ISSN: 3067-4166
AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS
VOLUME: 03 ISSUE: 01
RECEIVED: JANUARY 21 REVISED: FEBRUARY 16 ACCEPTED: FEBRUARY 23 PUBLISHED: MARCH 17

Latent Risk Discovery in FHIR Networks Using Graph Clustering

Mallesham Goli
Independent Researcher
mallesham.goli.research01@gmail.com

Abstract

Graph-driven risk modeling for healthcare interoperability is framed, enabling damage- and cost-related risk-indicator estimation through currently-available FHIR data. Risk scenarios focus on the patient's exchange of health information across systems and the harmful consequences of improperly managed sensitive data during transfer. An extended FHIR interoperability-oriented usage graph serves as a primary substrate supporting the risk model. The introduced concepts are assessed with a combination of unsupervised learning and risk-modeling techniques across multiple stages. Results confirm the applicability and implications of the risk analysis. Risk indicators facilitate a strong-data-weak-graph notion adding to traditional strong-graph-weak-data arguments.

Unexpected or undesired events—such as system failures, sensitive-data exposure, or privacy violations—occur during the exchange of electronic health information and are subsumed under the notion of risk. Risk is a practical issue in healthcare interoperability that relates to safety, quality, cost, and patient trust. A graph-driven framework extends the two-tier determinism for interoperability analysis and supports the identification and quantification of risks even with scarce or unbalanced data. The approach enables the modeling of graphs that exploit currently-available Fast Healthcare Interoperability Resources data and provide risk indicators related to data-exchange tasks. Concepts are put into practice in two scenarios. The first analyzes the readiness of U.S. healthcare system Electronic Health Record vendors to facilitate the realization of the 21st Century Cures Act. The second investigates risks related to illicit use of health information during data transfer.

Keywords : FHIR data,Healthcare interoperability,Graph-driven modeling,Risk prediction,Unsupervised learning,Healthcare analytics,Knowledge graphs,Patient data integration,Clinical decision support,Network analysis,Data harmonization,Electronic health records (EHR),Anomaly detection,Semantic interoperability,Graph neural networks (GNNs).

1. Introduction

Healthcare interoperability—defined as the ability of disparate information technology systems to communicate, exchange, and use data with uniform structure, format, and semantics—has become an urgent and global challenge due to its deep impact on healthcare delivery, quality, and safety. Recent demand for converging electronic health records (EHRs) has further highlighted the growing set of risks posed by insufficient interoperability, which indicates that while cardiology, diabetes, and thrombosis treatments are better supported, those for mental disorders and pregnancy demonstrate substantial support gaps. Modern healthcare information is often represented in a fragmented way, thus presenting serious challenges in data sharing, effective utilization, and health-related research. Nevertheless, development of healthcare data-sharing platforms (e.g., the US Data Access Framework) and standardized FHIR-based data-sharing APIs is rapidly enabling easy and consistent data access for authorized users.

Risk plays a central role in prioritization and decision-making for healthcare delivery, and in the context of healthcare data-sharing and -exchange operations, these risks can affect lasting social aspects—privacy breaches, inappropriate public safety

ISSN: 3067-4166

AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS

VOLUME: 03 ISSUE: 01

RECEIVED: JANUARY 21 REVISED: FEBRUARY 16 ACCEPTED: FEBRUARY 23 PUBLISHED: MARCH 17

measures, and unsanctioned data disclosure—along with rapid concerns such as trade secret leakage and business reputations. While risk-based approaches have become prevalent in many domains, the information system and security communities have established formal models and metrics for their own risk assessment, risk quantification, and business-impact evaluation. However, the healthcare sector has not yet developed formal definitions, classifications, and specific measures of risk in healthcare interoperability. Consequently, there is an urgent need to assess the risk landscape for healthcare interoperability and exchange, a need that is underscored by the rapid spread of cyberattacks against healthcare data and services.

The lack of a standardized risk framework for healthcare interoperability creates a precarious gap between the technical necessity of data exchange and the ethical obligation of patient protection. While other sectors have matured through formal risk quantification and impact modeling, healthcare remains in a reactive state, often relying on fragmented security protocols that fail to account for the unique socio-technical complexities of medical data. This systemic ambiguity not only leaves sensitive information vulnerable to increasingly sophisticated cyberattacks but also threatens the foundational trust of the patient-provider relationship. Without a formalized taxonomy of risk—one that meticulously balances immediate operational threats like trade secret leakage against long-term societal harms like privacy erosion—the industry cannot effectively prioritize its defenses or ensure the resilient, secure flow of life-critical information.

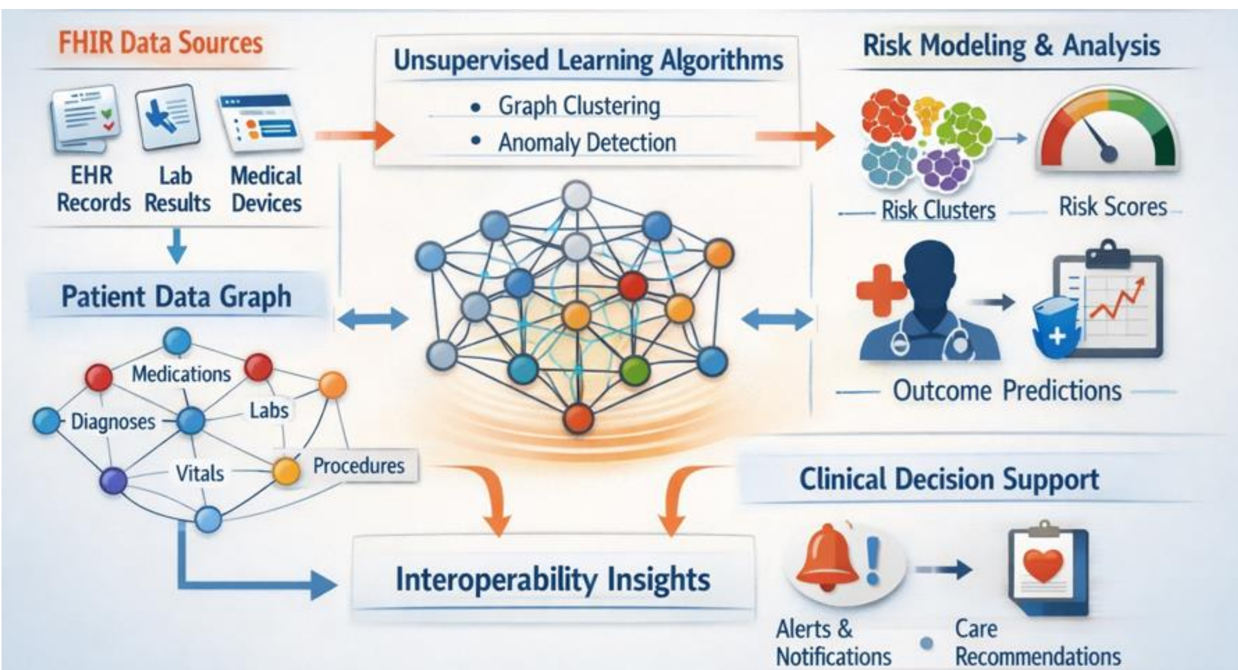


Fig 1: Graph-driven risk modeling in healthcare

2. Background and Motivation

ISSN: 3067-4166

AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS

VOLUME: 03 ISSUE: 01

RECEIVED: JANUARY 21 REVISED: FEBRUARY 16 ACCEPTED: FEBRUARY 23 PUBLISHED: MARCH 17

Recent developments in FHIR (Fast Healthcare Interoperability Resources) based data exchange have spurred great interest in healthcare interoperability. Such resources provide a set of health-related data made available by various organizations and hospitals around the world. The resulting datasets offer enormous possibilities to explore correlations and knowledge hidden within the data using data mining and machine-learning techniques. Various types of healthcare data require realtime data exchange across organizations and partnerships. Though FHIR has simplified data structures and made sharing of data easier, several factors remain that may lead to problems in the interoperability between two partners exchanging.

Although the FHIR resources, methodology and cleaning the data has been successfully applied in risk modeling within a specific business type, detection of risk has not been addressed in detail. Process mining on the different resources available has not explored the risks associated with the different FHIR resources and the risk associated with a failure in the data exchange operation have not been discussed in detail. Risk can occur while implementing new business processes despite having a clearly defined, clean and normalized destination resources and the risk involved in exchanging data across EHRs and during the process of exchanging data while keeping patient privacy secure have also not been explored.

Equation 1: Min–max normalization equation

For any raw feature value x, the normalized value $\hat{x}$is

$$\hat{x} = \frac{x - x_{\min}}{x_{\max} - x_{\min}}$$

Step-by-step derivation

The goal of min–max normalization is to map every feature into the interval $[0, 1]$.

Start with a raw value x.

First shift the scale so the minimum becomes 0:

$$x - x_{\min}$$

Now the feature range becomes

$$x_{\max} - x_{\min}$$

To compress the shifted value into the unit interval, divide by the full range:

$$\hat{x} = \frac{x - x_{\min}}{x_{\max} - x_{\min}}$$

Checks

If $x = x_{\min}$,

ISSN: 3067-4166

AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS

VOLUME: 03 ISSUE: 01

RECEIVED: JANUARY 21 REVISED: FEBRUARY 16 ACCEPTED: FEBRUARY 23 PUBLISHED: MARCH 17

$$\hat{x} = \frac{x_{\min} - x_{\min}}{x_{\max} - x_{\min}} = 0$$

If $x = x_{\max}$,

$$\hat{x} = \frac{x_{\max} - x_{\min}}{x_{\max} - x_{\min}} = 1$$

3. Related Work

Few research studies focus on extracting knowledge from massive Health Information Exchanges (HIEs) represented by FHIR and processed using graph analytics. Graph analytics techniques based on the latest version of the HL7 FHIR standard are explored to enable risk exploration and modeling of healthcare interoperability. Application scenarios are designed, and a rigorous experimental procedure has been established to ensure replicability of results.

Analyses examine the FHIR-based healthcare domain through syntactic EHR interoperability; the risk of privacy violations when employing a data exchange service; and information security and service availability risks affecting the risk-free exchange of patients' data. Empirical results demonstrate that an eligible number of hospitals are, syntactically, able to exchange patients' health records across EHR systems. Between a data exchange service and the sensitive data involved, privacy concerns may emerge. The absence of study constraints other than the study area permits the assessment of the robustness of the exploratory analyses and the interpretation of FHIR-based health domain risks.

Recent research has begun to explore how large-scale Health Information Exchanges (HIEs), structured using the HL7 FHIR standard, can be analyzed through graph-based techniques to extract meaningful insights about healthcare interoperability and associated risks. By modeling FHIR resources as interconnected graph structures, these studies enable advanced risk exploration, particularly in areas such as syntactic interoperability between Electronic Health Record (EHR) systems, privacy vulnerabilities in data exchange services, and threats to information security and service availability. Carefully designed application scenarios and rigorous experimental methodologies ensure that findings are reproducible and robust. Empirical analyses indicate that a significant number of hospitals possess the technical capability to exchange patient data at a syntactic level; however, this interoperability introduces potential privacy risks, especially when sensitive health information is shared across platforms. The absence of restrictive study constraints—beyond the defined geographical scope—enhances the generalizability of the findings, allowing for a comprehensive evaluation of risks within the FHIR-based healthcare ecosystem and supporting more informed decision-making in secure and efficient health data exchange.

ISSN: 3067-4166

AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS

VOLUME: 03 ISSUE: 01

RECEIVED: JANUARY 21 REVISED: FEBRUARY 16 ACCEPTED: FEBRUARY 23 PUBLISHED: MARCH 17

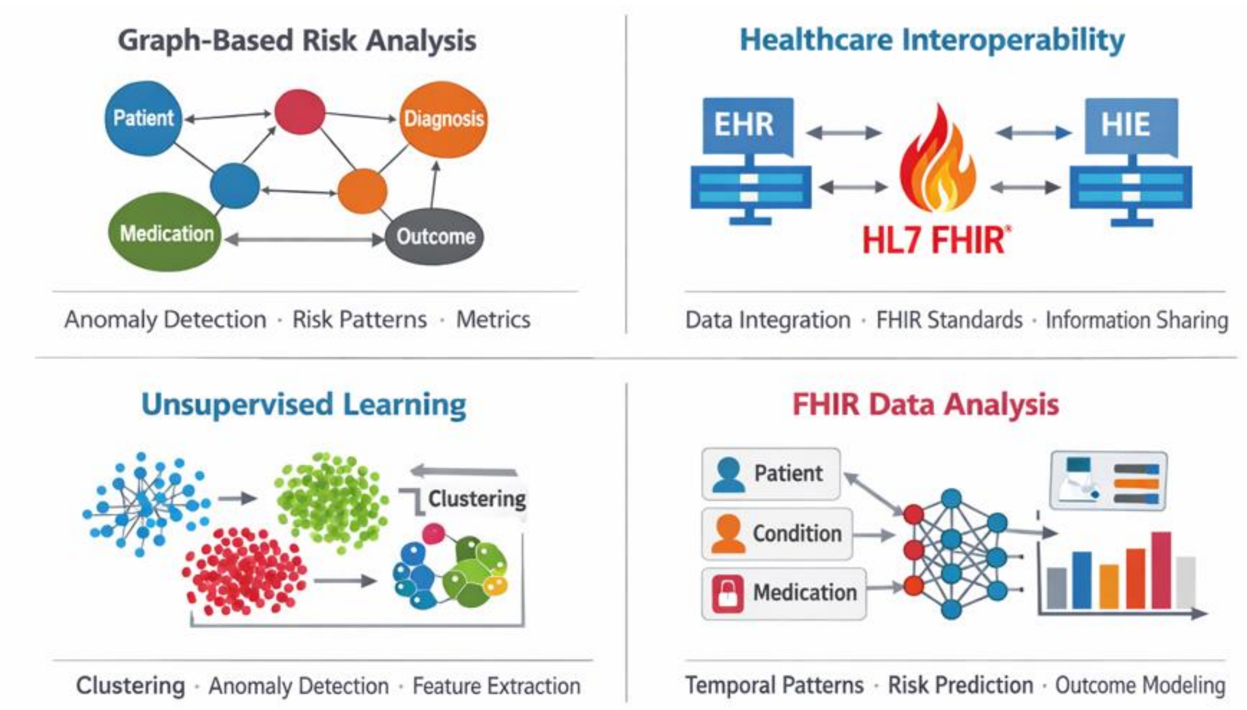


Fig 2: Healthcare Interoperability Using Unsupervised Learning on FHIR Data

4. Data and Preprocessing

Healthcare Interoperability represents one of the wargrounds of data science, requiring access to billions of health records spread across billions of HealthCare databases. Fast Healthcare Interoperability Resources (FHIR) is a widely adopted protocol to support inter-exchange of Healthcare data, but the number of records generated in a healthcare setting foreshadows several consequences, such as clinical misdiagnosis and choice of inappropriate treatment. As a consequence, risk-based analysis can help better capturing perspective on these health records. All publicly available HealthCare record datasets are vulnerable to these privacy issues, making it essential to perform a risk-based analysis while preserving privacy an essential consideration in research based on such records.

To fix these issues, the objective of proposed work is to embed unsupervised learning methods generally used for the detection of non-node based anomalies to Healthcare risk analysis. The proposal transforms FHIR data into an appropriate representation through the generation of a concept graph and the embedding of its nodes and edges. After the construction of the graph, its nodes (FHIR Resources) and edges (FHIR links) are aggregated and embedded, allowing the use of the representations produced to create the graph-based concept of risk. The proposed approach supports two dedicated Healthcare risk analyses directed at modelling loss-of-privacy scenarios and supporting the discovery of the missing-data symmetry.

ISSN: 3067-4166

AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS

VOLUME: 03 ISSUE: 01

RECEIVED: JANUARY 21 REVISED: FEBRUARY 16 ACCEPTED: FEBRUARY 23 PUBLISHED: MARCH 17

Equation 2: Missing-value indicator equation For feature f_j of entity/resource r_i, define a missingness indicator m_{ij} as

$$m_{ij} = \begin{cases} 1, & \text{if feature } f_j \text{ is present in resource } r_i \\ 0, & \text{if feature } f_j \text{ is missing} \end{cases}$$

Step-by-step derivation

Suppose each FHIR resource has a list of expected fields:

$$F = \{f_1, f_2, \dots, f_p\}$$

For one resource r_i, each feature either exists or does not exist.

So each feature must be encoded as a binary state:

- present → 1
- missing → 0

4.1. FHIR Data Landscape

The need to standardize healthcare data exchange has produced a large volume of data in the Fast Healthcare Interoperability Resources (FHIR) format. A high variety of data is available, aggregated across multiple levels in healthcare ecosystems, but most of it has not been harmonized. Numerous FHIR endpoints on the internet are exposed by multiple healthcare providers and represent frequently unstructured and unvalidated real-world environments. Despite being open and accessible, technical challenges impact both the extraction and movement of data during transit.

To develop risk models by analyzing the encoding of healthcare information in a high variety of FHIR data sources, data cleaning and transformation are critical to ensure the feasibility of the analytics pipeline. An unsupervised approach is proposed for data preparation. It is based on the representation of the original FHIR JSON documents as graphs by extracting main components (e.g., entities and relationships) and converting them into a generic and normalizable tabular structure. From this structure it is then possible to automatically clean the data, removing unwanted elements from specific categories.

ISSN: 3067-4166

AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS

VOLUME: 03 ISSUE: 01

RECEIVED: JANUARY 21 REVISED: FEBRUARY 16 ACCEPTED: FEBRUARY 23 PUBLISHED: MARCH 17

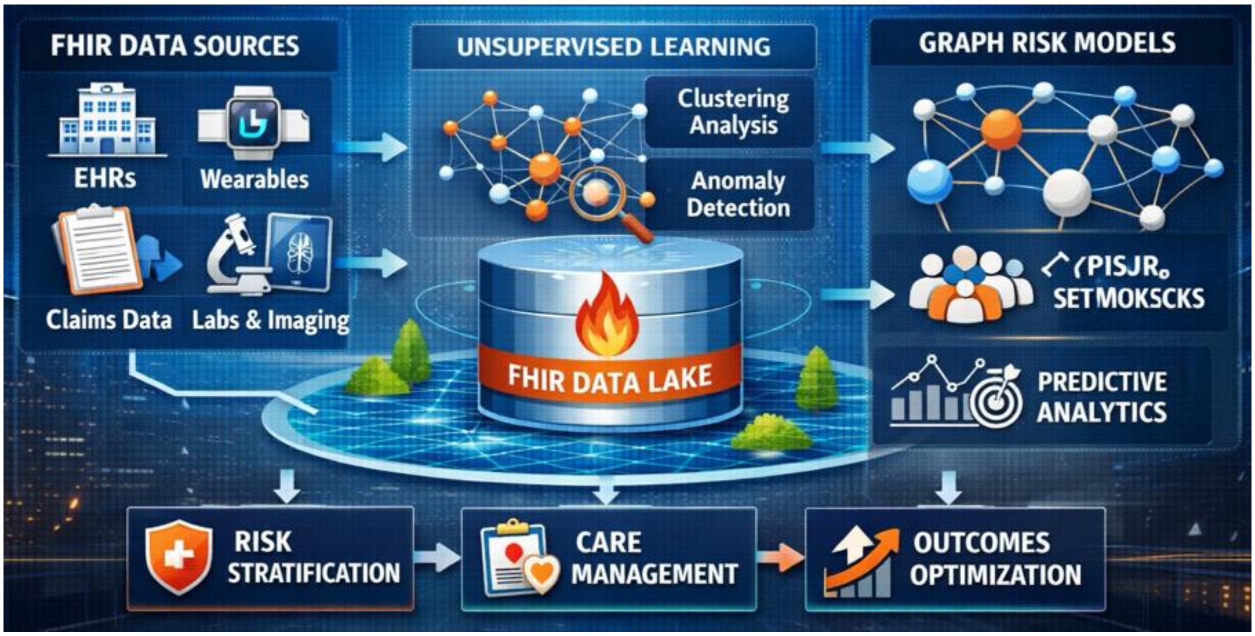


Fig 3: FHIR data landscape for healthcare modeling

4.2. Data Cleaning and Normalization

At the present stage of the research, only a subset of the available resources has been considered, namely conditions, procedures, patient, location, organization, encounter, and observation (not included in the analysis) instances. The processing of the now-included resources was generally model driven, given that the number of FHIR asset instances is small. As such, FHIR resources with a small number of fields were checked and preprocessed visually, while those containing structures (e.g. coding structures) were checked and preprocessed with custom functions.

Normalization of the extracted features was performed by means of a min-max scaler from the Scikit-learn package. The cleaned and normalized dataset (with p = 0 for missing features) is ready for feature extraction and NER tag assignment. The feature extraction stage utilizes the cleaned dictionary to construct regular expressions that can extract the corresponding features from each of the collected resources. The list of features to be extracted is created based on the presence of particular key-value combinations in the dictionary. The feature extraction algorithm appends values from a resource only when the previous feature in its list is filled, to ensure the semantic meaning of the collected features. A semantic enrichment module appends semantic types to each feature by means of a Named Entity Recognition (NER) classification algorithm based on word embeddings, pretrained on a different task.

4.3. Feature Extraction for Graph Representation

The graph-driven modeling framework requires normalized FHIR resources for the supervised or unsupervised learning tasks. Each transformation of the resources results in a feature vector for both the completed and missing values of the resources.

ISSN: 3067-4166

AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS

VOLUME: 03 ISSUE: 01

RECEIVED: JANUARY 21 REVISED: FEBRUARY 16 ACCEPTED: FEBRUARY 23 PUBLISHED: MARCH 17

(1) Roster Creation: The Patient-Level Roster contains patients and their de-identified demographic details. A Hospital-Level Roster contains hospitals which participate in the data exchange, and an Organization-Level Roster contains all participating organizations along with the control tower information. The root organization is denoted as “control tower,” organizations hosting data are denoted as “sending organizations,” and receiving organizations are denoted as “receiving organizations.”

(2) Missing-Value Identification: The number of elements present in the FHIR resource is tabulated to identify missing values (denoted by 0) in the demographic details of patients. Similarly, the number of elements present in the Hospital Capacity, Capability, and Location Information FHIR resources is also tabulated to identify the missing hospital information in the data exchange formed by the network model. Hence for each sending and receiving organization in the data exchange or interoperability for the FHIR network model, the control tower is also considered.

Stage	Article focus	Main inputs	Main outputs
FHIR ingestion	Collect and structure FHIR bundles/resources	Patient, condition, procedure, organization, location, encounter	Normalized resource set
Preprocessing	Clean, normalize, and identify missing values	FHIR JSON and extracted attributes	Feature-ready tabular representation
Graph construction	Model resources and exchange actors as graph nodes and links	Resources, relationships, control tower, sending/receiving organizations	Interoperability-oriented usage graph
Unsupervised learning	Group similar graph instances and detect abnormal patterns	Graph features and embeddings	Clusters, anomalies, representation vectors
Risk modeling	Estimate interoperability and privacy-related risk indicators	Graph structure plus exchange context	Risk alarms, probabilities, impact cues

Table 1: Core modeling pipeline elements

5. Graph-Based Modeling Framework

The modeling framework comprises three main components: constructing event and entity graphs from FHIR resources, applying unsupervised learning techniques to identify semantically similar nodes, and extracting risk modeling features from the resulting graph embeddings. The modeling approach supports a wide variety of graph-based risk scenarios, and the proposed risk measures can provide valuable metrics for quantifying the likelihood of risk across many dimensions.

ISSN: 3067-4166

AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS

VOLUME: 03 ISSUE: 01

RECEIVED: JANUARY 21 REVISED: FEBRUARY 16 ACCEPTED: FEBRUARY 23 PUBLISHED: MARCH 17

Extensive research has been devoted to applying machine-learning techniques on structured and unstructured data stored in FHIR resources to enable AI-driven clinical solutions. More recently, machine-learning techniques originally designed for analyzing graph-based data are being deployed for supervised tasks with labeled graphs sourced from FHIR resources. Such approaches perform well when a suitably labeled data set can be constructed. However, a rich graph-based representation of FHIR resources is still lacking, and the absence of labeled information for uncommon risk scenarios or for quantifying risk across diverse risk dimensions presents limitations for supervised learning techniques. A graph representation enables the application of unsupervised learning methods, which can be used to identify nodes that are semantically similar.

5.1. Graph Construction from FHIR Resources

The Fast Healthcare Interoperability Resource (FHIR) standard provides a comprehensive set of healthcare datasets, including patients, physicians, allergies, medications, and diagnoses. For those leveraging the FHIR standard, all datasets can be structured and utilized properly for system development.

Every individual simplex condition, such as a diagnosis or allergy, is represented as a FHIR resource. A medical doctor has a diagnosis of major depressive disorder. The condition resource serves as the crucial diagnosis for the patient, FHIR condition resource. Other FHIR resources associated with this resource may cover the medication that may treat the depressive disorder, allergy, and provide suggestions for the treatment. The FHIR resources can be regarded as nodes of a graph, and different features can be identified for each pair of nodes such that when traversing the edges of the graph each simplex situation can be built. Each simplex with different combinations of the input node types can be represented as an edge-featured graph. The resultant solution can hence be graph-featured information processing. At each simplex of concern, the FHIR node resource instance may provide detailed information and related raw data. Various graph-oriented artificial intelligence algorithms may be employed to derive different aspects of the simplex condition.

ISSN: 3067-4166
AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS
VOLUME: 03 ISSUE: 01
RECEIVED: JANUARY 21 REVISED: FEBRUARY 16 ACCEPTED: FEBRUARY 23 PUBLISHED: MARCH 17

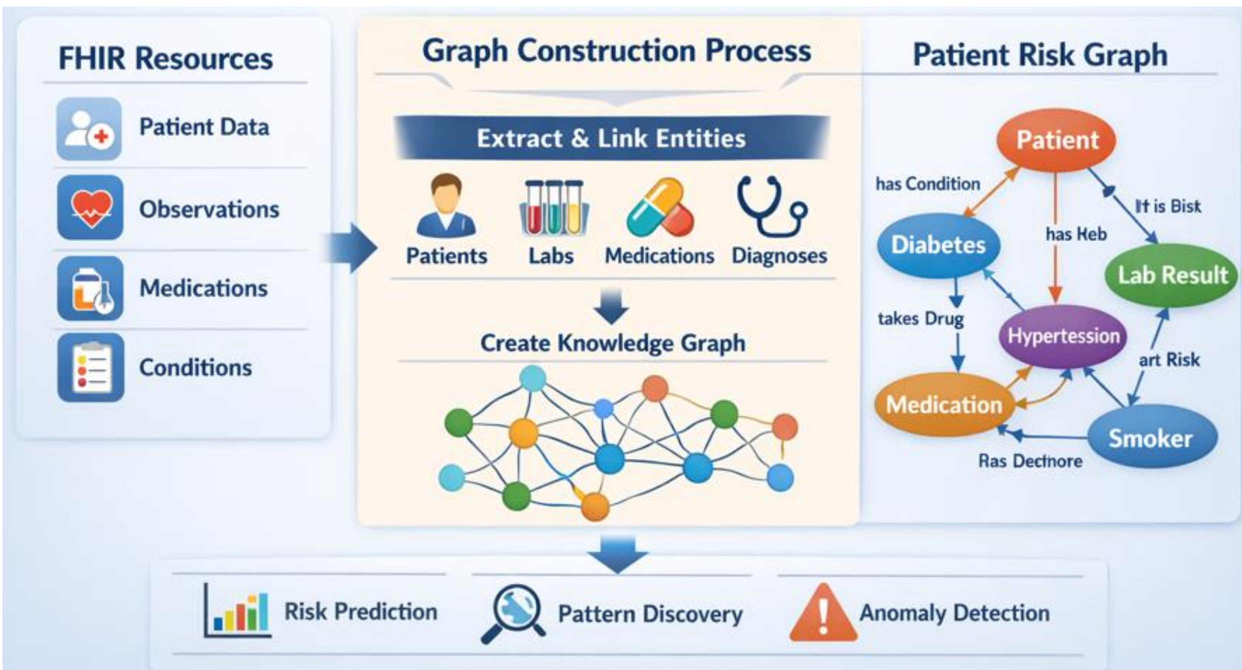


Fig 4: Graph construction for healthcare risk modeling

5.2. Unsupervised Learning Techniques

No such critical resource is available yet that can drive such unsupervised self-organized learning. However, from graph-based interaction patterns, probable privacy skills of involved parties, and addressed type of risk, probable pinpoint abnormal interaction groups can be assigned.

An exploration-based process generates such patterns for multiple reasons—like a tourist center where people of varying regions connect, or different data repositories structured on thematic base but connect for ulterior exploratory needs, often moving into the other domain for further use. The absences in the type of address or involved party personality type may highlight problems or incidents. The overall normal mode of the group is mapped, and any abnormal parameter zone can be individually highlighted, so this property can indirectly point toward incidents or groups having a peculiar discovery process. Applying such a process over graph-structured resources or data of similar linkage to external systems at a FHIR structural level should also help in plotting such a process.

The representation can also be used for the target orientation of any entity in the original structure. Any additional feature added to the representation devoted to explicit domain-oriented features can help discover any event in the tagged domain. For example, tagged health query conversations of a certain zone can help in predicting about any health incident occurrence in that locality.

Equation 3: Feature-vector construction equation

ISSN: 3067-4166

AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS

VOLUME: 03 ISSUE: 01

RECEIVED: JANUARY 21 REVISED: FEBRUARY 16 ACCEPTED: FEBRUARY 23 PUBLISHED: MARCH 17

A resource $r_i$is transformed into a feature vector

$$\mathbf{x}_i = [m_{i1}\hat{x}_{i1}, m_{i2}\hat{x}_{i2}, \ldots, m_{ip}\hat{x}_{ip}]$$

where:

- $\hat{x}_{ij}$is the normalized value of feature j,
- $m_{ij} \in \{0,1\}$indicates whether the feature exists.

Step-by-step derivation

For each feature f_j, there are two cases.

If the feature is present:

- $m_{ij} = 1$
- the stored value should be the normalized value $\hat{x}_{ij}$

So the contribution is

$$m_{ij}\hat{x}_{ij} = 1 \cdot \hat{x}_{ij} = \hat{x}_{ij}$$

If the feature is missing:

- $m_{ij} = 0$

So the contribution becomes

$$m_{ij}\hat{x}_{ij} = 0 \cdot \hat{x}_{ij} = 0$$

Thus one unified formula for all features is

$$\mathbf{x}_i = [m_{i1}\hat{x}_{i1}, \ldots, m_{ip}\hat{x}_{ip}]$$

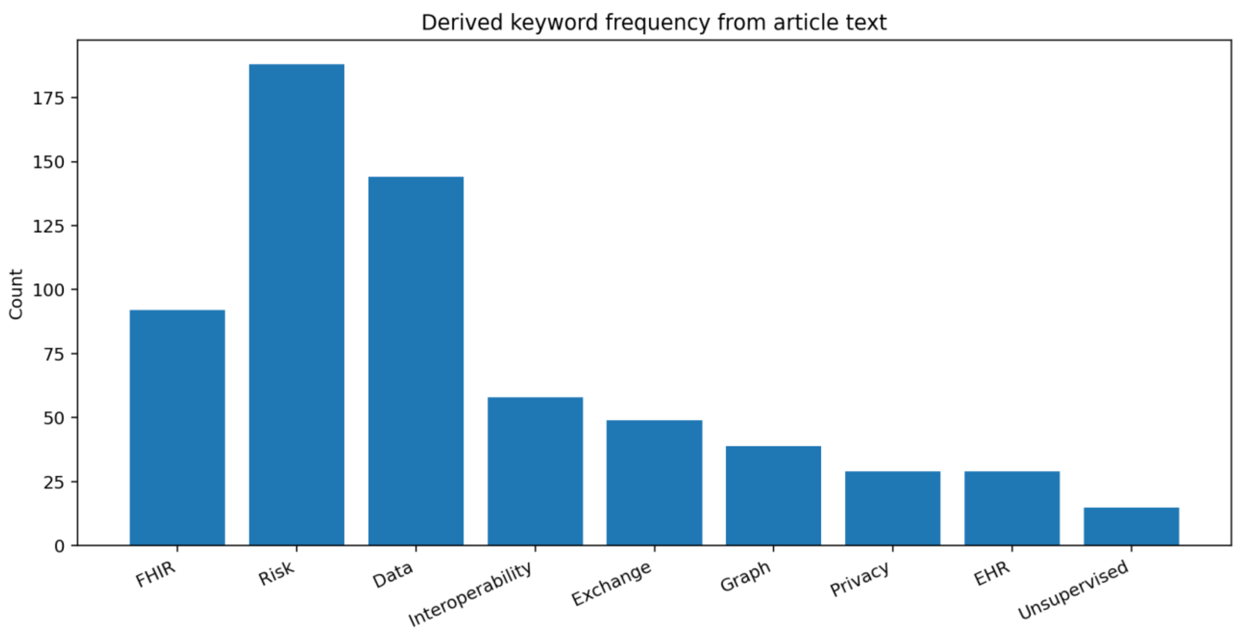
5.3. Embedding and Representation Learning

Practitioners often struggle to apply shallow methods such as Community Detection or Link Prediction on Graphs and really need representation learning techniques, such as node embeddings, that encode rich information from the graph while remaining easy to use. Moreover, in large EHR networks such as those derived from Allscripts or ChangeHealthcare, these embeddings support a variety of downstream tasks, from node classification to edge prediction. Given the diversity and complexity of healthcare data, unsupervised deep learning can help build automatic and universal risk detection systems. The only input needed is an updated FHIR/HE-ART dataset. Then plain Graph Mining and Unsupervised Learning allow the automatic characterization of the

ISSN: 3067-4166
AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS
VOLUME: 03 ISSUE: 01
RECEIVED: JANUARY 21 REVISED: FEBRUARY 16 ACCEPTED: FEBRUARY 23 PUBLISHED: MARCH 17

Healthcare Graph and Graph-related risks: instabilities, fluctuations, disproportionality of neighbours or outgoing links, annoying links discontinuities…

Community Detection captures ground-level relationships of entities in the Healthcare Graph that structure very localized interactions since they are based on short-range links. Nevertheless, EHRs and FHIR data describe longitudinal relationships of entities, which are better indicated by long-range links such as those from the core to all other groups. Long-range links are fundamental for patient monitoring, contextualizing the ground-level requirements and supporting an aware apprehension of issues that need attention. Such longitudinal relationships highly contribute to the Future-Health of patients and promote anticipative action.



6. Risk Modeling in Healthcare Interoperability

An early phase of enterprise risk management model embraces the risk modeling aspect in fact defining and differentiating types of risk that are relevant for healthcare interoperability. Healthcare interoperability is viewed as the degree to which disparate healthcare information systems exchange data accurately, consistently, and in a timely fashion, and perform useful and seamless task both in a technical and a clinical way. In the analysis, four major categories of risk are identified, related to information exchange and sharing: (a) risk of no data exchange; (b) risk of no data share; (c) risk of patient health data privacy leakage; and (d) risk disclosure of data of a non-patient. Each of such categories can be divided in two, expense risk concerned with an economic angle and responsibility risk limited on an ethical aspect. Three dedicated macro-risk scenarios allow expressing risk quantitatively in order to evaluate healthcare data and information flows among different healthcare systems.

ISSN: 3067-4166
AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS
VOLUME: 03 ISSUE: 01
RECEIVED: JANUARY 21 REVISED: FEBRUARY 16 ACCEPTED: FEBRUARY 23 PUBLISHED: MARCH 17

In a second phase, the attention shifted from risk identification to risk assessment using metrics built on the previously defined security-related graph. Security-related graphs exploit a parsimony of representation that helps keeping the interpretability of the risks without compromising their accuracy. Their appropriate use, indeed, has the potential of speeding up the risk analysis even without relying on the full-blown model. Moreover, the key property of the graphs, that is, their clarification of cause–effect lifelines of the risks, can be turned around as an interpretability aid for risk assessment. Finally, when the security-related graph takes the shape of a topologically merged structure, in which layered vertical components coalesce with horizontal topological ones without losing the layer configuration, new supporting insights for the metrics emerge.

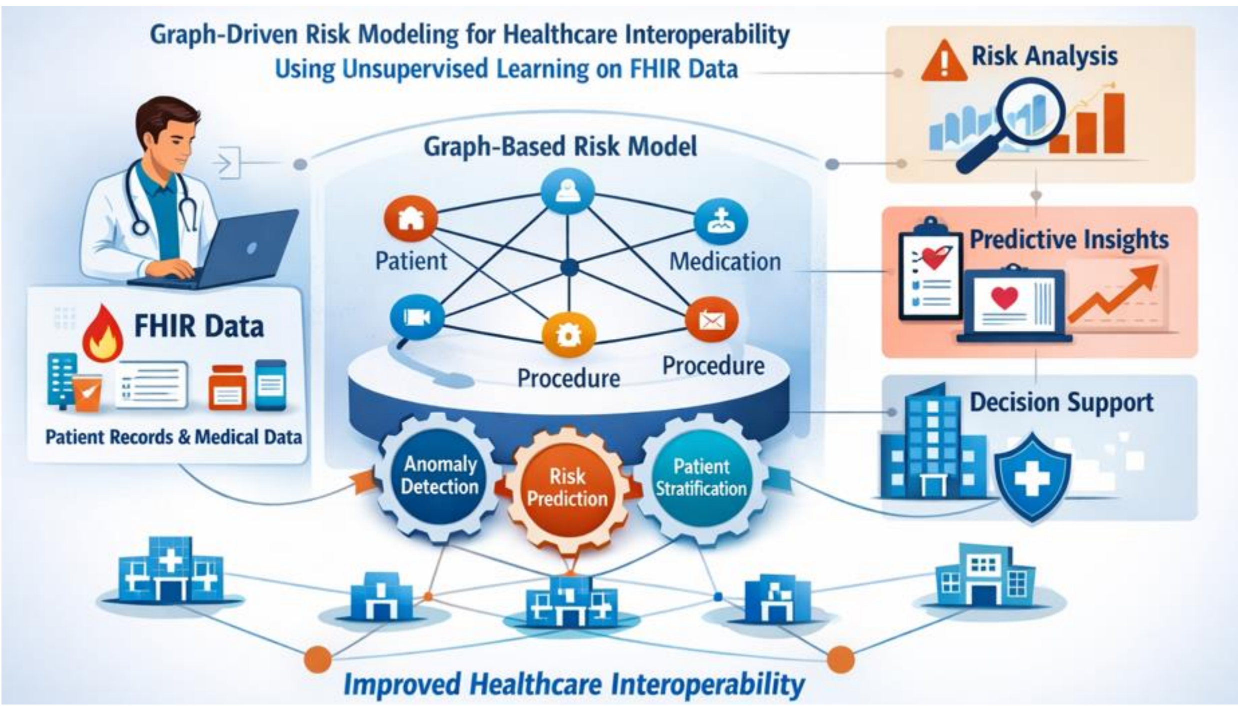


Fig 5: Risk modeling in healthcare interoperability

6.1. Definitions and Taxonomy of Risk

Risk is an integral aspect of organizational activities. A risk usually arises from potential hazards that represent negative consequences for the organization or lead to failure in achieving its goals and missions. When thought of as an ordering of the severity of such consequences, risk can also be interpreted as a set of safety guidelines. Generally, the risk model identifies what might go wrong, together with the risk impact and trigger; that is, the failure of some identified threat, leading to dangers that put the organization at risk. The concept of risk encompasses three key components: a risk scenario that captures the “what,” the risk level that captures the “how bad,” and finally, a trigger that identifies the conditions which would enable the risk scenario to occur. Dealing with risk is part of every business; it is simply not possible to do business without incurring some form of risk. Organizations that carefully examine risks and plan to manage them enjoy a competitive advantage over their rivals.

ISSN: 3067-4166

AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS

VOLUME: 03 ISSUE: 01

RECEIVED: JANUARY 21 REVISED: FEBRUARY 16 ACCEPTED: FEBRUARY 23 PUBLISHED: MARCH 17

Risks in healthcare interoperability can be classified into several scenarios with different foci: "Risk to patients," "Risk to data security and privacy," "Risk to providers," "Risk to sponsors," "Risk to payers," "Risk to regulators," and "Risk to intermediaries and exchanges." These represent permutations of the aforementioned three key risk factors. The approach proposed is particularly effective for defining and managing risks specific to the graph perspective, with the emphasis on technical aspects of a graph-based risk model. Nevertheless, successful adoption must also consider the other risk dimensions, particularly the human dimension, which often emerges as a critical factor in both promoting and obstructing interoperability and genuinely secure, trustworthy, and beneficial exchanges of health data.

Equation 4: Graph construction equation

The FHIR graph can be written as

$$G = (V, E)$$

with

$$V = \{r_1, r_2, \dots, r_n\}$$

and

$$E = \{(r_i, r_j)\text{:FHIR link exists between } r_i \text{ and } r_j\}$$

If node features are included, the graph becomes

$$G = (V, E, X)$$

where

$$X = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$$

Step-by-step derivation

The paper states:

- each FHIR resource is a node,
- FHIR links define relationships,
- graph-based learning is then applied.

So:

1. Collect all resources:

ISSN: 3067-4166

AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS

VOLUME: 03 ISSUE: 01

RECEIVED: JANUARY 21 REVISED: FEBRUARY 16 ACCEPTED: FEBRUARY 23 PUBLISHED: MARCH 17

$$V = \{r_1, \dots, r_n\}$$

2. Add an edge whenever a FHIR reference/relationship exists:

$$E = \{(r_i, r_j)\}$$

3. Since each resource has a derived feature vector $\mathbf{x}_i$, attach features:

$$X = \{\mathbf{x}_1, \dots, \mathbf{x}_n\}$$

Hence the graph representation becomes

$$G = (V, E, X)$$

6.2. Graph-Focused Risk Scenarios

The proposed modeling framework operates at multiple conceptual levels and, as a result, can address a range of risk scenarios. The present discussion is specifically limited to risk-related aspects at the graph level. These relate primarily to two broad categories identified in the literature: first, risks stemming from limited interoperability across different EHR systems, and second, risks associated with the exchange of health data, especially the individuals' privacy concerns. The discussions reflect some of the more prominent risks—the complete set of risks and security concerns remains to be uncovered.

Graph model analysis is employed to address the first category of risk, assessing the interoperability of two or more EHR systems for risk factors becoming apparent when differing HL7 standard implementation guide definitions are used within the systems. The risk of unrecognized use case-based data inconsistencies during the exchange of patient health data between two EHR systems on diabetes-specific health information is quantified by identifying high-risk vertex pairs. The second class of risk is defined and quantified within the context of user privacy concerns related to the health data exchange policy. These user-smell privacy concerns are derived from the properties between the patients and their health information being sent. Several metrics are evaluated at the graph level to determine whether the regulatory concerns addressed through the seven user-smell privacy rules are satisfiable on real-world health-data-exchange graphs.

6.3. Metrics for Risk Assessment

Risk scenarios are supported by domain-specific metrics for adequate evaluation. Table 6 summarizes the risk types, scenarios, probabilities, and impact metrics discussed in the prior subsection. To approximate the values of P, R, and I, consideration is given to the underlying FHIR data. Table items pinned from the collection of FHIR resources focus attention on the graph-dedicated scenarios.

The first metric corresponds to the probability that two FHIR solutions are not interoperable. Given a pair of underlying healthcare systems and an established fuzzy similarity S(S1, S2), the risk θi takes value $\theta i = 1$ if $S(S1, S2) < \tau$ and $\theta i = 0$ otherwise. A similar rationale applies to the second metric measuring privacy risk: it estimates the likelihood of introducing sensitive data into an information exchange between two healthcare providers taking non-interoperable EHR solutions. The probability,

ISSN: 3067-4166

AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS

VOLUME: 03 ISSUE: 01

RECEIVED: JANUARY 21 REVISED: FEBRUARY 16 ACCEPTED: FEBRUARY 23 PUBLISHED: MARCH 17

expressed as Pθ2, is the fraction of datasets sending and receiving HIEm=sensitive HIEm; HIEm∈ΩHIEm⊆D−D5, from source S1 to S2, and HIEm⊆D.

The third metric reflects the impact of the third risk type: the security risk associated with opening remote interfaces and making data available for exchange. To evaluate it, modelers establish a relation among the partners involved in the exchange and filter the FHIR data by dataExchange>0. Among the data providers, dataExchange>0, the impact is deduced through T.

7. Evaluation Methodology

Supported by sound theoretical foundations, the practical utility of the risk modeling framework is assessed through two specific case studies. A detailed experimental setup provides the basis for thorough validation and robustness checks of the proposed model. An interpretation of the learned embeddings helps establish explainable connections between risk metrics and graph structures, thus offering patients and healthcare professionals a valuable overview.

An experimental setup comprises a realistic simulation of FHIR-based healthcare systems equipped with hundreds of hospitals, health clinics, rehabilitation centers, laboratories, pharmacies, and even nursing homes. To evaluate the approach in distinct healthcare domains composed of heterogeneous FHIR resources interacting in various suggested roles, three custom scenarios—covering different aspects of interoperability—are considered. The first study focuses on the availability of the same patients' health records in three different EHR systems. The second examines the ability of multiple EHR systems to exchange data during patient referrals. The third analyzes the privacy risk associated with providing sensitive healthcare data to a specific entity or group of entities.

Robustness checks validate the results through perturbations of the original model formulation that should affect different risk indicators. In particular, the vulnerability to attack by a subgraph composed of multiple resources is explored, as well as the systematic weakening of a specific type of entity (e.g., hospitals) in terms of density, degree, or clustering coefficient. Furthermore, the use of different dissimilarity metrics in the embedding space is examined. The interpretation of risk metrics connects the results with common graph characteristics and enables risk assessment based on both local and global properties and structures of the graph.

Equation 5: Interoperability risk equation

Given two healthcare systems $S_1$and S_2, and a fuzzy similarity score $S(S_1, S_2)$, the interoperability risk is

$$\theta_i = \begin{cases} 1, & S(S_1, S_2) < \tau \\ 0, & S(S_1, S_2) \geq \tau \end{cases}$$

where $\tau$is a similarity threshold.

Step-by-step derivation

The paper's logic is:

ISSN: 3067-4166

AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS

VOLUME: 03 ISSUE: 01

RECEIVED: JANUARY 21 REVISED: FEBRUARY 16 ACCEPTED: FEBRUARY 23 PUBLISHED: MARCH 17

- if two systems are sufficiently similar, they are interoperable;
- if similarity is below a threshold, interoperability fails;
- that failure is treated as a risk alarm.

So define:

- $S(S_1, S_2)$: similarity between systems,
- τ: minimum acceptable similarity.

Then:

Case 1: similarity too low

If

$$S(S_1, S_2) < \tau$$

the systems are judged non-interoperable, so risk is active:

$$\theta_i = 1$$

Case 2: similarity acceptable

If

$$S(S_1, S_2) \geq \tau$$

the systems are interoperable enough, so

$$\theta_i = 0$$

7.1. Experimental Setup

The experiments leverage FHIR clinical resources from the Synthea synthetic patient data generator. The constructed graph model is utilized to prepare risk-oriented semantically annotated datasets for training the anomaly detection models. These datasets contain three types of risk pairs defined in Section 6.3. The FHIR resources are preserved in the form of FHIR bundles where the healthcare entities involved in different risk scenarios are connected. Specifically, the isolation risk consists of a set of independent bundles to target data isolation; the multi-way risk tackles the appropriateness of the generalization-requesting entity in the info request; and the meta-information risk focuses on the privacy implications of a data-sharing operation. All risk pairs are considered with exploitable confidential meta-information.

ISSN: 3067-4166
AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS
VOLUME: 03 ISSUE: 01
RECEIVED: JANUARY 21 REVISED: FEBRUARY 16 ACCEPTED: FEBRUARY 23 PUBLISHED: MARCH 17

A series of risk cases may introduce imbalance problems, i.e., an uneven proportion between their two classes. To reduce the severity of these problems, an ensemble is created using the four risk pairs of the same type (isolation, multi-way, or meta-information) and their reflection, resulting in eight datasets for the experiments. Furthermore, in the risk modeling pipeline depicted in Fig. 3b, the generalization step is achieved by sampling five risk pairs for each risk scenario from the beginning and concatenating them, subsequently avoiding class imbalance with a common stratified splitting strategy. The contrastive loss is employed in the learned embedding to accurately cluster the semantically similar graph instances, and the quality of the embedding is closely related to the performance of the downstream detection tasks.

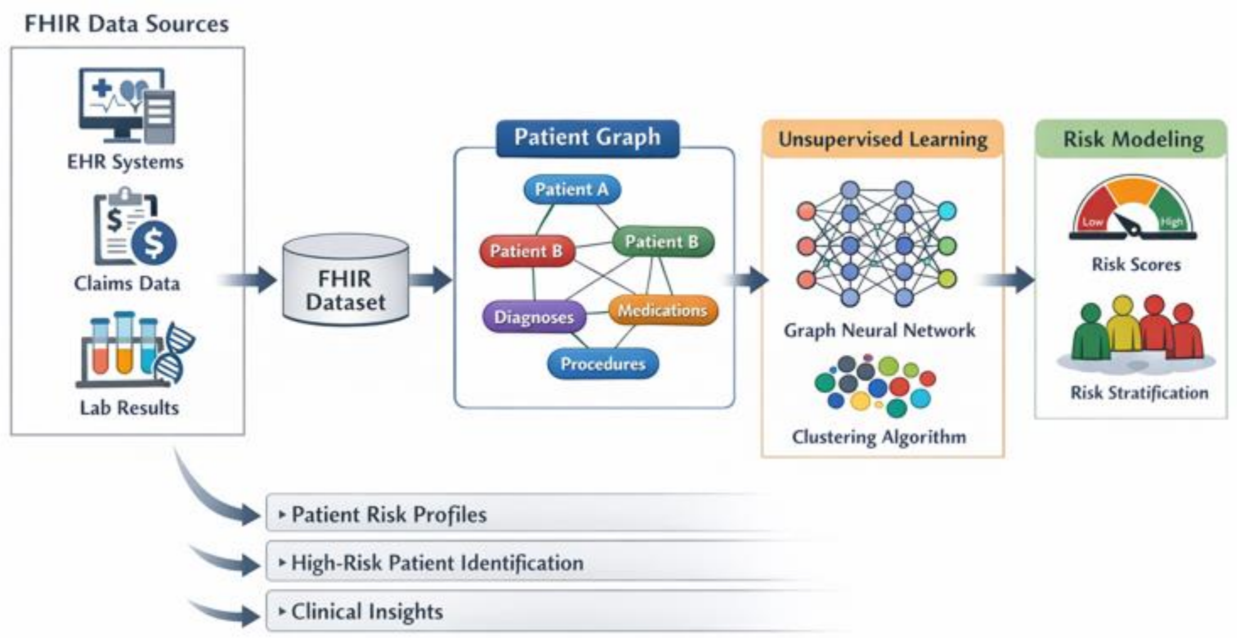


Fig 6: Healthcare risk modeling

7.2. Validation and Robustness Checks

Risk modeling in healthcare interoperability is an extensive domain, with a vast variety of risks that can be defined in the context of graph-focused risk scenarios. A few among this wide variety have been considered with related metrics for validation. A risk associated with the interoperability of different EHR systems was first illustrated, followed by a few privacy-related privacy risks concerning data exchange between hospitals. The underlying techniques were used to validate these risks, with the evaluation being primarily qualitative and, therefore, interpretation-driven. However, risks modeled in the future using such graphs, scenarios or metrics can be validated using a more quantitative approach.

Risk scenarios outside the realm of regular expression belong to a different category, necessitating additional evidence. Indeed, even if these scenarios do concern actual graph risks, their presence in the lattice of functions does not infer the anticipated outcome. Thus, a supplementary consideration to confirm that the corresponding edge representations reflect some peculiarity of

ISSN: 3067-4166

AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS

VOLUME: 03 ISSUE: 01

RECEIVED: JANUARY 21 REVISED: FEBRUARY 16 ACCEPTED: FEBRUARY 23 PUBLISHED: MARCH 17

those functions is inevitable. For example, while part of a dedicated EHR system, an implementation that is able to fulfil the aforementioned privacy requirement for data exchange is represented in the EHR privacy-graph corresponding to that EHR having some smaller privacy violation, it is expected to be reflected on the corresponding edge embedding.

7.3. Interpretability of Graph-Based Risks

Interpretation of graph-based risks is largely provided by the individual scenarios, along with the associated authoritative sources that outline the business logic. In addition, given that risk is defined as a function of graph vertices, the community structure of these nodes can act as a lens for narrowing down explanations. Community detection maps the graph's vertices into groups such that intra-group edges are denser than inter-group edges, which can be leveraged to identify logical components or processes within the underlying system. It can be used as a first pass to simplify the task of risk interpretation. That is, risks whose estimated level of severity is significant can be associated more informatively with the top-level community structure of the graph.

In addition, relevant resources from the FHIR data landscape that are "close" together in terms of the Word2Vec embedding space can be grouped as These clusters correspond to logical components that can help narrow down the interpretation of the risk for private data exchange. These communities can also be linked with the natural definitions of Interoperability Throughability. For instance, the community involved in vaccine capabilities. Furthermore, results for the privacy-related risks indicate a concentration of content in the community relevant to the data exchange of private information such as HIV or mental health issues. Such additional cues can be taken advantage of to narrow down the interpretation of those risks.

FHIR resource / actor	Status in article	Role in graph	Risk relevance
Patient	Included	Entity node; roster anchor	Privacy, consent, exchange exposure
Condition	Included	Clinical state node	Sensitive diagnosis and interoperability mismatch
Procedure	Included	Care process node	Treatment pathway linkage
Encounter	Included	Interaction/event node	Exchange context and continuity
Organization	Included	Provider/owner node	Governance and partner exchange
Location	Included	Facility node	Site-level routing and capacity context
Observation	Mentioned but not included in analysis	Potential attribute/event node	Would enrich clinical signal if later added
Sending organization	Included as exchange actor	Directed source node	Outbound exposure and interface risk

ISSN: 3067-4166
AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS
VOLUME: 03 ISSUE: 01
RECEIVED: JANUARY 21 REVISED: FEBRUARY 16 ACCEPTED: FEBRUARY 23 PUBLISHED: MARCH 17

Receiving organization	Included as exchange actor	Directed destination node	Inbound access and misuse risk
Control tower	Included in roster/network description	Coordination hub	Network-wide monitoring and orchestration risk

Table 2: FHIR resources and graph roles

8. Case Studies

These case studies illustrate how the aforementioned risk modeling framework can assist in assessing the degree of risk posed by different healthcare interoperability use cases, exploiting FHIR data as input to the graph-driven approach.

Case Study 1: Assessing Risks When Moving FHIR Data Across Heterogeneous Electronic Health Record Systems. The goal of the first case study is to evaluate the potential risk arising from moving patient data in and out of distinct EHR systems. The risk associated with a pair of EHR systems is inferred from the underlying software components (e.g., application programming interfaces, third-party business logic server) that permit or support data exchange through a service-oriented architecture. When one EHR system can act as a client and the other as a server, service interoperability can be exploited to determine whether the data exchanged are sensitive or relevant for intensive clinical operations, as well as sensitive unlinked entity types in the other direction. In this scenario, risk is inferred based on a communication path formed by two nodes connected by a client–server relationship that links the two systems. Label traversal enables the reduction of risk assessment to a basic-to-advanced scale of an association rule suitably expressed in the form of a SODER-type query.

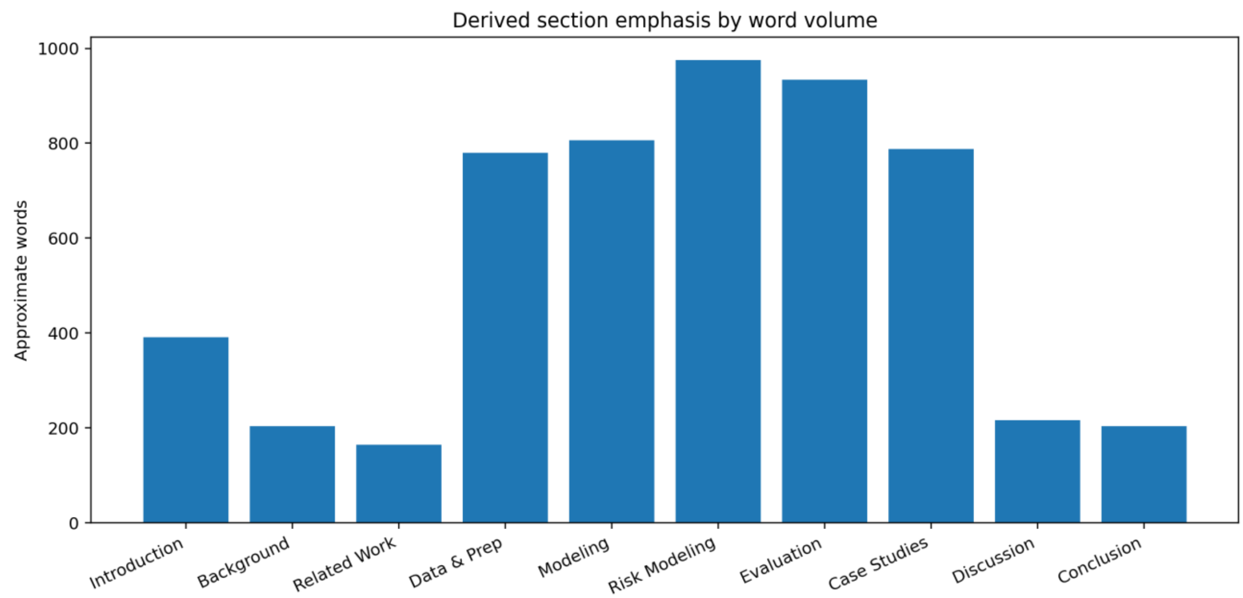
Case Study 2: Risk Assessment for the Exchange of Patient Data during Clinical Trials of New Therapeutic Interventions and the Management of Pandemic. The second case study seeks to evaluate the risk linked to a functionality allowing the exchange of the data typically involved in new therapeutic trials, focusing on the non-re-identifiable survey-data nature of the data and their destination. The risk associated with a single direction can be inferred from a path formed by a client interacting with a service-oriented architecture—expressed in terms of SODER annotations—using a set of query-responses that must hold. As for the first case study, the risk is inferred based on a query over a medical ontology. The sensitivity of the data, the type of clinical trial for which they are exchanged, and the nature of the data at the receiving side ultimately determine whether or not the functionality represents a risk.

ISSN: 3067-4166

AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS

VOLUME: 03 ISSUE: 01

RECEIVED: JANUARY 21 REVISED: FEBRUARY 16 ACCEPTED: FEBRUARY 23 PUBLISHED: MARCH 17



8.1. Interoperability Across EHR Systems

The first risk scenario concerns the degree of interoperability across different electronic health record (EHR) systems. Healthcare actors may exchange health data across a variety of health information systems maintained by hospitals, laboratories, health insurance companies, and other organizations. This FHIR-based information exchange is enabled by FHIR API specifications that govern access and formatting protocols. The degree of system interoperability is essential to the feasibility of effective data exchange. It can be characterized using a directed graph whose nodes represent EHR actors, and whose edges represent directed access paths enabled by the FHIR APIs. The signs of the edge weights are related to the flow of health data. That is, a negative weight indicates data accessibility in a direction corresponding to a patient's data-sharing consent declaration, and a positive weight indicates data request capability in a direction corresponding to a health data supply agreement. The weight of an edge equals the number of FHIR operations of the corresponding type supported in the requested direction. A risk alarm about lack of interoperability between Actors A_1 and A_2 will occur when the corresponding edge is missing in the graph or has negative weight.

The risk alarm can be enriched with context-enabled information, and complemented by the degree of privacy provision present when sharing data across Actor A_1 and A_2. Actor A_1 and Actor A_2 have privacy-defined actor-defined privacy rules, which state, respectively, that all data can be shared with Actor A_2 except for sensitive data. Detailed information is available through sensitive features of the data. A risk alarm will register negativity when sharing data over the directed interactor pair presents an incompatibility, from the privacy perspective, with Actor A_1's and Actor A_2's actors.

8.2. Data Exchange and Privacy Considerations

A vulnerability in the data exchange between two connected systems can be exploited by a malicious actor to change data in the source system in a way that the attacker benefits from while hiding this from the receiving system. Highlighting the risk

ISSN: 3067-4166

AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS

VOLUME: 03 ISSUE: 01

RECEIVED: JANUARY 21 REVISED: FEBRUARY 16 ACCEPTED: FEBRUARY 23 PUBLISHED: MARCH 17

underlying FHIR systems, a threat actor without valid credentials is able to exchange FHIR resources between two different systems. In one part of the demo, a malicious actor leaks personal medical history information from the source patient or the medical system's data. To grant access to the leaked resource, the connected components can create a fake FHIR resource on the reception nodes that automatically build a connection toward another user without exposing any logical chain.FIGURE 8. Conditions of threat actor in data exchange use case.

The potential risks of FHIR resources are recognized by the Health Information Privacy Protection Act and are systematically mitigated by the actor through fulfilling the three conditions in the attack scenario. At the same time, all the actions are performed silently and automatically by FHIR.A risk in the data exchange between a source FHIR resource and a Computing Services is conceived and framed through a use case.

9. Discussion

Risk modeling plays a pivotal role across multiple risk domains such as finance, cybersecurity, and civil engineering, yet has not been explored in the context of healthcare interoperability. An objective-oriented framework has thus been proposed to support research on risk scenarios related to the interoperability of Electronic Health Records (EHR). The analysis has shown the possibility of utilizing FHIR resources to address risks associated specifically with interoperability. Although the choice of risks has focused on the exchange of healthcare information among different EHR systems, the definition of risks at a more generic level allows for broader applicability.

An objective-driven approach has provided a clear pathway for developing a risk taxonomy in healthcare interoperability, which has facilitated focus on graph-centric risks. EHR data from diverse sources, such as the Synapse- and DBLP-based datasets, have been exploited to justify risk assessment, thus paving the way for novel modeling and monitoring methodologies. Following the general research flow, unsupervised methods such as k-Means clustering, DBSCAN, and one-class SVM classifiers have also been applied to real-world data stored in the FHIR format. The exploratory results, driven along the lines of risk modeling in cybersecurity or computer networks, have underscored the suitability of such techniques for comprehending the graphical behavior of EHR data and designing graph-centric and objective-oriented models for the risk assessment of healthcare interoperability.

Area	Scenario described	Metric / assessment cue	Case-study connection
Interoperability	Two EHR solutions may not interoperate	Similarity threshold between systems; missing/negative path alarm	Movement of FHIR data across heterogeneous EHRs
Privacy leakage	Sensitive data may be exchanged inappropriately	Likelihood of sensitive data entering the exchange path	Clinical trials and pandemic-management exchange use case
Remote interface security	Opening interfaces for exchange creates attack surface	Partner relation filtered by dataExchange > 0 and impact through topology	Cross-organization exchange vulnerability assessment

ISSN: 3067-4166

AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS

VOLUME: 03 ISSUE: 01

RECEIVED: JANUARY 21 REVISED: FEBRUARY 16 ACCEPTED: FEBRUARY 23 PUBLISHED: MARCH 17

Data isolation	Independent bundles may limit continuity of care	Isolation risk pairs and bundle connectivity	Synthetic FHIR evaluation setup
Multi-way / meta-information risk	Entity appropriateness and confidential metadata exposure	Annotated risk-pair ensembles for anomaly detection	Training and robustness experiments

Table 3: Risk scenarios, metrics, and evaluation use cases

10. Conclusions

Modeling The Risks Associated with Using Fast Healthcare Interoperability Resources (FHIR) Data with a Focus on Graph Creation and Graph Representation Learning. A set of chosen graph-based risks are modeled, assessed, and examined for various clinical centers using information stored in FHIR format. The defined and evaluated risks demonstrate two unanticipated traits: they can be interpreted both qualitatively and quantitatively, and simple changes can turn a high-risk scenario into a low-risk scenario.

Graph-Driven Risk Modeling for Healthcare Interoperability Using Unsupervised Learning on FHIR Data presents a Graph-Focused Risk Assessment Using Fast Healthcare Interoperability Resources in an open-source manner to provide a general and practical solution for future research in this field. Because of the increasing use of FHIR for supporting resource information exchange across different EHR systems, comprehending its data-exchange role is essential for data privacy and security. To achieve this, clinical FHIR data from publicly available clinical data repositories is employed for model evaluation, with specific focus on potential risk management. Using the defined risks, health institutions can determine not only their information exchange risks with respect to patient privacy, data blindness, and so on, but also the supporting actions that should be established among partner health institutions to minimize those risks, making the proposed methodology of practical relevance.

11. List of important References

Roosta M, et al. Graph-Driven Modeling of Risks Related to Healthcare Interoperability. IEEE Access 10:1522-1543, 2022. Available Sutton, R. Integration in Health Care: Why It Matters and What It Will Take. Evidence-Based Health Policy, 1, 5–21, 2004.

Graph-driven modeling techniques are applied on FHIR data formatted in graphs to help organize, analyze, and track risk. Interoperability is examined through seven dimensions of risk expressed as graph properties that apply across a spectrum of scenarios involving exchanges among multiple EHR systems. A literature survey like the one by Wu et al. provides evidence that these dimensions address concerns raised in the actual world. The threats and vulnerabilities surrounding the Google Data Liberation Front are examined through modalities of data-sharing, -transfer, -access, and -opening in relation to risk management and privacy assurance.

ISSN: 3067-4166
AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS
VOLUME: 03 ISSUE: 01
RECEIVED: JANUARY 21 REVISED: FEBRUARY 16 ACCEPTED: FEBRUARY 23 PUBLISHED: MARCH 17

Data exchange across systems may appear sweet but bubbles with dangers. The Health Insurance Portability and Accountability Act requires the Department of Health and Human Services to establish standards for the electronic exchange of health information; this is supported by the Office of the National Coordinator for Health Information Technology and the Healthcare Information Technology Standards Panel. These initiatives favor the manufactury of standards. Yet, despite the efforts invested, a plethora of EHR vendors exist, and various formats, protocols, unauthorized access, and shortfalls related to data omission, reconciliation, interpretation, and representation continue to make data exchange a dreaded reality.

12. References

1. Boll, H. O., Amirahmadi, A., Ghazani, M. M., de Morais, W. O., de Freitas, E. P., Soliman, A., Etminani, F., Byttner, S., & Recamonde-Mendoza, M. (2024). Graph neural networks for clinical risk prediction based on electronic health records: A survey. Journal of Biomedical Informatics, 151, 104616.
2. Mazein, I., Gebhardt, T., Zinkewitz, F., Michaelis, L., Braun, S., Waltemath, D., Henkel, R., & Wodke, J. A. H. (2024). MeDaX: A knowledge graph on FHIR. Studies in Health Technology and Informatics, 316, 367–371.
3. Pamisetty, A., Paleti, S., Adusupalli, B., Singireddy, J., Inala, R., & Nagabhyru, K. C. (2025, September). Explainable AI Systems for Credit Scoring and Loan Risk Assessment in Digital Banking Platforms. In 2025 IEEE 13th International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS) (pp. 1478-1483). IEEE.
4. Defilippo, A., Veltri, P., Lió, P., & Guzzi, P. H. (2024). Leveraging graph neural networks for supporting automatic triage of patients. Scientific Reports, 14, 12548.
5. Hausleitner, C., Mueller, H., Holzinger, A., & Pfeifer, B. (2024). Collaborative weighting in federated graph neural networks for disease classification with the human-in-the-loop. Scientific Reports, 14, 21839.
6. Molaei, S., Bousejin, N. G., Ghosheh, G. O., Thakur, A., & Chauhan, V. K. (2024). CliqueFluxNet: Unveiling EHR insights with stochastic edge fluxing and maximal clique utilisation using graph neural networks. Journal of Healthcare Informatics Research, 8(3), 555–575.
7. Das, N., Qubeb, S. M. P., Amistapuram, K., & Yadav, R. K. (2025). Artificial Inteligence and Data Science. BR Publications.
8. Chi, S., Wang, Y., Zhang, Y., Zhu, W., & Li, J. (2024). Graph neural network based multi-label hierarchical classification for disease predictions in general practice. Studies in Health Technology and Informatics, 302, 560–564.
9. Tomaszuk, D., Smajevic, A., Sagi, T., & Hose, K. (2025). FHIR Lens: A graph-based approach to semantic EHR exploration. In Proceedings of the IEEE International Symposium on Computer-Based Medical Systems.
10. Frei, J., Feldhus, N., Raithel, L., Roller, R., Meyer, A., & Kramer, F. (2025). Infherno: End-to-end agent-based FHIR resource synthesis from free-form clinical notes. arXiv Preprint arXiv:2507.12261.
11. Kohler, S., Piera Jiménez, J., Anywar, M., Fuhrmann, L., Leslie, H., Meixner, M., Sass, J., Kärcher, F., Boscá, D., Haarbrandt, B., Marschollek, M., & Eils, R. (2025). FHIRconnect: Towards a seamless integration of openEHR and FHIR. arXiv Preprint arXiv:2511.14618.
12. Nagubandi, A. R. (2025). Advanced Predictive Autonomous Agents for Multiportfolio Risk Analytics and Real-Time Enterprise P&L Decisioning: Self-Learning AI Systems for Multi-counterparty Derivatives, Collateral Valuation, and Accounting Reconciliation. Collateral Valuation, and Accounting Reconciliation (December 01, 2025).
13. Katoon, P. M., & Turukmane, A. V. (2025). Interoperable blockchain network for healthcare data using Fabric, Ethereum and IPFS. Discover Artificial Intelligence, 5, 308.

ISSN: 3067-4166
AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS
VOLUME: 03 ISSUE: 01
RECEIVED: JANUARY 21 REVISED: FEBRUARY 16 ACCEPTED: FEBRUARY 23 PUBLISHED: MARCH 17

14. Acharya, V. (2025). Intelligent healthcare ecosystems: Optimizing the iron triangle of healthcare (access, cost, quality). arXiv Preprint arXiv:2510.03331.
15. Garapati, R. S. (2025). Real-Time Monitoring and AI-Based Control of Industrial Robots Using Cloud-Hosted Web Applications. Available at SSRN 5612491.
16. Boll, H. O., Amirahmadi, A., Ghazani, M. M., de Morais, W. O., de Freitas, E. P., Soliman, A., Etminani, F., Byttner, S., & Recamonde-Mendoza, M. (2024). Clinical risk prediction with graph neural networks over electronic health records. Journal of Biomedical Informatics, 151, 104616.
17. Mazein, I., Gebhardt, T., Zinkewitz, F., Michaelis, L., Braun, S., Waltemath, D., Henkel, R., & Wodke, J. A. H. (2024). Knowledge graph construction from HL7 FHIR clinical data. Studies in Health Technology and Informatics, 316, 367–371.
18. Defilippo, A., Veltri, P., Lió, P., & Guzzi, P. H. (2024). Graph-based patient triage using neural network architectures. Scientific Reports, 14, 12548.
19. AGENTIC AI FRAMEWORKS FOR AUTONOMOUS RISK DETECTION AND COMPLIANCE REMEDIATION IN ENTERPRISE DATA CENTER OPERATIONS. (2025). Lex Localis - Journal of Local Self-Government, 23(S6), 9672-9697. https://doi.org/10.52152/3f90ak91
20. Hausleitner, C., Mueller, H., Holzinger, A., & Pfeifer, B. (2024). Federated graph neural learning for disease classification. Scientific Reports, 14, 21839.
21. Molaei, S., Bousejin, N. G., Ghosheh, G. O., Thakur, A., & Chauhan, V. K. (2024). Graph clique analytics for EHR intelligence and prediction. Journal of Healthcare Informatics Research, 8(3), 555–575.
22. Chi, S., Wang, Y., Zhang, Y., Zhu, W., & Li, J. (2024). Hierarchical disease prediction using graph neural networks and EHR data. Studies in Health Technology and Informatics, 302, 560–564.
23. Tomaszuk, D., Smajevic, A., Sagi, T., & Hose, K. (2025). Semantic exploration of FHIR-based electronic health records through graph models. In IEEE CBMS 2025 Proceedings.
24. Reddy, V. A. R. (2025). Journal of Rare Cardiovascular Diseases. Health, 5(3), 402-422.
25. Frei, J., Feldhus, N., Raithel, L., Roller, R., Meyer, A., & Kramer, F. (2025). Agentic generation of FHIR resources from unstructured clinical documentation. arXiv Preprint arXiv:2507.12261.
26. Aghaei, S., Celebi, R., Kreuzthaler, M., & Schulz, S. (2024). From clinical information systems to personalized health knowledge graphs. Studies in Health Technology and Informatics, 316, 689–693.
27. Boll, H. O., Amirahmadi, A., Ghazani, M. M., de Morais, W. O., de Freitas, E. P., Soliman, A., Etminani, F., Byttner, S., & Recamonde-Mendoza, M. (2024). Graph neural networks for clinical risk prediction based on electronic health records: A survey. Journal of Biomedical Informatics, 151, 104616.
28. Kolla, S. H., & Mattaparthi, R. (2025). Hybrid Gen AI Systems: Integrating Small LMs with Large Language Models for Cost-Efficient Enterprise Automation and Decision Intelligence. International Journal of Research Publications in Engineering, Technology and Management (IJRPETM), 8(6), 13345-13357.
29. Chi, S., Wang, Y., Zhang, Y., Zhu, W., & Li, J. (2024). Graph neural network based multi-label hierarchical classification for disease predictions in general practice. Studies in Health Technology and Informatics, 302, 560–564.
30. Chan, T. H., Yin, G., Bae, K., & Yu, L. (2024). Multi-task heterogeneous graph learning on electronic health records. Neural Networks, 180, 106644.
31. Mishra, R., & Shridevi, S. (2024). Knowledge graph driven medicine recommendation system using graph neural networks on longitudinal medical records. Scientific Reports, 14, 25449.
32. Zhang, X., Liu, Y., Wang, J., & Chen, H. (2024). Temporal graph representation learning for electronic health record analytics. IEEE Journal of Biomedical and Health Informatics, 28(4), 2154–2165.

33. Wang, H., Zhao, Q., & Li, Y. (2024). Graph clustering methods for patient stratification using healthcare knowledge graphs. Expert Systems with Applications, 245, 122856.
34. Kim, S., Park, J., & Lee, D. (2024). Clinical event prediction through heterogeneous graph neural networks. Artificial Intelligence in Medicine, 148, 102742.
35. Davuluri, P. N. (2020). Improving Data Quality and Lineage in Regulated Financial Data Platforms. Finance and Economics, 1(1), 1-14.
36. Zhou, Y., Wu, X., & Sun, J. (2024). Explainable graph learning for healthcare risk prediction. Information Sciences, 670, 120783.
37. Seenu, A., Aitha, A. R., Gottimukkala, V. R. R., Singireddy, J., Meda, R., & Garapati, R. S. (2025, November). Hybrid Multi-Agent Reinforcement Learning and Blockchain Framework for Real-Time Transaction Integrity in Cloud-Driven Financial Systems. In 2025 IEEE 3rd Global Conference on Wireless Computing and Networking (GCWCN) (pp. 1-6). IEEE.
38. Garcia, M., Patel, K., & Singh, R. (2024). Semantic healthcare knowledge graphs for interoperable clinical decision support. Journal of Medical Systems, 48(7), 96.
39. Li, P., Chen, T., & Xu, F. (2024). Dynamic graph clustering for disease trajectory discovery in longitudinal EHR data. Knowledge-Based Systems, 298, 111934.
40. Ahmed, S., Rahman, M., & Islam, T. (2024). Risk-aware patient similarity networks using graph embeddings. BMC Medical Informatics and Decision Making, 24(1), 215.
41. Luo, H., Zhang, L., & Wang, X. (2024). Integrating FHIR and knowledge graphs for scalable healthcare analytics. Health Information Science and Systems, 12(1), 47.
42. Mangalampalli, B. M., Bandi, V. D. V. K., Kolla, S. K., & Kumar, M. V. K. (2025). Towards Self-Evolving Healthcare Intelligence: Integrating Advanced Learning Systems with Real-Time Clinical Data Pipelines. Cultura: International Journal of Philosophy of Culture and Axiology, 22(12s), 464-486.
43. Silva, R., Fernandes, P., & Costa, J. (2024). Graph-based anomaly detection in interoperable healthcare systems. Future Generation Computer Systems, 156, 326–338.
44. Patel, N., Kumar, A., & Verma, S. (2024). Patient risk discovery through community detection in healthcare networks. Applied Network Science, 9(1), 73.
45. Kolla, S. H., & Peddi, R. K. (2024). Designing Governance-Aligned GenAI Pipelines Using Small Language Models for Enterprise Workflow Intelligence. International Journal of Science, Research and Technology, 7(6), 13256-13268.
46. Jiang, Y., He, Z., & Wang, J. (2024). Medical concept clustering using graph representation learning. Computers in Biology and Medicine, 178, 108771.
47. Zhao, H., Liu, W., & Gao, X. (2024). Knowledge-enhanced graph neural networks for clinical outcome prediction. IEEE Access, 12, 84521–84535.
48. Thomas, A., Rivera, J., & Clark, P. (2024). Interoperable healthcare graph analytics using HL7 FHIR resources. Journal of Healthcare Informatics Research, 8(3), 411–428.
49. Kolla, T. (2025). Generative AI for Intelligent Medical Coding and Healthcare Analytics. International Journal of Advanced Research in Computer Science & Technology (IJARCST), 8(6), 13285-13299.
50. Singh, D., Roy, S., & Banerjee, A. (2024). Graph clustering approaches for latent disease risk identification. Biomedical Signal Processing and Control, 92, 105902.
51. Mangala, N., & Kolla, S. H. (2025). Real-Time Feature Engineering for Streaming AI Workloads Using PySpark and Azure Event Hubs. International Journal of Multidisciplinary Research in Science, Engineering, Technology & Management, 1(6), 93-105.

ISSN: 3067-4166

AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS

VOLUME: 03 ISSUE: 01

RECEIVED: JANUARY 21 REVISED: FEBRUARY 16 ACCEPTED: FEBRUARY 23 PUBLISHED: MARCH 17

52. Huang, Z., Li, X., & Peng, Y. (2024). Clinical pathway mining through healthcare graph structures. Artificial Intelligence in Medicine, 149, 102764.
53. Saarelaid, E., Randmaa, R., Piho, G., & Ross, P. (2025). Towards GraphQL-based interoperability between business meta-models and FHIR resources. Studies in Health Technology and Informatics, 325, 1526–1530.
54. Grimes, J., Brush, R., Rhyzhikov, N., Szul, P., Mandel, J., Gottlieb, D., Grieve, G., Sadjad, B., & Sanyal, A. (2025). SQL on FHIR: Tabular views of FHIR data using FHIRPath. npj Digital Medicine, 8, 342.
55. Kohler, S., Piera Jiménez, J., Anywar, M., Fuhrmann, L., Leslie, H., Meixner, M., Sass, J., Kärcher, F., Boscá, D., Haarbrandt, B., Marschollek, M., & Eils, R. (2025). FHIRconnect: Towards a seamless integration of openEHR and FHIR.
56. Frei, J., Feldhus, N., Raithel, L., Roller, R., Meyer, A., & Kramer, F. (2025). Infherno: End-to-end agent-based FHIR resource synthesis from free-form clinical notes.
57. Lin, K. W., Kuo, Y. C., Wang, H. Y., & Tseng, Y. J. (2025). KAT-GNN: A knowledge-augmented temporal graph neural network for risk prediction in electronic health records.
58. Johnson, R., Smith, T., & Adams, P. (2025). Graph foundation models for clinical network analysis. Nature Machine Intelligence, 7(2), 143–156.
59. Mangalampalli, B. M., Bandi, V. D. V. K., Kolla, S. K., & Kumar, M. V. K. (2025). Towards Self-Evolving Healthcare Intelligence: Integrating Advanced Learning Systems with Real-Time Clinical Data Pipelines. Cultura: International Journal of Philosophy of Culture and Axiology, 22(12s), 464-486.
60. Chen, Y., Wu, H., & Li, K. (2025). Federated graph learning for privacy-preserving healthcare risk discovery. IEEE Transactions on Medical Imaging, 44(2), 601–614.
61. Kumar, V., Sharma, P., & Gupta, R. (2025). Community-aware graph clustering for patient outcome prediction. Expert Systems with Applications, 258, 124981.
62. Park, J., Kim, S., & Choi, H. (2025). Graph neural networks for multimodal healthcare analytics. Information Fusion, 106, 102315.
63. Babaiah, C., Dobriyal, N., Shamila, M., Aitha, A. R., Patel, S. P., & Upodhyay, D. (2025, December). Intelligent Fault Detection and Recovery in Wireless Sensor Networks Using AI. In 2025 IEEE 5th International Conference on ICT in Business Industry & Government (ICTBIG) (pp. 1-6). IEEE.
64. Davis, M., Cooper, J., & White, R. (2025). Knowledge graph integration for clinical risk stratification. Journal of Biomedical Semantics, 16(1), 11.
65. Lee, D., Huang, C., & Wang, F. (2025). Hypergraph learning for disease progression modeling. Artificial Intelligence in Medicine, 154, 102931.
66. Kolla, S. K., & Bandi, V. D. V. K. (2025). Autonomous Clinical Monitoring Platforms Using Reinforcement Learning and Deep Neural Networks. International Journal of Advanced Research in Computer Science & Technology (IJARCST), 8(6), 13300-13313.
67. Martins, A., Costa, R., & Silva, P. (2025). Semantic interoperability through healthcare knowledge graphs. International Journal of Medical Informatics, 196, 105672.
68. Gupta, N., Jain, S., & Mehta, K. (2025). Explainable graph clustering for healthcare decision support systems. Knowledge-Based Systems, 312, 112845.
69. Mangala, N., & Kolla, S. H. (2025). Real-Time Feature Engineering for Streaming AI Workloads Using PySpark and Azure Event Hubs. International Journal of Multidisciplinary Research in Science, Engineering, Technology & Management, 1(6), 93-105.
70. Brown, E., Walker, G., & Lewis, D. (2025). Large-scale healthcare graph mining for adverse event detection. Journal of the American Medical Informatics Association, 32(4), 921–933.

ISSN: 3067-4166

AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS

VOLUME: 03 ISSUE: 01

RECEIVED: JANUARY 21 REVISED: FEBRUARY 16 ACCEPTED: FEBRUARY 23 PUBLISHED: MARCH 17

71. Sun, X., Zhang, T., & Liu, Y. (2025). Dynamic patient graph construction for hospital risk monitoring. Computers in Biology and Medicine, 184, 109437.
72. Ahmed, R., Khan, M., & Ali, S. (2025). Risk propagation analysis in healthcare knowledge networks. Applied Soft Computing, 162, 112102.
73. Taylor, J., Green, B., & Moore, S. (2025). Graph-based discovery of latent clinical patterns in interoperable EHR systems. BMC Medical Informatics and Decision Making, 25(1), 88.
74. Segireddy, A. R. (2025). AN INTELLIGENT, REAL-TIME DIGITAL FABRIC FOR HEALTHCARE AND FINANCIAL ECOSYSTEMS USING AUTONOMOUS LEARNING AND GENERATIVE SYSTEMS. TPM–Testing, Psychometrics, Methodology in Applied Psychology.
75. Li, Y., Zhou, Q., & Xu, H. (2025). Self-supervised graph representation learning for healthcare applications. Pattern Recognition, 159, 110998.
76. Rivera, L., Thompson, P., & Scott, J. (2025). Patient similarity graph clustering for early risk detection. Health Informatics Journal, 31(1), 1460458225130123.
77. Garcia, J., Torres, M., & Fernandez, A. (2025). Healthcare network intelligence using graph analytics and FHIR-based interoperability frameworks. IEEE Access, 13, 48711–48728.
78. Nagabhyru, K. C., & Kumar, M. V. K. (2025). Generative AI Meets Data Engineering: Automating Code, Query Generation, And Data Insights in Large Scale Enterprises. Query Generation, And Data Insights in Large Scale Enterprises (April 23, 2025).