AMITY UNIVERSITY
—— UTTAR PRADESH ——

AMITY SCHOOL OF ENGINEERING & TECHNOLOGY

Department of Information Technology

Minor Project Synopsis

B. Tech IT

Group No: 19

Project Title: Integrating diverse medical datasets ensuring data privacy and compliance

Area: Minor Project, AI

Academic Session: 2025-26

Project Guide: Dr. Saru Dhir

Details of Project Team:

Programme: -		Year/Semester: -	
S. No.	Enrollment No.	Name	Signature
1.	A2305322060	Sarthak Jha	*Sarthak*
2.	A2305322044	Yash Swami	*ys*
3.	A230522036	Fazeelat Imam	*Fazeelat*

Abstract/Project summary (at least 250 words): Healthcare systems today generate vast and diverse datasets from different sources, including Electronic Health Records (EHRs), diagnostic systems, wearable devices, and medical imaging tools. These datasets are often stored in heterogeneous formats, creating challenges in interoperability and integration. Furthermore, sensitive health data requires strict adherence to privacy standards such as HIPAA and GDPR.

The proposed project aims to develop a secure and efficient framework to integrate diverse medical datasets into a unified system, ensuring privacy and regulatory compliance. The framework will leverage data anonymization techniques, secure

encryption protocols, and role-based access control (RBAC) to safeguard sensitive patient information. In addition, the project will implement interoperability standards such as HL7/FHIR for seamless data exchange among healthcare providers.

The solution will also include a data preprocessing layer for handling inconsistent formats, followed by a secure API-based integration module to merge datasets from various sources

Objectives of the Project: -
Major Objectives (35%)
Integrate diverse medical data from multiple sources (EHR, imaging, genomics) into a unified system.

Ensure data privacy and security using encryption, anonymization, and comply with laws like HIPAA and GDPR.

Maintain data quality and accuracy for reliable clinical and research use.

Support clinical decisions and medical research through integrated, high-quality data.

Research publication (Paper will be published).

Minor Objectives (65%)
Improve system scalability and performance to handle large datasets efficiently.

Manage user access and roles securely with proper authentication and audit trails.

Track data provenance to ensure accountability and reproducibility.

Standardize data using medical ontologies for consistent interpretation.

Facilitate secure collaboration between healthcare institutions.

Optimize costs and resources in data integration processes.

Research paper will be published to check validation.

Methodology to be adopted: -
Requirement Analysis – Identify data sources, privacy regulations, and integration goals.
- Design Phase – Create an architecture with modules for data ingestion, anonymization, encryption, and compliance checks.
- Implementation – Develop APIs, data pipelines, and RBAC mechanisms.
- Testing & Validation – Verify security, compliance, and system performance.
- Deployment & Documentation – Deploy the framework and prepare user manuals

Resource requirement (Hardware & software etc.): -

Hardware: Standard PC/Laptop with Internet access.

- Software:

  Python 3.9

- Stream lit

- Google Gemini API access

- Git/GitHub

- IDE (VsCode/Jupyter)

- Web Browser

Database: PostgreSQL / MongoDB

API Integration: RESTful APIs, FHIR libraries

Expected Outcomes/Justification of the project: -

A unified and secure platform for healthcare data integration.

- Improved data accessibility for clinical decision-making and research.

- Compliance with data privacy laws, reducing the risk of breaches.

- Foundation for future AI/ML-based healthcare analytics.

Major Topic Research:

Integrating Diverse Medical Datasets with Data Privacy and Compliance

Focuses on combining different types of medical data (like electronic health records, imaging, genetic info) into one system for better healthcare insights.

Explores privacy-preserving methods such as anonymization, encryption, and secure data sharing.

Studies with regulations compliance (HIPAA, GDPR) to protect patient information and legal requirements.

Aims to improve data quality for reliable clinical decisions and advanced medical research (e.g., AI in diagnosis).

Minor Topic Research:

Supporting Features and Optimization for Medical Data Integration

Investigates scalable architectures (cloud, distributed systems) for managing large medical datasets.

Examines user access controls, authentication, and audit trails to ensure secure data handling.

Looks into tracking data provenance to ensure transparency and trustworthiness of data.

Studies use of medical terminologies (SNOMED CT, LOINC) for consistent data interpretation.

Researches collaboration frameworks for data sharing between hospitals and research centers.

Considers cost-effectiveness and usability to optimize system adoption and resource use

)

PERT Chart/Schedule of project completion: -

Task	Duration
Requirement Analysis	Week 1-2
System Design	Week 3-4
Development	Week 5-8
Testing & Validation	Week 9-10
Documentation & Review	Week 11-12

References: Research papers / books / websites etc.: -
- HL7 FHIR Standard Documentation – https://www.hl7.org/fhir/
- HIPAA Privacy Rule – https://www.hhs.gov/hipaa
- GDPR Compliance – https://gdpr-info.eu/
- Research Papers on Medical Data Integration and Privacy Preservation

1) Sorthuk
2) Yog
3) Pozellab

Signature(s) of project team

Name and Signature of project guide
28/7/25

Date: 28/07/25