/364707363_Automating_the_Mapping_of_Course_Learning_Outcomes_to_Program_Learning_Outcomes_using_Natural_Language_Processing_for_Accurate_Educational_Program_Evaluation [Accessed 29 July 2024].

APPENDIX

1. Relevance of NLP data extraction development

1.1 Rating of Indicators Extraction

(c) Code 1: Rating of indicators Extraction available at: https://colab.research.google.com/drive/1ZghkF8kAHw61HlCWn6P0zmt0fxzM7g3Y?usp=sharing

(d) Extracted data of rating available at: https://drive.google.com/file/d/13dHxhQzvzte3G9JKYf85ZX8IigVwJYj3/view?usp=sharing

(e) ANNOTATED DATASET_Rating133.xlsx available at: https://docs.google.com/spreadsheets/d/1-pBbl1ON5EAL9syYK3F83i0h1wNYf1vE/edit?usp=sharing&ouid=110502415767573536935&rtpof=true&sd=true

1.2 Best practices Extraction

(a) Code 2.1-2.2: Best practices available at: https://colab.research.google.com/drive/1rOwk7urCKeoLOpEg9lumhZzB6aOLCUwz?usp=sharing

(b) Extracted data of Best practices available at: https://docs.google.com/spreadsheets/d/10SM0sUcNXMMqqMLso19Ct4S8I0Y1Gfq9/edit?usp=sharing&ouid=110502415767573536935&rtpof=true&sd=true

(c) ANNOTATED DATASET_BP available at: https://docs.google.com/spreadsheets/d/1AfLoLvVN1f4MWYSD1c-g_JZPB4Z4pZUO/edit?usp=sharing&ouid=110502415767573536935&rtpof=true&sd=true

2. Questionnaire

Survey on Opinions Toward an NLP Data Extraction Pipeline

for Comparative Educational Quality Assessment (EQA Benchmarking)

Instructions:
This survey aims to gather the opinions of ONESQA executives and staff on the analysis of educational quality assessment results through benchmarking. The survey is part of research on *Enhancing Educational Quality Assessment in Thailand: Leveraging NLP and Machine Learning for Effective Benchmarking and Decision-Making* (Researcher: Ms. Punyisa Phumiphol, Research and Knowledge Management Officer, ONESQA). The research has two main objectives: 1) To develop an NLP-based automated data extraction pipeline for EQA artifacts in Thai, and 2) To demonstrate the extracted data utility through cluster analysis and visualization for EQA benchmarking

Please kindly respond based on your actual experience and opinions. Your information will be kept confidential and used solely for research purposes. Your responses will be highly valuable for developing ONESQA's educational quality assessment analysis to better align with stakeholders' needs and support decision-making in improving the education system.

Section 1: Respondents' Personal Information

1.1 Position

○ Internal Stakeholders

○ External Stakeholders

1.2 For Internal Stakeholders

○ Leaders of Onesqa

○ Assessment & Certification Bureau officers

○ Development & Promotion Bureau officers

○ Policy & Strategy Bureau officers

1.3 For External Stakeholders

◯ Department of Local Administration (DLA), Ministry of Interior

◯ Office of the Basic Education Commission, Ministry of Education

◯ Office of the Higher Education Commission, Ministry of Education

◯ Office of the Private Education Commission, Ministry of Education

◯ Bangkok Metropolitan Administration (BMA)

◯ Department of Local Education Development, Ministry of Culture

1.4 Work experience (years)

◯ less than 1 years

◯ 1 - 5 years

◯ 5 - 10 years

◯ more than 10 years

Section 2: Opinions on the NLP Data Extraction Pipeline for Comparative Educational Quality Assessment (EQA Benchmarking)

Instructions: Since you work closely with stakeholders such as schools, governing agencies, the government, and the public, please share your opinion on the presentation format for external quality assessment analysis through benchmarking by selecting the option that best reflects your viewpoint.

1.How much do you think there is a need for ONESQA to implement a feature like the NLP data extraction pipeline that can extract relevant variables from existing EQA reports and school reports for comparative analysis for EQA benchmarking, as detailed below?

Objective of the analysis: To examine the relationship between the variables *teaching effectiveness* (Standard 3) and *school management effectiveness* (Standard 2) that influence *student outcome* (Standard 1). For example, the comparative analysis results will be presented using a scatter plot with the following details; The X-axis represents *teaching effectiveness* (Standard 3: Student-centered teaching methods). The Y-axis represents *school management effectiveness* (Standard 2: Management and administration processes). The color of the points in the chart represents the outcome, which is *student quality* (Standard 1).

- Strongly Needed
- Needed
- Neutral
- Slightly Needed
- Not Needed at All

1.1: What type of visualization would be most effective for examining the relationship between teaching effectiveness (Standard 3) and school management effectiveness (Standard 2) in relation to student quality (Standard 1)?

- Scatter Plot with annotations
- Heatmap
- Radar Chart
- Box Plot
- Geographic Map with annotations
- Other (Please specify)

1.2: What type of analysis would be most effective for examining the relationship between teaching effectiveness (Standard 3) and school management effectiveness (Standard 2) in relation to student quality (Standard 1)?

- Correlation Analysis
- Regression Analysis
- Factor Analysis
- Comparative Analysis
- Clustering Model
- Geographic Analysis
- Other (Please specify)

2.1 How clearly do you think the comparative analysis using a scatter plot as mentioned in item 1 can **compare the assessment results** of each standard?

- **(Please rate your level of agreement with the following statement):**
  "The scatter plot format effectively compares the assessment results of teaching effectiveness, school management effectiveness, and student quality."
- Strongly Disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

2.2: How much do you agree that ONESQA should implement an NLP data pipeline to extract variables for data analysis and clustering of educational institutions based on quality characteristics (Cluster Analysis), in order to provide more tailored recommendations that align with the specific characteristics of each group of institutions?

- Strongly Disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

2.3: How much do you agree that ONESQA should implement an NLP data pipeline to extract variables for data analysis and clustering of educational institutions based on quality characteristics (Cluster Analysis), which would be crucial for customizing strategies, tools, and questions to monitor school performance in future rounds, focusing on the areas that matter most for each group?

- Strongly Disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

2.4: To what extent do you agree that ONESQA should implement an NLP data pipeline to extract variables for data analysis and clustering of educational institutions based on quality characteristics? This could be crucial for accurately reflecting school performance and uncovering hidden patterns in school quality.

- Strongly Disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

2.5: What type of analysis techniques would be most effective for grouping educational institutions based on quality characteristics?

- Correlation Analysis
- Regression Analysis
- Factor Analysis
- Comparative Analysis
- Clustering Model
- Geographic Analysis
- Other (Please specify)

2.6: What type of visualization would be most effective for clustering educational institutions based on quality characteristics?

- Scatter Plot with annotations
- Heatmap
- Radar Chart
- Box Plot
- Geographic Map with annotations
- Other (Please specify)

3.1: To what extent do you agree that ONESQA should implement an NLP data pipeline to extract variables for clustering schools and use geographic analysis to identify high-performing schools and showcase their best practices? This approach could be beneficial for collaborative benchmarking, encouraging schools to improve their quality by comparing themselves with high performers, and driving overall quality enhancement.

- Strongly Disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

3.2: What type of analysis technique would be most effective for identifying the locations of high-performing schools and showcasing their best practices?

- Correlation Analysis
- Regression Analysis
- Factor Analysis
- Comparative Analysis
- Clustering Model
- Geographic Analysis
- Other (Please specify)

3.3: What type of visualization would be most effective for identifying the locations of high-performing schools and showcasing their best practices?

- Scatter Plot with annotations
- Heatmap
- Radar Chart
- Box Plot
- Geographic Map with annotations
- Other (Please specify)

4.1: To what extent do you agree that ONESQA should implement an NLP data pipeline to extract variables for clustering schools and use scatter plots with annotated performance ratings and clusters to visualize each school's performance? This approach could **enhance regulatory benchmarking** by providing a performance scoreboard and clearly identifying each school's quality performance across different standards.

- Strongly Disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

4.2: To what extent do you agree that ONESQA should implement an NLP data pipeline to extract variables for clustering schools and use scatter plots with annotated performance ratings and clusters to visualize each school's performance? This approach could enhance regulatory benchmarking by providing a performance scoreboard and identifying each school's performance ranking.

- Strongly Disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

4.2.1: What type of visualization would be most effective for displaying each school's performance ranking and scoreboard?

- Scatter Plot with annotations
- Heatmap
- Radar Chart
- Box Plot
- Geographic Map with annotations
- Other (Please specify)

4.3.1: To what extent do you agree that ONESQA should implement an NLP data pipeline to extract variables for clustering schools and use comparative analysis with radar charts to visualize school performance against overall benchmarks, high-performing schools, cluster benchmarks, and individual school performances? This approach could **enhance regulatory benchmarking** by providing a performance scoreboard that encourages continuous improvement, even for schools with excellent ratings, by **comparing themselves with high performers.**

- Strongly Disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

4.3.2: What type of visualization would be most effective for comparing each school's performance against overall benchmarks, high-performing schools, and cluster benchmarks, in a way that encourages continuous improvement, even for schools with excellent ratings?

- Scatter Plot with annotations
- Heatmap
- Radar Chart
- Box Plot
- Geographic Map with annotations
- Other (Please specify)

5: To what extent do you agree that ONESQA should implement an NLP data pipeline to extract sufficient variables for clustering schools based on their performance characteristics in a given evaluation cycle? The pipeline should capture key performance indicators for each standard and apply basic clustering techniques like quadrant-based clustering using medians to divide clusters. Additionally, if basic clustering does not adequately reveal school characteristics, more advanced techniques should be considered.

- Strongly Disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

6.1: To what extent do you agree that ONESQA should implement an NLP data pipeline to extract sufficient variables for clustering schools based on their performance characteristics in a given evaluation cycle? The pipeline should capture key performance indicators for each standard and provide enough data to apply advanced clustering techniques.

- Strongly Disagree

- Disagree
- Neutral
- Agree
- Strongly Agree

6.2: To what extent do you agree that implementing an NLP data pipeline to extract sufficient data for both basic and advanced clustering techniques would be beneficial for clearly revealing school performance characteristics? The pipeline should capture key performance indicators for each standard and provide comprehensive data to effectively utilize these techniques.

- Strongly Disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

6.3: What type of analysis technique would be most effective for clustering schools based on their performance characteristics to clearly reveal these characteristics?

- Basic Clustering Analysis
- Advanced Clustering Analysis
- Both
- Other (Please specify)

6.4: What type of visualization would be most effective for comparing school clusters based on their performance characteristics to clearly reveal these characteristics?

- Scatter Plot with annotations
- Heatmap
- Radar Chart
- Box Plot
- Geographic Map with annotations
- Other (Please specify)

7: To what extent do you agree that implementing an NLP data pipeline to extract sufficient data for both basic and advanced clustering techniques, and visualizing this data with box plot charts to compare performance across clusters, would be beneficial for clearly revealing school performance characteristics?

- Strongly Disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

8: To what extent do you agree that implementing an NLP data pipeline to extract sufficient data for both basic and advanced clustering techniques, and visualizing this data with Cross-Tab Analysis (Heatmaps) to compare school needs across clusters, would be beneficial for identifying which areas require immediate government support?

For example:

> The X-axis represents categories of needs in each standard (e.g., Standard 2: Facilities & Management Performance), showing specific areas for development such as digital infrastructure or emergency preparedness.
>
> The Y-axis represents clusters of schools.
>
> Colors in the heatmap indicate the level of need for improvement: red indicates that many schools in that cluster need help in that area, while blue indicates few or no needs.

- Strongly Disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

9: To what extent do you think ONESQA needs to implement an NLP data pipeline to extract sufficient data for comparative analysis and advanced clustering techniques?

- Strongly Needed
- Needed

- Neutral
- Slightly Needed
- Not Needed at All

10**:** To what extent do you agree that implementing an NLP data pipeline to extract variables for comparative analysis to set EQA benchmarks would be beneficial for clearly revealing school performance characteristics?

- Strongly Disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

11**:** To what extent do you agree that implementing an NLP data pipeline to extract variables for comparative analysis to set EQA benchmarks could enhance regulatory benchmarking by providing a performance scoreboard that encourages continuous improvement and allows schools to compare themselves with high performers?

- Strongly Disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

12: Do you have any suggestions for improving EQA comparative analysis and visualization for benchmarking? (open-ended question)

Section 3: Evaluation of Data Extraction Potential and Relevance of Assessment Artifacts

Instructions: In this section, please review and evaluate the identified artifact forms used in the external quality assessment process. Your task is to **identify relevant forms** by determining which data or variables in which form are valuable for representing school performance across different standards. Secondly, **to categorize data by standard,** you will specify which data or variables correspond to the following standards: Student Outcomes (Standard 1) Facilities & School Management Effectiveness (Standard 2) and Teaching Approach Effectiveness (Standard 3) Lastly, to **identify and evaluate key variables** from the existing EQA forms that could be used for setting EQA benchmarks. Your feedback will help us understand how effectively these forms support meaningful analysis and benchmarking, and how they can be improved to enhance the assessment process.

1. Which EQA forms contain valuable data or variables for representing overall school performance? Please list the forms and describe the data or variables they contain.

2. Which forms include data or variables related to Student Outcomes? Please list these forms and specify the relevant data.

3. Which forms include data or variables related to Facilities & School Management Effectiveness? Please list these forms and specify the relevant data.

4. Which forms include data or variables related to Teaching Approach Effectiveness? Please list these forms and specify the relevant data.

5. From the existing EQA forms, which variables do you believe are key for setting EQA benchmarks? Please list these variables and specify which forms they are found in.

6. Are there any additional variables that you think should be included or improved upon in the EQA forms to better support benchmark setting? Please provide details.

3. Official Request to ONESQA for Access to EQA Artifacts and Survey Authorization

รับเข้า

เลขที่รับ 729/2567
วันที่ 28 ส.ค. 67
เวลา 17:53 น.

บันทึกข้อความ

สำนักงานรับรองมาตรฐานและประเมินคุณภาพการศึกษา (องค์การมหาชน)
สำนักนโยบายและยุทธศาสตร์
ภารกิจวิจัยและจัดการความรู้ โทร.229

ที่ ภวจ. ๕๖/๒๕๖๗ **วันที่** ๒๗ สิงหาคม ๒๕๖๗
เรื่อง ขออนุญาตเก็บข้อมูลเพื่อทำการวิจัย

เรียน รักษาการผู้อำนวยการ (ดร.นันทา หงวนตัด)

เนื่องด้วย ข้าพเจ้า นางสาวปุญญิศา ภูมิผล นักวิชาการภารกิจวิจัยและจัดการความรู้ เป็นผู้ได้รับทุน กพ. ให้ศึกษาต่อในระดับปริญญาโท มหาวิทยาลัย Berlin School of Business and Innovation ซึ่งขณะนี้อยู่ระหว่างการดำเนินการวิจัยเรื่อง Enhancing Educational Quality Assessment in Thailand: Leveraging NLP and Machine Learning for Effective Benchmarking and Decision-Making โดยมีวัตถุประสงค์หลัก ๒ ข้อ คือ ๑. เพื่อพัฒนาโมเดล Natural Processing Language (NLP) ในการสกัดข้อมูลผลการประเมินคุณภาพการศึกษาจากไฟล์ PDF และ ๒. นำเสนอผลการวิเคราะห์ในลักษณะการเทียบเคียงกับคู่เทียบ (Benchmarking)

ในการนี้ จึงขออนุญาตเก็บข้อมูลเพื่อทำการวิจัยดังกล่าว ประกอบด้วย

- ๑. การสำรวจความคิดเห็นของผู้บริหารและเจ้าหน้าที่ สมศ. สำนักประเมินและรับรอง สำนักพัฒนาและส่งเสริม และสำนักนโยบายและยุทธศาสตร์ ที่มีต่อการประเมินคุณภาพภายนอกรูปแบบต่างๆ เนื่องจากเป็นผู้เชี่ยวชาญด้านการวัดและประเมินผลการศึกษา อีกทั้งยังมีส่วนในการนำผลการวิเคราะห์ไปใช้ จึงสามารถให้ความคิดเห็นและข้อเสนอแนะที่เป็นประโยชน์ต่อการพัฒนารูปแบบการวิเคราะห์ผลการประเมินคุณภาพภายนอกได้เป็นอย่างดี (แบบสำรวจฯ สามารถเข้าถึงได้ดังลิ้งนี้ https://forms.gle/yhdnVU๙๙nN๒uq๑๖g๗)
- ๒. การเข้าถึงข้อมูลรายงานการประเมินคุณภาพภายนอกของสถานศึกษาที่ได้รับการประเมินคุณภาพภายนอกรอบสี่ ประจำปีงบมาณ พ.ศ.๒๕๖๗ ประกอบด้วย แบบฟอร์ม BF-๐๒ และ BF-๐๓ ของสถานศึกษา จำนวน ๑๓๓ แห่ง (รายละเอียด ดังสิ่งที่แนบมาด้วย ๑) ข้อมูลในส่วนนี้จะนำมาใช้เพื่อทดสอบว่าโมเดลที่สร้างมีประสิทธิภาพในการดึงข้อมูลผลการประเมินฯ จากเอกสารที่เป็นไฟล์ PDF ได้ครบถ้วนและถูกต้องมากน้อยเพียงใด
- ๓. การสอบถามข้อมูลเกี่ยวกับระบบฐานข้อมูล Automated QA ของ สมศ. จากภารกิจเทคโนโลยีสารสนเทศและการสื่อสาร ซึ่งข้อมูลในส่วนนี้จะนำมาใช้เพื่ออภิปรายผลการวิจัยในส่วนของการประยุกต์ใช้โมเดลที่พัฒนาขึ้น ในการติดต่อกับฐานข้อมูลและดึงข้อมูลที่ต้องการมาใช้งานได้อย่างมีประสิทธิภาพ ดังนั้นจึงจำเป็นต้องเข้าใจโครงสร้างฐานข้อมูลและวิธีการเข้าถึงข้อมูลภายในระบบของ Automated QA ในปัจจุบัน (รายละเอียดข้อคำถาม ดังลิ้งค์นี้ https://forms.gle/๙NYFyXzAA๕๖Mu๒๑P๖)

จึงเรียนมาเพื่อโปรดพิจารณาให้ความอนุเคราะห์ด้วย จะเป็นพระคุณยิ่ง

(นางสาวปุญญิศา ภูมิผล)
นักวิชาการ
#๒๒๙ M.๐๙๐๙๐๑๓๗๗๐

๒๗ ส.ค. ๖๗ เวลา ๑๖:๕๐:๐๐ Non PKI Server Sign
Signature Code : RAA๒A DgANw BEADM AMAA๕

ความเห็น/ข้อสั่งการที่ ๑

เรียน รักษาการผู้อำนวยการ (ดร.นันทา หงวนตัด)

เนื่องด้วย นางสาวปุญญิศา ภูมิผล นักวิชาการ ภวจ. ผู้ได้รับทุน กพ. ให้ศึกษาต่อในระดับปริญญาโท ซึ่งขณะนี้อยู่ระหว่างการดำเนินการวิจัยเรื่อง Enhancing Educational Quality Assessment in Thailand: Leveraging NLP and Machine Learning for Effective Benchmarking and Decision-Making นั้น

ในการนี้ นางสาวปุญญิศา ภูมิผล ขออนุญาตเก็บข้อมูลเพื่อการวิจัยกับผู้บริหารและ จนท.ในสำนักงาน (สปร. สพส. สนย. และ ภทส.) ดังลิ้งค์ที่แนบ

จึงเรียนมาเพื่อโปรดพิจารณาอนุญาตให้เก็บข้อมูลเพื่อการวิจัยดังกล่าวด้วย จะขอบคุณยิ่งค่ะ

(นางอรนิศา เพชรผล)
หัวหน้าภารกิจวิจัยและจัดการความรู้
#๒๒๘ M.๐๙๒๘๖๘๒๕๖๔
๒๗ ส.ค. ๖๗ เวลา ๑๗:๕๘:๓๖ , Non PKI Server Sign , Signature Code : MQA๓A EMAMA AxAEM AQwA๐

ความเห็น/ข้อสั่งการที่ ๒

เรียน รักษาการผู้อำนวยการ (ดร.นันทา หงวนตัด)

ตามเรื่องที่เสนอ จึงเรียนมาเพื่อโปรดอนุญาตให้นางสาวปุญญิศา ภูมิผล เก็บข้อมูลเพื่อการวิจัยกับผู้บริหารและ จนท. ในสำนักงาน (สปร. สพส. สนย. และ ภทส.) ตามรายละเอียดที่เสนอข้างต้น ทั้งนี้เพื่อเป็นประโยชน์ต่อการนำผลการศึกษาวิจัยดังกล่าวมาประยุกต์ใช้ในการดำเนินงานของ สมศ. ต่อไป

(ดร.สมยศ ชี้แจง)
หัวหน้าสำนักนโยบายและยุทธศาสตร์
#๑๖๙ M.๐๘๙๘๐๘๐๒๒๔
๒๗ ส.ค. ๖๗ เวลา ๑๙:๒๖:๔๖ , Non PKI Server Sign , Signature Code : RgBCA DcAMA A๔ADc AMwA๒

ความเห็น/ข้อสั่งการที่ ๓

เห็นควรอนุมัติ

(นาวาตรีหญิง ดร.กิตติยา เอ็ฟฟานส)
รองผู้อำนวยการ
#๑๙๔ M.๐๙๒๐๕๗๘๕๑๑
๒๘ ส.ค. ๖๗ เวลา ๑๓:๒๖:๔๘ , Non PKI Server Sign , Signature Code : RAAyA DcAMA A๐ADQ ARABC

ความเห็น/ข้อสั่งการที่ ๔

อนุมัติ ตามรายละเอียดที่เสนอ และมอบหมายให้ผู้รับผิดชอบ ดำเนินการในส่วนที่เกี่ยวข้อง ให้ถูกต้องตามกฎหมาย ระเบียบ ข้อบังคับหรือประกาศที่กำหนดไว้อย่างเคร่งครัด ต่อไป

(ดร.นันทา หงวนตัด)

รักษาการผู้อำนวยการ

#๑๙๓ M.๐๘๑๖๔๗๖๖๙๙

๒๘ ส.ค. ๖๗ เวลา ๑๗:๔๙:๓๖ , Non PKI Server Sign , Signature Code : RAA๑A DUANw AyAEI ANgA๐

ความเห็น/ข้อสั่งการที่ ๕

รับทราบ

(ดร.ณัฐพล ธิตินานันทกูล)

หัวหน้าภารกิจประเมินและรับรองสถานศึกษาด้านการอาชีวศึกษา

#๒๗๐ M.๐๖๒๕๑๔๔๒๒๒

๒๘ ส.ค. ๖๗ เวลา ๑๙:๒๓:๔๓ , Non PKI Server Sign , Signature Code : QQA๕A DkANg AxAEE ARgBG

ความเห็น/ข้อสั่งการที่ ๖

รับทราบ

(ดร.พัฒนพงษ์ ทองเนื้อสุก)

นักวิชาการ

#๒๓๐ M.๐๙๕๘๒๘๘๘๙๙

๒๙ ส.ค. ๖๗ เวลา ๐๐:๐๕:๓๑ , Non PKI Server Sign , Signature Code : OQBCA DkAQw A๕AEY AQwAz

ความเห็น/ข้อสั่งการที่ ๗

รับทราบครับ

(ดร.สมยศ ชี้แจง)

หัวหน้าสำนักนโยบายและยุทธศาสตร์

#๑๖๙ M.๐๘๙๘๐๘๐๒๒๔

๒๙ ส.ค. ๖๗ เวลา ๐๖:๓๖:๔๗ , Non PKI Server Sign , Signature Code : MQAwA DAANQ BFADc ARABF

ความเห็น/ข้อสั่งการที่ ๘

รับทราบค่ะ

ภารดี เจียรนัยกูร

(นางสาวภารดี เจียรนัยกูร)

หัวหน้าภารกิจประเมินและรับรองสถานศึกษาระดับการศึกษาขั้นพื้นฐาน ๒

#๒๖๐ M.๐๘๔๑๓๐๒๖๖๕

๒๙ ส.ค. ๖๗ เวลา ๐๖:๕๔:๐๑ , Non PKI Server Sign , Signature Code : MwA๐A DIANg BFADM AMwAz

ความเห็น/ข้อสั่งการที่ ๙

รับทราบ และดำเนินการแสดงความคิดเห็นทางแบบฟอร์มที่กำหนดแล้ว ขอบคุณมากครับ

(นายชัยมงคล พุกสุวรรณ)

นักวิชาการภารกิจพัฒนาระบบประเมิน

#๑๔๓

๒๙ ส.ค. ๖๗ เวลา ๐๗:๒๖:๑๗ , Non PKI Server Sign , Signature Code : NQBBA DMAOQ A๒ADQ ARQA๔

ความเห็น/ข้อสั่งการที่

๑๐

รับทราบ

(นางวีรนุช สุขสว่าง)

หัวหน้าภารกิจพัฒนานโยบายและยุทธศาสตร์

#๒๒๒ M.๐๖๑๖๙๓๖๕๖๒

๒๙ ส.ค. ๖๗ เวลา ๐๘:๐๐:๕๕ , Non PKI Server Sign , Signature Code : NwA๑A DYAQg BEAEQ ANwA๒

ความเห็น/ข้อสั่งการที่

๑๑

รับทราบ

(นาวาตรีหญิง ดร.กิตติยา เอ็ฟฟานส)

รองผู้อำนวยการ

#๑๙๔ M.๐๙๒๐๕๗๘๕๑๑

๒๙ ส.ค. ๖๗ เวลา ๐๘:๒๒:๑๖ , Non PKI Server Sign , Signature Code : MwA๔A DIAMw AyAEE AMgA๑

ความเห็น/ข้อสั่งการที่

๑๒

รับทราบ

(ดร.วรวิชช ภาสาวสุวัศ)

รองผู้อำนวยการ

#๑๙๖ M.๐๙๔๕๕๑๘๓๘๘

๒๙ ส.ค. ๖๗ เวลา ๐๙:๓๓:๒๖ , Non PKI Server Sign , Signature Code : MwBGA EYANQ A๒ADU AOQBF

ความเห็น/ข้อสั่งการที่

๑๓

รับทราบค่ะ

(นางสาวมาลัย ศรีศิริรุ่งโรจน์)
หัวหน้าสำนักพัฒนาและส่งเสริม
#๑๕๙
๒๙ ส.ค. ๖๗ เวลา ๐๙:๕๔:๔๔ , Non PKI Server Sign , Signature Code : NwAyA DcAMw BBADA AOQA๔

ความเห็น/ข้อสั่งการที่
๑๔

เรียน หัวหน้างานระบบการประเมิน

เพื่อโปรดพิจารณาดำเนินการอนุเคราะห์ข้อ ๒ และข้อ ๓ (ดังรายละเอียดข้างต้น) ต่อไป

(นายสมพล จารุธนศักดิ์กูร)
หัวหน้าภารกิจเทคโนโลยีสารสนเทศและการสื่อสาร
#๑๗๐ M.๐๘๙๑๑๑๖๗๙
๒๙ ส.ค. ๖๗ เวลา ๑๓:๑๐:๔๗ , Non PKI Server Sign , Signature Code : OQAwA EIARQ A๐ADQ AMgA๑

ความเห็น/ข้อสั่งการที่
๑๕

รับทราบค่ะ

(ดร.มณีรัตน์ จันทนา)
ประธานกรรมการ
#๑๕๐ M.๐๘๑๘๒๕๙๗๖๘
๒๙ ส.ค. ๖๗ เวลา ๑๓:๒๓:๐๓ , Non PKI Server Sign , Signature Code : MwA๐A DYAMw A๔ADc AQgAz

ความเห็น/ข้อสั่งการที่
๑๖

รับทราบค่ะ

(นางสาวพรรณทิพย์ หนุนวงษ์)
หัวหน้างานนโยบายและแผน
#๒๑๘
๒๙ ส.ค. ๖๗ เวลา ๑๓:๒๕:๓๕ , Non PKI Server Sign , Signature Code : RQBDA EMANA AyADA AOQAz

ความเห็น/ข้อสั่งการที่
๑๗

รับทราบครับ

(นายกีรติ ฆารเจริญ)
หัวหน้าภารกิจประเมินและรับรองสถานศึกษาระดับการศึกษาขั้นพื้นฐาน๑
#๒๕๐
๒๙ ส.ค. ๖๗ เวลา ๑๓:๔๘:๕๒ , Non PKI Server Sign , Signature Code : NAA๔A EMARA BEAEI AOQA๐

ความเห็น/ข้อสั่งการที่
๑๘

รับทราบ และดำเนินการแสดงความคิดเห็นและข้อเสนอแนะทางแบบฟอร์มที่แจ้งไว้เรียบร้อยแล้วครับ

(นายธนัญชัย ไชยหงษ์)
นักวิชาการ
#๒๖๕
๒๙ ส.ค. ๖๗ เวลา ๑๔:๒๙:๒๓ , Non PKI Server Sign , Signature Code : NgA๓A EMAMQ A๐ADE AMwA๒

ความเห็น/ข้อสั่งการที่
๑๙

รับทราบ และแสดงความคิดเห็นตามแบบฟอร์มเรียบร้อยแล้วค่ะ

(นางสาวพรวลัย เบญจรัตนสิริโชติ)
นักวิชาการ
#๒๖๒ M.๐๘๘๒๐๕๕๔๗๗
๒๙ ส.ค. ๖๗ เวลา ๑๖:๐๔:๒๖ , Non PKI Server Sign , Signature Code : MAAxA DIARA AzADc AOQA๒

ความเห็น/ข้อสั่งการที่
๒๐

ดำเนินการกรอกข้อมูลเรียบร้อยแล้ว

(นายธนภัทร สัมพันธ์รัตนชัย)
หัวหน้างานระบบการประเมิน
#๑๗๗ M.๐๙๐๕๙๗๙๐๙๐
๒๙ ส.ค. ๖๗ เวลา ๑๖:๕๕:๑๔ , Non PKI Server Sign , Signature Code : MgA๕A EUAQQ BFADk ARgBE

ความเห็น/ข้อสั่งการที่
๒๑

ทราบ

(ดร.นันทา หงวนตัด)

รักษาการผู้อำนวยการ

#๑๙๓ M.๐๘๑๖๔๗๖๖๙๙

๒๙ ส.ค. ๖๗ เวลา ๑๘:๒๖:๓๔ , Non PKI Server Sign , Signature Code : RgAxA DkARg A๕ADI AMwBD

ความเห็น/ข้อสั่งการที่

๒๒

รับทราบและดำเนินการแสดงความคิดเห็นตามแบบฟอร์มเรียบร้อยแล้วค่ะ

(ดร.อมรทิพย์ สันตวิริยะพันธุ์)

นักวิชาการ

#๑๔๐

๐๒ ก.ย. ๖๗ เวลา ๑๐:๒๒:๔๓ , Non PKI Server Sign , Signature Code : MwAyA EQARg AwADk AQQA๒