

# Software Requirements Specification for Scanalyze AI: AI-Powered Diagnostic System for Chest X-Rays

Team 16, Ace  
Harrison Chiu  
Hamza Issa  
Ahmad Hamadi  
Jared Paul  
Gurnoor Bal

April 4, 2025

# Contents

<b>1</b>	<b>Purpose of the Project</b>	<b>4</b>
1.1	User Business . . . . .	4
1.2	Goals of the Project . . . . .	4
<b>2</b>	<b>Stakeholders</b>	<b>5</b>
2.1	Client . . . . .	5
2.2	Customer . . . . .	5
2.3	Other Stakeholders . . . . .	5
2.4	Hands-On Users of the Project . . . . .	6
2.5	Personas . . . . .	8
2.6	Priorities Assigned to Users . . . . .	8
2.7	User Participation . . . . .	8
2.8	Maintenance Users and Service Technicians . . . . .	9
<b>3</b>	<b>Mandated Constraints</b>	<b>9</b>
3.1	Solution Constraints . . . . .	9
3.2	Implementation Environment of the Current System . . . . .	10
3.3	Partner or Collaborative Applications . . . . .	11
3.4	Off-the-Shelf Software . . . . .	11
3.5	Anticipated Workplace Environment . . . . .	12
3.6	Schedule Constraints . . . . .	13
3.7	Budget Constraints . . . . .	14
3.8	Enterprise Constraints . . . . .	14
<b>4</b>	<b>Naming Conventions and Terminology</b>	<b>14</b>
4.1	Glossary of All Terms, Including Acronyms, Used by Stakeholders involved in the Project . . . . .	14
<b>5</b>	<b>Relevant Facts And Assumptions</b>	<b>17</b>
5.1	Relevant Facts . . . . .	17
5.2	Business Rules . . . . .	17
5.3	Assumptions . . . . .	18
<b>6</b>	<b>The Scope of the Work</b>	<b>19</b>
6.1	The Current Situation . . . . .	19
6.2	The Context of the Work . . . . .	20
6.3	Work Partitioning . . . . .	20

6.4	Specifying a Business Use Case (BUC)	22
<b>7</b>	<b>Business Data Model and Data Dictionary</b>	<b>23</b>
7.1	Business Data Model	23
7.2	Data Dictionary	23
<b>8</b>	<b>The Scope of the Product</b>	<b>24</b>
8.1	Product Boundary	24
8.2	Product Use Case Table	25
8.3	Individual Product Use Cases (PUC's)	25
<b>9</b>	<b>Functional Requirements</b>	<b>27</b>
9.1	Functional Requirements	27
<b>10</b>	<b>Look and Feel Requirements</b>	<b>30</b>
10.1	Appearance Requirements	30
10.2	Style Requirements	30
<b>11</b>	<b>Usability and Humanity Requirements</b>	<b>31</b>
11.1	Ease of Use Requirements	31
11.2	Personalization and Internationalization Requirements	32
11.3	Learning Requirements	32
11.4	Understandability and Politeness Requirements	32
11.5	Accessibility Requirements	32
<b>12</b>	<b>Performance Requirements</b>	<b>33</b>
12.1	Speed and Latency Requirements	33
12.2	Safety-Critical Requirements	33
12.3	Precision or Accuracy Requirements	33
12.4	Robustness or Fault-Tolerance Requirements	34
12.5	Capacity Requirements	34
12.6	Scalability or Extensibility Requirements	35
12.7	Longevity Requirements	35
<b>13</b>	<b>Operational and Environmental Requirements</b>	<b>35</b>
13.1	Expected Physical Environment	35
13.2	Wider Environment Requirements	36
13.3	Requirements for Interfacing with Adjacent Systems	36
13.4	Productization Requirements	36

13.5 Release Requirements . . . . .	37
<b>14 Maintainability and Support Requirements</b>	<b>37</b>
14.1 Maintenance Requirements . . . . .	37
14.2 Supportability Requirements . . . . .	38
14.3 Adaptability Requirements . . . . .	38
<b>15 Security Requirements</b>	<b>38</b>
15.1 Access Requirements . . . . .	38
15.2 Integrity Requirements . . . . .	39
15.3 Privacy Requirements . . . . .	39
15.4 Audit Requirements . . . . .	39
15.5 Immunity Requirements . . . . .	40
<b>16 Cultural Requirements</b>	<b>40</b>
16.1 Cultural Requirements . . . . .	40
<b>17 Compliance Requirements</b>	<b>40</b>
17.1 Legal Requirements . . . . .	40
17.2 Standards Compliance Requirements . . . . .	41
<b>18 Requirement Traceability</b>	<b>41</b>
<b>19 Open Issues</b>	<b>43</b>
<b>20 Off-the-Shelf Solutions</b>	<b>44</b>
20.1 Ready-Made Products . . . . .	44
20.2 Reusable Components . . . . .	44
20.3 Products That Can Be Copied . . . . .	44
<b>21 New Problems</b>	<b>45</b>
21.1 Effects on the Current Environment . . . . .	45
21.2 Effects on the Installed Systems . . . . .	45
21.3 Potential User Problems . . . . .	46
21.4 Limitations in the Anticipated Implementation Environment That May Inhibit the New Product . . . . .	46
21.5 Follow-Up Problems . . . . .	46
<b>22 Tasks</b>	<b>47</b>

22.1 Project Planning . . . . .	47
22.2 Planning of the Development Phases . . . . .	47
<b>23 Migration to the New Product</b>	<b>48</b>
23.1 Requirements for Migration to the New Product . . . . .	48
23.2 Data That Has to be Modified or Translated for the New System	48
<b>24 Costs</b>	<b>49</b>
<b>25 User Documentation and Training</b>	<b>49</b>
25.1 User Documentation Requirements . . . . .	49
25.2 Training Requirements . . . . .	49
<b>26 Ideas for Solution</b>	<b>50</b>
<b>27 Waiting Room</b>	<b>50</b>
<b>28 Appendix — Reflection</b>	<b>51</b>

## Revision History

Date	Author	Section
11 October 2024	Ahmed	7-9, 24-26
11 October 2024	Harrison	18-23
11 October 2024	Jared	1-2, 15-17
11 October 2024	Hamza	10-15
11 October 2024	Gurnoor	3-6

# 1 Purpose of the Project

## 1.1 User Business

Chest X-rays are a critical diagnostic tool used by radiologists to identify lung and heart diseases such as pneumonia, cardiomegaly, and pleural effusion. However, manual interpretation of chest X-rays is time-consuming and requires significant expertise, leading to potential diagnostic delays and human error.

The **Chest Scan** project aims to address these challenges by developing a multiclass convolutional neural network (CNN)-based system capable of classifying multiple diseases from chest X-ray images with high accuracy. Unlike traditional AI-assisted diagnostic tools that focus on binary classification, this system will provide multi-disease classification, confidence scores, and visual heatmap outputs to assist radiologists in decision-making.

By leveraging state-of-the-art deep learning techniques, this project aims to enhance diagnostic efficiency, reduce workload on radiologists, and ultimately improve patient care.

## 1.2 Goals of the Project

The objective of this project is to design and implement an AI-powered diagnostic tool that leverages a convolutional neural network (CNN) to analyze chest X-rays and assist radiologists in identifying a wide range of lung and heart conditions. The tool is intended to generate structured diagnostic reports that include relevant clinical information, such as identified diseases and confidence scores, thereby supporting more efficient and accurate medical decision-making.

To enhance interpretability, the system incorporates visual explanation techniques that highlight key regions of interest within the X-ray images, making model outputs more transparent and clinically useful.

Ultimately, this project aims to reduce radiologist workload, improve diagnostic precision, and streamline the medical imaging workflow. The system's effectiveness will be measured through model performance metrics, real-world usability evaluations, and alignment with technical and regulatory standards.

## 2 Stakeholders

This section outlines the primary stakeholders involved in the Chest Scan system. It identifies their roles, responsibilities, and levels of engagement, ensuring a structured approach to project development and implementation.

### 2.1 Client

The client represents the entity funding or overseeing the project. The client is responsible for ensuring the system aligns with its intended objectives and compliance requirements.

Potential clients include:

- **Healthcare Institutions (Hospitals, Radiology Clinics):** Interested in reducing radiologist workload and improving diagnostic speed.
- **Medical Research Institutions:** Seeking AI-driven advancements in chest X-ray diagnostics.
- **AI Healthcare Startups:** Exploring commercial applications of CNN-based diagnostic models.

### 2.2 Customer

- **Radiologists and Medical Imaging Specialists:** Use the system to interpret chest X-rays more effectively, providing input on system performance and ensuring alignment with diagnostic procedures.
- **Healthcare Institutions:** Implement the technology to improve patient care by delivering faster and more accurate chest X-ray results.
- **Patients:** Benefit from quicker and equally accurate X-ray results, leading to an improved treatment experience.
- **Medical Researchers:** Utilize AI-powered diagnostic tools to study the integration of AI in radiology.

### 2.3 Other Stakeholders

- **Healthcare Administrators:** Oversee implementation in clinical settings and ensure compliance with medical standards and regulations.



- **Biomedical and AI Researchers:** Benefit from new research insights and conduct studies based on this project.
- **Hospitals and Medical Centers:** Deploy the system for improved diagnostic efficiency.
- **Regulatory Bodies:** Ensure compliance with standards such as HIPAA[1], GDPR, and FDA guidelines.
- **Software Engineers and AI Developers:** Responsible for system development, model training, and maintenance.
- **Hospital IT Teams:** Handle system deployment, integration, and ongoing technical support.

## 2.4 Hands-On Users of the Project

### Radiologists

- **User Role:** Diagnose patients based on AI-generated reports.
- **Expertise:** Extensive experience in interpreting medical imaging, particularly chest X-rays. Familiar with identifying a wide range of thoracic conditions, and understanding the clinical implications of radiographic findings.
- **Technological Experience:** Proficient in PACS, RIS systems, and computer-aided detection tools. Some may require training in AI-assisted workflows and interpretation of probabilistic outputs (e.g., heatmaps).
- **Needs:** Seamless integration into existing imaging workflows, interpretability of predictions, ability to validate AI-generated results, and full diagnostic transparency.

### Medical Imaging Technologists

- **User Role:** Manage imaging hardware and data preparation pipelines.
- **Expertise:** Knowledgeable in radiation safety, image acquisition protocols, patient positioning, and quality assurance.
- **Technological Experience:** Regular users of imaging consoles and

image management platforms; comfortable with high-resolution medical file handling (DICOM)[3].

- **Needs:** A system that allows quick upload of scans, easy correction of metadata, and tracking of processed scans across patients.

### Healthcare IT Teams

- **User Role:** Configure, deploy, and maintain Chest Scan in secure hospital networks.
- **Expertise:** Network administration, data privacy compliance, systems security, and interoperability between clinical systems (e.g., HL7).
- **Technological Experience:** Skilled in cloud infrastructure, VPN tunnels, and containerized applications.
- **Needs:** Deployment documentation, regular system updates, API access, and audit logs for security monitoring.

### Nurses

- **User Role:** Support patient preparation and record annotation related to diagnostic imaging.
- **Expertise:** Trained in patient care protocols and general imaging workflows within clinical environments.
- **Technological Experience:** Comfortable using electronic health records and interfacing with diagnostic scheduling systems.
- **Needs:** Clear system alerts/notifications, support for documentation handoff to radiologists or physicians.

### Patients

- **User Role:** Recipients of diagnostic services.
- **Expertise:** Minimal to moderate health literacy; understanding of X-rays and diagnostic terms may vary.
- **Technological Experience:** Use patient portals or mobile health applications to access diagnostic summaries.
- **Needs:** Quick result turnaround, comprehensible explanations (as applicable), and assurance of data privacy.

## 2.5 Personas

**Dr. Emily Richards** is a 47-year-old radiologist working at a large regional hospital. She has over 15 years of experience in diagnostic imaging, specializing in thoracic and cardiovascular radiology. Emily is married with two children and leads an active lifestyle, enjoying hiking, book clubs, and attending medical conferences. She values efficient, accurate diagnostic tools that reduce her workload and minimize report-writing time. As someone who has adapted to emerging technologies in her department (e.g., CAD, digital PACS), she is cautiously optimistic about AI. Her main concerns include reliability, interpretability, and regulatory approval. She wants a tool that seamlessly integrates into her existing RIS/PACS workflow and supports her decision-making without replacing it. She appreciates interfaces that are intuitive and provide visual justifications for predictions, such as heatmaps and severity scores.

## 2.6 Priorities Assigned to Users

### Key Users:

- **Radiologists:** Primary users. Their feedback is critical for evaluating system accuracy, usability, and overall success.
- **Medical Imaging Specialists:** Key in ensuring correct image interpretation and model output validation.

### Secondary Users:

- **Patients:** Indirect users; benefit from system outcomes. Their satisfaction is considered, though not central to design decisions.

## 2.7 User Participation

- **Radiologists:** Provide insights on diagnostic workflows and usability requirements. (10 hours)
- **Medical Imaging Specialists:** Review medical images and validate AI outputs. (8 hours)
- **Patients:** Share user experience feedback. (4 hours)

- **AI Researchers and Engineers:** Assist in interface prototyping and model evaluation. (6 hours)

## 2.8 Maintenance Users and Service Technicians

- **Radiology Support Specialist:** Maintains and troubleshoots the AI diagnostic system.
- **Clinical Systems Technician:** Manages upkeep of clinical technology systems.
- **AI Systems Maintenance Engineer:** Supports ongoing AI model maintenance and infrastructure.

# 3 Mandated Constraints

This section defines the essential requirements the product must follow, including legal, regulatory, or system-related restrictions. These constraints ensure the product complies with necessary standards and functions within the project’s limitations.

## 3.1 Solution Constraints

This subsection outlines the constraints placed on the proposed solution for the chest X-ray analysis project. These constraints cover the technical environment, supporting systems, workplace and enterprise considerations, as well as the fit criteria required for each.

MC1: The product shall operate as a web application

Rationale: The web-based nature of the system will allow medical professionals, researchers, and other stakeholders to access the system easily across various platforms (Windows, MacOS, Linux, etc.) without requiring complex local installations.

Fit Criterion: The product will include back-end, front-end, and database integration to run as a web application accessible via a URL, optimized for different devices and browser types (Chrome, Firefox, Safari, etc.).

MC2: The product shall integrate publicly available chest X-ray datasets (NIH Chest X-Ray Dataset)

Rationale: Using public datasets will ensure that the project has access to a broad range of chest X-ray images needed for training and testing the neural network without violating patient privacy laws.

Fit Criterion: The solution will successfully use the CheXpert and MIMIC datasets for model training, with scripts to download and preprocess these datasets for CNN usage. It will comply with healthcare privacy standards, including the Health Insurance Portability and Accountability Act (HIPAA)[1].

MC3: The system shall provide a structured classification output with confidence scores and visual explanations (heatmaps using Grad-CAM).

Rationale: Interpretability is crucial for clinical integration and compliance with explainable AI (XAI) principles.

Fit Criterion: Each output must include disease labels, associated confidence values (e.g., "Pneumonia: 92"), and heatmap overlays on input images.

### **3.2 Implementation Environment of the Current System**

This section outlines the technological and physical environment in which the product will be installed and operated. It encompasses all relevant hardware, software, and organizational elements that will interact with the system. The goal is to provide a comprehensive understanding of the current infrastructure so that the product can be designed to effectively integrate and function within this environment. Understanding the operational requirements derived from this description will help the design team ensure that the solution meets all necessary criteria for successful implementation and usability in real-world settings.

MC4: The product shall use machine learning frameworks like TensorFlow or PyTorch

Rationale: Machine learning frameworks are essential for training the convolutional neural network (CNN) model. TensorFlow and PyTorch are widely used, industry-standard frameworks that offer extensive support for neural network development and GPU acceleration.

Fit Criterion: The product will use either TensorFlow or PyTorch for model development and training, with support for cloud-based GPU instances to speed up computation.

### **3.3 Partner or Collaborative Applications**

This section identifies and describes systems external to the product but necessary for its operation and integration. These systems may include internal databases, external software solutions, or third-party services. Understanding these relationships is essential for ensuring smooth integration and efficient product functionality within the broader technological environment.

MC5: The product shall integrate with external healthcare information systems

Rationale: These systems are essential for accessing and managing patient data, facilitating seamless communication between the product and existing infrastructures. Integrating with these systems will enhance the product's ability to deliver accurate results within the clinical workflow.

Fit Criterion: The product will support integration with healthcare systems through standardized interfaces and protocols such as HL7 and FHIR[4].

MC6: The system shall support external research platforms (e.g., PhysioNet, NIH Data Commons)

Rationale: Enabling data exchange with research repositories supports model evaluation and benchmarking.

Fit Criterion: Exported anonymized outputs must meet external platform standards (e.g., CSV or JSON formats).

### **3.4 Off-the-Shelf Software**

This section identifies and describes commercial, open-source, and other off-the-shelf software (OTS) products that will be incorporated into the solution. The use of OTS software is intended to leverage existing capabilities and minimize development time. Understanding the characteristics, behaviors, and interfaces of these products will help define the design constraints and inform decision-making throughout the project lifecycle. Additionally, potential conflicts between the requirements and the capabilities of the OTS

software must be considered to ensure that the product aligns with organizational needs.

MC7: The product shall integrate with a commercial machine learning framework

Rationale: Established frameworks like TensorFlow or PyTorch provide robust tools for developing and training machine learning models, allowing for efficient implementation of advanced algorithms.

Fit Criterion: The product will utilize the selected framework to train and deploy convolutional neural networks for analyzing chest X-ray images, ensuring compliance with model performance benchmarks.

MC8: The product shall use established machine learning and web development libraries such as TensorFlow, PyTorch, and React

Rationale: Leveraging mature software packages reduces development overhead and increases reliability.

Fit Criterion: The system must function using these components, and any modifications must comply with their respective licenses.

MC9: The database shall be built on a secure and scalable engine such as PostgreSQL or MongoDB

Rationale: These databases are known for stability and support encryption, indexing, and scalability for large imaging datasets.

Fit Criterion: The backend must support data CRUD operations using the selected database technology.

### **3.5 Anticipated Workplace Environment**

This section outlines the anticipated workplace environment where users will interact with the product. Understanding the features of this environment is critical for ensuring that the product is designed effectively to meet user needs and to compensate for any potential difficulties. The description will encompass physical aspects, such as space layout and equipment, as well as social and cultural factors that may impact the product's usability and acceptance.

MC10: The product is intended for use in healthcare settings

Rationale: Working in healthcare environments imposes strict regulations and standards for data security and patient privacy, which must be considered in the product design.

Fit Criterion: The product will implement robust security measures, including encryption and access controls, to ensure compliance with healthcare regulations such as HIPAA[1].

### 3.6 Schedule Constraints

This section outlines the critical timelines and deadlines associated with the project's development phases. Establishing a clear schedule is essential for managing resources, coordinating team efforts, and ensuring that all project components are delivered on time. These constraints help prioritize tasks and guide decision-making throughout the project lifecycle, ultimately contributing to the successful implementation of the proposed solution. By identifying key milestones, this section clarifies the expectations for progress and accountability, enabling the team to meet operational and stakeholder requirements efficiently.

MC11: The key milestones of the project's proposed solution must be completed by specified deadlines

Rationale: This ensures that each major phase of the project is delivered on time, allowing sufficient opportunities for testing, evaluation, and any necessary adjustments before the final release. Timely completion is critical for aligning with external dependencies and operational readiness.

Fit Criterion: All significant milestones, including the completion of the proof of concept, initial version, and final version demonstrations, must adhere to the following schedule:

- Proof of Concept Demonstration: November 11-22, 2024
- Revision 0 Demonstration: February 3-14, 2025
- Final/Revision 1 Demonstration: March 24-30, 2025
- Final Documentation Submission: April 2, 2025



### **3.7 Budget Constraints**

This section outlines the financial limitations associated with the project, detailing the budget allocated for developing the proposed solution. It emphasizes the necessity for cost-effective planning and resource allocation to ensure that the project remains within its financial boundaries.

MC12: The total budget for the project's proposed solution is limited to \$750

Rationale: This budgetary limit is established to prevent the acquisition of pre-existing solutions and to promote the development of a tailored solution. It emphasizes the importance of minimizing all associated costs to maintain financial efficiency throughout the project.

Fit Criterion: The combined expenses for all necessary hardware and software to execute the project's proposed solution must not exceed \$750.

### **3.8 Enterprise Constraints**

N/A

## **4 Naming Conventions and Terminology**

### **4.1 Glossary of All Terms, Including Acronyms, Used by Stakeholders involved in the Project**

Motivation: Names are very important. They invoke meanings that, if carefully defined, can save hours of explanations. Attention to names early in the project helps to highlight misunderstandings. As the detailed work progresses, the glossary provides input to the more precisely specified business/work data model and data dictionary.

Glossary of All Terms, Including Acronyms, Used By Stakeholders

**ML:**

Machine Learning: A branch of artificial intelligence that involves the use of algorithms to allow computers to learn from and make predictions based

on data. This is a core technology used in the project for analyzing chest X-rays.

**DL:**

Deep Learning: A subset of machine learning involving neural networks with many layers, used to analyze various types of data, including images.

**DICOM:**

Digital Imaging and Communications in Medicine: A standard for transmitting, storing, and sharing medical imaging information. It is used to manage medical images in the proposed solution.

**CNN:**

Convolutional Neural Network: A type of deep learning model specifically designed for processing structured grid data like images, used in the project for chest X-ray analysis.

**EHR:**

Electronic Health Record: A digital version of a patient's paper chart, used for storing patient information and history that may be integrated with the proposed solution.

**API:**

Application Programming Interface: A set of rules and protocols for building and interacting with software applications, enabling the integration of the proposed solution with other systems.

**MC:**

Mandated Constraints: Various constraints placed on the project's proposed solution that must be adhered to throughout the development process.

**FR:**

Functional Requirement: A requirement that specifies what functionality the project's proposed solution must provide to meet user needs.

**NFR:**

Nonfunctional Requirement: A requirement that specifies criteria that can be used to judge the operation of a system, rather than specific behaviors (e.g., performance, usability).

**BUC:**

Business Use Case: A scenario that describes how the proposed solution can

be used within a business context to achieve specific goals.

**PUC:**

Product Use Case: A scenario that details how an individual user will interact with the proposed solution to achieve specific tasks.

**MVP:**

Minimum Viable Product: A version of the proposed solution that includes only the essential features required to meet the core needs of the users and stakeholders.

**HIPAA:**

Health Insurance Portability and Accountability Act: A U.S. law that sets standards for protecting sensitive patient data. Systems that handle protected health information (PHI) must ensure that all necessary physical, network, and process security measures are in place and followed.

**XAI:**

Explainable Artificial Intelligence: A field of AI that focuses on methods and techniques to make the behavior of machine learning models transparent, understandable, and interpretable by humans.

**Grad-CAM:**

Gradient-weighted Class Activation Mapping: A visualization technique used to highlight regions of an input image that are most relevant for a neural network's prediction. Used in the project to provide visual explanations for predictions.

**CRUD:**

Create, Read, Update, Delete: The four basic operations of persistent storage, essential for managing data in backend databases.

**JSON:**

JavaScript Object Notation: A lightweight data-interchange format used for data exchange between systems.

**CSV:**

Comma-Separated Values: A common data format used for exporting and importing tabular data, especially in research platforms.

**PACS:**

Picture Archiving and Communication System: A medical imaging technol-

ogy used for storing, retrieving, and sharing digital images.

**RIS:**

Radiology Information System: A networked software suite for managing medical imagery and associated data.

## 5 Relevant Facts And Assumptions

### 5.1 Relevant Facts

**Content:**

Relevant facts that affect the development of the proposed solution but are not specified as mandated requirements include the following:

A1: Sufficient training data is available to support multiclass classification.  
Risk if False: If the dataset is imbalanced or too small for certain disease categories, model performance may degrade.

Technology Adoption: The target users (radiologists and medical technicians) are increasingly adopting AI-based tools in their workflow, indicating a positive reception and potential for successful implementation.

Compute Resources: Many modern development environments allow free or affordable access to GPU acceleration via platforms such as Google Colab or university-provided computing clusters, which facilitates model training and experimentation.

**Motivation:**

These relevant facts provide essential background information to stakeholders and development teams, helping them understand the business context in which the proposed solution will operate. They may also guide the formation of requirements as the project progresses and can impact design decisions to enhance user experience and regulatory compliance.

### 5.2 Business Rules

**Content:**

The following business rules have been identified that may impact the project related to chest X-ray analysis and the development of the proposed solution:

Patient Data Privacy: All patient data used for training and analysis must comply with HIPAA regulations[1], ensuring confidentiality and security in handling personal health information.

Reason for the Rule: To protect patient privacy and maintain legal compliance in the handling of sensitive information.

Authority for the Rule: Health Insurance Portability and Accountability Act (HIPAA)[1], enforced by the U.S. Department of Health and Human Services (HHS).

Integration with EHR Systems: The solution must integrate seamlessly with existing Electronic Health Record (EHR) systems used by healthcare facilities.

Reason for the Rule: To facilitate workflow efficiency and ensure accurate patient record-keeping.

Authority for the Rule: Institutional healthcare IT policies and HL7/FHIR[4] standards.

#### **Motivation:**

These business rules play a critical role in shaping the requirements for the proposed solution. By documenting and understanding these rules, the project team can ensure compliance and identify relevant requirements throughout the development process. Revisiting these rules as the project progresses will help maintain alignment with organizational and regulatory expectations.

### **5.3 Assumptions**

The following assumptions have been made by the project team regarding the development and implementation of the chest X-ray analysis solution:

Availability of Training Data: It is assumed that sufficient high-quality, labeled chest X-ray images will be available for training the machine learning model.

Effect if False: If adequate training data is not available, the accuracy and reliability of the AI model may be compromised, potentially delaying the project timeline.

Technical Resources Availability: It is assumed that the necessary technical resources, including hardware and software tools, will be available on schedule for development and testing phases.

Effect if False: Delays in resource availability may lead to schedule slippage and affect overall project milestones.

Clinical Workflow Alignment: It is assumed that the proposed system can be aligned with typical radiology workflows and does not require fundamental changes to hospital IT infrastructure.

Effect if False: Misalignment with clinical workflows may hinder adoption, requiring rework or system redesign.

**Motivation:**

By explicitly stating these assumptions, the project team aims to ensure all stakeholders are aware of the foundational beliefs guiding the project. Addressing these assumptions early helps manage expectations and facilitates proactive planning to mitigate risks associated with their potential inaccuracy.

## 6 The Scope of the Work

The scope of the work determines the boundaries of the business area to be studied and outlines how it fits into its environment. Our project, which focuses on AI-driven chest X-ray analysis, aims to create a convolutional neural network (CNN) model that assists radiologists in diagnosing lung and heart conditions. The boundaries of the work involve AI-driven diagnostics, healthcare applications, and web-based solutions for radiologists.

### 6.1 The Current Situation

The current diagnostic process for chest X-rays is primarily manual. Radiologists analyze images to detect abnormalities such as pneumonia, lung cancer, or cardiac conditions. While existing automated systems provide some assistance, they lack the accuracy and speed needed to fully integrate into a clinical workflow. Errors in diagnosis can occur due to the overwhelming number of X-rays radiologists must process daily, contributing to diagnostic delays and potential misinterpretations. In the current environment, radiologists perform these tasks with limited AI assistance, and the process is often time-consuming and subject to human fatigue.

By improving AI-based diagnostic tools, the project aims to reduce the workload of radiologists and increase diagnostic accuracy. This involves replacing

or augmenting manual processes with an advanced CNN model that can provide high-confidence predictions on chest X-rays. The study of the current situation allows us to understand the potential effects of such a change and select the most effective approach for integration.

The current process can be modeled using a swimlane diagram that outlines the roles of radiologists, existing software, and the proposed AI solution. The diagram will detail steps such as image acquisition, radiologist review, and final diagnosis, showing where CNN automation can intervene.

## 6.2 The Context of the Work

The work context diagram shows the interactions between the AI-driven chest X-ray diagnostic tool and other adjacent systems. Key external systems include the hospital’s radiology information system (RIS), the picture archiving and communication system (PACS), and medical imaging databases like CheXpert or MIMIC. The product will take input from chest X-ray datasets or directly from hospital systems and provide diagnostic outputs to radiologists via a user interface.

Defining the work context helps ensure the AI diagnostic tool integrates seamlessly with existing hospital systems, avoiding any disruptions to current workflows. Without understanding these systems, the product might fail to meet critical requirements, such as compatibility with existing image formats or HIPAA compliance[1].

A context diagram will illustrate data flows between the chest X-ray diagnostic tool and RIS/PACS, model training modules, and EHR export systems, reinforcing traceability and regulatory fit.

## 6.3 Work Partitioning

The work can be divided into the following key business events, each representing a specific response in the diagnostic workflow. Each business use case (BUC) describes how the AI system will interact with radiologists and adjacent systems:

Image Upload and Preprocessing

- Input: Chest X-ray images in DICOM[3]/JPG/PNG format

- Output: Normalized, resized images ready for model inference
- BUC Summary: The system prepares images for classification by ensuring format consistency and proper scaling.

#### CNN Classification

- Input: Preprocessed images
- Output: Probabilities for various disease labels (e.g., Pneumonia: 82%, Cardiomegaly: 67%)
- BUC Summary: The CNN model performs a multiclass prediction using the trained network.

#### Heatmap Generation

- Input: Image + model prediction layers
- Output: Grad-CAM heatmaps highlighting abnormal regions
- BUC Summary: The system produces interpretability aids for radiologists.

#### Report Generation

- Input: Prediction results and heatmaps
- Output: Structured diagnostic report
- BUC Summary: A summary of findings including predicted labels, confidence scores, and heatmaps.

#### Radiologist Review

- Input: CNN outputs and structured report
- Output: Verified/edited report ready for upload to hospital systems
- BUC Summary: The radiologist reviews AI results and finalizes the diagnosis.

#### Data Validation and Training Feedback

- Input: Confirmed diagnoses from radiologists
- Output: Feedback used to retrain and improve the CNN model



- BUC Summary: The system uses real-world feedback from radiologists to continually improve the model's performance.

## 6.4 Specifying a Business Use Case (BUC)

Each BUC responds to one of the business events described in the partitioning. Here's how a single BUC can be defined in detail:

Business Use Case: AI Diagnosis Processing

- Trigger Event: Chest X-ray image upload from PACS
- Primary Actor: Radiologist or system administrator
- Goal in Context: The AI system processes the X-ray to identify any signs of lung or heart conditions
- Preconditions: The X-ray image has been successfully uploaded and pre-processed for analysis
- Success End Condition: The AI system outputs an accurate diagnosis with a confidence score and explanation of findings
- Failed End Condition: The system fails to provide a diagnosis or outputs unreliable predictions (low confidence score)
- Primary Actor's Steps:
  1. The X-ray is processed by the CNN model
  2. The system evaluates the image, running it through a series of diagnostic layers
  3. The AI outputs its diagnosis with a corresponding confidence score
  4. The system flags uncertain or abnormal findings for human review

This BUC could be represented in multiple ways:

- Process Flow Diagram: Showing the flow of data from image ingestion to diagnosis output
- Sequence Diagram: Highlighting the interaction between the radiologist, PACS system, and AI model during diagnosis

## 7 Business Data Model and Data Dictionary

### 7.1 Business Data Model

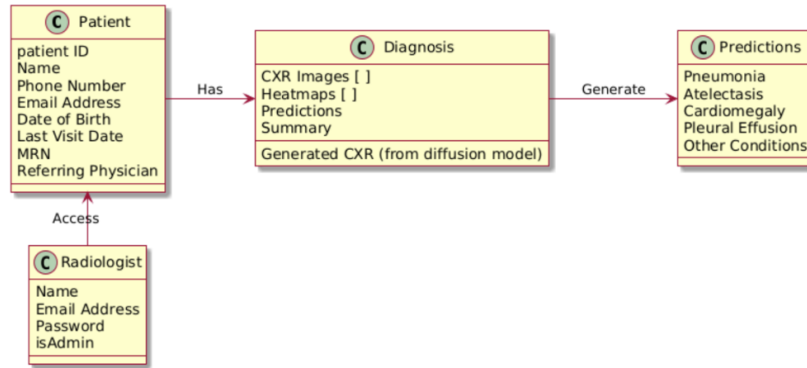


Figure 1: Business Data Model

### 7.2 Data Dictionary

- **Patient:** Stores relevant patient demographics and identifiers (e.g., Patient ID, Name, Age, Sex).  
Description: Accessed securely from the EHR system and stored for reference only; not used in AI inference.  
Related To: Radiologist, Diagnosis
- **Radiologist:** Contains user credentials, access level, and session logs.  
Description: Used for authentication, tracking actions, and system interaction.  
Related To: Diagnosis, Feedback Loop
- **Diagnosis:** A composite entity containing results derived from AI inference and human validation.
  - **CXR Images:** Raw uploaded chest X-rays in DICOM[3], PNG, or JPG format.  
Description: Original medical images uploaded for diagnostic processing.
  - **Heatmaps:** Grad-CAM visual overlays highlighting salient features.

Description: Explainable AI artifacts used to support radiologist interpretability.

Format: Image files (PNG overlay)

- **Predictions:** Probabilistic output of CNN classifications.  
Description: Confidence scores for each potential disease class (e.g., Pneumonia: 91%).  
Format: JSON list of label-score pairs
- **Summary:** Brief textual explanation aggregating predicted results.  
Description: Automatically generated plain-text summary to assist in report generation.
- **Generated CXR:** Optional synthetic X-rays produced using generative models.  
Description: Experimental feature for data augmentation and visualization in research.
- **Predictions:** Redundant — merged above. Removed to avoid information duplication.

## 8 The Scope of the Product

### 8.1 Product Boundary

The Chest Scan system encompasses the end-to-end pipeline for automated diagnostic assistance in chest X-ray analysis. The system includes image ingestion, preprocessing, disease classification using a CNN, Grad-CAM-based heatmap generation, structured diagnostic report creation, and secure delivery of results to radiologists. The product is focused exclusively on chest X-rays for identifying thoracic and cardiac abnormalities. It supports multi-label classification, allowing simultaneous detection of multiple conditions (e.g., pneumonia and cardiomegaly). Excluded from the product scope are real-time patient monitoring, support for other imaging modalities (e.g., CT scans, MRIs), and long-term data archiving.

## 8.2 Product Use Case Table

The following table summarizes the product use cases for the project’s proposed solution:

Table 2: Product Use Case Table

Use Case ID	Use Case Summary
PUC1	Process a chest X-ray image and prepare it for classification
PUC2	Classify the X-ray image using a trained CNN
PUC3	Generate heatmaps showing areas of interest related to predictions
PUC4	Convert classification and visual outputs into a structured report
PUC5	Display the diagnostic report and heatmaps in the web interface
PUC6	Securely store the image, prediction, and report for future access
PUC7	Authenticate radiologists and manage role-based access to data

## 8.3 Individual Product Use Cases (PUC’s)

The following are the individual product use cases (PUCs), with a description, lists of actors, preconditions, and postconditions provided for each.

- **PUC1: Process a chest X-ray image and prepare it for classification**

**Actors:** XRayImage, Preprocessing Pipeline

**Precondition(s):** Image is uploaded in supported format (DICOM[3], JPG, PNG)

**Postcondition(s):** Output is normalized and resized, stored for model input

- **PUC2: Classify the X-ray image using a trained CNN**

**Actors:** CNN Model, Processed XRayImage

**Precondition(s):** Image preprocessing is complete

**Postcondition(s):** Probabilities are assigned to relevant disease classes

- **PUC3: Generate heatmaps showing areas of interest related to predictions**

**Actors:** Grad-CAM, Prediction Output

**Precondition(s):** CNN classification is successful

**Postcondition(s):** Attention maps are created and linked to the input image

- **PUC4: Convert classification and visual outputs into a structured report**

**Actors:** Report Generator, Prediction, Heatmap

**Precondition(s):** Predictions and heatmaps are available

**Postcondition(s):** A summary diagnostic report is generated in structured format

- **PUC5: Display the diagnostic report and heatmaps in the web interface**

**Actors:** Web Frontend, DiagnosticReport, Heatmap

**Precondition(s):** Report is generated

**Postcondition(s):** Radiologist views findings via secure UI

- **PUC6: Securely store the image, prediction, and report for future access**

**Actors:** Backend Database, Patient Record

**Precondition(s):** Diagnostic process is completed

**Postcondition(s):** Encrypted records are stored and linked to radiologist session

- **PUC7: Authenticate radiologists and manage role-based access to data**

**Actors:** Authentication System, Radiologist Profile

**Precondition(s):** Radiologist account exists

**Postcondition(s):** Access granted with proper permissions and audit tracking

## 9 Functional Requirements

### 9.1 Functional Requirements

- **FR1: The system must allow users to upload chest X-ray images in supported formats (DICOM[3], JPG, PNG).**
  - **Rationale:** Enables ingestion of input data into the diagnostic pipeline. This is the initial interaction between the user and the system and is essential for enabling AI-driven analysis.
  - **Fit Criterion:** Image upload is successful, accepted by backend, and stored for further processing. System rejects unsupported formats and provides descriptive error messages.
  - **Product Use Case(s):** PUC1
- **FR2: The system must preprocess images by resizing, normalizing, and converting to grayscale if required.**
  - **Rationale:** Ensures uploaded images conform to the CNN model input specifications, improving consistency and prediction accuracy.
  - **Fit Criterion:** Preprocessed image meets CNN input specs (e.g., 224x224 resolution, normalized pixel values); system logs errors for invalid files and halts processing.
  - **Product Use Case(s):** PUC1
- **FR3: The system must classify diseases using a CNN model and return predictions with confidence scores.**
  - **Rationale:** Core diagnostic engine for AI-assisted classification of medical conditions based on image patterns.
  - **Fit Criterion:** Each prediction includes a disease label and confidence score; accuracy must meet or exceed 60% on the validation dataset used in testing.
  - **Product Use Case(s):** PUC2
- **FR4: The system must support multi-disease classification for a single image.**

- **Rationale:** Multiple pathologies can co-exist in a single X-ray. The model must return all applicable findings with individual confidence scores.
- **Fit Criterion:** When more than one disease is detected, each is listed with its confidence level. The system avoids false multi-label outputs on single-condition inputs.
- **Product Use Case(s):** PUC2
- **FR5: The system must display predicted labels, confidence scores, and heatmaps in a readable format.**
  - **Rationale:** Radiologists and users must be able to interpret model outputs clearly, including visual cues and statistical confidence.
  - **Fit Criterion:** Report is generated and viewable within 5 seconds; disease predictions, confidence scores, and (if available) heatmaps are structured for legibility.
  - **Product Use Case(s):** PUC4, PUC5
- **FR6: The system must generate Grad-CAM heatmaps and overlay them on the original image.**
  - **Rationale:** Provides visual explanations for model decisions by highlighting areas associated with detected conditions.
  - **Fit Criterion:** Heatmap overlays are rendered accurately, highlighting relevant regions without obscuring clinical information.
  - **Product Use Case(s):** PUC3, PUC5
- **FR7: The system must store image, prediction, heatmap, and report securely for user access.**
  - **Rationale:** Supports traceability, auditing, and follow-up analysis. Storage must comply with privacy and security standards.
  - **Fit Criterion:** Data is encrypted at rest using AES-256 and indexed by user session; user can retrieve prior uploads, predictions, and diagnostic visuals via dashboard.
  - **Product Use Case(s):** PUC6

- **FR8: The system must authenticate users and implement role-based access control.**
  - **Rationale:** Maintains data confidentiality by limiting system access based on user roles (e.g., radiologist, admin).
  - **Fit Criterion:** Only users with valid credentials and appropriate roles can access restricted features; failed attempts are logged for auditing.
  - **Product Use Case(s):** PUC7
- **FR9: The system must offer a secure API endpoint to receive image inputs and return JSON predictions.**
  - **Rationale:** Enables hospital systems and external tools to interface with the backend service securely.
  - **Fit Criterion:** API accepts JPG, PNG, DICOM[3]; returns JSON with labels and confidence scores; unauthorized or malformed requests are rejected with informative errors.
  - **Product Use Case(s):** PUC1, PUC2
- **FR10: The system must gracefully handle errors such as invalid uploads and display informative notifications.**
  - **Rationale:** Enhances user experience by preventing silent failures and offering guidance on next steps.
  - **Fit Criterion:** Error messages clearly indicate issues (e.g., unsupported file format, corrupted file); UI supports retries.
  - **Product Use Case(s):** PUC1, PUC4
- **FR11: The system must provide a user dashboard to view past uploads and diagnostic results with timestamps.**
  - **Rationale:** Allows users to track historical diagnoses and export information for further review.
  - **Fit Criterion:** Dashboard displays previous images, predictions, timestamps, and links to heatmaps and diagnostic reports.
  - **Product Use Case(s):** PUC6



- **FR12: The system must notify users when model confidence is low ( $\leq 50\%$ ) and suggest radiologist review.**
  - **Rationale:** Promotes safe usage of AI by alerting users to borderline or uncertain diagnoses that require expert review.
  - **Fit Criterion:** If model returns all confidence values  $\leq 50\%$ , the UI displays a warning message advising manual confirmation by a specialist.
  - **Product Use Case(s):** PUC4, PUC5

## 10 Look and Feel Requirements

### 10.1 Appearance Requirements

**NF-AR1:** The web application shall implement a calming and professional color scheme, consisting of white and soft blue tones.

- **Rationale:** Blue tones are commonly associated with trust and calm in healthcare environments, contributing to user comfort and lowering cognitive stress.
- **Fit Criterion:** At least 75% of surveyed users must indicate that the color scheme positively impacts their user experience.

**NF-AR2:** The system shall utilize consistent and readable fonts and layout spacing across all components.

- **Rationale:** Consistent typography and layout reduce visual strain and improve usability, especially for extended use.
- **Fit Criterion:** User interface consistency is validated in usability reviews; no reports of layout or text inconsistencies are accepted.

### 10.2 Style Requirements

**NF-SR1:** The application shall maintain a simple, minimalistic design with a focus on function over aesthetics.

- **Rationale:** Minimizing distractions helps users focus on diagnostic tasks without unnecessary cognitive load.

- **Fit Criterion:** At least 80% of users should complete core tasks without UI-related confusion or visual overload.

**NF-SR2:** The user interface shall maintain a clean and organized layout that avoids clutter and highlights important diagnostic results.

- **Rationale:** Clarity in layout supports quick decision-making in clinical settings.
- **Fit Criterion:** During usability testing, 80% of test users report the interface is easy to navigate and not cluttered.

## 11 Usability and Humanity Requirements

### 11.1 Ease of Use Requirements

**NF-EUR0:** Efficient Task Completion for Healthcare Professionals

- **Description:** The system should be designed so that healthcare professionals can efficiently complete essential tasks, such as uploading images and interpreting results, without requiring extensive training.
- **Rationale:** The system must be user-friendly for healthcare professionals, ensuring they can effectively use the tool without external assistance after initial exposure.
- **Fit Criterion:** At least 90% of users should be able to complete key tasks (uploading images, viewing results, interpreting model outputs) without requiring additional guidance after initial training.
- **Traceability:** Traces to FR1, FR3, FR4, FR5, FR6.

**NF-EU2:** The software shall be intuitive enough that users can remember how to operate it after 1-2 training sessions.

- **Rationale:** Ease of remembering ensures that even occasional users can operate it effectively without needing extensive retraining.
- **Fit Criterion:** One month after training, 80% of users shall report they can navigate the software and perform basic functions without assistance.

## 11.2 Personalization and Internationalization Requirements

N/A

## 11.3 Learning Requirements

**NF-LR0:** Context-Sensitive Help and Tooltips

- **Description:** The system should provide context-sensitive help features, including tooltips and a dedicated help section, to guide users through the application.
- **Rationale:** Users may require on-the-spot guidance when interacting with system functionalities, so immediate, contextually relevant assistance improves usability.
- **Fit Criterion:** Tooltips should display concise explanations when users hover over UI elements. A help section should be available from all primary screens of the application.
- **Traceability:** Supports learning and complements FR6 and UI-related requirements.

## 11.4 Understandability and Politeness Requirements

**NF-UPR0:** The chest X-ray images generated by the model and displayed should resemble real-world diagnostic visuals.

- **Rationale:** The visual output must meet a standard that is interpretable and familiar to radiologists for it to be clinically useful.
- **Fit Criterion:** In a double-blind review, at least 70% of radiologists should not be able to distinguish between real and AI-generated overlays or reports.
- **Traceability:** Traces to FR5 (heatmaps), FR4 (structured report).

## 11.5 Accessibility Requirements

N/A

## 12 Performance Requirements

### 12.1 Speed and Latency Requirements

**NF-SLR0:** The system shall process and return classification results within `MAX_IMAGE_GEN_TIME` (`MAX_IMAGE_GEN_TIME` = 60 seconds) for 95% of uploaded chest X-ray images.

- **Rationale:** Healthcare professionals require timely insights for clinical decisions. Delays beyond one minute may hinder usability.
- **Fit Criterion:** System shall be tested with 1024x1024 pixel images and meet the threshold under standard conditions.
- **Traceability:** Traces to FR1, FR3.

### 12.2 Safety-Critical Requirements

**NF-SCR0:** All AI-generated diagnostic outputs shall require radiologist approval before final submission.

- **Rationale:** Ensures that AI results are used as assistive tools and do not replace expert human judgment.
- **Fit Criterion:** The system logs a radiologist's review before finalizing a report and flags unapproved results as "Pending."
- **Traceability:** Traces to FR4, FR6, FR8.

### 12.3 Precision or Accuracy Requirements

**NF-PAR0:** The CNN model shall achieve at least `MIN_ACCURACY_REQUIREMENT` (`MIN_ACCURACY_REQUIREMENT` = 65%) average accuracy across primary disease labels.

**NF-PAR1:** The system shall provide confidence scores for all predictions and display a warning for low-confidence cases (`CONFIDENCE_WARNING_THRESHOLD` = 50%).

- **Rationale:** Reliable accuracy is essential for medical applications to ensure trust in AI-assisted diagnostics.

- **Fit Criterion:** Accuracy to be verified against a labeled validation dataset during testing and VnV evaluation.
- **Traceability:** Traces to FR3.

**NF-PAR1:** The system shall provide confidence scores for all predictions and display a warning for low-confidence cases (<50%).

- **Rationale:** Users must understand the certainty of predictions to make informed decisions.
- **Fit Criterion:** If top confidence <50%, a warning message appears.
- **Traceability:** Traces to FR3, FR5.

## 12.4 Robustness or Fault-Tolerance Requirements

**NF-RFTR0:** The system shall detect corrupted or invalid image files and reject them without affecting application stability.

- **Rationale:** Prevents errors and misdiagnoses caused by unsupported input.
- **Fit Criterion:** Uploads of corrupted or invalid formats trigger a clear error message without system crash.
- **Traceability:** Traces to FR1.

## 12.5 Capacity Requirements

**NF-CR0:** The system shall support concurrent processing of up to 20 images without significant performance degradation.

- **Rationale:** Enables radiologists or institutions to batch process scans.
- **Fit Criterion:** No individual image processing time exceeds 45 seconds under concurrent load.
- **Traceability:** Traces to FR1, FR3.

## 12.6 Scalability or Extensibility Requirements

**NF-SER0:** The system shall be designed with a modular architecture to allow future model upgrades and dataset expansions.

- **Rationale:** Scalability ensures the system remains useful as new diseases or imaging formats are introduced.
- **Fit Criterion:** New models can be integrated without refactoring the entire application structure.
- **Traceability:** Traces to FR1, FR3, FR7.

## 12.7 Longevity Requirements

**NF-LR0:** The system shall maintain 99.9% uptime over a continuous one-week testing period.

- **Rationale:** High availability is necessary for consistent clinical support.
- **Fit Criterion:** Uptime is measured by automated monitoring tools and confirmed during final VnV evaluation.
- **Traceability:** Traces to FR1, FR6, FR7.

# 13 Operational and Environmental Requirements

## 13.1 Expected Physical Environment

**NF-EPE0:** The system shall be accessible through standard clinical workstations in hospital or research environments.

- **Rationale:** Medical professionals typically operate on institutional desktop machines with limited administrative privileges.
- **Fit Criterion:** The application shall run on commonly used browsers (Chrome, Firefox, Safari) on Windows/macOS systems without needing installation.
- **Traceability:** Traces to FR1, FR6.

## 13.2 Wider Environment Requirements

N/A

## 13.3 Requirements for Interfacing with Adjacent Systems

**NF-RIAS0:** The system shall integrate with hospital PACS and EHR systems via secure APIs.

- **Rationale:** Interoperability with radiology workflows and patient health records is critical for clinical deployment.
- **Fit Criterion:** JSON responses follow HL7 FHIR standards[4] and can be pulled by external systems without manual conversion.
- **Traceability:** Traces to FR9.

**NF-RIAS1:** The system shall support importing data from public chest X-ray datasets (e.g., CheXpert, MIMIC-CXR).

- **Rationale:** Facilitates model training and evaluation on standardized datasets.
- **Fit Criterion:** Dataset import scripts accept metadata and image formats used in these datasets and validate inputs.
- **Traceability:** Traces to FR1, FR3.

## 13.4 Productization Requirements

**NF-PR0:** The system shall be packaged for deployment on institutional servers or approved cloud platforms (e.g., AWS, Azure).

- **Rationale:** Flexibility in deployment enables clinical trials, research use, or pilot integration.
- **Fit Criterion:** Dockerized deployment scripts and environment variables provided in final deliverables.
- **Traceability:** Traces to FR1, FR6, FR7.

## 13.5 Release Requirements

**NF-RELO:** The system shall include a minimum viable product (MVP) with core classification, heatmap, and reporting features for Rev 1.

- **Rationale:** Demonstrates key value to stakeholders and allows phased feedback.
- **Fit Criterion:** The Rev 1 deliverable includes working functionality for image upload, prediction display, report generation, and access control.
- **Traceability:** Traces to FR1, FR4, FR5, FR6, FR8.

**NF-REL1:** The system shall include configuration options for logging level, timeout settings, and API keys.

- **Rationale:** Enables customization per deployment setting and eases system administration.
- **Fit Criterion:** Configuration file or environment variable schema must be documented in the deployment guide.
- **Traceability:** Traces to FR6, FR7, FR9.

## 14 Maintainability and Support Requirements

### 14.1 Maintenance Requirements

**NF-MR0:** The system shall be developed using a modular architecture with well-documented components.

- **Rationale:** A modular structure supports easier debugging, updates, and testing of individual components without system-wide disruption.
- **Fit Criterion:** System architecture diagram and module documentation must be available and reviewed for clarity and independence.
- **Traceability:** Traces to FR1, FR3, FR6.

**NF-MR1:** The codebase shall be version-controlled and stored in a shared repository with structured commit messages.



- **Rationale:** Version control ensures traceability of changes and supports collaborative development.
- **Fit Criterion:** Git repository with readable commit logs, clear branching strategy, and change history audit.
- **Traceability:** Traces to FR6, FR7.

## 14.2 Supportability Requirements

**NF-SR0:** The system shall maintain at least 80% code coverage through automated tests.

- **Rationale:** High coverage improves reliability and eases future refactoring.
- **Fit Criterion:** Coverage reports must verify  $\geq 80\%$  across core functionalities.
- **Traceability:** Traces to FR3, FR6.

**NF-SR1:** The system shall provide a detailed setup and troubleshooting guide.

- **Rationale:** Comprehensive support documentation accelerates onboarding for new developers and helps diagnose issues quickly.
- **Fit Criterion:** A markdown-based deployment guide must be included with the repository and verified during handoff.
- **Traceability:** Traces to FR6, FR7.

## 14.3 Adaptability Requirements

N/A

# 15 Security Requirements

## 15.1 Access Requirements

**NF-AR0:** Unauthorized User Access Prevention

- **Rationale:** Unauthorized access to patient data and medical diagnoses could lead to privacy breaches and regulatory non-compliance.
- **Fit Criterion:** The system should deny login attempts with incorrect credentials. Unauthorized users should not be able to retrieve or modify any patient data. Logs of failed access attempts should be maintained.
- **Traceability:** Traces to FR8, FR9.

## 15.2 Integrity Requirements

**NF-IR0:** Uploaded chest X-rays and diagnostic results shall only be accessible to the authenticated user who uploaded them.

- **Rationale:** Limiting access to sensitive data minimizes breach impact and ensures compliance with privacy laws.
- **Fit Criterion:** Data access is scoped to user sessions, and cross-user access attempts are blocked and logged.
- **Traceability:** Traces to FR7, FR8.

## 15.3 Privacy Requirements

**NF-PR0:** AES-256 Encryption for Data at Rest and in Transit

- **Rationale:** Encryption is essential for maintaining patient confidentiality and meeting data security regulations. AES-256 provides industry-standard security to protect sensitive information.
- **Fit Criterion:** The system should encrypt stored images, patient data, and diagnostic results using AES-256. Network communication should be encrypted using TLS (HTTPS). Security audits should confirm that unauthorized access to encrypted data is prevented.
- **Traceability:** Traces to FR7, FR8, NF-WER0.

## 15.4 Audit Requirements

**NF-AUR0:** The system shall log authentication attempts, access to sensitive data, and system configuration changes.

- **Rationale:** Maintaining logs provides accountability and enables investigation of suspicious behavior or breaches.
- **Fit Criterion:** Audit logs must include timestamp, user ID, action performed, and status. Logs must be tamper-evident and stored securely.
- **Traceability:** Traces to FR8, NF-AR0, NF-IR0.

## 15.5 Immunity Requirements

N/A

# 16 Cultural Requirements

## 16.1 Cultural Requirements

**NF-CR0:** The system shall support English as the default and only language.

- **Rationale:** Due to scope constraints and user base expectations, multilingual support (e.g., French) has been removed. English remains the standard language for radiology documentation in the primary target regions.
- **Fit Criterion:** All content, tooltips, messages, and reports are presented in English. No language toggle or alternative localization paths are included.
- **Traceability:** Traces to FR5, FR6.

# 17 Compliance Requirements

## 17.1 Legal Requirements

**NF-LR0:** Compliance with HIPAA and PIPEDA

- **Rationale:** Compliance with legal and regulatory frameworks is necessary to protect patient confidentiality and prevent unauthorized ac-

cess to sensitive medical data. Ensuring adherence to HIPAA[1] and PIPEDA[2] helps avoid legal consequences and builds trust with users.

- **Fit Criterion:** The system must implement security and privacy measures that meet HIPAA[1] and PIPEDA[2] standards. A documented compliance review must confirm adherence to required regulations. All data handling, storage, and transmission processes must align with the legal frameworks.
- **Traceability:** Traces to FR7, FR8, NF-PR0.

## 17.2 Standards Compliance Requirements

**NF-SCR0:** The system shall comply with medical imaging standards and HL7 FHIR[4] specifications.

- **Rationale:** Diagnostic image outputs and structured medical data must follow established standards for interoperability with hospital systems and for training clinical models.
- **Fit Criterion:** Outputs must meet DICOM[3] formatting and diagnostic reports must be mappable to FHIR-compatible JSON structures[4]. Compliance is verified during integration and system validation.
- **Traceability:** Traces to FR4, FR5, FR9.

## 18 Requirement Traceability

NFR ID	Related FR(s)	Rationale
NF-SLR0	FR1, FR3	Uses <code>MAX_IMAGE_GEN_TIME</code> instead of the magic number “60 seconds” to define timely image processing for 95% of cases.
NF-SCR0	FR4, FR6, FR8	Radiologist approval tied to reporting, display, and access.
NF-PAR0	FR3	Uses <code>MIN_ACCURACY_REQUIREMENT</code> instead of “65%” to enforce accuracy across primary disease labels.

NF-PAR1	FR3, FR5	Uses <code>CONFIDENCE_WARNING_THRESHOLD</code> instead of “50%” to trigger UI warnings for low-confidence predictions.
NF-RFTR0	FR1	Invalid files blocked at upload stage.
NF-CR0	FR1, FR3	Concurrent processing in upload + classifier.
NF-SER0	FR1, FR3, FR7	Modular data flow enables extensions.
NF-LR0	FR1, FR6, FR7	Uptime depends on core services.
NF-AR1	FR5	UI appearance affects readability.
NF-AR2	FR5	Layout consistency for usability.
NF-SR1	FR5, FR6	Minimal design helps diagnosis.
NF-SR2	FR5, FR6	Layout clarity aids interpretation.
NF-EUR0	FR1, FR3, FR4, FR5, FR6	End-to-end task completion required.
NF-EU2	FR5	Users should remember operation.
NF-UPR0	FR5, FR4	Output must resemble clinical norms.
NF-EPE0	FR1, FR6	Accessible via typical hospital workstations.
NF-RIAS0	FR9	Needs integration with hospital systems.
NF-RIAS1	FR1, FR3	Training on external datasets.
NF-PR0	FR1, FR6, FR7	Portable, deployable product required.
NF-REL0	FR1, FR4, FR5, FR6, FR8	MVP must demonstrate key features.
NF-REL1	FR6, FR7, FR9	Configurable deployment settings.
NF-MR0	FR1, FR3, FR6	Modular design supports updates.
NF-MR1	FR6, FR7	Git-based change tracking.
NF-SR0	FR3, FR6	Tests verify key paths.
NF-SR1	FR6, FR7	Dev guide for deployment.

NF-AR0	FR8, FR9	Block unauthorized access.
NF-IR0	FR7, FR8	Only authorized users see data.
NF-PR0	FR7, FR8, NF-WER0	AES encryption protects privacy.
NF-AUR0	FR8, NF-AR0, NF-IR0	All sensitive actions logged.
NF-CR0	FR5, FR6	Interface in English only.
NF-LR0	FR7, FR8, NF-PR0	Compliance with HIPAA, PIPEDA.
NF-SCR0	FR4, FR5, FR9	Standards compliance via FHIR/DICOM.

## 19 Open Issues

- **Data Privacy and Security:** Implementation of security measures in order to protect patient information. The data stored should be encrypted and anonymized.
- **Neural Network Optimization:** Accurate models are often large and take a long time to produce output. The performance of the algorithms needs to be optimized in order for the model to be usable for chest x-ray analysis.
- **Web Application Concurrent Requests:** The interface to use the neural network is a web application. There may be multiple instances of the web application from different users to access the same neural network in the backend. It will be difficult to make the model support multiple concurrent users.
- **Explanation of Disease:** Making the model explain why it identifies a disease at a specific location will help transparency and let radiologists make a more informed decision on whether to trust it or if it is a false positive.

## 20 Off-the-Shelf Solutions

### 20.1 Ready-Made Products

- **TensorFlow / PyTorch:** These open-source machine learning libraries will be used for model development, training, and inference. Both frameworks support GPU acceleration and are compatible with Python.  
Advantages: Industry standard, large community, extensive documentation.  
Disadvantages: Requires tuning and customization to achieve optimal medical classification performance.  
Fit Criterion: The CNN model shall be developed and deployed using one of these libraries without licensing issues.
- **CheXpert / MIMIC-CXR / NIH Datasets:** These publicly available datasets provide labeled chest X-ray images for training and evaluating the CNN model.  
Advantages: Large, well-annotated, publicly accessible.  
Disadvantages: Dataset bias or label inconsistencies may require preprocessing.  
Fit Criterion: The system shall load, preprocess, and train on these datasets using provided metadata.

### 20.2 Reusable Components

- **Grad-CAM:** This open-source visualization tool will be reused to highlight the regions in X-ray images most relevant to the CNN's predictions.  
Advantages: Well-established, integrates easily with CNNs.  
Disadvantages: Can produce less interpretable results if improperly configured.  
Fit Criterion: Grad-CAM heatmaps must be generated for all classified images and displayed in the user interface.

### 20.3 Products That Can Be Copied

- **UI Styling Components (e.g., Bootstrap, Tailwind CSS):** Standard UI styling frameworks may be used to replicate functional and

aesthetic components from established design systems.

Advantages: Reduces time spent on custom styling; ensures responsive design.

Disadvantages: May need customization to meet healthcare UI conventions.

Fit Criterion: UI components must be responsive and consistent with medical software standards.

- **Open Source Authentication Libraries (e.g., JWT Auth):** Authentication systems can be implemented using reusable open-source modules for secure login and session management.

Advantages: Proven security practices and active community maintenance.

Disadvantages: Requires integration with frontend and backend systems.

Fit Criterion: User authentication must conform to system access requirements defined in Section 15.

## 21 New Problems

### 21.1 Effects on the Current Environment

- Improvements in neural networks and their ability to identify diseases may lead to an over-reliance on automated tools. Healthcare professionals may trust outputs without verifying them, which could erode traditional diagnostic skills.
- CNN-based models can be difficult to interpret. Without visual explanation tools like Grad-CAM, accountability may be diminished when errors occur.
- Biases present in training data (e.g., overrepresentation of certain conditions) may lead to uneven model performance across patient populations.

### 21.2 Effects on the Installed Systems

N/A



### **21.3 Potential User Problems**

- Users may hesitate to upload chest X-ray images due to concerns about data privacy and potential breaches.
- Non-chest X-ray images may be uploaded by mistake, leading to unreliable outputs from the model.
- Some radiologists may be skeptical of AI performance and resist integrating the system into their workflow.

### **21.4 Limitations in the Anticipated Implementation Environment That May Inhibit the New Product**

- The system backend currently supports single-image classification at a time. A queue system may be required to support larger volumes.
- Dependence on stable internet access may impact usability in lower-bandwidth settings, such as rural hospitals.

### **21.5 Follow-Up Problems**

The system may store uploaded X-ray images. If encryption or server access is compromised, patient images may be leaked. This represents a significant privacy concern and may expose the development team to legal risks.

## 22 Tasks

### 22.1 Project Planning

Deliverable	Deadline
Team Formed, Project Selected	September 18
Problem Statement, POC Plan, Development Plan	September 25
Requirements Document Revision 0	October 11
Hazard Analysis 0	October 20
V&V Plan Revision 0	November 3
Proof of Concept Demonstration	November 13-24
Design Document Revision 0	January 17
Revision 0 Demonstration	February 5-February 16
V&V Report Revision 0	March 6
Final Demonstration (Revision 1)	March 18-March 29
EXPO Demonstration	April 8
Final Documentation (Revision 1)	April 4

### 22.2 Planning of the Development Phases

1. Team Formation and Project Setup (September 18)
  - Form the project team and contact supervisor
  - Set up GitHub repository and documentation workflow
2. Project Planning and Definition (October 11)
  - Develop roadmap and scope
  - Identify key requirements, stakeholders
  - Complete initial SRS draft
3. Research (October)
  - Investigate CNN architectures and applications in chest X-ray classification

- Evaluate public datasets (CheXpert, MIMIC-CXR)
  - Select explainability tools (e.g., Grad-CAM)
4. Initial Implementation (November)
    - Clean and preprocess training data
    - Build baseline model pipeline and prototype frontend
  5. Feature Development (December)
    - Implement classification interface, Grad-CAM visualizations
    - Build web backend with database and security controls
  6. Testing (January)
    - Run model evaluations
    - Conduct user acceptance testing
    - Finalize documentation for V&V
  7. Finalization
    - Deliver final demo
    - Compile final reports and documentation
    - Prepare for Expo presentation

## **23 Migration to the New Product**

### **23.1 Requirements for Migration to the New Product**

N/A. The product is developed from the ground up.

### **23.2 Data That Has to be Modified or Translated for the New System**

Training data (chest X-ray images and associated radiology reports) must be preprocessed before use. Preprocessing steps include standardizing image resolution, normalization, grayscale conversion, and extracting structured labels from radiology reports.

## 24 Costs

There are no monetary costs anticipated. All model training and deployment will occur on local hardware. Web hosting, if required, will use university-provided infrastructure. The project will operate within a \$500 contingency budget if any costs arise.

## 25 User Documentation and Training

### 25.1 User Documentation Requirements

- **User Interface Guide:** Explains how radiologists or clinicians interact with the platform.
  - How to upload and classify CXR images
  - How to interpret confidence scores and heatmaps
  - How to download structured reports
- **System Overview:** Overview of how the model processes data from input to output.
- **Frequently Asked Questions (FAQ):** Troubleshooting, file format guidance, and system usage tips.

### 25.2 Training Requirements

- **Minimal Training:** Designed for ease-of-use; minimal learning curve expected.
- **Tutorial Session:** A 1-2 hour session to familiarize healthcare professionals with:
  - Uploading X-rays, viewing results
  - Understanding Grad-CAM visualizations
  - Reviewing model confidence scores
- **Online Help:** Integrated help features including tooltips and links to documentation or video guides.

## 26 Ideas for Solution

Potential system improvements and future enhancements:

- **Batch Processing:** Support classification of multiple CXR images in parallel.
- **Expanded Classification:** Increase disease label set supported by model.
- **Mobile App:** Lightweight version of platform for phones or tablets.
- **Advanced AI Models:** Evaluate transformers or ensembles for improved performance.
- **EHR Integration:** Export reports to hospital records via API.
- **Explainability Enhancements:** Display prediction rationale and confidence analysis.

## 27 Waiting Room

Possible future features or enhancements:

- **Batch Processing:** Enable uploading and processing of multiple CXR images.
- **Additional Disease Detection:** Extend the system to identify more conditions.
- **Mobile App:** Consider developing a mobile version for accessibility.
- **Explainability:** Provide AI explanations for the model's predictions.

## 28 Appendix — Reflection

The information in this section will be used to evaluate the team members on the graduate attribute of Lifelong Learning. Please answer the following questions:

1. What knowledge and skills will the team collectively need to acquire to successfully complete this capstone project? Examples of possible knowledge to acquire include domain specific knowledge from the domain of your application, or software engineering knowledge, mechatronics knowledge or computer science knowledge. Skills may be related to technology, or writing, or presentation, or team management, etc. You should look to identify at least one item for each team member.

To successfully complete the chest X-ray disease classification and heatmap interpretation project, the team must acquire a range of technical and collaborative skills. The AI/ML developer needs strong knowledge in convolutional neural networks (CNNs), medical image classification, and experience using libraries such as TensorFlow or PyTorch. The data scientist must be proficient in preprocessing large imaging datasets (e.g., CheXpert, MIMIC-CXR), statistical data cleaning, and transforming radiology report annotations. The frontend developer should have experience with frameworks such as React and Tailwind CSS to design a clean, professional interface for displaying image uploads, heatmaps, and structured outputs. The backend developer will require expertise in secure API development and integration with the model serving layer, including user authentication and encrypted storage. Finally, the project manager will be responsible for stakeholder communication, timeline coordination, and organizing meetings and documentation. Each team member will combine their respective technical discipline with a collaborative mindset to ensure successful project delivery.

2. For each of the knowledge areas and skills identified in the previous question, what are at least two approaches to acquiring the knowledge or mastering the skill? Of the identified approaches, which will each team member pursue, and why did they make this choice?

The AI/ML developer will take specialized online courses on CNNs (e.g., fast.ai or DeepLearning.AI on Coursera) and study open-source

implementations on GitHub to learn practical model optimization strategies. The data scientist will review Kaggle notebooks involving CheXpert preprocessing and also consult documentation for radiology-specific datasets. The frontend developer will follow the official React and Tailwind CSS documentation alongside YouTube walkthroughs to gain rapid prototyping skills. The backend developer plans to build small API services locally for practice and study role-based access controls via tutorials and open-source Flask/Django projects. The project manager will study Agile project workflows through platforms like Atlassian and supplement this with hands-on scheduling in shared tools like Trello and Slack. This multifaceted learning strategy enables both foundational understanding and immediate project application.

### **What went well while writing this deliverable?**

The team developed a much stronger understanding of how to formalize and structure software requirements clearly and thoroughly. Collaborating on the functional and non-functional requirements brought clarity to system expectations and use cases. Using symbolic constants also helped us eliminate ambiguity from our testable metrics. Additionally, cross-referencing requirements with traceability links improved our ability to write modular, maintainable specifications.

### **What pain points did you experience during this deliverable, and how did you resolve them?**

Aligning our functional requirements with non-functional requirements while preserving clarity was initially challenging. We often discovered vague phrasing or redundant requirements during peer reviews. To resolve this, we did multiple revision cycles focused on rewriting, scoping, and clearly labeling each requirement. Creating traceability mappings and symbolic constants also required restructuring our LaTeX file, which we resolved by distributing editing tasks by section.

**How many of your requirements were inspired by speaking to your client(s) or their proxies (e.g. your peers, stakeholders, potential users)?**

Approximately 70-80% of our requirements were directly inspired by user needs communicated by stakeholders and informed by proxy conversations with peers and domain experts. Features such as uploading diagnostic images, displaying confidence scores, and visualizing heatmaps originated from feedback during elicitation. Others, like security and audit logging, emerged from inferred expectations around clinical deployment.

**Which of the courses you have taken, or are currently taking, will help your team to be successful with your capstone project?**

- **SFWRENG 3XA3 (Software Requirements)** - Directly influenced the structure and rigor of this deliverable. - **SFWRENG 4G06 (Capstone Project)** - Provides end-to-end guidance and iterative planning methodology. - **Machine Learning** - Helped us define accuracy, precision, and confidence requirements. - **Human-Computer Interaction** - Influenced UI/UX related requirements like clarity, layout, and usability. - **Web Programming and Databases** - Enabled us to shape the backend and deployment-related requirements with real-world feasibility in mind.



## References

- [1] U.S. Department of Health and Human Services. *Health Insurance Portability and Accountability Act of 1996 (HIPAA)*. Available at: <https://www.hhs.gov/hipaa/>
- [2] Office of the Privacy Commissioner of Canada. *Personal Information Protection and Electronic Documents Act (PIPEDA)*. Available at: <https://www.priv.gc.ca/en/privacy-topics/privacy-laws-in-canada/the-personal-information-protection-and-electronic-documents-act-pipeda/>
- [3] National Electrical Manufacturers Association. *Digital Imaging and Communications in Medicine (DICOM) Standard*. Available at: <https://www.dicomstandard.org/>
- [4] HL7 International. *FHIR - Fast Healthcare Interoperability Resources*. Available at: <https://www.hl7.org/fhir/>