

The Design and Optimization of a Scalable National Product Information Management System
for the NHS

A thesis submitted in partial fulfillment of the University's requirements for the
degree of

Master of Science

in

Data Science

By

Ayokunle James Olagunju

2217353

to the

School of Mathematics and Computer Science

University of Wolverhampton

Supervised by: Dr Ingo Frommholz

May 10, 2024

Abstract

- Concise summary of the thesis objectives, methods, findings, and conclusions.

Acknowledgments:

- Gratitude towards individuals, institutions, or organizations that contributed to my research or supported during the research process.

Special recognition goes to Afrobeat music for keeping me awake throughout the process of writing this thesis.

Table of Contents

Abstract	2
Acknowledgments:	3
List of Figures	1
Chapter 1	2
Introduction	2
1.1 Introduction to research	2
1.2 Overview of the research problem	2
1.4 Problem Statement	3
1.5 Motivation for research	3
1.6 Aims and Objectives of research	4
1.6.1 Research Aims.....	4
1.6.2 Research Objectives	4
1.7 Research Questions	4
1.8 Significance of the study	5
1.9 Structure of the Thesis	5
Chapter 2	7
Literature Review	7
2.1 Introduction to literature review	7
2.2 Product Information Management	7
2.2.1 What is Product information?.....	7
2.2.2 Product Information management	7
2.3 Background of Product Information Management in the NHS: As-Is	8
2.4 Centralization of Product Information Management in the NHS: To-Be	10
2.5 Product Information Management (PIM) Systems	14
2.5.1 Benefits of a PIM System.....	14
2.5.2 Benefits of a PIM System to the NHS	15
2.5.3 Product Information Management (PIM) vs similar Information Systems.....	15

2.6 Information Management System Development Life Cycle.....	17
2.7 Information Management System Design Tools and Technologies.....	18
2.7.1 Databases and Database Management Systems	18
2.7.2 Advantages and Disadvantages of Database Management Systems	19
2.7.3 Types of Databases.....	19
2.7.4 Database System environment	20
2.8 Optimization strategies for relational database systems	21
2.9 Case Studies and Related Works	23
2.10 Summary of key findings and gaps in the literature.....	23
<i>Chapter 3</i>	25
<i>Methodology</i>	25
3.1 Introduction to Methodology	25
3.2 Research Design.....	25
3.3 Planning Phase.....	25
3.4 Requirements Analysis.....	26
3.5 Design Phase	27
3.6 Development and Evaluation	27
3.6.1 System Development.....	28
3.6.2 Hardware Requirements	28
3.6.3 Software Requirements	28
3.6.4 System Evaluation	29
3.7 Data Collection Methods.....	30
3.7.1 Qualitative data.....	30
3.7.2 Quantitative data.....	30
3.8 Ethical Considerations.....	31
3.9 Limitations and Assumptions.....	31
3.10 Summary of research methodology	31
<i>Chapter 4</i>	33
<i>Requirements Analysis.....</i>	33

4.1 Introduction to Requirements Analysis	33
4.2 Identification of Stakeholders and Information Needs Assessment	33
4.4 System Requirements Analysis	35
4.4.1 Functional Requirements	35
4.4.2 Non-functional Requirements	35
4.5 Legal and Regulatory Frameworks	36
4.5.1 Relevant legal and regulatory compliance standards and obligations	36
4.5.2 Influence of legal and regulatory considerations on the design and implementation of the NHS PIM system.....	37
4.6 Summary of Requirements Analysis	39
<i>Chapter 5</i>	<i>41</i>
<i>Design and Modelling.....</i>	<i>41</i>
5.1 Introduction to system design	41
5.2 Database Design & Modelling.....	41
5.2.1 Business Rules.....	41
5.2.2 Entities and Attributes	42
5.2.3 Relationships	44
5.2.4 Entity Relationship Models	47
5.3 Database Normalization.....	49
5.3.1 The Normalization Process.....	49
5.3 Conceptual Design of User View and Interface Layout.....	55
5.4 Summary of System Design	58
<i>Chapter 6</i>	<i>59</i>
<i>Implementation and Development.....</i>	<i>59</i>
6.1 Introduction to Implementation.....	59
6.2 System Architecture.....	59
6.3 Database Implementation.....	60
6.3.1 Implementation Steps	60
6.3.2 Limitations of Local Database Server	62
6.3.3 Aiven for MySQL.....	63

6.3.4 Deploying MySQL database to the Aiven for MySQL service	64
6.4 Application Program Development	65
6.4.1 Implementation of Streamlit	66
6.4.2 User Interface (UI).....	67
6.4.3 User View(s).....	68
6.4.4 Limitations of local Application server	73
6.4.5 Application deployment – Streamlit Community Cloud.....	74
6.5 Scalability and Optimization Strategies	76
6.6 Data Security and Integrity Measures.....	77
6.6.1 Data Security	77
6.6.2 Data Integrity	79
6.7 Summary of Implementation	79
<i>Chapter 7</i>	<i>80</i>
<i>System Testing and Evaluation</i>	<i>80</i>
7.1 Introduction to Evaluation	80
7.2 User Testing and Feedback	80
7.3 Evaluation of Database Performance Metrics.....	80
7.4 Summary of System Evaluation.....	82
Chapter 8: Discussion and Conclusion	83
8.1 Interpretation of the findings in relation to the research objectives.	83
8.2 Discussion of implications, limitations, and future directions.	83
8.3 Summary of key findings and contributions of the thesis.	84
8.4 Recommendations for practice, policy, or further research.....	84
Database Management System for Student Admissions.....	84
An information management system for a large-scale biological collaboration	84
Web-based forensic information management system	85
Development of a Database-Driven Management System for Retail Food Packaging Eye Tracking Studies.....	85
design and development of web based biomedical dbms for health services multispecialty	85

Evaluation Of Performance of NoSQL Database Management System: -MongoDB And Relational Database Management System: - PostgreSQL for Efficient Management Of Healthcare Data.....	85
DESIGN AND IMPLEMENTATION OF A WEB-BASED HEALTH CARE RECORDS MANAGEMENT SYSTEM	
Database System for Medical Record Keeping and Retrieval	85
Implementation of product data management system.....	86
EVALUATION OF DATABASE MANAGEMENT SYSTEMS – PostgreSQL better than MySQL & SQLite in terms of speed and info retrieval	
Design and Implementation of a Web Shop System	86
<i>Bibliography.....</i>	89
<i>Appendices.....</i>	94

List of Figures

Figure 1: How Trusts can buy products	8
Figure 2: Typical product data management process in the NHS.	10
Figure 3:Framework for centralization of Product Information	11
Figure 4:Proposed NHS GS1 master data exchange network.	13
Figure 5: Information System Development lifecycle	17
Figure 6: Database System Application Processing	18
Figure 7: Summary of Research Methodology	32
Figure 8: Logical ER model for the PIM system.	48
Figure 9:Relational ER model for the PIM system.....	48
Figure 10: PIM System landing page layout.....	55
Figure 11: Conceptual Design of the System's landing page	56
Figure 12: PIM System data access layout	57
Figure 13: Flow chart of 'Search Trade Item' page	57
Figure 14:3-Tier Client-Server architecture.....	60
Figure 15: CREATE TABLE trade_item code snippet	61
Figure 16:MySQL database instance running on Aiven for MySQL	64
Figure 17: NHS_PIM cloud database connection credentials	64
Figure 18: MySQL Workbench connection to Aiven for MySQL service.....	65
Figure 19: NHS PIM System User Interface	67
Figure 20: Search trade Items page of the PIM system	68
Figure 21: Search Trade Items by GTIN	69
Figure 22: Search Trade Items by Supplier GLN	69
Figure 23: Search Trade Items by Manufacturer GLN	70
Figure 24: Search Trade Items by Provider GLN	71
Figure 25: 'Medical Devices' Screen and functionalities	72
Figure 26: 'Suppliers' Screen and functionalities	72
Figure 27: Database Connection Credential Secrets in TOML format.....	75
Figure 28:Python code establishing connection to MySQL database	75
Figure 29: Indexes on the NHS PIM database	76

Figure 30: NHS PIM system access control and privileges in MySQL	78
Figure 31: CPU usage	81
Figure 32: Disk space usage	81
Figure 33: Memory Usage	81
Figure 34: Disk IOPS (reads).....	82
Figure 35: Network receive rate	82
Figure 36: Network transmit rate	82

Chapter 1

Introduction

1.1 Introduction to research

This chapter contains an introduction to the area of focus of this thesis project. It focuses on the main research problem, the motivation for embarking on this research,, research aims and objectives, and the research questions that guide this study. The final section of this chapter presents the overall organizational structure of the thesis.

1.2 Overview of the research problem

In modern day healthcare systems, efficient and standardized procurement and supply chain processes are necessary in ensuring availability of medical supplies, medicines, devices, and equipment essential for the delivery of high-quality care (Steer-Stephenson, 2022). It is important to accurately manage healthcare product data because these products have a direct impact on patient health and safety. It is critical to enable the supply chain to deliver the right products to the right place at the right time (Department of Health & Social Care, 2023).

The primary healthcare provider in the United Kingdom, the National Health Service (NHS), serving millions of patients across the UK, is plagued with challenges in effectively managing

the product information sharing process between its suppliers and its vast network of trusts, hospitals, clinics, and healthcare facilities. Due to the absence of a unified and scalable system, product information management within the NHS is characterized by lack of standardization, inconsistency, inefficiency, and fragmentation, with disparate infrastructure and processes essentially leading to difficulties accessing accurate and up-to-date product information (Procurement, Investment & Commercial Division, DHSC, 2014). This leads to delays in procurement, disruptions in the supply chain, and poor decision-making, ultimately affecting patient care and outcomes. Hence, it is imperative to address these pressing challenges and improve the management of product information and data sharing within the NHS' procurement ecosystem.

1.4 Problem Statement

The NHS is faced with challenges in managing medical product information across its supply chain. These challenges include but are not limited to, a glaring inefficient data sharing process between manufacturers/suppliers and NHS trusts, a lack of standardized data formats and identifiers, disparate information scattered across repositories in different departments and healthcare trusts, manual processes for accessing and updating product information.

Consequently, healthcare practitioners face difficulties in finding the right products at the right time, causing delays in care and inefficient resource allocation.

The absence of a centralized and scalable system for managing product information poses an obstruction to interoperability and efficient exchange of data with external stakeholders e.g. suppliers, regulatory bodies, patients, and other healthcare organizations. This lack of integration and interoperability brings to fore, the challenges the NHS faces in maintaining accurate and reliable product information throughout its procurement lifecycle.

1.5 Motivation for research

This research is primarily motivated by the pressing need to address the challenges the NHS faces in product information management within its procurement process and catalogue management system. Efficient management of the supply chain is essential for an effective and functioning healthcare system. For a healthcare system such as the NHS, where resources are stretched thin, optimizing the supply chain system is essential for ensuring access to products, medical supplies, and equipment.

The implementation of a scalable national product information management system will enhance supply chain efficiency, ensure patient safety, and care quality, meet regulatory requirements, foster collaboration, and drive innovation within the UK's health system.

1.6 Aims and Objectives of research.

1.6.1 Research Aims

The primary aim of this research is:

1. To design and optimize a scalable national product information management system for the NHS.

1.6.2 Research Objectives

To achieve the project's aim, the following objectives were set:

1. To analyse the existing product information management practices within the National Health Service (NHS) procurement ecosystem, identifying key challenges and assessing the needs and requirements of stakeholders.
2. To design the conceptual, logical, and physical database frameworks for a scalable national product information management (PIM) system tailored to the needs and requirements of the NHS.
3. To develop a prototype of the proposed PIM system, leveraging advanced relational database management technologies and methodologies to ensure scalability, optimization, and usability in a real-world NHS setting.
4. To evaluate the developed PIM system application through testing, user feedback and performance monitoring.
5. To provide recommendations and guidelines for the implementation, adoption, and improvement of the national product information management system within the NHS.

1.7 Research Questions

This research will be guided by the following questions:

1. What are the key challenges faced by the NHS in managing product information within its procurement processes?

2. How can the implementation of a national product information management system improve supply chain efficiency & product data management within the NHS?
3. What are the essential features and functionalities required in a scalable product information management system tailored to the needs of the NHS?

1.8 Significance of the study.

The significance of a study on the design and optimization of a scalable national product information management (PIM) system for the NHS cannot be overemphasized. By streamlining the supply chain and hospital catalogue management systems through the implementation of a national PIM system, the NHS can increase its efficiency, minimize administrative burdens, and improve the overall supply chain process. A centralized and up-to-date product information management system will provide healthcare practitioners with reliable information about medical devices and products thereby minimizing the risk of errors, and further ensuring patient safety.

Furthermore, a scalable national product information system has the potential to ensure optimization of the supply chain by enhancing interoperability and collaboration between healthcare providers, suppliers, and other stakeholders critical to ensuring the delivery of service within the NHS. A seamless data exchange and real-time access to accurate product information will foster transparency and accountability across the procurement ecosystem.

Additionally, embracing the implementation of innovative technologies and digital solutions such as a product information management system can help the NHS leverage opportunities in advanced analytics, artificial intelligence, and automation to optimize the procurement process, identify cost saving opportunities, and position itself at the forefront of healthcare innovation. Overall, this study has the potential to revolutionize the healthcare supply chain and information management process, improve patient outcomes, and advance healthcare delivery not just within the NHS but globally.

1.9 Structure of the Thesis

This section explains the organizational structure of this thesis. This thesis is organized into 7 chapters, each one serving a specific purpose to providing an overall understanding of the work. Chapter 1 introduces the study. It serves as a foundational framework for the study, by providing

context, outlining the background, the research problem, aims & objectives, significance of study and structure.

Chapter 2 presents a comprehensive review of relevant literature to healthcare procurement, and product information management. It discusses key concepts, methodologies, and findings in this field, highlighting existing gaps in the literature which this research seeks to address.

Chapter 3 outlines the research methodology employed for the design, development, and evaluation of the proposed system.

Chapter 4 presents a thorough assessment of the needs and requirements for the design of a product information management system for the NHS. It employs the mixed-methods approach to collecting data from primary and secondary sources on the information needs of the relevant stakeholders within the NHS procurement ecosystem.

Chapter 5 discusses the design and modelling of a conceptual framework for the proposed product information management system, including the discussion of relevant entities, database normalization process, optimization strategies employed, and conceptual design of the user interface.

Chapter 6 focuses on the implementation and development of the system. It discusses the implementation of the database system, the user interface, limitations encountered and how they were mitigated.

Chapter 7 discusses the evaluation of the developed product information management system. Specifically, it discusses the evaluation criteria employed, and the evaluation of the system in comparison with set objectives.

In conclusion, chapter 8 discusses the result of the research, draws conclusion based on the results, and presents recommendations for future research.

Chapter 2

Literature Review

2.1 Introduction to literature review

In today's dynamic healthcare landscape, efficient information management systems and procurement practices are essential for ensuring access to medical products, optimizing resource allocation, and improving care outcomes. Though there is limited literature in recent times on the design and implementation of a product information management system especially for a healthcare organization such as the NHS, this literature review aims to explore existing literature on the design, optimization and application of Product Information Management (PIM) systems, the As-Is To-Be Analysis of the PIM system scoping framework, which is a gap analysis activity which aims to provide an understanding and assessment of the processes and systems currently in place in the organization including challenges, and the proposed future systems and processes (Battistello, 2020). This chapter aims to inform the research objectives of this study by providing a thorough understanding of the current state of knowledge by synthesizing key concepts, methodologies, and findings from relevant sources.

2.2 Product Information Management

2.2.1 What is Product information?

Product Information is any information about a product which a client or customer uses to make an informed decision about purchasing a product. (Palmer, 2024)

2.2.2 Product Information management

The concept of Product Information Management (PIM) began relatively circa 2003 (Abraham, 2014). Product information management is sometimes referred to as Product Data Management (Vedapudi, 2000).

So, what is Product Information Management (PIM)? To put it simply, PIM is the management of Product information. To further expatiate on this definition, PIM may be defined as the

processes and technologies set up to manage product information in one shared place – “a single source of truth”, to further distribute that information into different systems without having to manually re-enter it. (Abraham, 2014).

2.3 Background of Product Information Management in the NHS: As-Is

Prior to the set-up of the new NHS Supply Chain in 2019, NHS procurement activities were done through a fully outsourced operating model also called NHS Supply Chain. The new NHS Supply Chain was set up following the Lord Carter report which cited wide variations in products and supplier data across NHS trusts (Carter, 2016). There are several ways for NHS trusts to buy its products as expressed by (Davies, 2024) in Figure 1 below.

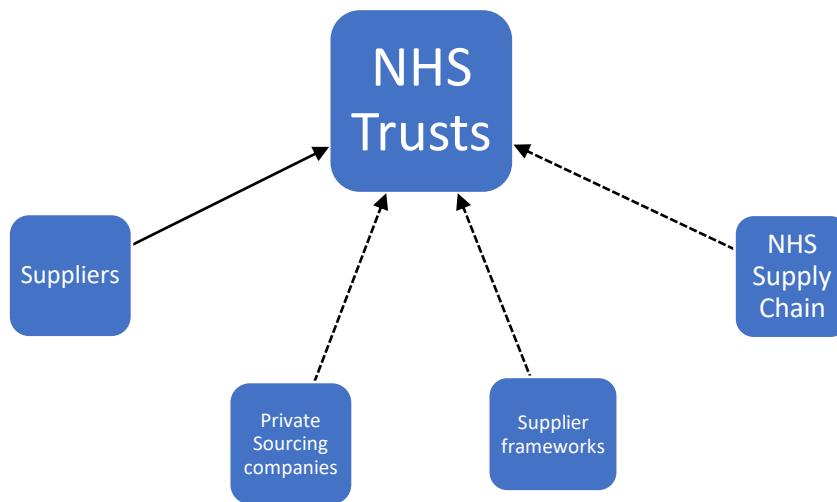


Figure 1: How Trusts can buy products

Under the new operating model, NHS trusts may buy medical products through the NHS Supply Chain, or directly from suppliers. Although nearly all trusts buy at least some specialized high-value products through the NHS Supply Chain.

The use of master data across the supply chain is necessary for effective supply chain management, especially in long and complex supply chains such as the NHS'. In addition to accurate analysis of expenditure, hospital catalogue management, and requisition exchange, the use of master product data is critical for patient care management in terms of decision making, and product traceability (Procurement, Investment & Commercial Division, DHSC, 2014).

Unfortunately, the use of master product data is very limited within the NHS, from manufacturer to patient. Suppliers of products and medical supplies to the NHS respond to multiple requests for master product data by several NHS providers who use their supplied products as shown in Figure 2 below (Commercial Division, DHSC, 2017). This information is often shared manually (e.g. emailing Excel files) which means it can easily become wrong, out-of-date, or inconsistent with other sources. This means that the same medical items are described and coded differently by NHS providers and suppliers (Procurement, Investment & Commercial Division, DHSC, 2014). The application of NHS eClass, the primary classification standard currently used by teams within the NHS, is inconsistent across the NHS and is not up to the global standards of product classification. (Procurement, Investment & Commercial Division, DHSC, 2014)

Historically, product information management within trusts has been plagued with several issues. These issues are outlined in a report by (Department of Health & Social Care, 2018):

- i. Out-of-date pricing.
- ii. Manual data entry at trust level.
- iii. Product descriptions being designed by trusts and not recognized by suppliers.
- iv. Instances of incorrect data including product codes.
- v. Product data being held in various systems within a given trust, thus creating several versions of the 'truth'.

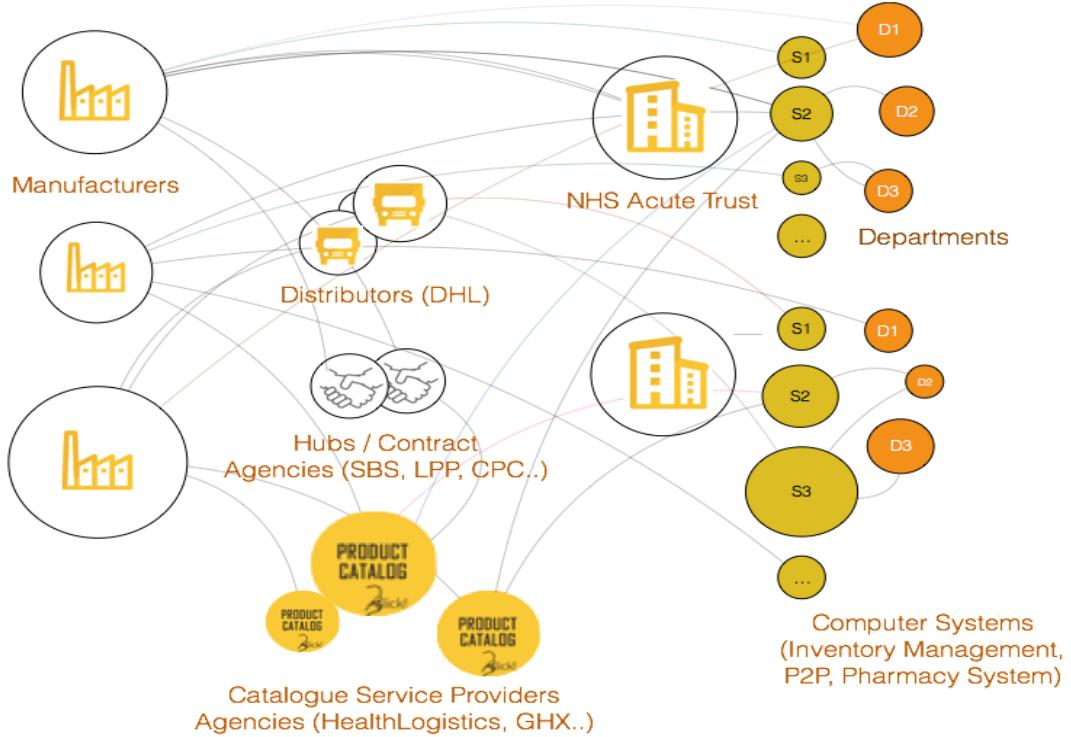


Figure 2: Typical product data management process in the NHS.

Source: Department of Health & Social Care

This inefficient product data management and exchange across the NHS results in duplication of efforts, increased costs, delays, and compromised data quality (Commercial Division, DHSC, 2017).

Introduction of a new product information management system is desperately needed in the NHS, and it is necessary to involve the end users of that technology in its development process. Stakeholder involvement is key to the successful development and implementation of any technology within the NHS (Ahmad, et al., 2012).

2.4 Centralization of Product Information Management in the NHS: To-Be

There are 3 sequential phases for the centralization of product information management (Battistello, 2020):

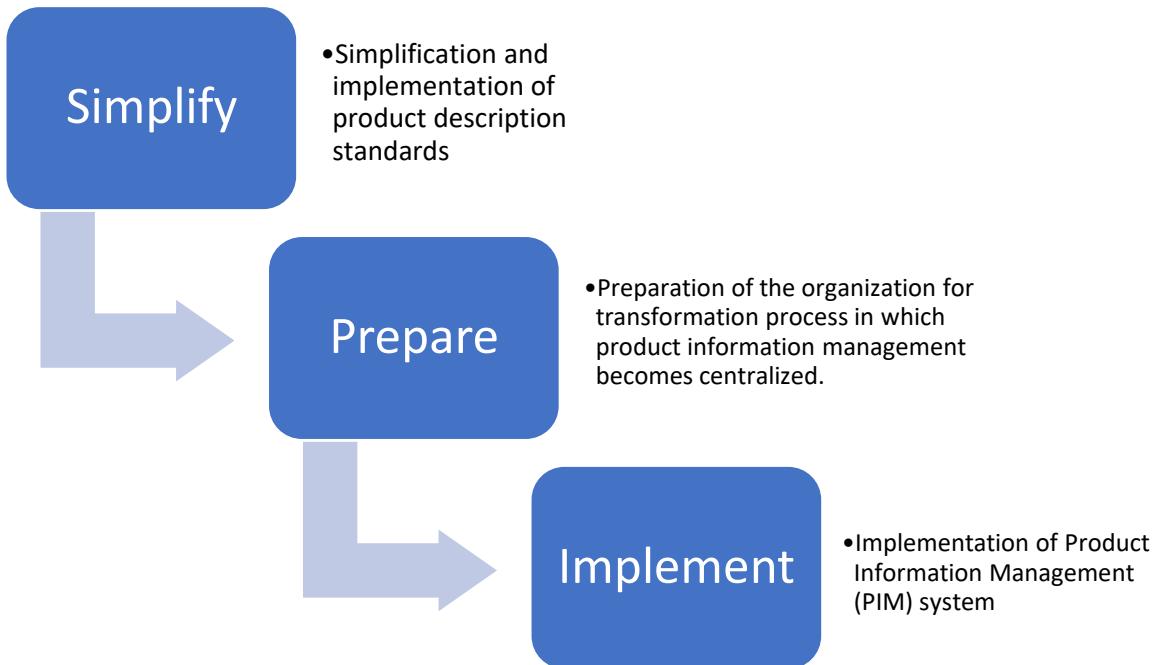


Figure 3: Framework for centralization of Product Information

In this section, the progress of the NHS towards the centralization of product information management is analyzed against these phases.

As the author (Battistello, 2020) proposed, the NHS has begun the sequential phases for the centralization of product information management.

1. Phase 1: Simplification and implementation of product description standards

Upon recognizing the need for, and benefits of improving its systems to improve data quality and data-sharing processes, the NHS launched the *NHS eProcurement Strategy*, to guide the adoption of GS1 standards as the supply chain coding standard by the NHS and its supplier base (Department of Health, 2013). GS1 standards are global coding standards used for the unique identification of products and locations, to enable data synchronisation and end-to-end traceability from manufacturer to patient.

The NHS recognized the adoption of GS1 coding standards by NHS providers and suppliers as the building block for improving data quality and enable interoperability between NHS provider and supplier systems (Procurement, Investment & Commercial Division, DHSC,

2014). The use of GS1 standards will also facilitate the use of Automatic Identification and Data Capture (AIDC) technology, which is used to identify a product at the point of use.

The NHS mandated both NHS providers and suppliers to adopt GS1 standards, including a requirement to place master product data in a GS1-certified data pool, through conditions of contract such as:

- i. the NHS Terms and Conditions for the Supply of Goods and the Provision of Services,
- ii. the 2014/2015 NHS Standard Contract between healthcare commissioners and NHS providers.

This adoption plan was mandated to encompass all goods supplied to the NHS providers, except for medicines which already had an established management system in place.

2. Phase 2: Preparation of the organization for transformation process

The preparation phase involves the organization preparing for the transformation process by establishing an understanding of the benefits of the project. The NHS embarked on a *Demonstration of Technology (DoT)* project to show how the adoption of GS1 can solve the problems of the NHS (Commercial Division, DHSC, 2017).

The main purpose of the DoT was to use the GS1 & Global Data Synchronization Network (GDSN) in a live environment to show the ability to seamlessly facilitate the exchange of master product data and understand the impact of adopting GS1 standards on interoperability between NHS trusts systems and supplier systems, data quality assurance, operating process, scalability from the DoT to an NHS-wide implementation.

DoT methodology

The DoT involved the participation of 6 medical device suppliers, 6 NHS Trusts and their catalogue service providers (CSP), the Department of Health, and GS1. The demonstration followed the following methodology (Commercial Division, DHSC, 2017):

- i. creating a product data dictionary,

- ii. sourcing data for 20 products from select suppliers,
- iii. the suppliers loading the data into a GS1-certified GDSN data pool of their choice,
- iv. publishing the supplier data on the 20 products to an NHS GS1 data pool - *this is done by subscribing to the suppliers' data pools using their GS1 Global Location Number (GLN)*,
- v. data validation,
- vi. maintenance updates, and finally,
- vii. providing the output from the master product data to the catalogues of the participating NHS organizations.

The DoT was successful in showing that the adoption of GS1 & GDSN for implementation of a PIM system for the exchange of master product data between suppliers and NHS providers would result in a more efficient and effective health care and ultimately improve patient safety.

3. Phase 3: Implementation of PIM system

The success of the DoT informed the decision to fully implement an NHS-wide product information management system for the seamless exchange of master product data between suppliers and NHS providers. Figure 4 below shows the proposed framework for the near real-time exchange of master product data between suppliers and NHS providers using the Global Data Synchronization Network (GDSN). The GDSN is a set of global data pools created in compliance with GS1 standards to share structured product data (GS1 UK, 2021).

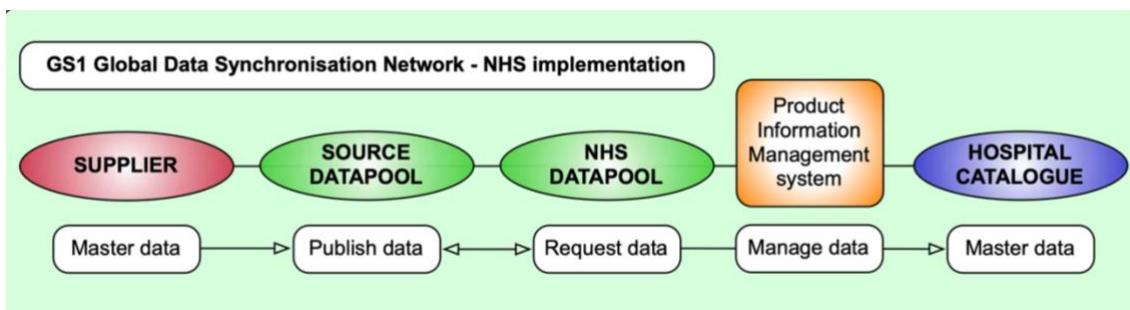


Figure 4:Proposed NHS GS1 master data exchange network.

Source: Department of Health & Social Care

The department of health establishes a GS1-certified data pool for the NHS which will take product data from suppliers' GS1-certified data pools to become the master product data repository for all medical devices and products supplied to the NHS. Manufacturers may load their entire product catalogues in their chosen GS1 data pool. (Procurement, Investment & Commercial Division, DHSC, 2014).

A National Product Information Management (PIM) system will then serve as the national infrastructure for the integration of the NHS data pool with NHS providers, where NHS providers can access master product data, thereby ensuring data integrity, quality, and accuracy across multiple systems.

Research Gap 1: *The lack of a national infrastructure for the management of master product data constitutes the gap for my research.*

2.5 Product Information Management (PIM) Systems

A Product Information Management (PIM) system is an information system used to centrally store, enrich, manage, and distribute product information across several different units of an organization, thus alleviating the need to manually re-enter the data in a different system. (Battistello, et al., 2021). Product Information management systems are necessary for the unification and synchronization of disparate product information.

2.5.1 Benefits of a PIM System

A Product Information Management (PIM) System offers several benefits across various industries. The following are some of the benefits of a PIM system as identified by the authors (Abraham, 2014; Battistello, 2020):

1. Centralized Data Management: a PIM system provides a centralized repository for storing, organizing, and managing product information. This ensures data consistency and accuracy by eliminating errors, duplicates, and/or redundancy.

2. Data Quality Improvement: a PIM system ensures the maintenance of a high data quality, by enforcing standardized data formats and data validation rules. Good quality data ensures reliability and compliance with regulatory requirements.
3. Enhanced Operational Efficiency: a PIM system helps to streamline product information management processes by removing manual workloads, minimizing errors, and improving data accuracy.
4. Increased Productivity: implementing role-based access controls and workflow management capabilities in PIM systems lead to improved productivity and faster decision-making.
5. Legal & Regulatory Compliance: a PIM system helps to ensure compliance with regulations on data privacy protections, standardization, and other data security measures.
6. Scalability: PIM systems offer the ability for organizations to expand product catalogs and support large volumes of product data.

2.5.2 Benefits of a PIM System to the NHS

- a. traceability of medical devices, especially for field safety corrective actions,
- b. adequate identification of medical devices through distribution and use,
- c. identification of medical devices in adverse events,
- d. reduction of medical errors,
- e. documenting and longitudinal capture of data on medical devices.
- f. Effective data retrieval systems

(IMDRF UDI Working Group, 2013)

2.5.3 Product Information Management (PIM) vs similar Information Systems

A Product Information Management (PIM) system is comparable to several other information systems that manage product data (Battistello, 2020). These systems include Product Data Management (PDM), Product Lifecycle Management (PLM) and Master Data Management (MDM).

Product Lifecycle Management (PLM) and Product Data Management (PDM) systems are internal-facing systems which focus on the manufacturing and developmental lifecycle of a

product from ideation till after the product is no longer being sold (for example, product not yet on the market, product discontinued etc.). (Abraham, 2014)

Master Data Management is the comprehensive management and maintenance of master data within an organization. A Master Data Management (MDM) system focuses on providing solutions to problems of data fragmentation, incoherent processes, and disparate systems (Nurminen, 2022).

Master data is the definitive single source of truth for all information which an organization holds about its core entities such as its products, employees, suppliers, accounts etc. (Nurminen, 2022). Master data serves as the foundational data which may be shared and reused by different information systems and business process applications in an organization as a source for accurate reporting, and for reduction of errors and redundancy (Edel & Sutedja, 2023; Pansara, 2021). This is further supported by (Nurminen, 2022) who posited that Master data must be accurate, relevant (proper data attribute selection), timely (data synchronization), complete, and accessible.

Master Data Management (MDM) may be said to be the first step in a Product Information Management (PIM) process (Abraham, 2014). Product Information Management (PIM) is a subset of Master Data Management that deals with product-related information for sales and marketing purposes (Battistello, 2020). In essence, a Master Data Management process must first be implemented to be able to implement an outward-facing Product Information Management system.

Both PIM and MDM systems, to some extent, aim to solve the same product data management challenges. However, where a Product Information Management system focuses on the outward view of Product information for sales & marketing activity, a PMDM focuses on the management of product data and leveraging it to improve business process and decisions that are dependent on product data (Sheldon & Goetz, 2014).

In the context of the NHS, the product data as supplied to the GS1 certified data pools by the suppliers of medical device products forms the basis of **Product Information Management** for sales and marketing from the perspective of the suppliers, while the product data as received by

the NHS Providers through the GDSN forms the basis of **Product Master Data Management** within the NHS.

2.6 Information Management System Development Life Cycle

The lifecycle of an information or database system includes all the steps necessary for the design and implementation of the system. According to (Langer, 2008) , system developments projects must include the following steps, regardless of which design methodology is used:

1. Identifying the need for a business process improvement/support system e.g. a Product Information management system
2. Defining the goals for that system
3. Gathering the business requirements
4. Converting the business requirements to system requirements
5. Designing the system
6. Building, testing, and deploying the system.

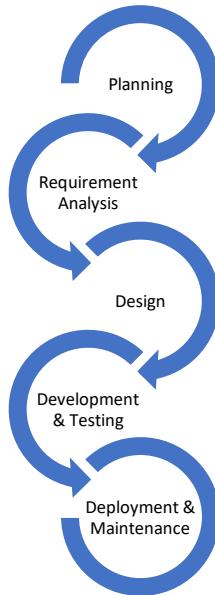


Figure 5: Information System Development lifecycle

These outlined steps will serve as a guide for the execution of this research project in achieving its outlined research objectives.

2.7 Information Management System Design Tools and Technologies

2.7.1 Databases and Database Management Systems

A Database Management System (DBMS) is a software which manages and controls access to a collection of logically related data called a *database*, designed to meet the information needs of an organization. (Connolly & Begg, 2015). A database is said to resemble a well-organized electronic filing cabinet whose content is managed by a powerful software known as the database management system (Coronel & Morris, 2016).

The DBMS is used to define and manipulate the database using Data Definition Language (DDL) and Data Manipulation Language (DML) respectively (Connolly & Begg, 2015). The interaction between a relational DBMS and a database is facilitated by a query language called **Structured Query Language (SQL)**.

An application program is a computer program written in a programming language which is used for interacting with the database by issuing SQL statement requests to the DBMS. An application program could be produced by designing a software application or by using an **application generator** instead (Connolly & Begg, 2015).

The relationship between an application program, a DBMS, a database, and users of the database system is shown in Figure 6 below.

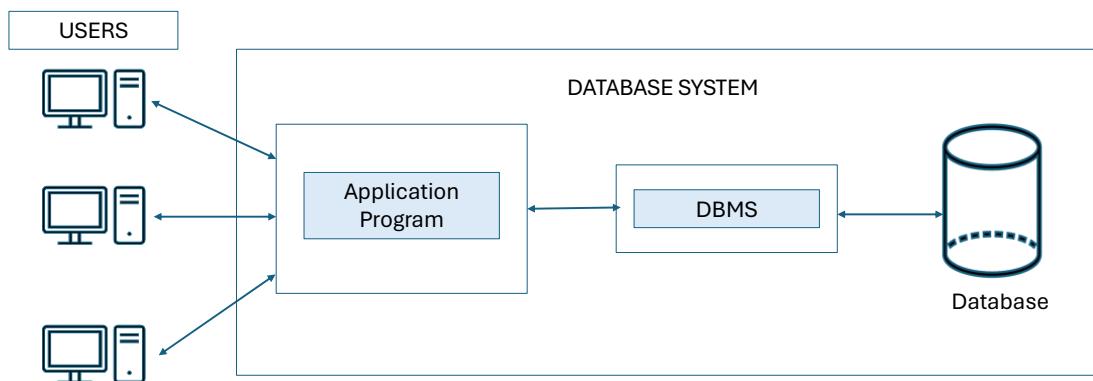


Figure 6: Database System Application Processing

2.7.2 Advantages and Disadvantages of Database Management Systems

The advantages and disadvantages of a database management system are listed in the tables below:

Table 1: Advantages of a DBMS

ADVANTAGES
i. Improved data sharing
ii. Improved data security
iii. Minimized data inconsistency.
iv. Improved data integration
v. Improved data access
vi. Improved decision making
vii. Increased productivity

Table 2: Disadvantages of a DBMS

DISADVANTAGES
i. Complexity of DBMS software
ii. Size
iii. Cost of DBMSs
iv. Additional hardware costs
v. Cost of conversion
vi. Performance
vii. Greater impact of a failure

2.7.3 Types of Databases

There are several methods of classifying databases such as by the number of users, location of data, data type, data usage and data structure. They are summarized by (Coronel & Morris, 2016) as follows:

- i. A **single-user database** (desktop database) supports only one user at a time while a **multiuser database** supports multiple users at the same time.
- ii. A **centralized database** supports data located at a single site while a **distributed database** supports data distributed across different sites.
- iii. A **general-purpose database** contains a variety of data used across several disciplines while a **discipline-specific database** contains data focused on a particular subject area.
- iv. An **operational database** primarily supports a company's day-to-day operations such as sales, payments, purchases, etc., while an **analytical database** focuses on maintaining historical data used for strategic decision making.
- v. A **SQL database** supports the management of structured data while a **NoSQL (Not-only-SQL) database** has emerged in recent years for the management of vast amounts of data

with varying structures (semi-structured, unstructured, structured) and velocity. The choice between SQL and NoSQL databases depends on the needs of the organization and the type of data being managed. Organizations with a need for structured and consistent data will find SQL databases more preferable while organizations with large amounts of unstructured data will prefer NoSQL (Dhasmana, et al., 2023).

Adequate infrastructure is required to support the implementation and operation of databases. The organization typically maintains this infrastructure, but the use of cloud databases has grown in recent years. (Coronel & Morris, 2016). A cloud database is a database created and maintained using cloud data services which provide defined performance measures for the database like speed, scalability, storage capacity (Google Cloud, 2024).

The product information needs of the NHS are structured, relational data, most of which are stored on spreadsheets across different parts of the supply chain. Hence, a **discipline-specific multiuser cloud relational database system (Cloud SQL)** would be best utilized.

2.7.4 Database System environment

A typical database system environment comprises 5 major components. According to (Connolly & Begg, 2015) and (Coronel & Morris, 2016), the major components of the DBMS environment are:

- i. **Hardware:** this refers to all the physical device components of the system such as computers, storage devices, servers, printers, network devices and other devices which the DBMS and its applications require to run.
- ii. **Software:** this consists of the DBMS software itself, the application program(s), and the operating system software on which the hardware components function, allowing all other software to run. Examples of DBMS software include Microsoft SQL Server, Oracle MySQL, and IBM's DB2 while examples of operating system software include Microsoft Windows, Linux, Mac OS etc.
- iii. **Data:** This is the bridge between the machine components and the human components of the DBMS environment. The database contains both the data stored in the database and the metadata.
- iv. **Procedures:** these are the instructions and guidelines that govern the design and use of the database system.

- v. **People:** This includes all the users of the database system. Database users may include database administrators, system analysts, database designers, and end users of the system.

2.8 Optimization strategies for relational database systems

In the context of a relational database system, optimization is the process of improving the performance, efficiency, and reliability of the system (Györödi, et al., 2021).

Two major strategies for optimizing relational database systems have been proposed (Šušter & Ranisavljević, 2023). They include:

1. **Physical programming:** this is the optimization of the physical storage of a database system to reduce the number of Input/Output (I/O) operations required to access data. Some physical programming techniques include:
 - a) **Storage engine:** the performance and scalability of a relational database can be improved by the suitability of the storage engine selected. In MySQL databases, InnoDB storage engine is more suitable for workloads with high concurrency needs because of its support for row-level locking and foreign key constraints which enforce referential integrity between tables (Šušter & Ranisavljević, 2023).
 - b) **Indexing:** Indexing is the act of ordering the content of a data table by making indexes on one or more columns of the data table, reducing the amount of data the system scans through to retrieve data (Maesaroh, et al., 2022). It has been shown that indexing results in a higher performance, with the most significant impact on the reduction of query response times. In relational databases, an index called a *Clustered Index* is automatically created on the primary key (Györödi, et al., 2021). Other indexing strategies for performance optimization in relational databases include *Non-Clustered Indexes*, *Covering Indexes* and *Filter Indexes* (Praveena & Chikkamannur, 2021). SELECT operations are slowed down by too few indexes while UPDATE and INSERT operations may be slowed down by too many indexes (Šušter & Ranisavljević, 2023).
 - c) **Field optimization-based methods:** According to a study by (Györödi, et al., 2021), implementing field optimization-based methods resulted in an 80% faster query response time in SELECT operations. The methods employed include some of the following:
 - i. **Data types:** choosing the right data types help to reduce database storage needs and improve performance. This is supported by (Šušter & Ranisavljević, 2023)

- who suggested the use of TINYINT instead of INT or BIGINT for columns that only store values between 0 and 255 or the use of VARCHAR instead of TEXT for fields which do not need a lot of text data.
- ii. **Null Fields:** reducing the number of nullable fields means the database will be cleaner and the database system can better search the database tables
 - iii. **Field Values:** reducing the possible values of a data field results in less null values and reduce the length of a query making it more efficient.
 - iv. **Field length:** when creating a table, the length of fields is set to the maximum that type of field can take. For instance, the column of a telephone number field known to not exceed 13-14 characters is strictly defined to not take longer than that threshold.
2. **Data Tuning:** also known as performance optimization, it refers to a set of activities used to ensure databases run efficiently, to locate and get rid of performance bottle necks. (Šušter & Ranisavljević, 2023).
- There are several methods employed for data tuning including:
- a) **Query optimization:** This is a critical part of data tuning, for ensuring that SQL queries perform better. It may involve improving the structure of database queries, adding, or removing indexes, or rewriting the query. In a study done by (Mahajan, et al., 2019), it was observed that query optimization is capable of improving the energy efficiency of databases. Examples of strategies employed include improving joins and subqueries, indexing frequently queried columns, using EXISTS instead of IN in WHERE clauses, and calling out the column names in SELECT statements rather than using the SELECT * wildcard.
 - b) **Server configuration optimization:** parameters such as size of buffer pool, query cache size and thread concurrency may be changed to optimize memory utilization and boost query performance. The max_connections parameter may be increased to manage high user concurrency (Šušter & Ranisavljević, 2023).
 - c) **Hardware tuning:** Upgrading the CPU, memory and storage of the database could increase the performance of the MySQL database server.

- d) **Normalization:** reduction of data redundancy and duplication can enhance query performance and lower storage costs (Šušter & Ranisavljević, 2023). Normalization is a database design technique used to achieve database optimization by structuring the database in a way to reduce redundancy and improve data consistency.

2.9 Case Studies and Related Works

<https://www.gs1.org/industries/healthcare/udi> -UDI implementation around the world

UDI database (UDID)- EU and FDA

European database on medical devices EUDAMED

ACCESS GUDID

The work also aligns with the 'collect once, use often' approach to data as recommended by the Independent Medicines and Medical Devices Safety Review (<https://www.immdsreview.org.uk/Report.html>) and will help bring the UK MedTech system in line with international efforts, as both the USA (<https://accessgudid.nlm.nih.gov/>) and the EU (<https://ec.europa.eu/tools/eudamed/#/screen/search-device>) have launched similar solutions

<https://www.federalregister.gov/documents/2013/09/24/2013-23059/unique-device-identification-system>

2.10 Summary of key findings and gaps in the literature.

INSERT TABLE OF RELEVANT PAPERS AND THEIR CONTRIBUTION. IDENTIFY
GAP in literature
empirical review

- domain
- database management system/database infrastructure
- optimization techniques
- user interface

Author(s)	Methodology	Finding(s)	Gap
-----------	-------------	------------	-----

			<ul style="list-style-type: none"> - Can't download query results
			<ul style="list-style-type: none"> - Not specific to the uk health system which is different
			<ul style="list-style-type: none"> - Cannot easily trace the location of the recipient as its not provided in the system.
GUDiD			<ul style="list-style-type: none"> - Few search options - Doesn't include trade item to improve traceability

There is a dearth of literature on the design of product information management systems especially in the UK healthcare space, hence the need for this project to fill that gap.

Chapter 3

Methodology

3.1 Introduction to Methodology

This chapter outlines the methodology employed in this project for the design, development, and evaluation of the National Product Information Management (PIM) System for the NHS. It encompasses the information system development lifecycle stages as outlined in the previous chapter, in addition to ethical considerations.

3.2 Research Design

This study leverages existing research findings and data collected by the NHS and the Department of Health & Social Care (DHSC) to inform the design and optimization of the PIM system, ensuring alignment with user needs and requirements while maximizing the utilization of available resources.

A sequential mixed-methods approach is adopted, combining qualitative and quantitative methods, to address the research problems and objectives appropriately (Office for Health Improvement and Disparities, 2020). The qualitative approach involves the research and review of literature, documents, and publications to understand the research problem, and goals of the intended PIM system using secondary data collected by the NHS through interviews, surveys, and stakeholder engagement sessions. The quantitative approach involves the performance testing and evaluation of the PIM system after its design and implementation to evaluate how the system meets defined requirements through statistical analysis of primary quantitative data from user survey responses.

3.3 Planning Phase

The planning phase of the development lifecycle involves an organization reaching a high-level understanding of their current situation, assessment of improvements needed, and a plan for reaching the improvement goals (Oppel, 2009). This phase forms the basis from which the project is launched. Once a business improvement need is identified, a feasibility study is done to evaluate how the project meets the organizations set goals.

The Department of Health & Social Care embarked on an in-house discovery phase which involved interactions (interviews, site visits, focus group sessions, resource sharing etc.) with

over 25 internal and external stakeholders from NHS trusts, the Medicines and Healthcare products Regulatory Agency (MHRA), NHS Supply Chain, NHS England, Medical Device Outcome Registry, manufacturers, and medical device data bodies (Crown Commercial Service, 2024).

A Freedom of Information (FOI) request, empowered by the Freedom of Information Act (FOIA) 2000 which empowers the public to access information held by public authorities, was sent to the Department of Health & Social Care (DHSC) - **FOI-1498477**, and the NHS - **FOI-2403-2087423**, to gain access to the full report of this in-house discovery phase which is not publicly accessible. The effort proved abortive as the DHSC claimed not to have that report and referred the author to the NHS, who had not responded to the request at the time of writing this report. An access request was also sent to Crown Commercial Service, the executive trading fund agency and public procurement organization of the Cabinet Office of the UK Government. Copies of these FOI requests and responses are attached in the appendix.

Hence, all information on this phase is gotten from research of publicly available publications, website documents, and reports.

The findings of this planning and discovery phase forms a foundational basis for this project, its scope, and requirements analysis. The primary data source for this consists of research reports, publications and documentations produced to this effect by the NHS, DHSC and other major players within the UK health system.

3.4 Requirements Analysis

In the requirements gathering and analysis phase, a high-level description of what the project must accomplish are gathered and documented. It focuses on the *what* rather than the *how*, which is developed in the design phase. The requirements analysis phase captures the existing and expected proposed system, business rules, and entities (Oppel, 2009).

It involves an identification of the stakeholders, and their needs, forming the basis of the analysis of the user views. A user view is a method for presenting a set of data to the user of a system in a way which meets the identified needs of that user (Oppel, 2009).

The requirements analysis of this study builds upon the insights gathered during the planning phase to define and refine users' needs, and system functionalities. It also establishes non-functional and the legal requirements which the system must meet.

3.5 Design Phase

The design phase involves translating the findings from the requirement analysis into a comprehensive design framework for the PIM system. This framework encompasses data model, system components and its user interface (UI).

The design phase may be broken down into the following 3 frameworks:

1. **Conceptual Design:** this involves the creation of a model of the data used in the organization independent of all physical considerations such as DBMS, application programs and programming languages (Connolly & Begg, 2015). It is built using the information from the requirements analysis as foundation. It also involves the design of the externals of the system i.e., the layout of reports, screens, and web pages etc. This may be documented in the form of flowcharts or screen diagrams, providing an understanding of the logical flow of the system (Savelios, 2021).
2. **Logical Design:** this is the technical design phase. It involves the design of the internals of the system, a blueprint for the design of the relational database. The major task in this phase is *Normalization*, a technique for producing a set of relations with desirable properties, given the data requirements of an organization. The main purpose of normalization is to remove data redundancy, and eliminate *Insert*, *Update*, and *Deletion* anomalies (Olagunju, 2023).
3. **Physical Design:** in this phase, the logical design is converted into the hardware and system software which will be used for the implementation of the PIM system (Oppel, 2009). The normalized relations from the logical design phase are implemented in the Database Management System (DBMS) using Data Definition Language (DDL).

3.6 Development and Evaluation

This phase builds on the system design specifications to develop, test and evaluate the PIM system. It follows the process of system development, testing and evaluation procedures to ensure the system's functionality, performance, and usability align with the defined user needs and objectives.

3.6.1 System Development

The system is developed using the appropriate tools and technologies, while adhering to best practices and coding standards. The tools and technologies used are discussed briefly in subsections 3.6.2 Hardware Requirements and 3.6.3 Software Requirements.

3.6.2 Hardware Requirements

The following machine features were used to build and implement the frontend and backend of the prototype PIM system:

Table 3: Hardware Requirements

Chip:	Apple M1
CPU:	8-core CPU (4 performance and 4 efficiency)
Memory:	8 GB
Storage Capacity:	256 GB SSD
Operating System:	MacOS Sonoma Version 14.2.1

The following hardware features will be sufficient to build and implement this PIM system: at least 8GB RAM, 256 GB Solid-State Drive (SSD) storage, Windows 10, and an intel core i5 processor or higher.

3.6.3 Software Requirements

The backend of the PIM system is built using the following software tools and technologies:

- i. **Oracle SQL Developer Data Modeler:** this is a free graphical tool by Oracle which facilitates efficient data modelling, ensuring the integrity and consistency of the database structure. This tool is used in this study for designing and creating the conceptual Entity-Relationship Diagram (ERD) for the PIM system.
- ii. **MySQL:** one of the world's most advanced open-source object-relational database system, MySQL will serve as the relational database management system for storing and managing all the entities, fields, and data within the PIM system. Its performance, scalability and security features make it an ideal choice for a PIM system handling a large volume of structured data (Wang, et al., 2023).
- iii. **MySQL Workbench:** this is the most popular and advanced open-source administration and development platform for MySQL. It is used in this study as the user interface tool for administration and interaction with the MySQL database.

- iv. **Aiven for MySQL:** this is an open-source cloud data platform for hosting, managing, and deploying MySQL databases. It is used in this study as the cloud data service for deploying the MySQL database to the cloud.
- v. **Python:** Python is a free high-level, powerful, and easy-to-use object-oriented programming language. Python is used in this study to generate dummy data for the PIM system and to develop the backend logic and data processing scripts due to its versatility, extensive library, and ecosystem of frameworks.
- vi. **GitHub:** this is a web-based software development platform used for storing and tracking software projects. GitHub is used in this study as a version control system and platform for storing, managing, and documenting SQL and python scripts used in the development of this PIM system.

The software requirement for building and designing the front-end of the PIM system is:

- i. **Streamlit:** this is an open-source Python framework for building and deploying dynamic web-based data applications. It is used in this study as the application generator for producing the front end of the prototype PIM system, for implementing the functional requirements and interacting with the backend SQL database.

3.6.4 System Evaluation

The process of evaluating the developed system to ensure its functionality, usability and alignment with requirements was structured as follows:

1. **Functionality Evaluation:** The PIM system's functionality is evaluated against the system requirements to ensure the system was correctly implemented.
2. **Performance evaluation:** performance evaluation is done to evaluate the system's performance in terms of its responsiveness, and efficiency under varying loads. The system's CPU % usage, disk space % usage, disk IOPS (reads), memory % usage, network receive (bytes/s) and network transmit (bytes/s) were the major performance metrics that were analyzed.
3. **Usability evaluation:** usability testing is conducted to evaluate the effectiveness and efficiency of the user interface and experience. Feedback is collected to identify any usability issues, and areas of improvement.

4. **Alignment with User Requirements:** deviations between the developed system and the documented user requirements are identified and evaluated.

3.7 Data Collection Methods

Data collection for this study is done through review and thematic analysis of secondary data sources like research publications, and documentations, and through analysis of system evaluation survey responses.

The system evaluation involved distribution of a questionnaire to individuals identified as potential users of the system. The potential users were required to respond to the questionnaire after they had used the prototyped PIM system, to gain insights into their perception of the system's performance and to identify areas of improvement.

The questionnaire was designed using the online questionnaire tool, **Google Forms**. The survey employs close-ended and open-ended questions to receive responses from users on their perception of the system.

3.7.1 Qualitative data

The qualitative data collection method was employed initially to understand the business problem, and requirements of the PIM system as defined by the NHS, and to evaluate the performance of the PIM system after development.

- i. Literature review findings: qualitative findings from the review of secondary data sources like research publications, and documentation produced by the NHS, and the department of health.
- ii. Open-ended survey response: Qualitative feedback provided by survey respondents through open-ended questions. Open-ended responses offer detailed insights into user perceptions, suggestions, and areas for improvement regarding the PIM system.

3.7.2 Quantitative data

- i. Close-ended survey responses: quantitative data obtained through surveys. responses provide numerical metrics and ratings on aspects such as usability, satisfaction levels and perceived performance.
- ii. System performance metrics: Quantitative measures of the PIM system's performance, including response times, data processing speeds, and resource utilization. These metrics are collected using monitoring tools and performance testing methodologies.

3.8 Ethical Considerations

Ethical considerations were of paramount importance throughout this research process. Measures were taken to ensure data privacy, confidentiality, and informed consent of participants in the user testing and evaluation of the system.

Appropriate ethical considerations in relation to the use of secondary data were also carefully considered. The research contributions of secondary sources such as the NHS, Department of Health & Social Care, and all other relevant sources were properly cited and acknowledged to maintain the integrity of this research work.

In general, this study complied with all ethical guidelines and principles by respecting the rights of participants involved in the data collection and system evaluation process.

3.9 Limitations and Assumptions

This study adopts a transparent approach to address potential bias and mitigate risks by acknowledging the limitations associated with the methodology, including resource, time, and data access constraints.

The use of secondary data for analysis of requirements is subject to limitations such as a potential bias in the original data collection, limited control over data quality, and constraints on scope and depth of analysis & insights.

These limitations are acknowledged and addressed in the interpretation and discussion of research findings.

3.10 Summary of research methodology

The chart below summarizes the methodology employed in this study.

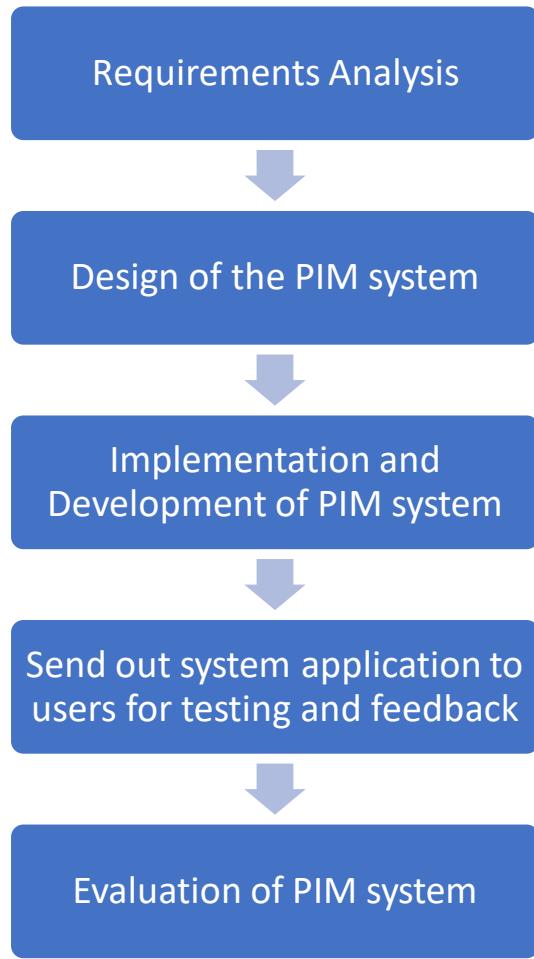


Figure 7: Summary of Research Methodology

Chapter 4

Requirements Analysis

4.1 Introduction to Requirements Analysis

This chapter covers the requirements analysis section of this information system design project.

This is a foundational step in the system development lifecycle of a National Product

Information Management (PIM) system for the NHS.

This phase is necessary to capture all the business needs in relation to the proposed PIM system (Langer, 2008).

4.2 Identification of Stakeholders and Information Needs Assessment

Stakeholder identification is the first step in the Requirements analysis phase (Simplilearn, 2023). Identifying the key stakeholders whose interests must be considered in the design and implementation of this national PIM system is crucial.

This section builds on the findings of the in-house planning and discovery phase for this project, done by the Department of Health as the foundational basis for identifying the problem, and business improvement need. As discussed earlier, this phase involved interviews, focus group sessions, resource sharing and site visits with internal and external stakeholders.

The main question asked during this phase was:

"How can we make basic information about medical products more available and of better quality?" (Crown Commercial Service, 2024)

This section presents a comprehensive overview of insights gained in this discovery phase as reported by (Crown Commercial Service, 2024), providing a foundational understanding of major stakeholders and an assessment of their information needs in relation to this PIM system.

1. Stakeholder: Suppliers/Manufacturers

- Information Need(s): Suppliers and manufacturers of medical devices and products in the UK need to share information about their product with the NHS for the trusts to

have accurate, consistent information about them. The product information provided by suppliers is used as the basis of the Product master data management system within the NHS. This data forms the basis of the PIM system.

2. Stakeholder: Data Systems Leads

- Information Need(s): Data systems leads in NHS trusts need to have accurate product information so that this information may be used in other systems within the trust e.g. financial records, electronic health records etc.

3. Stakeholder: Procurement Leads

- Information Need(s): Procurement Leads in NHS trusts need to have adequate, accurate information about medical products or devices to fully understand the characteristics of different products and be able to make the right decision on which products to buy, and/or alternatives.

4. Stakeholder: Patient Safety Leads

- Information Need(s): Patient Safety leads in NHS trusts are responsible for ensuring products and devices being used in patients' care are safe. They need to have accurate and up-to-date information about products to facilitate product traceability in case of recalls, and to meet other national safety initiatives like the Scan4Safety.

5. Stakeholder: Data Analysts

- Information Need(s): Data analysts in core sectors of the UK Health System like the NHS England, the MHRA, and the Department of Health & Social Care (DHSC), need to have information about products being used in patient care for easy export and joining with product information with other datasets, to analyse the market and understand patient safety trends.

6. Stakeholder: Patients

- Information Need(s): Patients need to have information about products or devices being used on them to ensure their confidence in the safety of the products.

4.4 System Requirements Analysis

The system requirements analysis builds on the findings of the needs assessment to define the specific requirements which the proposed PIM system must meet. The problem which the PIM system aims to solve is the inefficient data sharing process and inconsistent product information management throughout the supply chain from supplier to patient. This system aims to improve accessibility and quality of medical product data distributed across the UK healthcare system.

4.4.1 Functional Requirements

The following functions are required from the national PIM infrastructure according to analysis of findings from (Boulding & Hinrichs-Krapels, 2021) and (Procurement, Investment & Commercial Division, DHSC, 2014):

1. Users must be able to access to the existing dataset of medical device products supplied to NHS Trusts.
2. Users must be able to use the system to view product data,
3. Users must be able to search for product information,
4. Users must be able to apply appropriate filters,
5. Users must be able to export product information from the system,
6. Users must be able to identify obsolete items.

4.4.2 Non-functional Requirements

Non-functional requirements of the national PIM system are as follows, as outlined in a report by (Commercial Division, DHSC, 2017):

1. **Scalability:** The system must be scalable i.e. it must be able to accommodate increasing volumes of data and users.
2. **Quality assurance:** The system must provide consistent, standardized data to ensure the quality of information gotten from it.
3. **Interoperability:** The system must ensure interoperability between systems used by NHS providers, suppliers, and other relevant bodies.
4. **Security:** The database system must be secure and password-protected.
5. **Legal Compliance:** the system must comply with all legal and regulatory frameworks which govern the creation of such system, globally and in the UK.

This national PIM system will act as a shared reference point for exchanging information about medical products for NHS stakeholders. It will make data sharing more efficient, and the use of accurate, consistent data is going to improve the decision-making process, thus improving patient care outcomes.

4.5 Legal and Regulatory Frameworks

The design and implementation of a national PIM system for a healthcare organization like the NHS requires strict adherence to legal and regulatory requirements. This section provides an examination of the standards, guidelines and regulations governing healthcare product data management in the UK.

4.5.1 Relevant legal and regulatory compliance standards and obligations

The NHS operates within complex legal & regulatory frameworks which are aimed at ensuring data privacy and security across all information management systems. It is necessary to comply with the legal and regulatory requirements to avoid litigation and ensure trust and reliability among all stakeholders.

The following are some legal and regulatory standards which the UK health system must comply with in the development of a product information management infrastructure:

- 1. International Medical Device Regulators Forum (IMDRF) guidance for Unique Device Identification (UDI) of Medical Devices:** this is a supporting guidance for the unique identification of medical devices (GMDN, 2023). It is a framework for developing a UDI system which achieves harmony with the global standard (IMDRF UDI Working Group, 2013).
- 2. European Commission (EC) framework for a unique identification system of medical devices in the union:** This framework applies to medical devices, implantable medical devices, and in-vitro diagnostic medical devices. The rationale for which is improving incident reporting, facilitating actions by national bodies, enabling queries in several data systems, reducing medical errors (European Commission, 2013).

3. **The Medical Devices Regulations 2002:** information provided by manufacturers and suppliers of medical device products to the NHS are subject to data validation and scrutiny by the MHRA in accordance with the Medical Device Regulation of 2002 (Medicines and Healthcare products Regulatory Agency, 2020). Global Medical Device Nomenclature (GMDN) terms and codes must be implemented by suppliers/manufacturers as part of the regulations for registering devices for the UK market (GMDN, 2023).
4. **Data Protection Act (DPA 2018):** since the UK's exit from the European Union, the DPA 2018 has replaced the EU General Data Protection Regulation (GDPR) for operations inside the UK. The provisions of the GDPR have been directly incorporated into the UK law (Spencer & Patel, 2019). This law imposes rules on the collection, storage, and processing of personal data in the UK, ensuring the protection of rights of individuals, and the responsible handling of data. Personal data in this context includes information about any persons associated with products e.g., suppliers or employees.

4.5.2 Influence of legal and regulatory considerations on the design and implementation of the NHS PIM system.

All the requirements set out by the Department of Health for the NHS' National Product Information Management System align with the GDSN standard for healthcare across Europe, based on a common European data dictionary (Commercial Division, DHSC, 2017), which also aligns with the standard set internationally by the IMDRF (European Commission, 2013). This ensures interoperability, not only within the NHS, but across global healthcare organizations.

The internationally agreed standards for unique device identification management which have influenced the UK Department of Health's Product Information Management system requirements outlined by (European Commission, 2013) and (IMDRF UDI Working Group, 2013), are summarized as follows:

Product Information

- i. A Unique Device Identifier (UDI) will be assigned to a medical device/trade item or its package. Higher levels of packaging will be assigned their own UDI;
- ii. the UDI will contain 2 parts: a UDI Device Identifier (UDI-DI) and a UDI-Production Identifier (UDI-PI);

- a. a UDI-DI is a globally unique identifier specific to a device trade item which is represented by the Global Trade Item Number (GTIN)
 - b. a UDI-PI is a unique identifier which identifies the unit of production. It comprises a device's serial number, lot/batch number, manufacturing and/or expiry date.
- iii. A new UDI-DI must be assigned if any of the following data attributes change; brand name, device model/version, size, labelled as single use, sterilization, quantity per package, warnings, and contraindications etc.;

Product Information Management System

- iv. the system must incorporate all core data attributes; the minimum required to identify and describe a product throughout the supply chain;
- v. the system will contain no commercially sensitive or confidential information.
- vi. the core data attributes must be publicly accessible;
- vii. Product lifecycle information e.g., information about product discontinuation or suspension, must be included in the system.

Based on these guidelines, the following are the core data attribute requirements for the design and implementation of a National PIM database system as recommended by the EC and the IMDRF:

- a) Product brand/trade name;
- b) UDI-Device Identifier – GTIN
- c) Unit of Issue e.g. each, case, pallet
- d) Quantity per package configuration e.g. 1 each, 10 each(es)
- e) Manufacturer's name, address, and customer service contact information;
- f) Authorized representative's name, details and contact information;
- g) Global Medical Device Nomenclature (GMDN) code and description;
- h) UDI-Production Identifier
- i) Global Model Number (GMN) - Device model or version;
- j) Manufacturer Reference and/or Catalogue number (if applicable)
- k) Clinical size (volume, length, diameter, gauge);
- l) additional product description or information;

- m) Storage and/or handling conditions;
- n) License or Registration number
- o) Is device labelled as Single Use? (Yes/No)
- p) If reusable, restricted number of reuse(s);
- q) Is device packaged sterile? (Yes/No)
- r) Does device need to be sterilized before use? (Yes/No)
- s) Sterilization method – if device needs to be sterilized;
- t) Critical warnings or contraindications associated with device use:
 - i. Is device labelled as containing latex? (Yes/No)
 - ii. Is device labelled as containing DEHP? (Yes/No)
 - iii. Is device compatible with MRI? (Yes/No)
- u) Product lifecycle information and date

These core data elements will be incorporated into the design of the proposed NHS Product Information Management System.

4.6 Summary of Requirements Analysis

It has been established that the product information management ecosystem within the NHS suffers from major burdens such as manual data requests resulting in outdated, inconsistent, and even incorrect product information. Hence, there is need for a new system to collect standardized and consistent data once and use often, leading to improved accessibility and data quality for all stakeholders and initiatives.

The major stakeholders and users of the system have been identified as suppliers and manufacturers, data analysts, data systems leads, procurement leads, patient safety leads, and patients. The goal is for them to have access to accurate product data and information. The system must be accessible, accurate, secure etc. ensuring quality of product data, interoperability, and scalability.

The work complies with EC and IMDRF UDI frameworks, MDR 2002 and DPA2018.

Minimum data attribute requirements for the NHS PIM system have been established and will be used in the design and modelling of the system in the following chapters.

Finally, this chapter provides adequate answers to my third research question.

Chapter 5

Design and Modelling

5.1 Introduction to system design

This chapter covers all phases of the design of the PIM system including the conceptual, logical, and physical designs. This chapter covers the techniques employed in these phases including the definition of business rules, entity-relationship modelling, normalization, and the conceptual design of the user interface. This is a critical phase in the developmental lifecycle of the PIM system as it lays the foundation for the implementation of the functional and non-functional requirements as discussed in the requirements analysis phase.

5.2 Database Design & Modelling

The data modelling process is the first step in designing the PIM system. It involves creating a data structure representation of the complex, real-world environment to solve a specific problem (Coronel & Morris, 2016). The data model represents the structures of the data involved, their attributes, relationships, and constraints.

5.2.1 Business Rules

The business rules are brief descriptions of the policies or procedures within an organization which govern the use and storage of data (Coronel & Morris, 2016). These rules inform the definition of entities, attributes, relationships and constraints during the design and modelling of an information system for the organization.

In the context of the NHS National Product Information Management System, the following business rules may be defined in accordance with standards established in the requirements analysis:

- Each trade item is supplied by only one supplier, and a supplier may supply many trade items.
- Each trade item is supplied to an NHS provider/trust, and each NHS provider may have many trade items supplied to them.
- A trade item contains a unique medical device, but a medical device may be contained in many trade items (IMDRF UDI Working Group, 2013).

- A Global Medical Device Nomenclature (GMDN) code may be assigned to many medical devices, but a medical device can only have one GMDN code (GMDN, 2023).
- An NHS product classification (eClass) code may be assigned to many medical devices, but a medical device can only have one eClass code. (Procurement, Investment & Commercial Division, DHSC, 2014)
- A medical device may be assigned a Basic UDI-DI, a Global Model Number (GMN) to uniquely identify the model or product family of that medical device, and a GMN may be assigned to multiple medical devices. (GS1, 2024)
- A medical device is manufactured by only one manufacturer, but a manufacturer may manufacture many medical devices.
- A medical device may be assigned a classification reference number by the manufacturer, and a classification number may be assigned to many medical devices of the same family.
- A medical device must have a manufacturer assigned authorized representative, and an authorized representative may be assigned to many medical devices or family of devices.
- A medical device is assigned a regulatory device classification according to its risk class, and a risk class may be assigned to many medical devices or medical device models (Medicines and Healthcare products Regulatory Agency, 2020).

5.2.2 Entities and Attributes

The PIM system is designed to accommodate the information needs of the identified stakeholders while considering the minimum data requirements defined in the requirements analysis, and the business rules defined in the preceding subsection. To meet these needs, the relevant entities and attributes are modelled as follows, with the primary key underlined:

1. Medical Device: This entity provides information about medical devices in the UK market. Attributes include: (GTIN, brand_name, unit_of_use, quantity_of_uou, item_length, item_height, item_width, item_weight, item_volume, unit_of_dimension, product_description, storage_handling, single_use, restricted_no_of_use, sterile, sterilize_before_use, sterilization_method, item_contains_latex, item_contains_dehp, item_mri_compatible, item_model_gmn, gmdn_code, nhs_eclasse_code, manufacturer_reference_no, risk_class_name)

- Foreign key(s): **item_model_gmn, gmdn_code, nhs_eclasse_code, manufacturer_reference_no, risk_class_name**

2. Trade Item: This entity provides information about unique trade items supplied to the NHS healthcare providers. Attributes include: (**UDI, gtin, serial_number, batch_number, manufacturing_date, expiry_date, udi_pi, unit_of_issue, unit_of_use_udi, supplier_gln, nhs_provider_gln**)

- Foreign key(s): **gtin, supplier_gln, nhs_provider_gln**

3. Supplier: This entity provides information about the suppliers who trade medical device items with the NHS organization and its healthcare providers. Attributes include: (**supplier_gln, supplier_name, supplier_address, company_registration_no, customer_service_phone, customer_service_email**)

4. Manufacturer: This entity provides information about the manufacturers of the medical devices on the UK market. Attributes include: (**manufacturer_gln, manufacturer_name, manufacturer_address, company_registration_no, customer_service_phone, customer_service_email**)

5. NHS Provider: This entity provides information about the healthcare providers within the NHS organization who are at the receiving end of the medical device trade items supplied by suppliers. Attributes include: (**provider_gln, provider_name, provider_address, provider_registration_no**)

6. Manufacturer Catalog: This entity provides information about products listed in a manufacturers catalog. Attributes include: (**manufacturer_reference_no, product_name, product_model, product_category, market_availability_date, lifecycle_status, last_status_update, manufacturer_gln, authorized_rep_id**)

- Foreign key(s): **manufacturer_gln, authorized_rep_id**

7. Risk Class: This entity provides information about the classification of medical devices based on the level of risk they pose to the user. Attributes include: (class_name, class_description, regulatory_requirements)

8. Item Model: This entity provides information about the GS1 identification system of a product model or product family. Attributes include: (GMN, model_name, device_type, market_availability_date, lifecycle_status, last_status_update)

9. Authorized Rep: This entity provides information about the representatives authorized by the manufacturer for their product catalog. Attributes include: (rep_id, rep_name, contact_number, email)

10. GMDN: This entity provides information about the global standard for identification and classification of medical devices. Attributes include: (gmdn_code, gmdn_term_name, gmdn_term_definition)

11. NHS Product Classification: This entity provides information about the classification codes used within the NHS. Attributes include: (eclasse_code, description)

A comprehensive data dictionary describing all these terms and attributes is provided in the appendix of this report.

5.2.3 Relationships

A relationship describes the association between entities of a database. The following relationships are identified between entities of this PIM system:

a) Medical_Device and Trade_Item

- Relationship: Medical_Device (0..*) is contained in Trade_Item (1..1)

Each trade item contains a unique medical device on the UK market. Each medical device on the UK market may be contained in zero to many trade items.

- Cardinality: One-to-Many (1:N)
- Participation: Source Optional, Target Mandatory

b) Trade_Item and Supplier

- Relationship: Supplier (1..*) **supplies** trade_item (1..1)
Each registered supplier supplies one or more trade items. Each trade item can only be supplied by one supplier.
- Cardinality: One-to-Many (1:N)
- Participation: Source & Target Mandatory

c) Trade_Item and NHS_Provider

- Relationship NHS_Provider (1..*) **receives** trade_item (1..1)
Each NHS_Provider receives one or more trade items. Each trade item can only be received by one NHS_Provider.
- Cardinality: One-to-Many (1:N)
- Participation: Source & Target Mandatory

d) Medical_Device and Manufacturer_Catalog

- Relationship: Manufacturer_Catalog (1..*) **references** Medical_Device (1..1)
Each manufacturer catalog reference number references one or more medical devices. Each medical device is referenced by only one reference number.
- Cardinality: One-to-Many (1:N)
- Participation: Source & Target Mandatory

e) Manufacturer and Manufacturer_Catalog

- Relationship: Manufacturer (1..*) **manufactures** Manufacturer_Catalog(1..1)
Each manufacturer manufactures one or more products and references in a manufacturer catalog. Each product in a manufacturer catalog is referenced by only one manufacturer who manufactured the product.
- Cardinality: One-to-Many (1:N)
- Participation: Source & Target Mandatory

f) Manufacturer_Catalog and Authorized_Rep

- Relationship: Authorized_Rep (1..*) **controls** Manufacturer_Catalog(1..1)
Each authorized rep controls one or more products in a manufacturer catalog. Each product in a manufacturer catalog can only be controlled by one authorized rep.
- Cardinality: One-to-Many (1:N)
- Participation: Source & Target Mandatory

g) Risk_Class and Medical_Device

- Relationship: Risk_Class (1..*) **regulatorily classifies** Medical_Device (1..1)
Each risk class regulatorily classifies one or more medical devices. Each medical device is classified by only one risk class
- Cardinality: One-to-Many (1:N)
- Participation: Source & Target Mandatory

h) Item_Model and Medical_Device

- Relationship: Item_Model (1..*) **globally classifies** Medical_Device (1..1)
Each item model globally classifies one or more medical devices. Each medical device is classified by only one item model (GMN).
- Cardinality: One-to-Many (1:N)
- Participation: Source & Target Mandatory

i) Medical_Device and NHS_product_classification

- Relationship: NHS_product_classification (1..*) **nationally classifies** Medical_Device (1..1)
Each NHS eClass code nationally classifies one or more medical devices.
Each medical device is classified by only one NHS eClass code.
- Cardinality: One-to-Many (1:N)
- Participation: Source & Target Mandatory

j) Medical_Device and GMDN

- Relationship: GMDN (1..*) **globally names** Medical_Device (1..1)

Each GMDN code globally names one or more medical devices. Each medical device is named by only one GMDN code.
- Cardinality: One-to-Many (1:N)
- Participation: Source & Target Mandatory

5.2.4 Entity Relationship Models

Entity-Relationship models are a way of representing data structures graphically, generally accepted to be more effective than describing these structures in text. They are typically represented using entity relationship diagrams (ERD) which use graphical representations to model database components (Coronel & Morris, 2016). This study presents the ERD with logical and relational ER models for the PIM system.

The following steps were taken to design the ER diagrams for the system:

1. **Installing Oracle SQL Developer Data modeller:** the latest version for my operating system at the time of installation Version: 22.2.0.165.1149 was installed from
<https://www.oracle.com/database/sqldeveloper/technologies/sql-data-modeler/download/>
2. **Creating Entities:** 11 key entities, with the appropriate entity names were created in the logical model of the design.
3. **Creating attributes:** attributes of the established entities were defined and assigned appropriate data types and integrity constraints.
4. **Defining relationships:** relationships between the key entities were created and defined. The relationships between the entities are illustrated using the Crow's Foot notation. Figure 8 below shows the logical model for the PIM system designed in Oracle SQL Data Modeller by defining the entities, attributes, and relationship between attributes
5. **Engineering to relational model:** the logical model was then engineered into a relational model based on the defined relations and constraints as shown in Figure 9.

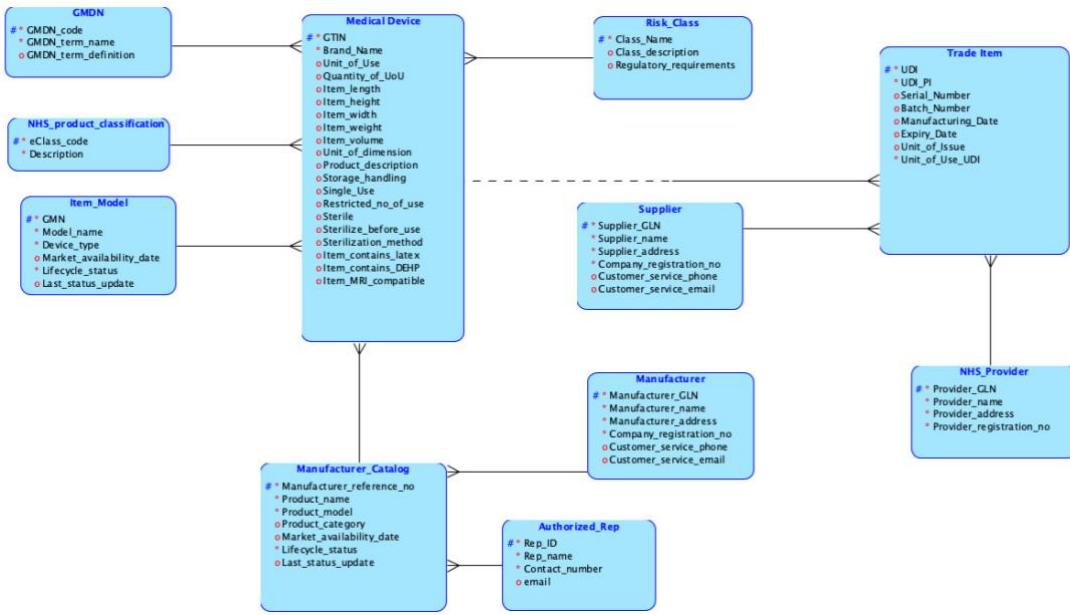


Figure 8: Logical ER model for the PIM system.

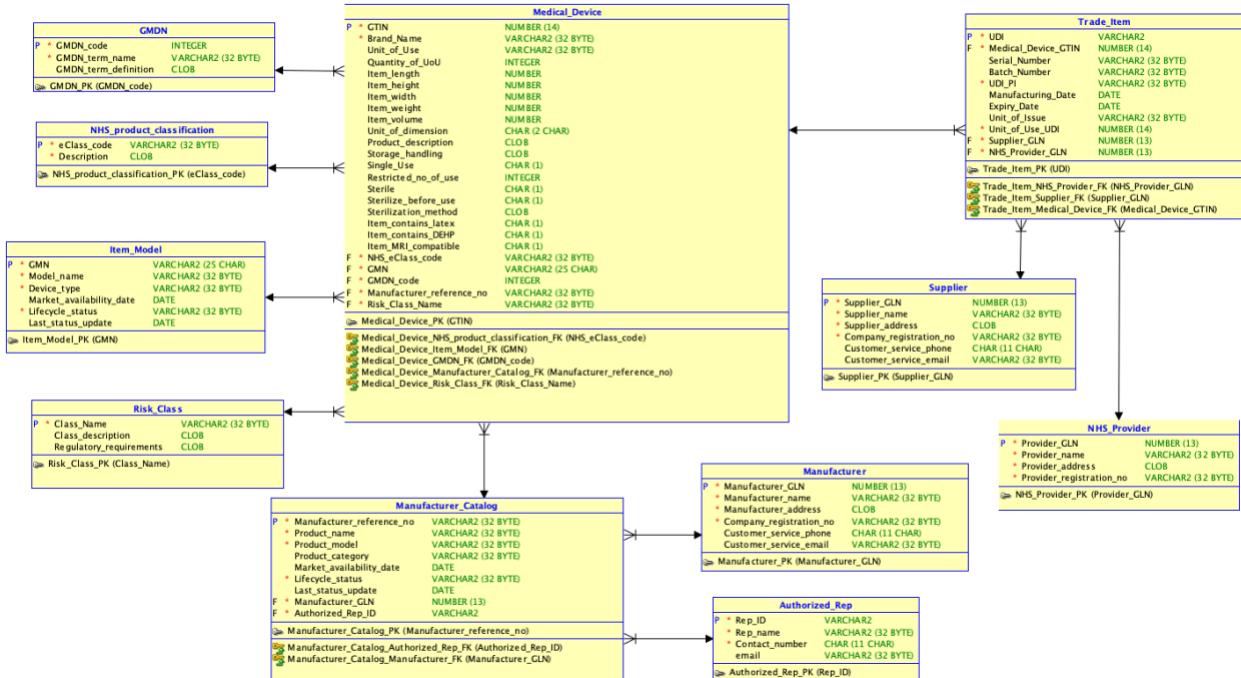


Figure 9: Relational ER model for the PIM system.

5.3 Database Normalization

Normalization is a process of organizing the database structure according to the requirements and need of an organization, minimizing redundancy which may cause update or delete anomalies.. From an Unnormalized form, there are 3 main normalization steps to achieve acceptable database normalization for this study.

- i. **Unnormalized Form (UNF):** this is a normalization form where the database entities contain repeating groups. For every row of an identified primary key, there are cells which contain more than one value at the intersection of the row and column. To take a database from UNF to 1NF, values for the repeating groups must be duplicated.
- ii. **First Normal Form (1NF):** This is when a database and its entities contain no repeating groups. At the intersection of a unique row and column, there is only one value representing the mapping of a unique identifier key on the database attribute.
- iii. **Second Normal Form (2NF):** This is when a database and its entities are in first normal form and contain no partial dependencies.
- iv. **Third Normal Form (3NF):** this is when a database and its entities are in second normal form and contain no transitive dependencies. A transitive dependency exists when there is a functional dependency among nonprime attributes such that if there exists $X \rightarrow Y$, and $Y \rightarrow Z$, and X is the primary key, then $X \rightarrow Z$ is a transitive dependency because the value of Z is determined by X through Y.

5.3.1 The Normalization Process

Step 1: *Eliminating the repeating groups and identifying the primary key*

In unnormalized form, the following functional dependency exists between the attributes of the **Trade_Item** table, with the identified primary key being the UDI:

$UDI \rightarrow GTIN, serial_number, batch_number, manufacturing_date, expiry_date, udi_pi, unit_of_issue, unit_of_use_udi, brand_name, unit_of_use, quantity_of_uou, item_length, item_height, item_width, item_weight, item_volume, unit_of_dimension, product_description, storage_handling, single_use, restricted_no_of_use, sterile, sterilize_before_use, sterilization_method, item_contains_latex, item_contains_dehp,$

item_mri_compatible, eclass_code, description, GMDN_code, GMDN_term_name, GMDN_term_definition, supplier_GLN, supplier_name, supplier_address, company_registration_no, customer_service_phone, customer_service_email, provider_GLN, provider_name, provider_address, provider_registration_no, manufacturer_reference_no, product_name, product_model, product_category, market_availability_date, lifecycle_status, last_status_update, manufacturer_GLN, manufacturer_name, manufacturer_address, company_registration_no, customer_service_phone, customer_service_email, authorized_rep_id, rep_name, contact_number, email, GMN, model_name, device_type, market_availability_date, lifecycle_status, last_status_update, risk_class_name, class_description, regulatory_requirements

The Unique Device Identifier key – the **UDI**, functionally determines every attribute in the database. This means that for every value of the UDI, there is only a single value of every other attribute i.e., there are no repeating groups. Therefore, the database is already in **1NF**.

Step 2: *Identifying and eliminating partial dependencies*

For the database to be in 2NF, partial dependencies must be identified and eliminated. A partial dependency exists when a table has a composite primary key and some of the non-key attributes fully functionally depend on only part of the primary key i.e., if $\{A, B\} \rightarrow C, D, E$, where $\{A, B\}$ is a composite primary key, a partial dependency exists if $A \rightarrow C$ or $B \rightarrow D, E$. Since there is no composite primary key i.e., there is only one identified primary key (UDI), no partial dependency exists. Hence, the database is already in **2NF**.

Step 3: *Identifying and eliminating transitive dependencies*

It is important to eliminate transitive dependencies and making new tables from a 2NF database table to take it to 3NF. To do this, the transitive dependencies must be identified.

Transitive dependency 1:

GTIN → brand_name, unit_of_use, quantity_of_uou, item_length, item_height, item_width, item_weight, item_volume, unit_of_dimension, product_description, storage_handling, single_use, restricted_no_of_use, sterile, sterilize_before_use, sterilization_method, item_contains_latex, item_contains_dehp, item_mri_compatible, eclass_code, description, GMDN_code, GMDN_term_name, GMDN_term_definition, manufacturer_reference_no, product_name, product_model, product_category, market_availability_date, lifecycle_status, last_status_update, manufacturer_GLN, manufacturer_name, manufacturer_address, company_registration_no, customer_service_phone, customer_service_email, authorized_rep_id, rep_name, contact_number, email, GMN, model_name, device_type, market_availability_date, lifecycle_status, last_status_update, risk_class_name, class_description, regulatory_requirements

The determinant of this transitive dependency is the **GTIN**. This is used to create the new **Medical_Device** table, which is uniquely identified by the GTIN, with its dependent attributes on the right side of the arrow.

Transitive dependency 2:

provider_GLN → provider_name, provider_address, provider_registration_no

The determinant of this transitive dependency is **provider_GLN**. This leads to the creation of the **NHS_Provider** table for identifying the NHS organization who receives the trade items from the suppliers. The dependent attributes of this new table are shown on the right side of the arrow.

Transitive dependency 3:

supplier_GLN → supplier_name, supplier_address, company_registration_no, customer_service_phone, customer_service_email

The determinant of this transitive dependency is **supplier_GLN**. This leads to the creation of a new table, **Supplier**. The dependent attributes of this new table are shown on the right side of the arrow.

Finally, the original **Trade_Item** table is left with the following attributes:

UDI → serial_number, batch_number, manufacturing_date, expiry_date, udi_pi, unit_of_issue, unit_of_use_udl, GTIN, provider_GLN, supplier_GLN

It can be observed that the **Medical_Device** table is still in 2NF as it still contains transitive dependencies.

Transitive dependency 4:

eclass_code → description

The **NHS_product_classification** table is created with the **eclasse_code** as the determinant attribute.

Transitive dependency 5:

GMDN_code → GMDN_term_name, GMDN_term_definition

This results in the creation of the **GMDN** table with **GMDN_code** as the determinant attribute.

Transitive dependency 6:

manufacturer_reference_no → product_name, product_model, product_category, market_availability_date, lifecycle_status, last_status_update, manufacturer_GLN, manufacturer_name, manufacturer_address, company_registration_no, customer_service_phone, customer_service_email, authorized_rep_id, rep_name, contact_number, email

A new **Manufacturer_Catalog** table is derived with the **manufacturer_reference_no** as the determinant attribute.

Transitive dependency 7:

GMN → model_name, device_type, market_availability_date, lifecycle_status, last_status_update

A new **Item_Model** table is created with the **GMN** as the determinant attribute and the dependent attributes on the right side of the arrow.

Transitive dependency 8:

risk_class_name → class_description, regulatory_requirements

A new **Risk_Class** table is created from transitive dependency 8 with the risk_class_name as the determinant attribute.

Medical_Device table is now void of any transitive dependencies and the corresponding dependent attributes are reassigned as such:

GTIN → brand_name, unit_of_use, quantity_of_uou, item_length, item_height, item_width, item_weight, item_volume, unit_of_dimension, product_description, storage_handling, single_use, restricted_no_of_use, sterile, sterilize_before_use, sterilization_method, item_contains_latex, item_contains_dehp, item_mri_compatible, eclass_code, GMDN_code, manufacturer_reference_no, GMN, risk_class_name

The derived **Manufacturer_Catalog** still contains the following transitive dependencies:

Transitive dependency 9:

manufacturer_GLN → manufacturer_name, manufacturer_address, company_registration_no, customer_service_phone, customer_service_email,

A new **Manufacturer** table is created out of transitive dependency 9 with the manufacturer_GLN as the determinant attribute.

Transitive dependency 10:

authorized_rep_id → rep_name, contact_number, email

A new **Authorized_Rep** table is created out of transitive dependency 10 with the authorized_rep_id as the determinant attribute.

After the elimination of the identified transitive dependencies 9 & 10, the

Manufacturer_Catalog is left with the following attributes:

manufacturer_reference_no → product_name, product_model, product_category, market_availability_date, lifecycle_status, last_status_update, manufacturer_GLN, authorized_rep_id

At the end of the normalization process, the database is now in **3NF**, and the following entities were derived:

- a) NHS_Provider: (provider GLN, provider_name, provider_address, provider_registration_no)
- b) Supplier: (supplier GLN, supplier_name, supplier_address, company_registration_no, customer_service_phone, customer_service_email)
- c) Trade_Item: (UDI, serial_number, batch_number, manufacturing_date, expiry_date, udi_pi, unit_of_issue, unit_of_use_udl, GTIN, provider_GLN, supplier_GLN)
- d) NHS_product_classification: (eclasse code, description)
- e) GMDN: (GMDN code, GMDN_term_name, GMDN_term_definition)
- f) Item_Model: (GMN, model_name, device_type, market_availability_date, lifecycle_status, last_status_update)
- g) Risk_Class: (risk class name, class_description, regulatory_requirements)
- h) Medical_Device: (GTIN, brand_name, unit_of_use, quantity_of_uou, item_length, item_height, item_width, item_weight, item_volume, unit_of_dimension, product_description, storage_handling, single_use, restricted_no_of_use, sterile, sterilize_before_use, sterilization_method, item_contains_latex, item_contains_dehp, item_mri_compatible, eclass_code, GMDN_code, manufacturer_reference_no, GMN, risk_class_name)

- i) Manufacturer: (**manufacturer_GLN**, manufacturer_name, manufacturer_address, company_registration_no, customer_service_phone, customer_service_email)
- j) Authorized_Rep: (**authorized_rep_id**, rep_name, contact_number, email)
- k) Manufacturer_Catalog: (**manufacturer_reference_no**, product_name, product_model, product_category, market_availability_date, lifecycle_status, last_status_update, manufacturer_GLN, authorized_rep_id)

5.3 Conceptual Design of User View and Interface Layout

The PIM system adopts a centralized user view approach which merges all the requirements of all user groups into a single view. This is employed due to the overlap in user requirements (Connolly & Begg, 2015). The system will be publicly accessible, and hence does not require login credentials to access the data.

The landing page will consist of a home page which welcomes the user to the PIM system and provides instructions on how to use and navigate the system as shown in the page layout and conceptual design in Figure 10 and Figure 11 respectively.

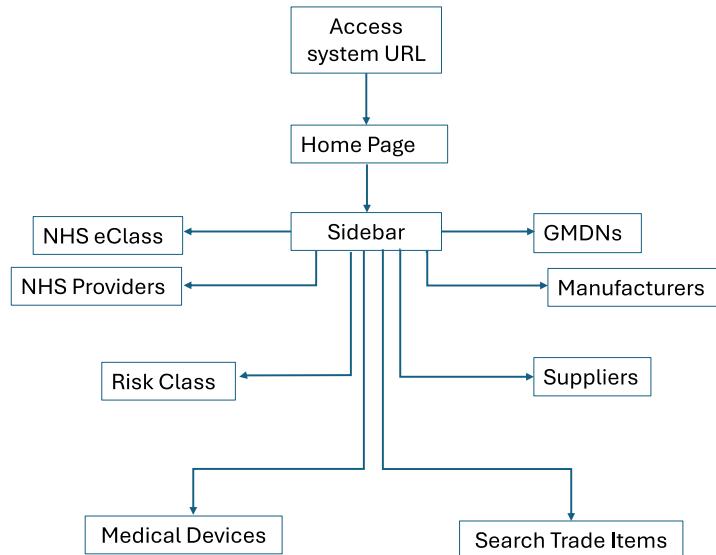


Figure 10: PIM System landing page layout

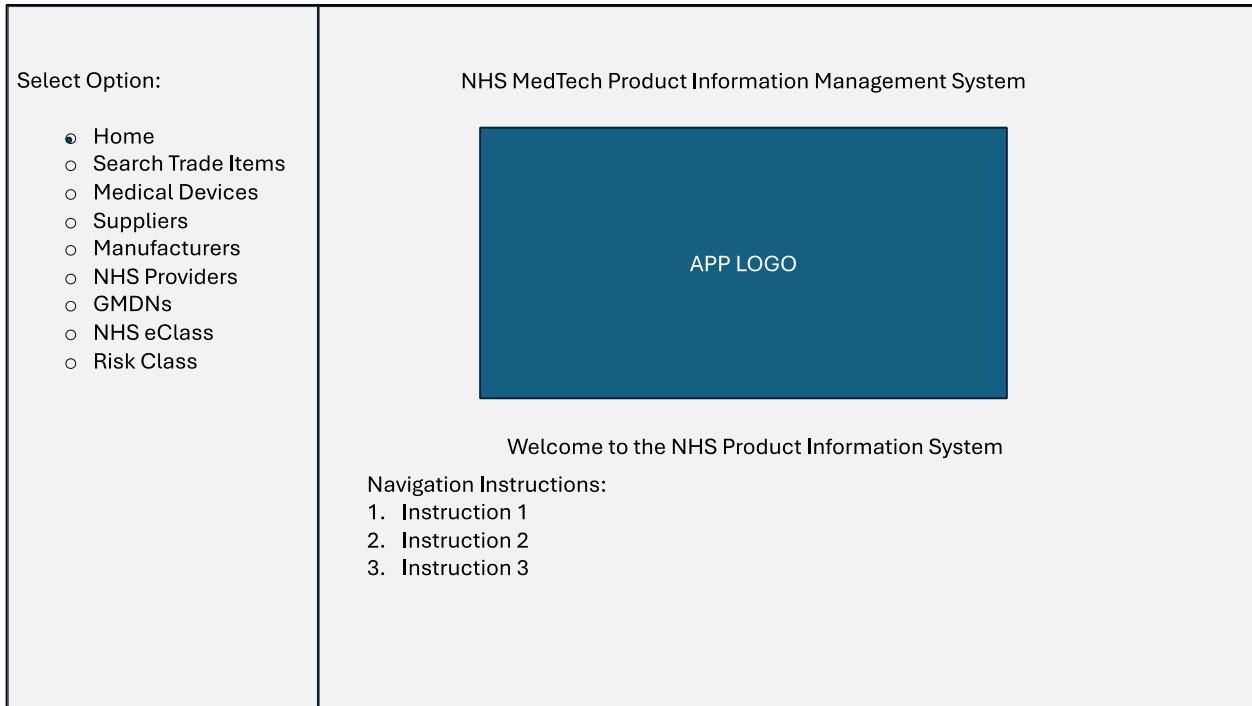


Figure 11: Conceptual Design of the System's landing page

There will be a side pane providing access to custom search and view trade items in the NHS supply chain and all information relating to the item, custom search and view medical devices on the UK market, suppliers and manufacturers of medical devices, NHS Providers, GMDNs, NHS eClass, and Risk Class. The system will also allow for the download of query results according to the user's custom search filter as shown in the data access layout in Figure 12.

The information needs of all stakeholders have been incorporated into the data model and design of the system layout. Depending on the user, the trade items can be searched using various search options as shown in Figure 13. For example, a data systems lead within an NHS trust can search for trade items by the unique identifier of the NHS trust which they belong to (Provider GLN) and access information about all unique medical devices that have been supplied to that NHS Provider, OR a patient safety lead could trace a medical device in the case of a product recall by searching the GTIN of that device.

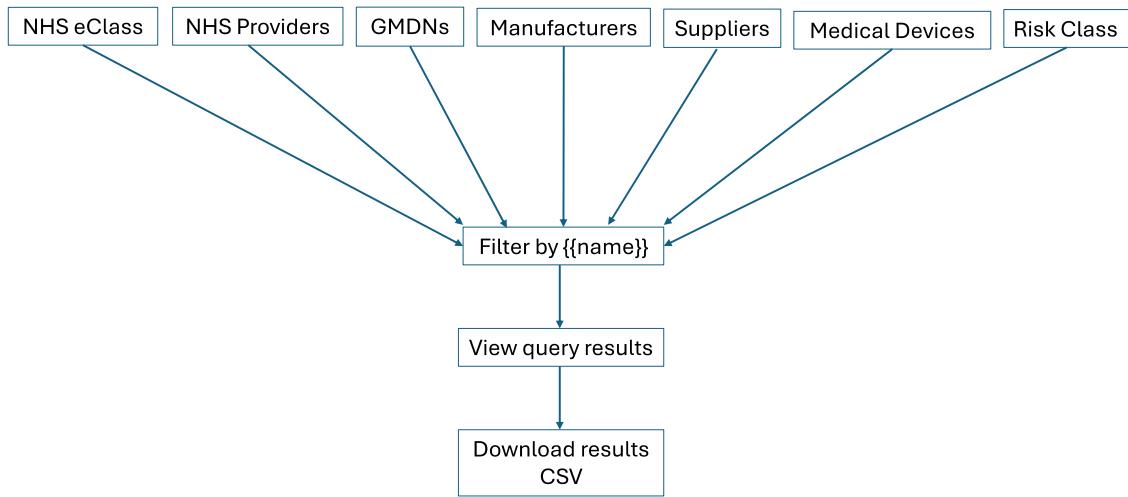


Figure 12: PIM System data access layout

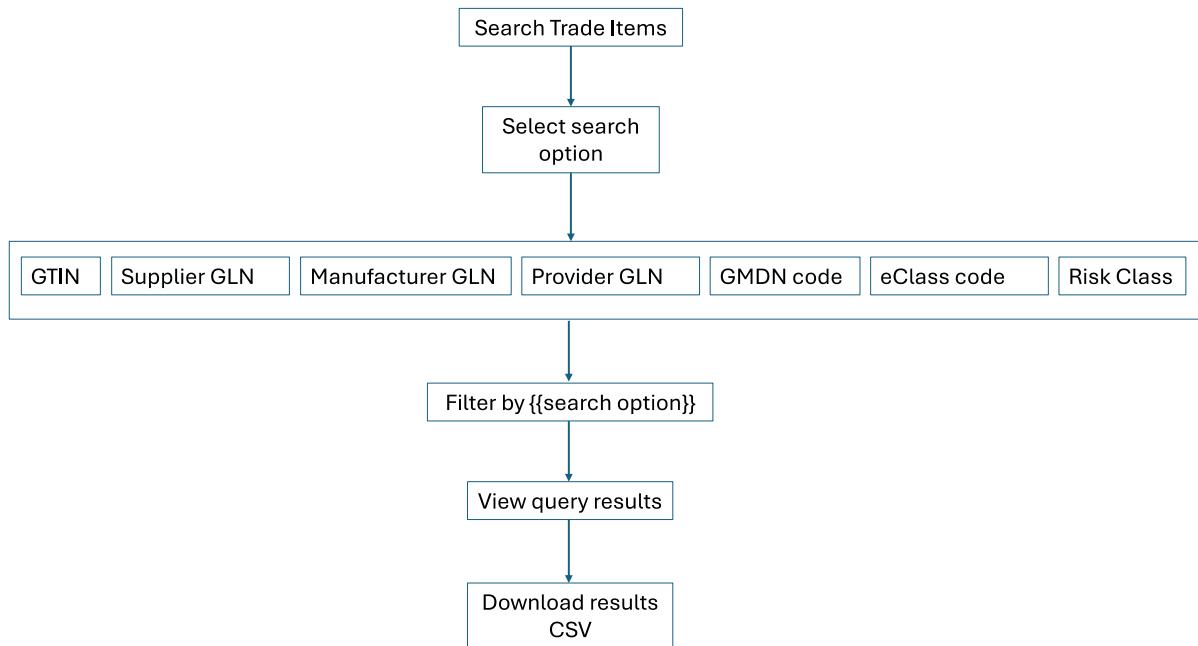


Figure 13: Flow chart of 'Search Trade Item' page

5.4 Summary of System Design

In this chapter, the requirements of the PIM system were adapted into a conceptual design and logical design in preparation for implementation. This included defining the business rules for the organization's data environment, defining key entities and attributes, establishing relationships between those attributes, developing logical and relational entity-relationship diagrams in Oracle SQL Developer Data Modeller, and normalization of the database to the third normal form (3NF). Finally, a conceptual design of the system's user view(s) and web page layout was created, illustrating the flow of navigation and features of the system.

In the next chapter, the physical design of the system will be covered, and the PIM system will be fully implemented.

Chapter 6

Implementation and Development

6.1 Introduction to Implementation

In this chapter, the implementation phase of this project is detailed, focusing on the development and deployment of the national PIM system for the NHS. It covers the system architecture, the database implementation process, deployment to a cloud infrastructure, development of the application program, scalability and optimization of the system, and data security and integrity measures taken.

6.2 System Architecture

The design of the PIM system adopts a Three-Tier Client-Server Architecture. This means that the system runs on 3 layers running on different platforms. The following are the layers of the 3-Tier client-server architecture:

1. **First Tier: *Client***

This represents the user interface layer. This is the device or computer (*client*) on which the PIM system user views the application. This application serves as the front-end for users to access the data and information, and to interact with the system's functionalities. The client application for this system is built using **Streamlit**.

2. **Second Tier: *Application Server***

This is the layer with deals with data processing and business logic. The Streamlit *application server* is responsible for hosting and running the Streamlit application. It uses Python to process user queries and communicate with the database server.

3. **Third Tier: *Database Server***

This is the database management system (DBMS) which stores and manages the data required by the second tier, which typically runs on the *database server*. In the case of this project, the database server is the **MySQL** database which handles data storage, retrieval, and manipulation.

Figure 14 below graphically illustrates this architecture at a higher level.

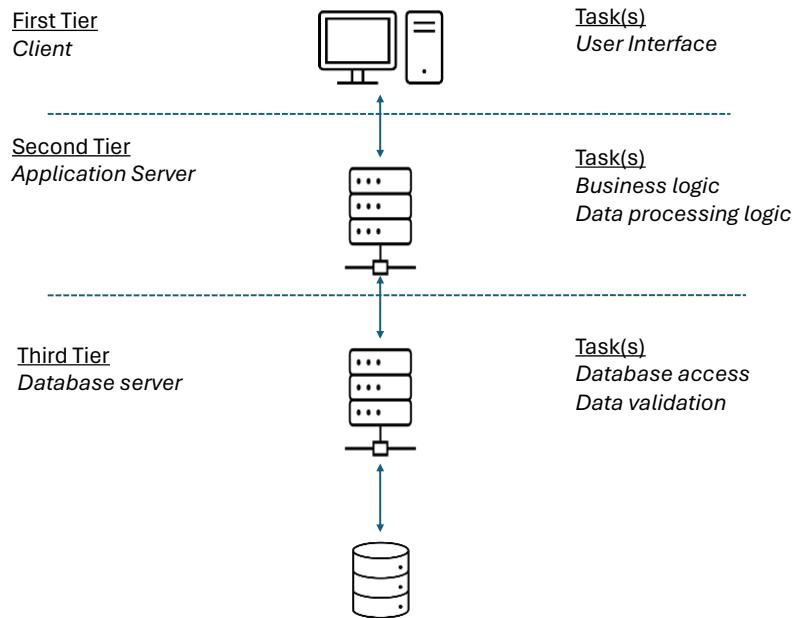


Figure 14:3-Tier Client-Server architecture

6.3 Database Implementation

The database implementation phase follows the output from the design phase to create the entities, attributes, indexes, integrity constraints and other guidelines (Coronel & Morris, 2016).

6.3.1 Implementation Steps

The following steps were taken to implement the database:

1. **Installation of DBMS:** The first step in implementing the database involved installing MySQL Community Server version 8.3.0-arm64, and its integrated development environment (IDE), MySQL Workbench version 8.0.36 which provides an environment to design, model and develop the MySQL database. These were downloaded from <https://dev.mysql.com/downloads>.
2. **Connecting to MySQL server from MySQL Workbench:** the next step was creating a connection for the NHS_PIM instance to the MySQL server. This is where all the database development takes place.

3. **Creation of ‘NHS_PIM’ database schema:** a new database schema is created inside the NHS_PIM database. This is done using the following SQL query:

```
CREATE SCHEMA NHS_PIM;
```

4. **Creation of tables:** tables are created and modified to the necessary entity, semantic and referential integrity constraints using SQL Data Definition Language (DDL). Tables are created using CREATE TABLE statements. For example, to create the **trade_item** table with its attributes and key constraints, the following MySQL code is run:

```
CREATE TABLE `trade_item` (
  `UDI` varchar(32) NOT NULL,
  `gtin` char(14) NOT NULL,
  `serial_number` varchar(32) DEFAULT NULL,
  `batch_number` varchar(32) DEFAULT NULL,
  `manufacturing_date` date DEFAULT NULL,
  `expiry_date` date DEFAULT NULL,
  `udi_pi` varchar(32) NOT NULL,
  `unit_of_issue` varchar(32) DEFAULT NULL,
  `unit_of_use_udt` char(14) NOT NULL,
  `supplier_gln` char(13) NOT NULL,
  `nhs_provider_gln` char(13) NOT NULL,
  PRIMARY KEY (`UDI`),
  KEY `trade_item_medical_device_fk` (`gtin`),
  KEY `trade_item_nhs_provider_fk` (`nhs_provider_gln`),
  KEY `trade_item_supplier_fk` (`supplier_gln`),
  CONSTRAINT `trade_item_medical_device_fk` FOREIGN KEY (`gtin`) REFERENCES `medical_device` (`gtin`),
  CONSTRAINT `trade_item_nhs_provider_fk` FOREIGN KEY (`nhs_provider_gln`) REFERENCES `nhs_provider` (`provider_gln`),
  CONSTRAINT `trade_item_supplier_fk` FOREIGN KEY (`supplier_gln`) REFERENCES `supplier` (`supplier_gln`)
) ;
```

Figure 15: CREATE TABLE trade_item code snippet

5. **Dataset generation:** a dummy dataset was generated to populate the created database tables. This was done using the python libraries *csv*, *random*, *string*, *faker*, *datetime*. The data generated matches the data specifications and standards for the data attributes. The dummy data generator python script is attached in the appendix.
6. **Population of tables:** The generated dummy data was inserted into the tables using the SQL Data Manipulation Language (DML) **INSERT** statements.

The full SQL DDL and DML code for creating tables and inserting data into the created tables of the NHS_PIM database are attached in the appendix.

6.3.2 Limitations of Local Database Server

The database server was originally set up on a local machine and run on localhost, but this came with certain limitations.

The following are some limitations of running database on local server:

- i. **Scalability Issues:** the database running on a localhost environment creates limitations in the scalability of the system because the database server is tied to the computer's resources such as the computer's CPU, memory, and disk space in comparison with cloud or dedicated servers. This impacts the performance of the database when handling increasing volumes of data and concurrent users.
- ii. **Limited Accessibility:** access to the database is limited to the local machine on which it is installed. This restricts remote access from other machines or over the internet. Connection cannot be established between the local database server and the PIM system's web application program.
- iii. **Backup and Recovery challenges:** due to challenges implementing backup and recovery strategies for a local database, there is a huge potential for data loss in the event of hardware failure.
- iv. **Security:** ensuring the security of a database on a localhost environment like access controls, data encryption and security updates can be a complex task, requiring advanced expertise to implement
- v. **Maintenance:** running the database on localhost requires regular hardware maintenance and software updates which takes time and effort.

To mitigate these limitations and have a scalable PIM system which can be accessed by multiple users in different locations, the database was deployed on a fully managed data cloud platform, **Aiven for MySQL**.

6.3.3 Aiven for MySQL

Aiven for MySQL is a fully managed MySQL database service which allows users to deploy their database in the cloud of choice. With *Aiven for MySQL*, database developers can get a fully managed MySQL database running in less than 10 minutes.

Key features of *Aiven for MySQL* include:

- i. **Managed service:** as a fully managed service, *Aiven for MySQL* handles all aspects of database administration and maintenance including automatic updates, backups, security patches and performance optimization.
- ii. **Scalability:** it provides the opportunity to scale horizontally and vertically to accommodate increasing data and workload volumes by allowing to adjust CPU and increasing storage as needed.
- iii. **Security:** : *Aiven for MySQL* boasts of robust security features including encryption, network isolation, access controls, and regular security updates to protect data and address any vulnerabilities.
- iv. **High availability:** *Aiven for MySQL* also boasts of a 99.99% uptime, using replication and failover mechanisms to automatically recover from disruptions or failures. This ensures the accessibility and reliability of the database.
- v. **Performance Monitoring:** *Aiven for MySQL* provides performance monitoring and alerting capabilities to help track the health and performance of the database
- vi. **Compatibility:** it provides compatibility with standard MySQL database clients like MySQL Workbench and other libraries to ensure ease of migrating existing MySQL apps to the Aiven service
- vii. **Multi-cloud support:** it provides the opportunity to choose a cloud platforms like Amazon Web Services (AWS), Google Cloud Platform (GCP), DigitalOcean, and Microsoft Azure to suit users requirements.
- viii. **Free Plan:** *Aiven for MySQL* offers a free plan which allows users to create one free service. It provides a single node, 1 CPU per virtual machine, 5GB disk storage, 1 GB RAM, and DigitalOcean hosting in certain regions. The plan can be upgraded at any time.

6.3.4 Deploying MySQL database to the Aiven for MySQL service

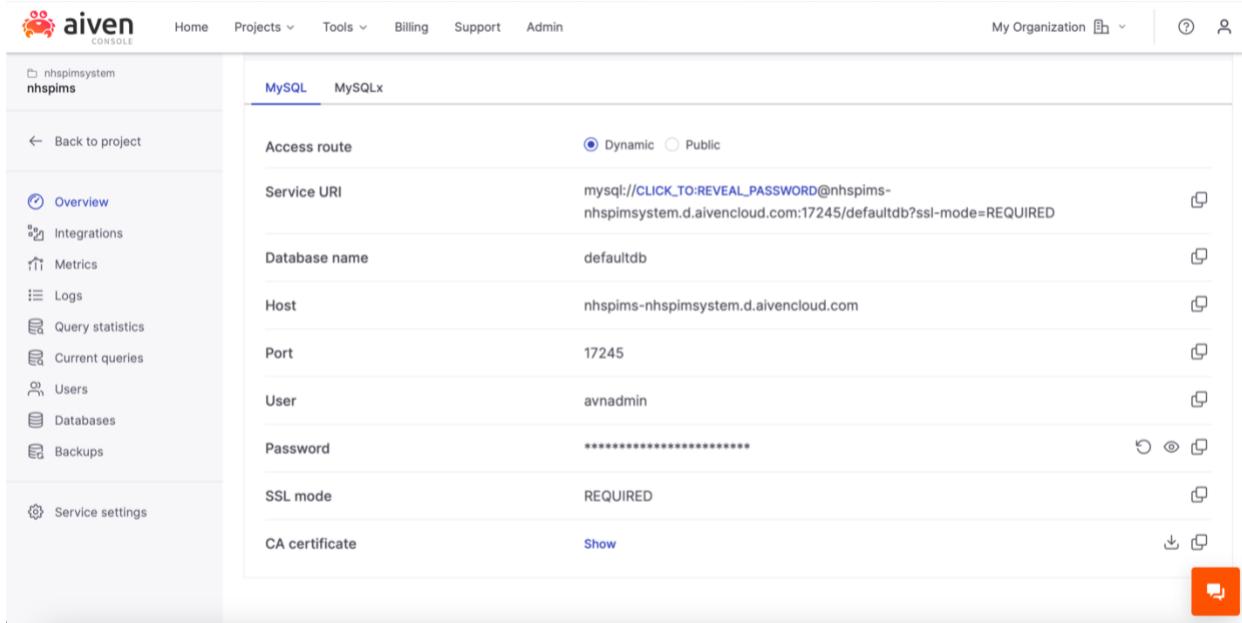
The MySQL database for the NHS_PIM system was created on the Aiven for MySQL by following steps similar to the local database. The following steps were taken:

1. Sign up on the Aiven console, and login.
2. **Create Aiven for MySQL service:** A free plan MySQL database instance was created on *Aiven for MySQL* for implementing and developing a working prototype of the PIM system, as shown in **Error! Reference source not found..**

Service	Nodes	Plan	Cloud	Created	Action
 nhspims MySQL • Running	Nodes 1	Free-1-5gb 1 CPU / 1 GB RAM / 5 GB storage	DigitalOcean: Ion Europe, England	5 days ago	...

Figure 16:MySQL database instance running on Aiven for MySQL

3. **Create connection parameters:** Database connection credentials were created for the admin user as shown in Figure 17 below.



The screenshot shows the Aiven Console interface for the 'nhspims' service. On the left, there's a sidebar with project navigation (Back to project, Overview, Integrations, Metrics, Logs, Query statistics, Current queries, Users, Databases, Backups, Service settings). The main area is titled 'MySQL' and shows the following configuration details:

Setting	Value	Actions
Access route	<input checked="" type="radio"/> Dynamic <input type="radio"/> Public	
Service URI	mysql://[REVEAL_PASSWORD]@nhspims-nhspimsystem.d.aivencloud.com:17245/defaultdb?ssl-mode=REQUIRED	Copy
Database name	defaultdb	Copy
Host	nhspims-nhspimsystem.d.aivencloud.com	Copy
Port	17245	Copy
User	avnadmin	Copy
Password	*****	Copy, Copy with mask, Copy with hex
SSL mode	REQUIRED	Copy
CA certificate	Show	Download, Delete

Figure 17: NHS_PIM cloud database connection credentials

- 4. Connecting to Aiven for MySQL service from MySQL Workbench:** the next step was creating a connection for the NHS_PIM instance to the Aiven for MySQL service using the connection credentials generated in step 3. This is made possible due to the compatibility feature earlier mentioned.

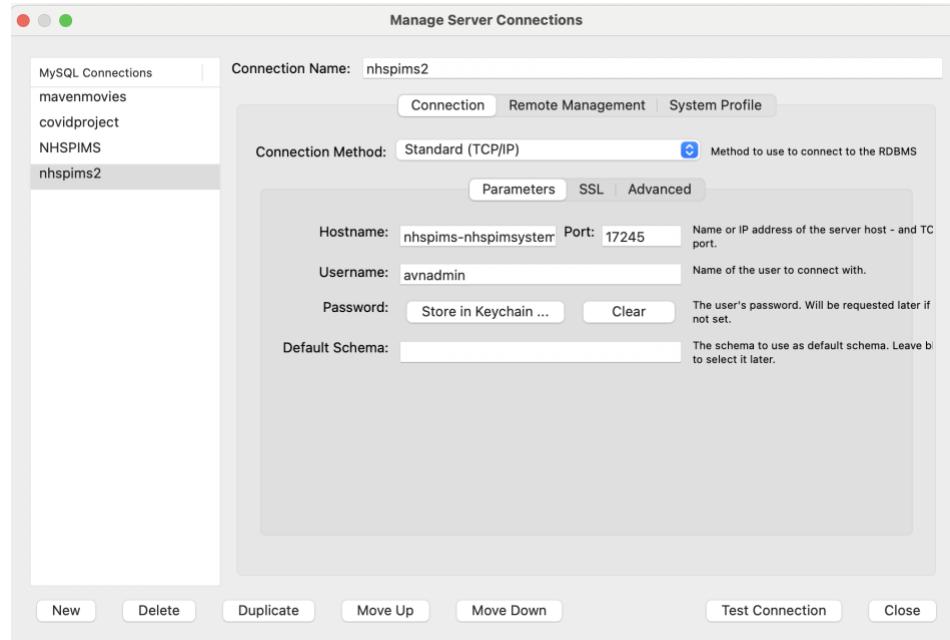


Figure 18: MySQL Workbench connection to Aiven for MySQL service

6.4 Application Program Development

The application program for this project is built using Streamlit. Streamlit is an open-source python library which allows developers to build web applications with Python scripts without knowledge of web development languages.

Streamlit is chosen because of its features which include:

- i. Ease of use,
- ii. Rapid prototyping capabilities,
- iii. Integration with other python libraries such as Pandas, Matplotlib etc.
- iv. Ease of deployment
- v. Built-in components for data science tasks like data visualization, model training etc.
- vi. Interactive widgets such as sliders, button text input etc. providing interactive interface.

Streamlit communicates with the MySQL database by establishing a connection to the database with the connection credentials. Streamlit uses Python to start up the Streamlit server which interprets user queries and sends to the database server for data access and processing, and the response is sent back to the client for display on the user interface

6.4.1 Implementation of Streamlit

The following steps were taken to implement the application program using Streamlit:

1. **Installation of required libraries:** The first step involved installing the necessary Python libraries which required the installation of:
 - Python version 3.11.2
 - Streamlit
 - Pandas: this is a python package which is used in this project for its flexibility and reliability in working with relational data structures
 - MySQL Connector: this enables python programs to access MySQL databases.These libraries were installed using Python's package installer, pip.
2. **Application set up:** an *nhspim.py* python script was created for building the Streamlit app.
3. **Importing requirements:** the required libraries *Streamlit*, *Pandas* and *mysql.connector* (app dependencies) are then imported.
4. **Establishing connection to MySQL database:** connection to the MySQL database is established using *mysql.connector.connect()* and the database connection credentials.
5. **Building app:** the PIM system app is built using Streamlit's python API to write code which define the user interface and components as laid out in the conceptual design of the app.

6.4.2 User Interface (UI)

The User Interface was designed to be clear and easy to use, using a combination of elements to present the PIM system to be user friendly and visually.



Figure 19: NHS PIM System User Interface

The following are some of the visual elements adopted in the UI design:

- i. **Title:** the system was designed with a clear and legible title instantly conveying the purpose
- ii. **Images and graphics:** the system was designed with minimal images and graphics, ensuring the reduced distraction from the purpose of the system. A logo for the PIM system was designed in Canva, to mirror the visual of the NHS logo while adding a spin of originality to it as shown in Figure 19.
- iii. **Colour:** there is a minimal and consistent use of colour throughout the system. making the system more accessible to sufferers of visual impairments such as colour blindness.
- iv. **Typography and Visual Hierarchy:** the typography of the system's UI ensures the legibility of texts within the app. It also makes good use of visual hierarchy to indicate the importance of text elements, like titles, subheadings, and normal text.
- v. **Widgets:** the PIM system adopts the use of interactive widgets like a dropdown to select search options, sidebar, buttons, and text input fields.

- vi. **Navigation Instructions:** the PIM system was designed with clear navigation instructions and use of proper widget labels
- vii. **Icons and Layout:** The system provides a clear layout of functionalities which can be navigated from the sidebar by clicking each page option. Data fields and query results are logically grouped, and familiar field labels are used.

6.4.3 User View(s)

The user views for the PIM system application were designed to meet all the minimum system functional requirements as outlined in the requirements analysis phase.

1. **'Search Trade Items' Screen:** This view allows the user to access product information about medical device products which have been supplied to NHS providers. It allows the user to search for specific trade items using a specific parameters of their choice which they can select from a dropdown as shown in Figure 20 below.

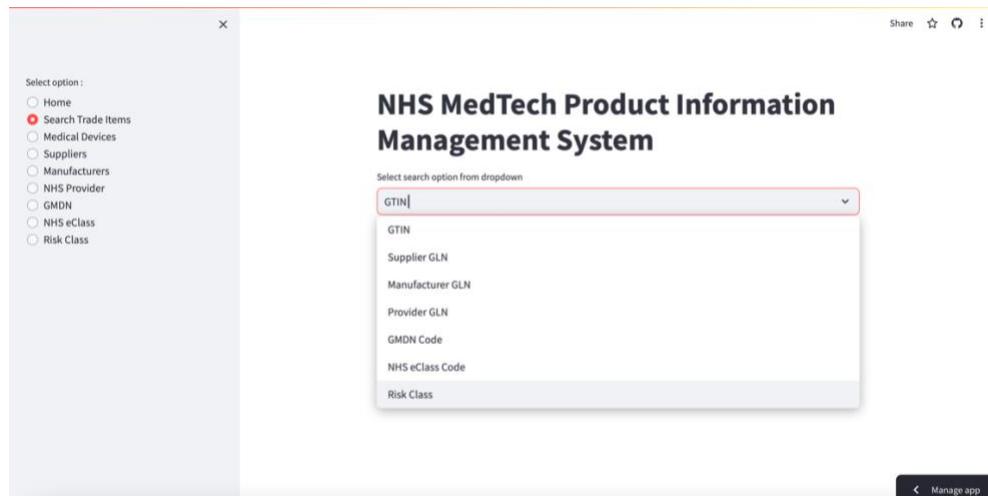


Figure 20: Search trade Items page of the PIM system

The functionalities of each of the trade item search options are summarized below:

- **Medical Device:** users can access information on trade items that have been delivered to the NHS by searching for specific medical devices. Users can perform this search by the GTIN of the medical device and view the minimum required product and supply chain information on where these devices are across the NHS organization. Users can search for medical device GTINs from the 'Medical Devices' page. Figure 21 shows a view of

the dataset of all trade items of the medical device with GTIN 25220261342140 and brand name *BioInnovations*.

The screenshot shows a search interface for medical devices. On the left, a sidebar lists options: Home, Search Trade Items (which is selected), Medical Devices, Suppliers, Manufacturers, NHS Provider, GMDN, NHS eClass, and Risk Class. The main area has a title 'NHS MedTech Product Information Management System'. It includes a dropdown menu for 'Select search option from dropdown' set to 'GTIN', a search input field containing '25220261342140', and a 'Search' button. Below this, a table titled 'Trade Items' displays one row of data:

UDI	gtin	brand_name	serial_number	batch_number
0 252202613421402604758KKDSFNE5L0	25220261342140	BioInnovations	FNESL0	58KKDS

A 'Download Trade Items as CSV' button is present. At the bottom right is a 'Manage app' button.

Figure 21: Search Trade Items by GTIN

- Item Supplier:** Users can also access at least the minimum required product information about medical devices that have been supplied to the NHS by a specific supplier. Figure 22 below shows product information about medical device trade items that have been supplied across the NHS organization by the hypothetical supplier, *XYZ Tech* with a unique Global Location Number (GLN) – 9175017029116. Users can search for Supplier GLNs from the ‘Suppliers’ page of the system.

The screenshot shows a search interface for medical devices. The sidebar and main area are identical to Figure 21, with the 'Supplier GLN' dropdown set to '9175017029116'. The 'Trade Items' table displays four rows of data:

UDI	gtin	brand_name	serial_number	batch_number
0 0284880237308133820237RBCVCSQW	02848802373081	BioDevices	VVC5QW	Z3TRBC
1 1316724013640986825X3Z2WT878Z	13167240136409	HealthLab	TL878Z	XV3Z2W
2 34996982473446444427TGHZAOOF9	34996982473444	BioDevices	A0OQF9	Z17GNZ
3 731980268556440428709EXEZVX9EB	73198026855644	MediSolutions	VX9EB	09EXEZ

Figure 22: Search Trade Items by Supplier GLN

- **Device Manufacturer:** Users can also access product information about medical devices manufactured by a specific manufacturer. Figure 23 below shows product information about medical device trade items that have been supplied across the NHS organization manufactured by the hypothetical manufacturer, *Orange Solutions* with a unique Global Location Number (GLN) – 4592397953624.

The screenshot shows a web-based application titled "NHS MedTech Product Information Management System". On the left, there is a sidebar with a "Select option:" dropdown containing several choices, with "Search Trade Items" being the selected option. The main content area has a heading "NHS MedTech Product Information Management System" and a sub-section "Search Trade Items by Manufacturer GLN". A search bar contains the value "4592397953624". Below the search bar is a "Search" button. The main table area is titled "Trade Items" and displays a list of trade items with columns: UDI, gtin, brand_name, serial_number, and batch_number. The table contains 7 rows of data. At the bottom of the table, there is a "Download Trade Items as CSV" button. In the bottom right corner, there are "Manage" and "Export" buttons.

UDI	gtin	brand_name	serial_number	batch_number
0 5528866445043652336DA6G4PQGXQQ8	55288664450436	MediTech	QGXQQ8	D66G4P
1 5615055244153050693G9WP51BCL7X1	56150552441530	MediTech	BCL7X1	G9WP51
2 5810107056862521823JUVK6THO2OO1	58101070568625	HealthTech	NO2OO1	IUVK6T
3 5819174479803541309LAQ859H4CUCN	58191744798035	LifeSolutions	I4CUCN	LAQ859
4 58378071471426320968K0RQQY4B8JU0	58378071471426	MediCare	Y4B8JU0	8K0RQQG
5 5948172469728983174OOIDBOONFF00	59481724697289	MediCare	GNFF00	O0IDB0
6 6025424091002187989YGYL3TJ51K	60254240910021	LifeTech	T13G1K	YGYL3B

Figure 23: Search Trade Items by Manufacturer GLN

- **NHS Provider:** users, especially stakeholders within an NHS Provider Trust can access product information about all medical devices that have been supplied to their NHS trust by searching the trade items page by the unique GLN of that NHS trust. For example, trade items that have been supplied to *Ashford and St Peter's Hospitals NHS Foundation Trust* with a unique GLN – 5355706845251 are shown in Figure 24 below. The full dataset of these trade items can be viewed in the app and exported as CSV.

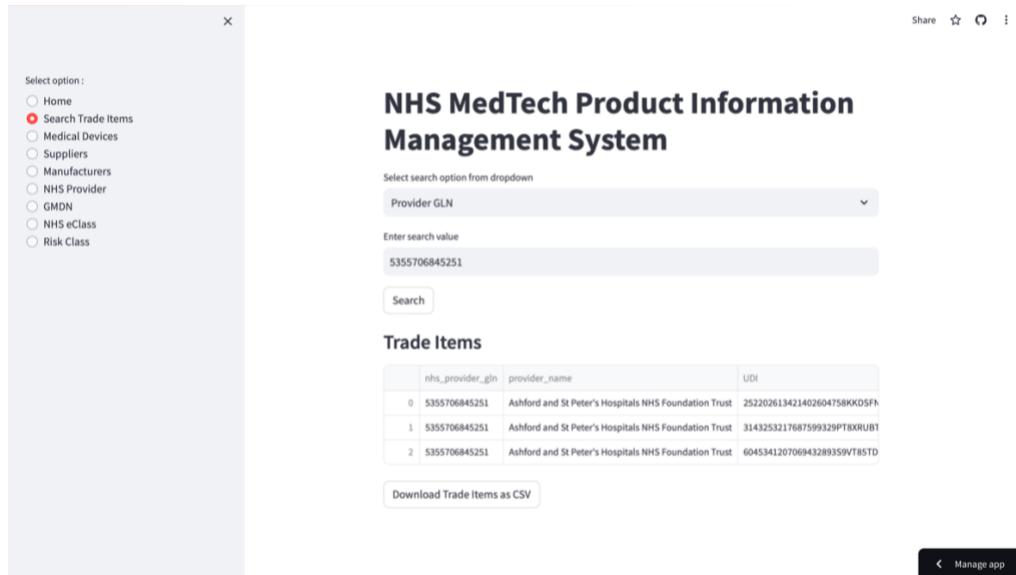


Figure 24: Search Trade Items by Provider GLN

- **GMDN:** Users can access product information of medical device trade items that have been supplied to the NHS by searching for specific GMDN codes. This returns a dataset of all medical trade items across all manufacturers, suppliers, NHS Trusts with the same GMDN code.
 - **NHS product classification:** Users can also access product information of medical device trade items that have been supplied to the NHS by their NHS eClass product classification. This returns a dataset of all supplied medical trade items with the same eClass code.
 - **Medical Device Risk Class:** users of the system can access product information of medical devices by the regulatory risk class names e.g. a user may have the need to find specific information about Class III devices due to the high risk they pose to the patient.
2. **'Medical Devices' Screen:** from the 'Medical Devices' page, users may access the dataset of product information about all medical devices on the UK market. It provides the opportunity to filter the dataset by medical device brand names as shown in figure below. It provides the GTIN of all devices and users may use this to search trade items.

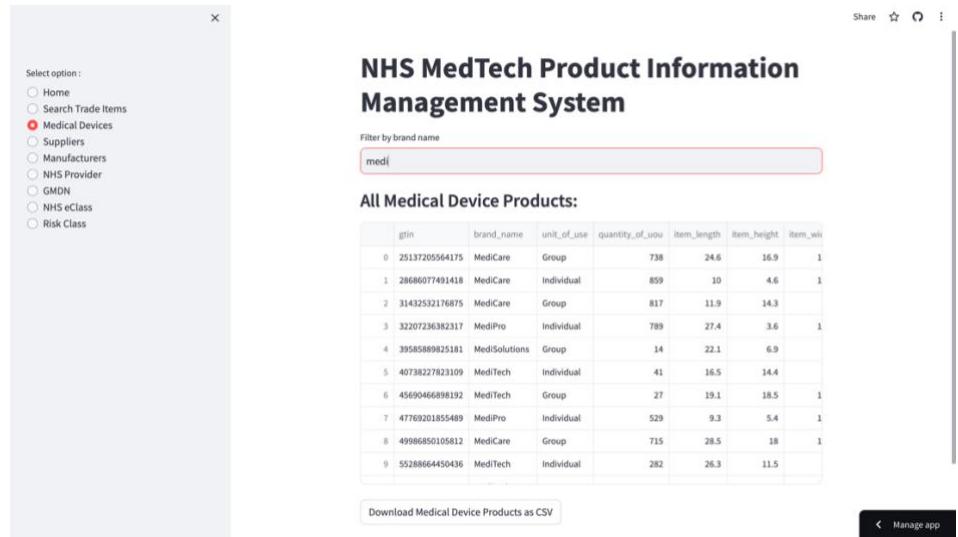


Figure 25: 'Medical Devices' Screen and functionalities

3. **'Suppliers' Screen:** users may access information about all suppliers of medical devices to the NHS from this page. They may also use information gotten from this page i.e. the supplier GLN, to search the trade items data to view all trade items that supplier has supplied to the NHS.

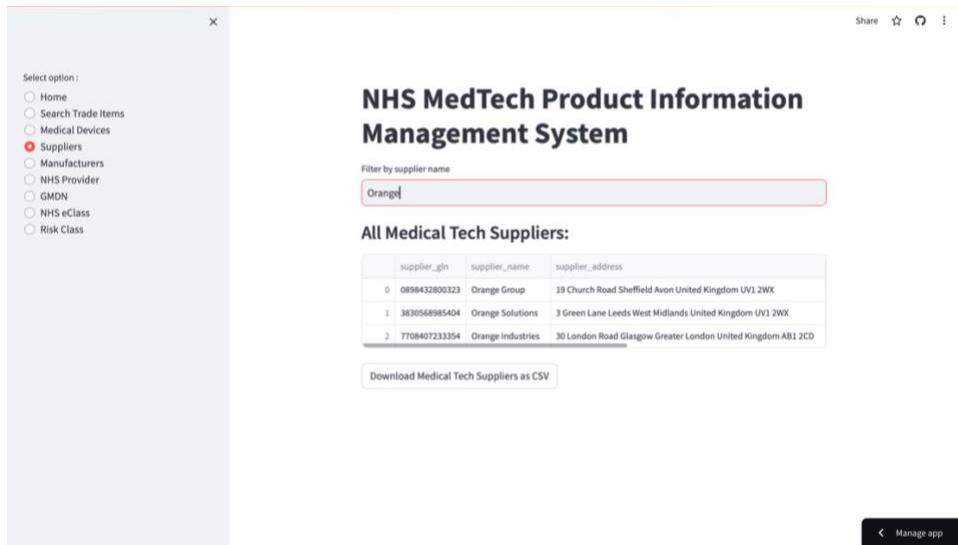


Figure 26: 'Suppliers' Screen and functionalities

4. **'Manufacturers' Screen:** users may access information about all manufacturers of medical devices on the UK market. Users may also search for specific manufacturers by the manufacturer's name. They may use the Manufacturer GLN gotten from this page to search the trade items data to view information about the manufacturer's products that have been supplied to the NHS.

5. **'NHS Provider' Screen:** stakeholders with interest in specific NHS provider trusts may lookup information about these trusts from this page. Users may filter the data by the name of that NHS Provider and get information with which they can search the trade items page to find all the medical devices supplied to that trust.
6. **'GMDN' Screen:** this provides users access to lookup the full dataset of Global Medical Device Nomenclature terms and apply filters to search for specific terms.
7. **'NHS eClass' Screen:** this screen provides users the access to view all NHS product classification codes and their descriptions. Users may also filter the dataset by searching keywords of products with which they get the eClass codes to search the corresponding trade items.
8. **'Risk Class' Screen:** users may look up medical device regulatory risk class information and use the risk class name to search the trade items of medical devices of a specific risk class.

The Medical Devices, Suppliers, Manufacturers, NHS Provider, GMDN, NHS eClass and Risk Class Screens serve the purpose of providing lookup tables for searching the 'Search Trade Items' page.

In addition, all data queries are downloadable in CSV format.

6.4.4 Limitations of local Application server

The application was initially deployed on a local machine, just like the database server. It was quickly realized that this also posed limitations similar to the limitations of the database server on local host.

The major limitation in this case being **limited accessibility**. The front-end application can only be accessed from the local machine on which the Streamlit service is running. Hence, this creates limitations on the accessibility of the system to all other users who need to access product information or test the application remotely, making it difficult to make the app public and gather user feedback.

This limitation creates a potential impact on one of the objectives of this study which is to develop a prototype of the system and evaluate its performance against user requirements. It

became imperative to deploy the application as a web application allowing the system to be accessed by any user through an internet web browser.

6.4.5 Application deployment – Streamlit Community Cloud

Streamlit provides a feature called *Streamlit Community Cloud* which allows users to deploy, manage, and share their Streamlit applications on the web for free. It connects directly to the user's GitHub repository and handles all the containerization, allowing the user to just focus on writing the python code for creating, exploring, and updating the app.

The following steps were taken to deploy the Streamlit to the cloud using *Streamlit Community Cloud*:

1. A Streamlit Community Cloud account was set up after setting up a GitHub account.
2. The Streamlit account was then connected to the GitHub account. This allows users to deploy apps directly from the files stored in connected GitHub repositories.
3. The *nhspim.py* python script was created in the connected GitHub repository by cloning the script from the local machine.
4. A *requirements.txt* file containing a list of the required libraries and app dependencies was included in the repository. These are all the external libraries included in the import statements of the *nhspim.py* python script. This ensures that the deployment environment is properly setup and configured allowing the Python server running on a remote machine have access to the necessary app dependencies.
5. In the *Streamlit Community Cloud* workspace, a new app was created using the connected GitHub repository.
6. A custom subdomain was created for the app URL - <https://nhspim.streamlit.app>
7. **Secrets Management:** it is not secure to store database connection credentials on a public GitHub repository. *Streamlit Community Cloud* provides a feature to store these credentials in a secure file not committed to the repository access them in the application as environment variables. This step is shown in Figure 27 below.

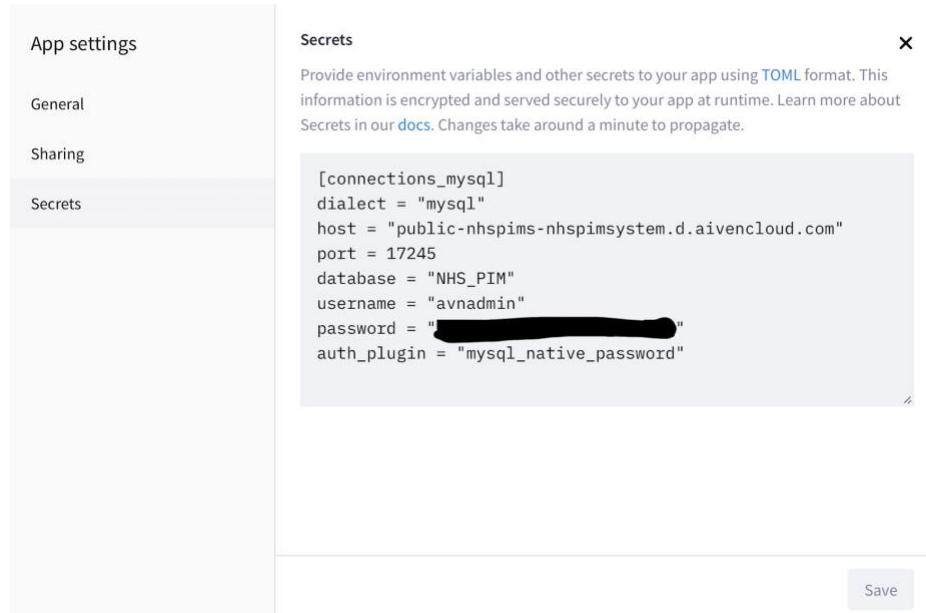


Figure 27: Database Connection Credential Secrets in TOML format

- Connection to database secrets established in *nhspim.py* Streamlit app using `st.secrets` as shown below.

```
# Access MySQL credentials from streamlit secrets
mysql_config = st.secrets["connections_mysql"]

def create_connection():
    """Create connection to MySQL database."""
    db = mysql.connector.connect(
        user=mysql_config["username"],
        password=mysql_config["password"],
        host=mysql_config["host"],
        port=mysql_config["port"],
        database=mysql_config["database"],
        auth_plugin=mysql_config["auth_plugin"]
    )
    return db
```

Figure 28: Python code establishing connection to MySQL database

- The app was deployed successfully to the *Streamlit Community Cloud*.

6.5 Scalability and Optimization Strategies

The following optimization techniques were employed to enhance system performance and scalability.

1. **Cloud Infrastructure:** the use of a cloud database solution enhances the system's performance and scalability through the provision of automated scaling, high availability, and managed infrastructure while ensuring the security and compliance of the system. The use of a cloud application server also ensures accessibility to an increasing number of users.
2. **Storage Engine:** The InnoDB MySQL data storage engine which is the more popular storage engine for high concurrency use-cases, and enabling scalability was used in this study.
3. **Indexing:** The indexing strategies proposed by (Praveena & Chikkamannur, 2021) were followed through the creation of indexes on primary and foreign keys and on columns used in JOINS and WHERE clauses as shown in Figure 29 below

Table	Name	Unique	Index...	Column	Seq in Index	Packed	Collat...
authorized_rep	PRIMARY	Yes	BTREE	rep_id	1	A	
gmdn	PRIMARY	Yes	BTREE	gmdn_code	1	A	
gmdn	gmdn_name	No	BTREE	gmdn_term_name	1	A	
item_model	PRIMARY	Yes	BTREE	gmn	1	A	
manufacturer	PRIMARY	Yes	BTREE	manufacturer_gln	1	A	
manufacturer	man_name_idx	No	BTREE	manufacturer_name	1	A	
manufacturer_catalog	PRIMARY	Yes	BTREE	manufacturer_reference_no	1	A	
manufacturer_catalog	authorized_rep_fk	No	BTREE	authorized_rep_rep_id	1	A	
manufacturer_catalog	manufacturer_fk	No	BTREE	manufacturer_manufacturer_gln	1	A	
medical_device	PRIMARY	Yes	BTREE	gtin	1	A	
medical_device	medical_device_gmdn_fk	No	BTREE	gmdn_code	1	A	
medical_device	medical_device_item_model_fk	No	BTREE	gmn	1	A	
medical_device	medical_device_nhs_product....	No	BTREE	nhs_eiclass_code	1	A	
medical_device	medical_device_manufacturer...	No	BTREE	manufacturer_reference_no	1	A	
medical_device	device_name	No	BTREE	brand_name	1	A	
medical_device	device_risk_fk	No	BTREE	risk_class_name	1	A	
nhs_product_classification	PRIMARY	Yes	BTREE	eclass_code	1	A	
nhs_provider	PRIMARY	Yes	BTREE	provider_gln	1	A	
risk_class	PRIMARY	Yes	BTREE	class_name	1	A	
supplier	PRIMARY	Yes	BTREE	supplier_gln	1	A	
supplier	sup_name_idx	No	BTREE	supplier_name	1	A	
trade_item	PRIMARY	Yes	BTREE	UDI	1	A	
trade_item	trade_item_medical_device_fk	No	BTREE	gtin	1	A	
trade_item	trade_item_nhs_provider_fk	No	BTREE	nhs_provider_gln	1	A	
trade_item	trade_item_supplier_fk	No	BTREE	supplier_gln	1	A	

Figure 29: Indexes on the NHS PIM database

4. **Normalization:** the database and all its tables were normalized to the third normal form (3NF) to reduce data redundancy and improve the database optimization.

5. **Query optimization:** query optimization techniques such as calling out all column names in SELECT statements, using appropriate JOINS, and minimizing redundant data retrieval, etc. were employed to optimize the database system.
6. **Field Optimization:** field optimization methods were adopted such as reducing nullable fields, setting the field length to the maximum allowable field length for standard fields like the GTIN, GLN, GMN, eClass Codes etc., and choosing the appropriate data types for all data fields.

6.6 Data Security and Integrity Measures

Data security and integrity are essential especially in a database system used by a healthcare organization such as the NHS which deals with sensitive personal information. It is critical to ensure the protection and reliability of data.

6.6.1 Data Security

These are the measures that protect an organization's database from deliberate or accidental threats. These measures apply to not just the data in the database, but every component of the database environment i.e., hardware, software, data, and people. The following data security measures were taken in the implementation of the PIM system for the NHS.

- **Encryption:** the cloud database infrastructure ensures data security by implementing encryption mechanisms to protect data both at rest and in transit.
- **Access control:** role-based access controls are used to manage access to database resources ensuring only authorized personnel can access sensitive data. Access controls can be set up using SQL GRANT statement.

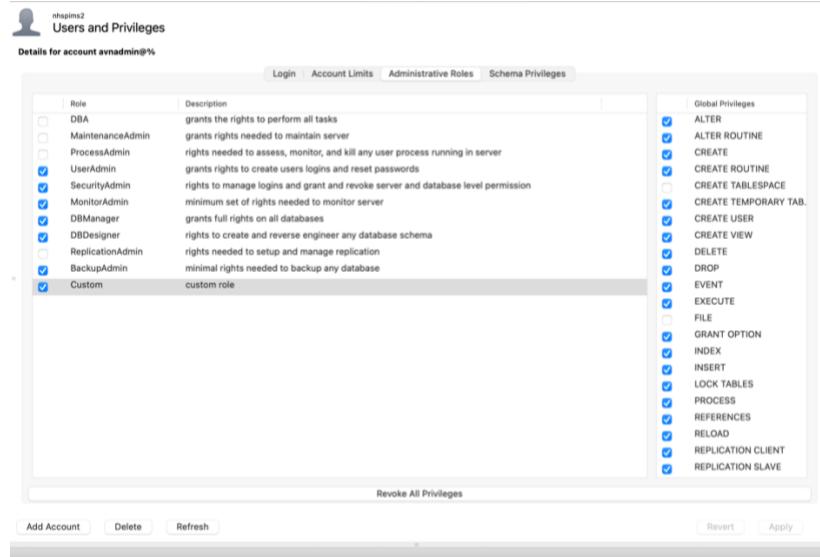


Figure 30: NHS PIM system access control and privileges in MySQL

- **Networking security:** access to the database services is provided over TLS encrypted connections and firewall rules that control connections from IP addresses within internal or external public networks.
- **Backup and recovery:** automated backup and recovery features of the managed database service ensures data security in the event of accidental data loss or system failures.
- **User views:** creation of user views ensures users only create virtual tables based on their queries without altering the underlying database schema.
- **Web application permissions and authentication:** authentication through GitHub is necessary to deploy or administer the web application. Only users with write and administrative access to the GitHub repository can make changes to, deploy, or delete apps. It uses single sign-on, multifactor authentication and strong password policies to protect access to cloud services.
- **Secrets Management:** the secrets management option provided by the Streamlit application allows the database connection credentials to be stored and transmitted to the application server in a secure manner.
- **CA Certificate:** the use of a digital CA certificate establishes a secure connection to the database over the internet.

6.6.2 Data Integrity

Data Integrity refers to the reliability, accuracy and consistency of data stored in a database. It is made up of the following parts:

- **Entity Integrity:** it was ensured that each row in all tables of the database represent a unique record. This was enforced by the use of primary keys (PK) to uniquely identify each record in the table. This means that there are no duplicate or null values in the primary key columns of any of the NHS PIM database records. Ultimately, this ensures the integrity and accuracy of data in the database.
- **Referential Integrity:** this refers to the logical dependency of a foreign key on a primary key. It was ensured through normalization that foreign key values in child tables correspond to primary keys in parent tables. This ensured consistency and validity of relationships between the database tables.
- **Semantic integrity:** it was ensured that data stored in the database reflects the intended meaning of the data. This was done by implementing domain constraints such that the value of a field must be within the allowable range, and implementation of defined business rules.

6.7 Summary of Implementation

The designed NHS PIM system was fully implemented in line with the specifications of the design. A 3-Tier client-server database architecture was adopted, resulting in the implementation of the database, the application server business and data processing logic, and the user interface of the system.

In addition, limitations encountered in the implementation of the database on a local server were mitigated by deploying the database to a fully-managed cloud data platform while similar limitations with the front-end application on a local server were mitigated by deploying the application to a cloud server making the application fully accessible via the web at <https://nhspim.streamlit.app>.

Finally, database optimization, data security and data integrity strategies employed in the study to improve database performance and enhance scalability were discussed.

Chapter 7

System Testing and Evaluation

7.1 Introduction to Evaluation

This chapter presents the user testing and evaluation process for validating the performance and effectiveness of the developed Product Information Management (PIM) system for the NHS against the requirements. It provides a description of the metrics measured and the results obtained.

Testing and evaluation is a critical stage of the system development lifecycle with a purpose to ensure that the system meets all functional requirements and identify any areas of the system which require improvement.

7.2 User Testing and Feedback

Participants were presented with a questionnaire and were asked to evaluate the PIM system application. The system was evaluated based on the users perception of its functionality, speed, response rate, and usability. **10 users participated in the evaluation** and the results of the evaluation are shown below.

1. On a scale of 1-5, how well do you understand the functionality of the system?
{{INSERT VIZ}}
2. Are you able to access the system and view information about medical devices?
3. Are you able to apply the filters to search for product information?
4. Are you able to export your query results using the download button?
5. On a scale of 1 to 5, how would you rate the speed and responsiveness of the system?
6. On a scale of 1 to 5, how easy was it to use the system?
7. Did you find the user interface intuitive and user-friendly?

7.3 Evaluation of Database Performance Metrics

Aiven for MySQL provides a service metrics interface in its console for evaluating the performance and monitor the health of the database service. The interface is able to show admins the performance metrics over time range options of 1 hour, 1 day, 1 week, 1 month, and 1 year. The following section provides an overview of the performance metrics are used for the evaluation of the PIM database system:

- CPU usage:** this metric shows the percentage of CPU capacity being used by the database server. This helps to evaluate the system's ability to handle concurrent users and execute complex queries. Figure 31 shows the CPU % usage is relatively low, indicating the capacity for the system to handle additional workload without performance issues.

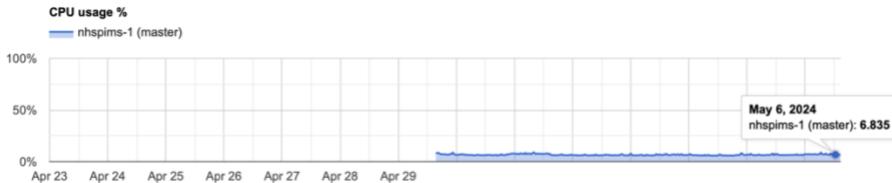


Figure 31: CPU usage

- Disk Space Usage:** this metric shows the percentage of available disk space being used by the database server. This helps evaluate the system's storage capacity and ensures continuous database operation. shows the disk space usage is low, indicating the system will not run out of disk space anytime soon, and ensuring optimized performance and resource utilization.

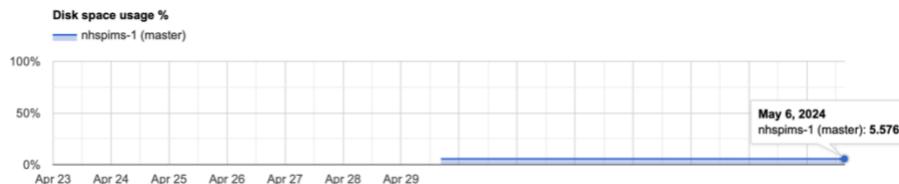


Figure 32: Disk space usage

- Memory Usage:** this metric shows the percentage of total memory being used by the database server. This helps to evaluate the system's memory requirements and identify memory utilization issues. Figure 33 shows the system memory usage is moderate with only about 53% being used at the time of writing this, indicating sufficient opportunity to accommodate additional processes.



Figure 33: Memory Usage

4. **Disk IOPS (reads):** this metric, measured by input/output operations per second IOPS, shows the rate at which the database server reads data from the disk storage. It helps evaluate the efficiency of data retrieval operations on the database. The low disk reads IOPS shown in Figure 34 evidences the read-optimized database design of the system resulting in faster data retrieval.

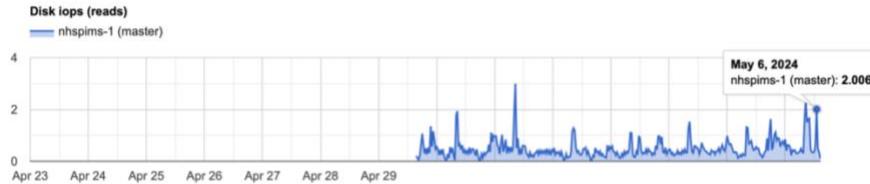


Figure 34: Disk IOPS (reads)

5. **Network receive (bytes/s):** this metric shows the rate at which the database server receives data over the network interface. It helps evaluate network bandwidth utilization. The network receive rate at the time of this report was relatively low as shown in Figure 35, suggesting the system is not overwhelmed by inbound data traffic, resulting in minimal latency.

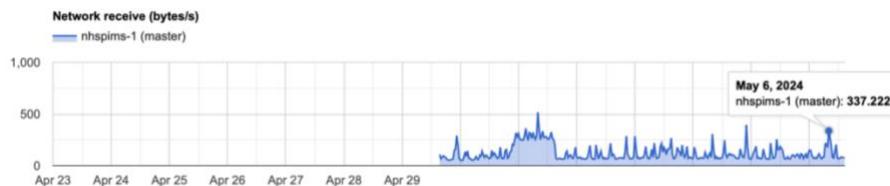


Figure 35: Network receive rate

6. **Network transmit (bytes/s):** this metric shows the rate at which the database server transmits data over the network interface. It also helps evaluate network performance and outbound traffic. The network transmit rate of the system shows a moderate outbound data traffic, allowing for efficient data exchange with external systems.

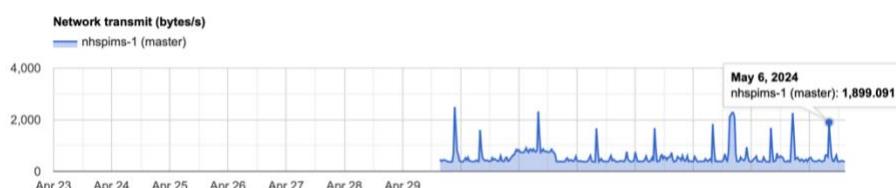


Figure 36: Network transmit rate

7.4 Summary of System Evaluation

- Summary of key findings, insights and conclusions from evaluation phase

- Overall, based on the provided system metrics, the developed PIM system performs well, with low CPU usage, moderate memory usage, and minimal disk IOPS. The network traffic rates are also within acceptable ranges, suggesting that the system is adequately handling inbound and outbound data traffic.
- Summary of recommendations for enhancing system performance, usability and scalability

Chapter 8: Discussion and Conclusion

8.1 Interpretation of the findings in relation to the research objectives.

8.2 Discussion of implications, limitations, and future directions.

Use of relational limits the db to only relational, structured data. future research should look into the implementation of NoSQL databases

These limitations of methodology are acknowledged and addressed in the interpretation and discussion of research findings.

8.3 Summary of key findings and contributions of the thesis.

8.4 Recommendations for practice, policy, or further research.

Evaluating Database Management Systems: A Framework And Application To The Veteran's Administration Hospital <https://dspace.mit.edu/bitstream/handle/1721.1/61034/06564848-MIT.pdf;sequence=2>

Database Management System for Student Admissions

https://soar.suny.edu/bitstream/handle/20.500.12648/10380/Savelios%20Aslanidis%20-%20Thesis%20Project_DBMS%20for%20Student%20Admissions.pdf?sequence=1&isAllowed=y

DESIGN AND IMPLEMENTATION OF SPECIAL EDUCATION APPS INFORMATION MANAGEMENT SYSTEM

https://etd.ohiolink.edu/acprod/odb_etd/ws/send_file/send?accession=kent1492704386514278&disposition=inline

An information management system for a large-scale biological collaboration

https://trace.tennessee.edu/cgi/viewcontent.cgi?article=6723&context=utk_gradthes

Web-based forensic information management system

<https://researchrepository.wvu.edu/cgi/viewcontent.cgi?article=3386&context=etd>

Development of a Database-Driven Management System for Retail Food Packaging Eye Tracking Studies

https://tigerprints.clemson.edu/all_dissertations/2171/

design and development of web based biomed dbms for health services multispecialty

https://web.archive.org/web/20210428045611id_/_https://jusst.org/wp-content/uploads/2020/12/Design-and-Development.pdf

Evaluation Of Performance of NoSQL Database Management System: -MongoDB And Relational Database Management System: - PostgreSQL for Efficient Management Of Healthcare Data

<http://sifisheressciences.com/journal/index.php/journal/article/view/2331>

DESIGN AND IMPLEMENTATION OF A WEB-BASED HEALTH CARE RECORDS MANAGEMENT SYSTEM

<http://ir.mtu.edu.ng/xmlui/bitstream/handle/123456789/178/OIBO%20JOY%20OPEYE MI%20-%20FULL%20PROJECT.pdf?sequence=1&isAllowed=y>

Database System for Medical Record Keeping and Retrieval

<http://jae-tech.com/index.php/jaet/article/view/101/86>

Implementation of product data management system

<https://oulurepo.oulu.fi/bitstream/handle/10024/15456/nbnfiolu-202009162929.pdf?sequence=1&isAllowed=y>

EVALUATION OF DATABASE MANAGEMENT SYSTEMS – PostgreSQL better than MySQL & SQLite in terms of speed and info retrieval

<https://hh.diva-portal.org/smash/get/diva2:367006/FULLTEXT01.pdf>

Design and Implementation of a Web Shop System

https://www.theseus.fi/bitstream/handle/10024/22377/Shen_Yeyin.pdf?sequence=1&isAllowed=y

Analysis and Design of Information Systems

By Arthur M. Langer

https://books.google.co.uk/books?hl=en&lr=&id=fHZBQZkp-TYC&oi=fnd&pg=PR2&dq=design+of+information+management+system&ots=zszXu4RAdr&sig=lOBhWdu0FQwBFPce-mknLWjkivw&redir_esc=y#v=onepage&q=design%20of%20information%20management%20system&f=false

<http://ndl.ethernet.edu.et/bitstream/123456789/55870/1/44%202014.pdf> ABRAHAM 2014

The digitalization journey of PIM (Battistello, 2020)

https://www.utupub.fi/bitstream/handle/10024/153916/Nurminen_Arttu_opinnayte.pdf?sequence=1 NURMINEN, 2022

Classification standards

Classification standards support

- procurement intelligence, by enabling similar items to be grouped together for spend analysis and demand aggregation;
- the traceability requirements of regulatory agencies.
 - Work will be undertaken at the national level to determine the most appropriate procurement classification standards for use by the NHS. The outcome of this work may be to maintain the status quo or to recommend an alternative approach. A likely outcome will be different classification standards for different categories. (eProcurement strategy)
 -
 - https://scan4safety.nhs.uk/wp-content/uploads/2022/11/gs1_uk_inventory_management_systems_guidance-1.pdf inventory management system implementation guidance scan4safety
 -
 -

- Data/Attribute selection: proper master data management entails an organization identifying and understanding data that is most meaningful to the business. Quality over quantity. (Nurminen, 2022)
- Master data vs transactional data: master does not change as often as transactional which changes continuously. (Nurminen, 2022) (talk about not using price data since its dynamic i.e. transactional)

Evaluation of existing processes, technologies, and systems within NHS

- Overall, it is a burden for the health system to provide, manage, and use medical product information. The inaccurate and incomplete data collected poses a risk to patient safety and effective procurement and administrative processes. There is a risk that manufacturers will not correct or verify the quality of their data which would restrict the usefulness of the proposed solution as trusts would have limited reason to trust and use the data.
- https://www.gs1uk.org/sites/default/files/MDE_Demonstration_of_Technology_Case_Study.pdf master data exchange demonstration of technology - PREPARATION
 - https://scan4safety.nhs.uk/wp-content/uploads/2022/11/Scan4Safety_Product_How_To_Guide-1.pdf
 - scan4safety/ catalogue mgt/product information
 - <https://scan4safety.nhs.uk/how-to-get-started/first-steps-for-suppliers/>
 - supplier how to get started
 - https://www.gs1uk.org/sites/default/files/GS1_UK_Healthcare_Terms_2024.pdf

<file:///Users/ayokunlejames/Downloads/4004202021guidance.pdf>

<https://www.imdrf.org/sites/default/files/docs/imdrf/final/technical/imdrf-tech-131209-udi-guidance-140901.pdf>

([https://eur-](https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2013:099:0017:0024:EN:PDF)

[lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2013:099:0017:0024:EN:PDF](https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2013:099:0017:0024:EN:PDF))

Bibliography

Abraham, J., 2014. *Product Information Management: Theory and Practice*. Cham: Springer International Publishing.

Ahmad, R., Kyratsis, Y. & Holmes, A., 2012. When the user is not the chooser: learning from stakeholder involvement in technology adoption decisions in infection control. *The journal of Hospital Infection*, 81(3), pp. 163-168.

Šušter, I. & Ranisavljević, T., 2023. Optimization of MySQL Database. *Journal of Process Management and New Technologies*, 11(1-2), pp. 141-151.

Battistello, L., 2020. *The digitalization journey of product information management*, Lyngby: Management Science.

Battistello, L., Haug, A., Suzic, N. & Hvam, L., 2021. Implementation of product information management systems: Identifying the challenges of the scoping phase. *Computers in Industry*, 133(103533).

Boulding, H. & Hinrichs-Krapels, S., 2021. Factors influencing procurement behaviour and decision-making: an exploratory qualitative study in a UK healthcare provider. *BMC Health Services Research*, 21(1087).

Carter, P., 2016. *Operational productivity and performance in English NHS acute hospitals: Unwarranted variations*. [Online]

Available at:

https://assets.publishing.service.gov.uk/media/5a80bd5ae5274a2e87dbb8f5/Operational_productivity_A.pdf

[Accessed 23 March 2024].

Commercial Division, DHSC, 2017. *GS1 and PEPPOL Adoption Case Study: Master Data Exchange Demonstration of Technology*. [Online]

Available at:

https://www.gs1uk.org/sites/default/files/MDE_Demonstration_of_Technology_Case_Study.pdf

[Accessed 2 April 2024].

Connolly, T. & Begg, C., 2015. *Database Systems: A Practical Approach to Design, Implementation, and Management*. 6th Edition ed. Harlow(Essex): Pearson education Limited.

Coronel, C. & Morris, S., 2016. *Database Systems: Design, Implementation, and Management*. 12th Edition ed. Boston(Massachusetts): Cengage Learning.

Crown Commercial Service, 2024. *Procurement details: DHSC: Medical Technology - Product Information Management Data Platform - Alpha phase*. [Online]

Available at: <https://redirect.contractawardservice.crowncommercial.gov.uk/digital-outcomes/opportunitiesDetails?projectId=38863&lot=1>

[Accessed 2 April 2024].

Davies, G., 2024. *NHS Supply Chain and efficiencies in procurement: NHS England, NHS Supply Chain*, London: National Audit Office.

Department of Health & Social Care, 2023. *Medical Technology Strategy*. [Online]

Available at: <https://assets.publishing.service.gov.uk/media/63dbe1f68fa8f57fbfff3db3/medical-technology-strategy.pdf>

[Accessed March 2024].

Department of Health & Social Care, 2018. *A Trust's guide to Product and catalogue management*. [Online]

Available at: https://scan4safety.nhs.uk/wp-content/uploads/2022/11/Scan4Safety_Product_How_To_Guide-1.pdf

[Accessed 2 April 2024].

Department of Health, 2013. *Better Procurement Better Value Better Care: A Procurement Development Programme for the NHS*. [Online]

Available at:

https://assets.publishing.service.gov.uk/media/5a7c267840f0b61a825d6c48/procurement_development_programme_for_NHS.pdf

[Accessed 1 April 2024].

Dhasmana, G., Gujjar, P. J., Prasad, G. M. & Kumar, P. H., 2023. *SQL and NOSQL Databases in the Application of Business Analytics*. Bangalore, 2023 International Conference on Computer Science and Emerging Technologies (CSET), pp. 1-5.

Edel, W. & Sutedja, I., 2023. Master Data Management Analysis For Today's Company: A Literature Review System. *Journal of Theoretical and Applied Information Technology*, 30 April.101(8).

Essig, M. & Arnold, U., 2001. Electronic Procurement in Supply Chain Management: An Information Economics-Based Analysis of Electronic Markets. *The Journal of Supply Chain Management*, September, 37(4), pp. 43-49.

European Commission, 2013. COMMISSION RECOMMENDATION of 5 April 2013 on a common framework for a unique device identification system of medical devices in the Union. *Official Journal of the European Union*, 9 April, 172(99), pp. 17-24.

GMDN, 2023. *DAPB4004 Global Medical Device Nomenclature (GMDN) for medical device identification*, s.l.: GMDN Agency Ltd.

Google Cloud, 2024. *What is a Cloud Database*. [Online]

Available at: cloud.google.com/learn/what-is-a-cloud-database

[Accessed 30 April 2024].

GS1 UK, 2021. *GS1 UK compliance specification for the NHS*, s.l.: GS1 UK.

Györödi, C. A. et al., 2021. Performance Impact of Optimization Methods on MySQL Document-Based and Relational Databases. *Applied Sciences*, 11(15).

IMDRF UDI Working Group, 2013. *UDI Guidance Unique Device Identification (UDI) of Medical Devices*. [Online]

Available at: <https://www.imdrf.org/sites/default/files/docs/imdrf/final/technical/imdrf-technote-131209-udi-guidance-140901.pdf>

[Accessed April 2024].

Langer, A. M., 2008. *Analysis and Design of Information Systems*. 3rd Edition ed. London: Springer Science & Business Media.

- Maesaroh, S. et al., 2022. Query Optimization in MySQL Database Using Index. *International Journal of Cyber and IT Service Management (IJCITS)*, 2(2), pp. 104-110.
- Mahajan, D., Blakeney, C. & Zong, Z., 2019. Improving the energy efficiency of relational and NoSQL databases via query optimizations. *Sustainable Computing: Informatics and Systems*, June, Volume 22, pp. 120-133.
- Medicines and Healthcare products Regulatory Agency, 2020. *Register medical devices to place on the market*. [Online]
- Available at: <https://www.gov.uk/guidance/register-medical-devices-to-place-on-the-market> [Accessed April 2024].
- Nurminen, A., 2022. *Master data management in industry*, Turku: University of Turku.
- Office for Health Improvement and Disparities, 2020. *Mixed methods study*. [Online]
- Available at: <https://www.gov.uk/guidance/mixed-methods-study#:~:text=A%20mixed%20methods%20study%20combines,with%20more%20in%2Ddepth%20findings>.
- [Accessed 5 April 2024].
- Olagunju, A. J., 2023. *Humanitarian Organization Information System*, Wolverhampton: University of Wolverhampton.
- Oppel, A. J., 2009. The Database Life Cycle. In: *Databases: A Beginner's Guide*. Chicago: McGraw-Hill.
- Palmer, B., 2024. *What is Product Information?*. [Online]
- Available at: <https://www.plytix.com/blog/what-is-product-information> [Accessed March 2024].
- Pansara, R., 2021. Master Data Management Challenges. *International Journal of Computer Science and Mobile Computing*, October, 10(10), pp. 47-49.
- Praveena, M. V. & Chikkamannur, A. A., 2021. Indexing Strategies for Performance Optimization of Relational Databases. *International Research Journal of engineering and Technology (IRJET)*, 08(05).
- Procurement, Investment & Commercial Division, DHSC, 2014. *NHS eProcurement Strategy*. [Online]
- Available at:
- https://assets.publishing.service.gov.uk/media/5a7ebfa3e5274a2e8ab47f34/NHS_eProcurement

Strategy.pdf

[Accessed March 2024].

Savelios, A., 2021. *Database Management System for Student Admissions*, New York: SUNY Polytechnic Institute.

Sheldon, P. & Goetz, M., 2014. *The Forrester WaveTM: Product Information Management (PIM), Q2 2014*, Cambridge, MA: Forrester Research, Inc..

Simplilearn, 2023. *What is Requirement Analysis*. [Online]

Available at: <https://www.simplilearn.com/what-is-requirement-analysis-article>

[Accessed March 2024].

Spencer, A. & Patel, S., 2019. Applying the Data Protection Act 2018 and General Data Protection Regulation principles in healthcare settings. *Nursing management*, 28 January, 26(1), pp. 34-40.

Steer-Stephenson, C., 2022. *Procurement strategies in the healthcare industry*. [Online]

Available at: <https://procurementmag.com/procurement-strategy/procurement-strategies-in-the-healthcare-industry>

[Accessed 8 March 2024].

Vedapudi, M., 2000. *Requirements for a Product Information Management (PIM) Infrastructure to support partner programs*, Massachusetts: Massachusetts Institute of Technology (MIT).

Wikipedia, 2023. *Database scalability*. [Online]

Available at: https://en.wikipedia.org/wiki/Database_scalability

[Accessed 2 April 2024].

Wyatt, J. C., 1995. Hospital information management: the need for clinical leadership. *The BMJ*, 15 July, 311(175).

Appendices

- Supplementary materials such as raw data, survey, or technical documentation.
- Evaluation Form
- Data Dictionary MUST
- link to github

Appendix A

- Data Dictionary MUST

Appendix A PIM System User Evaluation Form



UNIVERSITY OF
WOLVERHAMPTON

NHS MedTech Product Information Management System User Evaluation

PIM system Evaluation Form

jamesayokunle03@gmail.com [Switch account](#)



Not shared

Informed Consent

Dear Respondent,

My name is Ayokunle Olagunju, an MSc. Data Science student at the University of Wolverhampton. I am working on my Master's thesis titled "*The Design and Optimization of a Scalable National Product Information Management System for the NHS*".

I would appreciate if you could contribute to the research by completing this questionnaire which will help enhance the quality of my research work. It will take less than 1 minute of your time.

Please rest assured that your response will be handled with utmost confidentiality and will be used strictly for the purpose of this research.

Thank you in anticipation.

[Next](#)

[Clear form](#)

Never submit passwords through Google Forms.

This content is neither created nor endorsed by Google. [Report Abuse](#) - [Terms of Service](#) - [Privacy Policy](#)

Section A

System Functionality

Access the system [here](#). Explore the system's functionalities by going through the system home page and using the access guide for instructions.



How well do you understand the functionality of the system? *

- 1 - Not At All
- 2 - Somewhat
- 3 - Moderately
- 4 - Mostly
- 5 - Completely

Are you able to access the system and view information about medical devices? *

- Yes
- No

Are you able to apply the filters to search for product information? *

- Yes
- No

Are you able to export your query results using the download button? *

- Yes
- No

Back

Next

Clear form

Section B

System Performance

On a scale of 1 to 5, how would you rate the speed and responsiveness of the system? *

- 1 - Very Slow and Unresponsive
- 2 - Slow and Unresponsive
- 3 - Neutral
- 4 - Fast and Responsive
- 5 - Very Fast and Responsive

Back

Next

Clear form

Section C

Usability

On a scale of 1 to 5, how easy was it to use the system? *

- 1 - Very difficult to use
- 2 - Somewhat difficult to use
- 3 - Moderately
- 4 - Somewhat easy to use
- 5 - Very easy to use

Did you find the user interface intuitive and user-friendly? *

- Yes
- No

Please provide any suggestions for improving the user interface

Your answer

Do you have any additional feedback or suggestions for improving the PIM system?
Please provide your comments below.

Your answer

Back

Submit

Clear form

Never submit passwords through Google Forms.

This content is neither created nor endorsed by Google. [Report Abuse](#) - [Terms of Service](#) - [Privacy Policy](#)

Appendix A

FOI Request to the Department of Health & Social Care

Freedom of Information Request: GS1 Certified Datapools Attributes for NHS National PIM



OA

✓ Olagunju, Ayokunle J. <A.J.Olagunju@wlv.ac.uk>

Monday, February 26, 2024 at 11:05 AM

To: dhsc.publicenquiries@dhsc.gov.uk; Bcc: jamesayokunle03@gmail.com

Dear Department of Health and Social Care,

I hope this message finds you well. My name is James, and I am a master's degree (MSc) student at the University of Wolverhampton conducting research on the integration of GS1 certified data pools with, and the creation of the National Product Information Management (PIM) system within the NHS.

I am reaching out to request information pertaining to the set of attributes required for suppliers to create GS1 certified data pools for seamless integration with the NHS National PIM. As this information is critical for my research, I kindly request access to the relevant documents or guidelines available on the Department of Health and Social Care (DHSC) workspace.

Specifically, I am seeking details on the data attributes or data standards that suppliers must adhere to when setting up GS1 certified data pools to ensure compatibility and interoperability with the NHS National PIM system. This includes any documentation, specifications, or guidelines provided by DHSC to assist suppliers in this process.

Furthermore, I am interested in any additional information or resources available on the DHSC workspace related to the integration of GS1 standards with the NHS National PIM, including best practices, case studies, or implementation guides.

I understand that this request falls under the Freedom of Information Act (FOIA), and I would appreciate your assistance in providing access to the requested information within the statutory timeframe. If any fees are associated with processing this request, please inform me in advance.

I assure you that any information provided will be used solely for research purposes and will be handled in accordance with data protection regulations. I am committed to maintaining confidentiality and respecting the sensitivity of the information provided.

Thank you for your attention to this matter. I look forward to your prompt response.

Yours sincerely,

Ayokunle James Olagunju,
MSc Data Science,
University of Wolverhampton, UK.
[+44\(0\)7901782906](tel:+4407901782906)
A.J.Olagunju@wlv.ac.uk

Appendix B

Response to FOI Request from DHSC



**Department
of Health &
Social Care**

Mr Ayokunle James Olagunju
By email to: A.J.Olagunju@wlv.ac.uk

Freedom of Information Team
Department of Health and Social Care
39 Victoria Street
London SW1H 0EU
www.gov.uk/dhsc

26 March 2024

Dear Mr Olagunju,

Freedom of Information Request Reference FOI-1498477

Thank you for your request dated 26 February to the Department of Health and Social Care (DHSC), a copy of which can be found in the accompanying annex.

Your request has been handled under the Freedom of Information Act 2000 (FOIA).

DHSC does not hold the information you have requested.

However, you may wish to contact NHS England (NHSE), which may hold information relevant to your request. NHSE can be contacted at: england.contactus@nhs.net.

Please write "Freedom of Information" in the subject line of any FOI request.

If you are not satisfied with the handling of this request, you have the right to appeal by asking for an internal review. This should be sent to freedomofinformation@dhsc.gov.uk or to the address at the top of this letter and be submitted within two months of the date of this letter.

Please remember to quote the reference number above in any future communication.

If you are not content with the outcome of your internal review, you may complain directly to the Information Commissioner's Office (ICO). Generally, the ICO cannot make a decision unless you have already appealed our original response and received our internal review decision. You should raise your concerns with the ICO within three months of your last meaningful contact with us.

Guidance on contacting the ICO can be found at <https://ico.org.uk/global/contact-us> and information about making a complaint can be found at <https://ico.org.uk/make-a-complaint>.

Yours sincerely,

Freedom of Information Team
freedomofinformation@dhsc.gov.uk

Appendix C

FOI Request to NHS England

Freedom of Information Request: Data Attributes for NHS National PIM system

😊 ⏪ ⏴ ⏵



● Olagunju, Ayokunle J. <A.J.Olagunju@wlv.ac.uk>

Wednesday, March 27, 2024 at 3:15 PM

To: england.contactus@nhs.net

❗ This message is high priority.

Dear NHSE,

I hope this message finds you well. My name is James Olagunju, and I am a master's degree (MSc in Data Science) student at the University of Wolverhampton conducting research on the integration of GS1 certified data pools with, and the creation of the National Product Information Management (PIM) system within the NHS.

I am reaching out to request information pertaining to the **core set of data attributes or data fields** required for the creation of the NHS National PIM. As this information is critical for my research, I kindly request access to the relevant documents or guidelines containing this information.

Furthermore, I am interested in any available additional information or resources related to the integration of GS1 standards with the NHS National PIM, including best practices, case studies, or implementation guides.

I made a FOIA request (Reference FOI – 1498477) to the Department of Health and Social Care (DHSC), and I was informed that they do not hold this information, but that the NHS England might.

I understand that this request falls under the Freedom of Information Act (FOIA), and I would appreciate your assistance in providing access to the requested information within the statutory timeframe. If any fees are associated with processing this request, please inform me in advance.

I assure you that any information provided will be used solely for research purposes and will be handled in accordance with data protection regulations. I am committed to maintaining confidentiality and respecting the sensitivity of the information provided.

Thank you for your attention to this matter. I look forward to your prompt response.

Yours sincerely,

Ayokunle James Olagunju,
MSc Data Science,
University of Wolverhampton, UK.
[+44\(0\)7901782906](tel:+44(0)7901782906)
A.J.Olagunju@wlv.ac.uk

Appendix B

Discovery Phase report access request to Crown Commercial Service

Request for Documents: National Product Information System for the NHS

OA  Olagunju, Ayokunle J. <A.J.Olagunju@wlv.ac.uk>

To: info@crowncommercial.gov.uk

Tuesday, March 12, 2024 at 5:03 PM

Screenshot 2024-03... 989.1 KB

Download • Preview

Dear Crown Commercial Service,

I hope this email finds you well. I am writing to request the following documents referenced on your [website](#) regarding the National Product Information System for the NHS (*attached is an image of the exact webpage*):

1. Discovery Report: This document provides additional detail of the findings from your discovery work to complement the material included in the advert. It describes the problem, aims of the proposed solution, your view of long-term success, outputs expected from alpha, user stories, and learnings from others.
2. Intro to PIM Slides: These slides are historically used for onward sharing to new and interested stakeholders to briefly summarize the key features of the proposed PIM, including reference to the data fields.

These documents are essential for my ongoing research into the implementation and optimization of the National Product Information System for the NHS.

I have spoken with one of your customer service representatives, Dal, who advised me to reach out via email to request these documents.

Would it be possible for you to provide these documents at your earliest convenience? Your assistance in this matter would be greatly appreciated and invaluable to my research.

Thank you very much for your attention to this request. I look forward to your prompt response.

Best regards,

Ayokunle James Olagunju
MSc Data Science
University of Wolverhampton

Appendix C

Response from Crown Commercial Service

RE: Request for Documents: National Product Information System for the NHS

😊 ⏪ ⏴ ⏵



○ CCS Enquiries Service Desk <info@crowncommercial.gov.uk>

Thursday, March 14, 2024 at 9:31 AM

To: Olagunju, Ayokunle J.



To protect your privacy, some external images in this message were not downloaded.

[Download external images](#)

[Go to Settings](#)

Good morning Ayokunle,

Thank you for your email. Your case has been transferred over to me as a member of the DOS team.

Apologies, these documents will be held by the buyer for the linked opportunity and shared as part of the tendering process.

We at CCS do not have access to these documents, and as a result are unable to share them with you.

In order to apply for this opportunity, and view the files that are attached, you will need to be an approved / registered DOS 6 supplier.

I hope this helps, please let me know if there's anything else you need.

Kind regards,

Joel Stevenson
Pronouns - he / him
Commercial Support - Digital Capability and Delivery
Crown Commercial Service
Rosebery Court, St Andrews Business Park, Norwich, NR7 0HS
Customer Service Centre: info@crowncommercial.gov.uk 0345 410 2222

<https://url6.mailanyone.net/scanner?m=1rkhQw-00082C-3u&d=4%7Cmail%2F90%2F1710408600%2F1rkhQw-00082C-3u%7Cin6p%7C57e1b682%7C10977208%7C9441127%7C65F2C3F244637346972ACB74E99EFC69&o=w.wwcronwcaeom.cilmrku.vog&s=CW9kEYjPgTPqrK4-A5j0XZVlsM0> | follow us on Twitter | connect with us on LinkedIn

----- Original Message -----

From: Olagunju, Ayokunle J. [a.j.olagunju@wlv.ac.uk]
Sent: 12/03/2024 17:03

Appendix C

Tables, Attributes, Data Types and Null constraints of the NHS_PIM database

Table	Column	Type	Default Value	Nullable
authorized_rep	rep_id	varchar(32)		NO
authorized_rep	rep_name	varchar(32)		NO
authorized_rep	contact_number	char(11)		NO
authorized_rep	email	varchar(32)		YES
gmdn	gmdn_code	varchar(32)		NO
gmdn	gmdn_term_name	varchar(32)		NO
gmdn	gmdn_term_definition	text		YES
item_model	gmn	char(25)		NO
item_model	model_name	varchar(32)		NO
item_model	device_type	varchar(32)		NO
item_model	market_availability_date	date		YES
item_model	lifecycle_status	varchar(32)		NO
item_model	last_status_update	date		YES
item_model	risk_class_class_name	varchar(32)		NO
manufacturer	manufacturer_gln	char(13)		NO
manufacturer	manufacturer_name	varchar(32)		NO
manufacturer	manufacturer_address	text		NO
manufacturer	company_registration_no	varchar(32)		NO
manufacturer	customer_service_phone	char(11)		YES
manufacturer	customer_service_email	varchar(32)		YES
manufacturer_catalog	manufacturer_reference_no	varchar(32)		NO
manufacturer_catalog	product_name	varchar(32)		NO
manufacturer_catalog	product_model	varchar(32)		NO
manufacturer_catalog	product_category	varchar(32)		YES
manufacturer_catalog	market_availability_date	date		YES
manufacturer_catalog	lifecycle_status	varchar(32)		NO
manufacturer_catalog	last_status_update	date		YES
manufacturer catalog	manufacturer manufacturer aln	char(13)		NO

Table	Column	Type	Default Value	Nullable
manufacturer_catalog	manufacturer_manufacturer_gin	char(13)		NO
manufacturer_catalog	authorized_rep_rep_id	varchar(32)		NO
manufacturer_catalog	risk_class_class_name	varchar(32)		NO
medical_device	gtin	char(14)		NO
medical_device	brand_name	varchar(32)		NO
medical_device	unit_of_use	varchar(32)		YES
medical_device	quantity_of_uou	int		YES
medical_device	item_length	double		YES
medical_device	item_height	double		YES
medical_device	item_width	double		YES
medical_device	item_weight	double		YES
medical_device	item_volume	double		YES
medical_device	unit_of_dimension	char(2)		YES
medical_device	product_description	text		YES
medical_device	storage_handling	text		YES
medical_device	single_use	varchar(3)		NO
medical_device	restricted_no_of_use	int		YES
medical_device	sterile	varchar(3)		NO
medical_device	sterilize_before_use	varchar(3)		NO
medical_device	sterilization_method	text		YES
medical_device	item_contains_latex	varchar(3)		YES
medical_device	item_contains_dehp	varchar(3)		YES
medical_device	item_mri_compatible	varchar(3)		YES
medical_device	nhs_eclass_code	varchar(32)		NO
medical_device	gmn	char(25)		NO
medical_device	gmdn_code	varchar(32)		NO
medical_device	manufacturer_reference_no	varchar(32)		YES
nhs_product_classific...	eclass_code	varchar(32)		NO

Table	Column	Type	Default Value	Nullable
medical_device	gtin	varchar(32)		NO
medical_device	manufacturer_reference_no	varchar(32)		YES
nhs_product_classific...	eclasse_code	varchar(32)		NO
nhs_product_classific...	description	text		NO
nhs_provider	provider_gln	char(13)		NO
nhs_provider	provider_name	text		NO
nhs_provider	provider_address	text		NO
nhs_provider	provider_registration_no	varchar(32)		NO
risk_class	class_name	varchar(32)		NO
risk_class	class_description	text		YES
risk_class	regulatory_requirements	text		YES
supplier	supplier_gln	char(13)		NO
supplier	supplier_name	varchar(32)		NO
supplier	supplier_address	text		NO
supplier	company_registration_no	varchar(32)		NO
supplier	customer_service_phone	char(11)		YES
supplier	customer_service_email	varchar(32)		YES
trade_item	UDI	varchar(32)		NO
trade_item	gtin	char(14)		NO
trade_item	serial_number	varchar(32)		YES
trade_item	batch_number	varchar(32)		YES
trade_item	manufacturing_date	date		YES
trade_item	expiry_date	date		YES
trade_item	udi_pi	varchar(32)		NO
trade_item	unit_of_issue	varchar(32)		YES
trade_item	unit_of_use_udi	char(14)		NO
trade_item	supplier_gln	char(13)		NO
trade_item	nhs_provider_gln	char(13)		NO

