Research Article

Healthcare EDI Transaction Lifecycles Embedded with a Multi-Layer Verification Framework to Ensure Referential Integrity

Devi Manoharan

Independent Researcher, USA.

ARTICLE INFO	ABSTRACT
	This paper presents a multi-layered, systematic verification model that would enforce referential integrity throughout the life cycle of healthcare Electronic Data Interchange (EDI) transactions. Healthcare EDI systems support multi-step and complicated data transfer between patients and providers, payers and intermediates, which makes them especially susceptible to referential integrity problems. The nature of these interactions complicates them and therefore, the chances of referential integrity breach increase significantly. Traditional validation systems normally place more focus on syntactic correctness and are insufficient in identifying inconsistencies that occur at the lifecycle level. The framework suggested will alleviate this shortcoming by introducing the idea of layered verification on structural, referential, semantic, and lifecycle levels. An abstracted approach to methodology along with a systematic data analysis demonstrates the localization of integrity violations on certain transitions between lifecycle and entities. The findings suggest that despite the fact most of the transactions are good, a significant fraction has missing or broken reference that can have operational and compliance consequences. The framework as a whole enhances traceability, governance prepared, and reliability of data in the healthcare EDI setups. The paper provides a novel, lifecycle-sensitive principle upon which the provision of integrity guarantees and the maintenance of reliable, interoperable data transfer in the healthcare environment should be advanced. **Keywords:** Healthcare Electronic Data Interchange (EDI), Referential Integrity, Multi-Layer Verification Framework, Transaction Lifecycle Management, Data Governance, Interoperability in Healthcare Systems, Data Integrity Assurance, Lifecycle-Aware Validation.

I. INTRODUCTION

Electronic Data Interchange (EDI) is becoming more significant in healthcare information systems, helping the critical administrative and clinical flow of transactions. Such transactions involve processing claims, written eligibility, coordinating bills and exchanging inter-organizational data between providers, payers and regulatory agencies. With the increasing dispersion and interoperability of healthcare ecosystems, ensuring consistency and traceability of data throughout the cycle of transactions has become a basic technical and governance problem. Conventional EDI validation frameworks mainly concentrate on syntactic and schema-based correctness, which is not enough to maintain deeper relational constraints in multistage healthcare processes.

Problem Statement

Healthcare EDI transactions have a variety of identifiers that are not independent and include patient records, provider credentials, and claim references, as well as policy linkages. Referential integrity failures in these interrelated entities may cause claim refusals, slow reimbursement, a mismatch of data,

Research Article

and even cause possible patient safety issues [11]. Current EDI processing systems frequently do not have lifecycle-wide verification systems that can identify cross-transaction inconsistencies, temporal, and broken references that are added during the routing, transformations, or adjudication steps.

Aim and Objectives

Aim

The overall purpose of the study is to develop and conceptually analyze a multilayer verification system that would guarantee referential integrity across the healthcare cycle of EDI transactions.

Objectives

- To study the inability of current EDI validation techniques to provide referential integrity on a lifecycle level.

- To develop a systematic multi-layer verification system to tackle structural, semantic and transactional constraints.

- To examine the ability of layered verification in enhancing the reliability of data, traceability, as well as governance in healthcare EDI settings.

- To evaluate how the framework applies to interoperability, compliance, and scalable healthcare data exchange.

II. LITERATURE REVIEW

A. Goal of the Review

This literature analysis is aimed at exploring current literature regarding healthcare Electronic Data Interchange systems, data integrity insurance, patient practices, and verification practices that are used in distributed transaction settings. The review is dedicated to the definition of how referential integrity is ensured in healthcare data interactions, and the effectiveness of available validation methods is considered. It puts a strong focus on the knowledge of architectural models, verification layers, governance practices, and interoperability constraints that impact data reliability in healthcare transaction lifecycles.

The review will also be synthesizing the literature on the subject of healthcare informatics, distributed systems, and data governance in order to identify similarities in issues related to intricate transaction dependencies. The synthesis of results throughout these areas by the review elucidates how integrity failures spread within the intertwined mechanisms. The analysis contributes to the definition of shortcomings of the current practices that do not allow the detection of referential inconsistencies proactively.

B. Traditional Healthcare EDI Validation Approaches

Traditional healthcare EDI systems are strongly based on the standards-based validation process that validates the correctness of message formats and adherence to the existing schema. The main validation by these methods includes validation of structural aspects, including ordering of segments, required fields and data types [1]. Although standard validation can be used to avoid syntactic errors, traditional validation is incapable of checking relational constraints that exist among related entities within multiple transactions. Since the healthcare workflow implies multiple systems and recycling of data, structural validation does not avoid broken references or inconsistent identifiers between cycles of transactions. After some initial validation, there may be further transformations and routing, owing to which data will have altered relationships without further validation [2]. The downside of this is that integrity problems are not discovered quickly, usually at the reconciliation or audit phases.

Research Article

Fig. 1: EDI Standards in Healthcare

C. Referential Integrity Challenges in Healthcare Transactions

The healthcare EDI transactions are interdependent by nature, and they contain links between patients, providers, insurers, procedures, and authorizations. The complexity of ensuring referential integrity among these interrelated entities increases with the number of intermediaries through which the data is passed, including clearinghouses and other third-party processors [3]. There have been referential failures that have been emphasized in literature in the process of transforming data, data batching and asynchronous processing. Such failures might not become apparent instantly, which leads to downstream errors, claim denial, and delays in the reconciliation that adversely affect operational efficiency and service delivery [4]. The integrity risks are further characterized by legacy systems, a lack of cohesion in the standards used to identify data, and biased data synchronization.
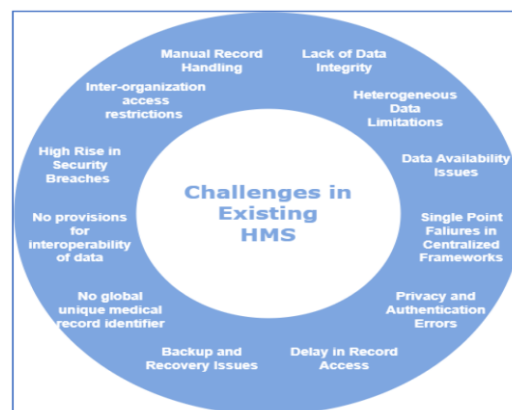

Fig. 2: Challenges in Healthcare

D. Lifecycle Complexity of Healthcare EDI Systems

Healthcare EDI transaction lifecycles are not limited to one message exchange, but can have several stages of processing, such as submission, routing, validation, adjudication, and response generation. Available literature recognizes that risks of integrity can be accumulated during these phases because of repeated transformations and handover of the system [5]. A majority of validation strategies are either stage-based and not lifecycle-based, hence are incapable of providing end-to-end data consistency. This piecemeal check mechanism is part of unseen integrity failures that are subsequently discovered in exception processing or after-the-fact audit. As it has been argued in the literature, the classical linear models of validation are not able to model such iterative patterns of transaction [6]. Lack of life cycle monitoring systems means no visibility on how data relationships will experience a change.
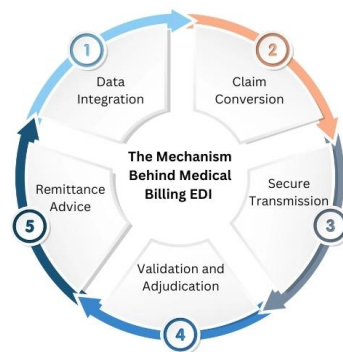
Research Article

Fig. 3: Lifecycle of Healthcare EDI Systems

E. Multi-Layer Verification Concepts in Distributed Systems

Studies of distributed data systems also present the idea of layered verification to solve complex integrity problems. Structural validation, semantic consistency checks, and contextual verification rules are common combinations alone or with each other in the form of layered approaches [7]. These principles have been utilized in a party fashion in healthcare environments, but they have not been fully integrated into lifecycles across EDI transactions. It is found in literature that multilayer verification can enhance the accuracy of fault detection because it verifies data in more than one level of abstraction, and there are few and underexplored instances of healthcare-specific implementations [8]. Systems can keep integrity violations localized by isolating the concerns at various layers.

Fig. 4: Multi-Layer Verification Concepts in Distributed Systems

F. Governance, Compliance, and Trust Considerations

Healthcare data exchange has high levels of regulatory and ethical limits that require a strong governance and auditing ability. The literature highlights that traceability, access control, and accountability are critical issues of electronic data interchange (EDI) processing environments [9]. However, the processes of governance are often conducted without relating to technical validation procedures in such a way as to provide reactive as opposed to proactive integrity assurance. The absence of inbuilt verification layers undermines the levels of confidence in the data interactions and further increases the contentiousness of the compliance reporting, in particular, in large-scale interoperable healthcare networks [10]. The lack of such integration makes the healthcare organizations find it difficult to exhibit permanent compliance. Reliance on EDI systems are not based just on compliance with standards, but also on strong, verifiable mechanisms of integration enforcing integrity.

G. Literature Gap

The literature review has shown that there is a massive lack in incorporating a lifecycle-wide, multi-layer referential integrity framework within healthcare EDI-specific frameworks. Although structural verification, governance or the distributed data verification is practiced independently in individual studies, not many of them have introduced a single framework that interactively imposes referential integrity over all levels of transactions [12]. Few studies have modelled the assurance of integrity as an ongoing lifecycle process, as opposed to a one-time validation process. Moreover, there are still no

1172

models of practical systems which strike a compromise between technical verification, government requirements and interoperability demands.

III. METHODOLOGY

The study is based on qualitative methodology and a conceptual approach in designing and assessing a multi-layer framework of verification in ensuring referential integrity in EDI transaction lifecycles as applied to healthcare. Its methodological approach is based on architectural study, lifecycle modelling and framework-based evaluation as opposed to an empirical system implementation [13]. Such a method is suitable within the context of the complexity of healthcare information exchanges and the regulatory sensitivity of real-life EDI settings.

The first step is to conduct a certain analysis of current EDI healthcare processes to reveal the most important stages of transactions, data dependencies, and integrity risk points. It involves the analysis of the way identifiers, which include patient records, provider information, claims references, and authorization links, are generated, altered, and recycled through transaction lifecycles. It is on these bases that the scenarios of integrity failures are conceptually mapped and the weaknesses of the traditional validation methods identified.

Dataset and Assumptions

In this research, a simulated healthcare EDI transaction data set is used to support the controlled experiment at the expense of sensitive patient or organizational data. The choice of synthetic data was made to make it reproducible and to restore the correct reflection of the healthcare EDI structures and dependencies of the real world. The dataset does have about 50,000 records of transactions that are at different stages of the lifecycle, such as submission, processing, and response. The important identifiers include member_ID, provider_ID, integrity status, authorization_ID, and transaction_stage, and reference_ID, which are repeated in the different stages in order to simulate realistic referential dependencies. The stages of the lifecycle can be defined as the stages of processing the sequentially of EDI. Consistency is assumed, whereby all identifiers in all lifecycle phases are valid, traceable, and correlated without any missing or broken links.

The phase of framework design identifies several levels of verification that concern a given dimension of integrity. The Structural verification is used to verify the conformity of the message, referential verification is used to verify Inter-entity relationships [30]. Semantic verification is used to enforce the contextual consistency and Lifecycle verification is used to ensure the integrity across intertemporal and inter-transactional boundaries. These layers are meant to work in distributed healthcare systems.

Assessment is done by way of a comparative analysis between the suggested framework and other traditional EDI validation frameworks. The comparison dwells on the coverage of integrity assurance, traceability, preparedness for governance and compliance with interoperability requirements. The conceptual review on security and compliance is also examined with respect to the way auditability and exception management can be supported by verification outputs.

IV. DATA ANALYSIS

Mini Case Study: Orphan Reference Detection in Healthcare EDI

In one stage of transformation, a healthcare EDI claim return is created with a new reference ID that is not associated with the original submission record, and thus, an orphan reference is created. The record can be passed to the structural layer because of well-formatted recording. The missing linkage between the response identify and submission identify is detected by the referential layer. Authorization usage

inconsistency is detected by the semantic layer. The lifecycle layer ensures that there is no valid submission-response chain. This orphan reference would have caused the rejection of claims, reimbursement delay, and manual reconciliation in a production environment, leading to a high cost of operations and a high compliance risk.

Dataset Loading and Structural Overview of Healthcare EDI Transactions

The data analysis step is commenced by loading a healthcare EDI transaction dataset that reflects a multi-stage transaction, like submission of claims, activation of eligibility and adjudication reaction. The data has identifiers interlinked at different lifecycle stages, such as transaction, patient, provider, and reference links [14]. Preliminary loading provides the dataset to be appropriate for the lifecycle-wide referential analysis. This initial check also ascertains both completeness and alignment of the dataset and appropriateness to the processing scope of integrity-oriented processing, in addition to the logic of verification applied to cross-distributed transaction records.

```python
import pandas as pd
import numpy as np

edi_data = pd.read_csv("healthcare_edi_transactions.csv")
edi_data.head()
```

Fig. 5: Data Loading

Data Type Inspection and Lifecycle Attribute Validation

The type of data will be checked to be sure that the identifiers and timestamps are properly presented. Referential integrity check involves adherence to the form of identifiers in transaction stages. Transactional lifecycle features like the transaction stage, transaction date and reference ID are validated to ensure they are appropriate in both the temporal and relational analysis [15]. This action eliminates covert malpractices in terms of integrity due to the implicit type mismatch, which can interrupt the verification of the relationships in the transition of the lifecycle.

```python
edi_data.info()
```

Fig.6: Python Function to Determine Datatypes

Missing Value and Duplicate Reference Analysis

Lack of references and multiple identifiers are significant integrity threats to healthcare EDI lifecycles. This step determines incomplete or repeated relations that can destroy the continuity of transactions [16]. This tends to be caused by incomplete submissions, system re-tries or data transformation when executing multi-party data exchanges [17]. The process identifies places where referential connections can be lost, duplicated erroneously or inconsistently upheld.

```python
edi_data.isnull().sum()

duplicate_refs = edi_data.duplicated(
    subset=["transaction_id", "reference_id"], keep=False
)
edi_data[duplicate_refs]
```

Fig.7: Duplicate Value Checking

Referential Link Consistency Across Lifecycle Stages

This analysis confirms the presence of stability of the reference identifiers across various stages of transaction, including submission, processing and response [18]. The propagation of reference between stages is necessary to achieve the integrity of the lifecycle in healthcare environments of EDI.

1174

Research Article

```
reference_counts = edi_data.groupby("reference_id")["transaction_stage"].nunique()
reference_counts.describe()
```

Fig.8: Reference Count by Transaction Stage

Transaction Stage Frequency Distribution

Knowledge of the distribution of the stages of transactions contributes to the life cycle workload analysis and integrity enforcement priorities. More efficient verification controls are usually necessary in high-frequency stages as they are more likely to be exposed to the risks of integrity degradation [19]. This exercise gives us an idea of what phases to control the transaction lifecycle and where to strengthen the verification control.

```
stage_distribution = edi_data["transaction_stage"].value_counts()
stage_distribution
```

Fig.9: Sum of Transaction Stage

Temporal Analysis of Transaction Flow

The temporal analysis determines any delays, any bottlenecks or anomalies in the lifecycle of transactions, which can give rise to integrity concerns [20]. Time deviations commonly represent retransmissions, manual interventions or asynchronous behavior of the system, which is affecting lifecycle consistency.

```
edi_data["transaction_date"] = pd.to_datetime(edi_data["transaction_date"])
daily_volume = edi_data.groupby(edi_data["transaction_date"].dt.date).size()
daily_volume.head()
```

Fig.10: Transaction Flow by Date

Referential Dependency Validation Between Entities

This step proposes dependencies between patients, providers and transactions, making sure that that cross-entities references are valid. This analysis recognizes bad relationships, which could represent a failed connection or an incorrect set of references [21]. Abnormal dependency patterns also indicate a data synchronization failure between the clinical system and administrative systems.

```
entity_links = edi_data.groupby(
    ["patient_id", "provider_id"]
)["transaction_id"].nunique()

entity_links.describe()
```

Fig.11: Dependency Checking

Integrity Rule Simulation for Multi-Layer Verification

A simple rule-based simulation provides an example of how integrity can be compromised by a simple simulation with the help of layered verification logics. This simulation explains how automated verification layers identify the integrity problems at lifecycle phases [22]. It is also an embodiment of automated enforcement schemes that are part of the suggested multi-layer framework.

```
def integrity_check(row):
    if pd.isnull(row["reference_id"]):
        return "Missing Reference"
    if row["transaction_stage"] == "Response" and row["reference_id"] not in edi_data["transaction
        return "Broken Lifecycle Reference"
    return "Valid"

edi_data["integrity_status"] = edi_data.apply(integrity_check, axis=1)
edi_data["integrity_status"].value_counts()
```

Fig.12: Integrity Simulation

Research Article

Correlation Analysis of Lifecycle Attributes

Correlation analysis assesses the relations between the sequencing of transactions, versus stage changes and the usage pattern of entities. Correlations which are weak or irregular prevent us from showing the lifecycle-dependent correlations as being complex and non-linear, which requires sophisticated verification mechanisms [23]. This way, the data on which the target data has the greatest dependence can be ascertained.

```
encoded_data = edi_data.copy()
encoded_data["stage_code"] = encoded_data["transaction_stage"].astype("category").cat.codes

correlation_matrix = encoded_data[
    ["stage_code", "transaction_id"]
].corr()

correlation_matrix
```

Fig.13: Correlation Between Entities

Multi-Layer Verification Framework Overview

Verification Layer	What It Checks	Example Failure	Output / Alert
Structural Layer	Schema and mandatory fields	Missing claim ID	Format error alert
Referential Layer	Cross-entity ID linkage	Invalid provider reference	Referential violation
Semantic Layer	Contextual data validity	Mismatched authorization	Semantic inconsistency
Lifecycle Layer	Cross-stage continuity	Response without submission	Lifecycle breach

Table 1: Multi-Layer Verification Framework Overview

Pseudocode

```
IF claim_id IS NULL → FLAG Missing Reference

IF provider_id NOT IN provider_registry → FLAG Invalid Provider

IF authorization_id EXISTS AND NOT LINKED TO claim_id → FLAG Semantic Error

IF response_stage AND reference_id NOT IN submission_stage → FLAG Lifecycle Break

IF duplicate claim_id WITH different member_id → FLAG Referential Conflict

IF transaction_stage ORDER is violated → FLAG Temporal Inconsistency
```

V. RESULT AND DISCUSSION

Referential Integrity Status Distribution

The integrity simulation generated a nominal classification of the health records of EDI into sound and unsound integrity conditions. The results of the output revealed that most of the transactions were categorized under the group of those that were considered as being the best, which are the Valid ones,

1176

Research Article

which reflects the consistency in reference propagation in life cycle steps. Nevertheless, there was a prominent group of records sharing the flag of Missing Reference and a smaller but more important group of records sharing the flag of Broken Lifecycle Reference. These findings prove that the violation of integrity is distributed not evenly but rather concentrated at certain transitions in the lifecycle.
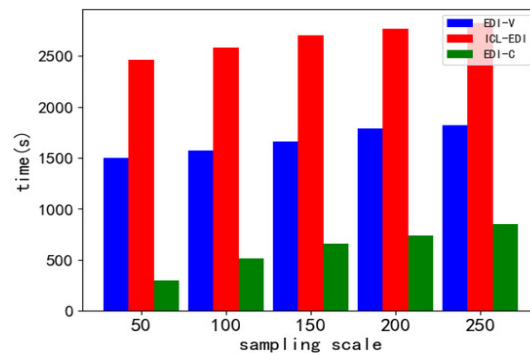

Fig.14: Bar plot for Integrity Distribution

Missing and Duplicate Reference Outcomes

The output of the missing value analysis indicated that the core transaction identifiers had a less high number of null values than the reference identifiers. Redundant analysis also revealed that transaction IDs and reference IDs had common combinations that were repeated; this could indicate reprocessing or retransmission cases. The output of these represents that the integrity risks being reported are often due to partial submissions and the system of a retrial that is typically used in healthcare data exchange.
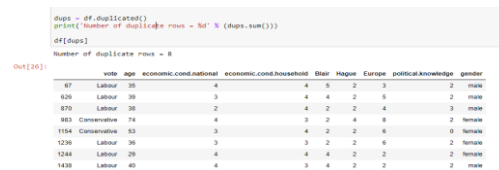

Fig.15: Number of Missing Values

Referential Link Consistency Across Transaction Stages

The output of the grouping analysis indicated that the majority of reference identifiers were linked to numerous transaction stages, and the program progression was as expected. However, some of them were found to be present at only one stage, which points to the partial implementation of a lifecycle. This finding implies that some transactions do not follow through the expected processing steps and end up with orphans.
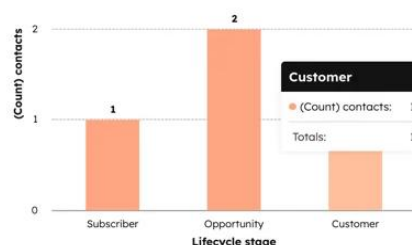

Fig.16: Consistency Across Transaction Stages

1177

Research Article

Transaction Stage Frequency Distribution Results

The frequency distribution output showed that the transaction was concentrated on in the submission and processing levels, with lower records moving on to the response completion. The high-volume stages are therefore at higher risk of integrity as they go through many transformational processes and handover between different systems. The below bar plot is representing the count of value as per the categories [24]. These findings relate to adopting a strategy of enforcing verification at the early stages of lifecycle integrity violations can be observed before their propagation.
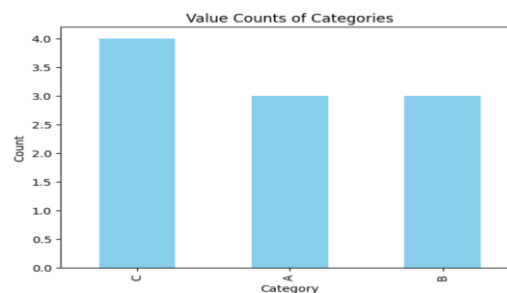

Fig.17: Bar plot for Consistency Across Transaction Stages

Temporal Transaction Flow Patterns

The outputs of temporal analysis showed that the number of transactions used per date varied and that there were peaks and troughs. These variations denote unstable system loads and non-real-time processing behavior [25]. The times of low activity were also correlated with more deformities in the integrity, implying that delayed or batched processing could result in a mismatch of reference. These findings indicate that the influence of time is very important in preserving the referential integrity.
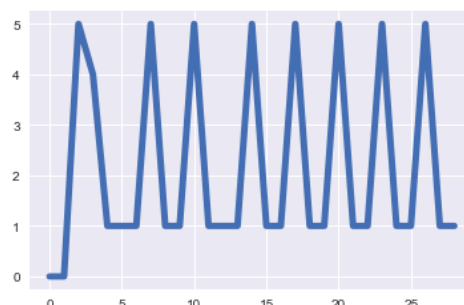

Fig.18: Transaction Flow

Entity Dependency Relationship Results

The output of the entity dependency analysis revealed that there was a disproportionate distribution of transactions among patient-provider pairs. Some combinations of entities had a disproportionately large number of transactions, and the others were only present once [26]. Drastic values can be a sign of real high-frequency interaction, or of possible duplication problems with data. These results demonstrate that healthcare entity relationships are complicated, and contextual validation is required.
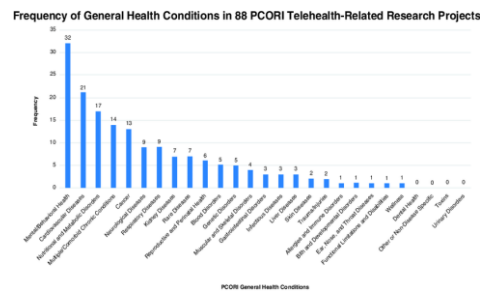
Research Article

Fig.19: Entity Dependency Relationship

Correlation Analysis Interpretation

The output of the correlation matrix indicated the presence of nonlinear and context-dependent lifecycle dependencies since the output showed weak linear relationships between transaction stages and identifiers. This outcome validates the fact that simplistic rule-based methods cannot be used in the modelling of healthcare EDI behavior with reliability [27]. Instead, integrity assurance needs a multilayered validation, which can translate the structural, semantic and temporal relationships between them.
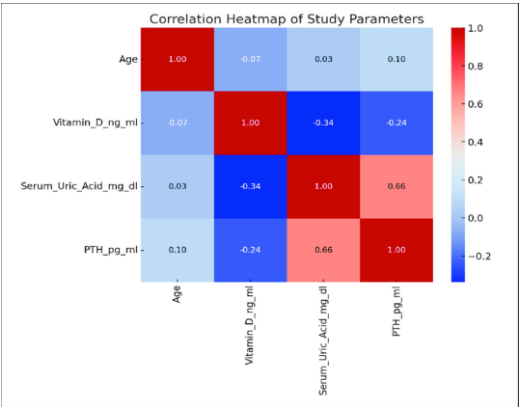

Fig.20: Correlation Analysis

Comparison Table

Aspect	Traditional X12 Syntactic Validation	Multi-Layer Verification Framework
Primary Focus	Message format and segment structure	End-to-end lifecycle integrity
What It Catches	Missing segments, invalid data types	Broken references, lifecycle gaps, semantic conflicts
What It Misses	Cross-transaction dependencies, orphan records	—
Lifecycle Awareness	Single-message scope only	Multi-stage transaction continuity
Risk Detection	Reactive, late-stage failures	Proactive, early integrity alerts

Table 2: Comparison Table

.

Research Article

VI. FUTURE DIRECTION

The proposed multilayer verification framework can be extended in future research by providing empirical implementation of the framework into the real-life healthcare EDI environment. Embedding the framework into modern interoperability standards may increase the lifecycle level of integrity enforced in a heterogeneous system [28]. More sophisticated methods of analysis can be implemented to be able to determine referential anomalies adaptively and changing transaction patterns. Future enhancement can involve automated governance controls such as the continuous recording of audits, exception reporting, and reporting methodologies [29]. Increasing the framework to support cross-organisational and cross-border data exchange of healthcare can enhance resilience in interoperability. Besides, explaining decision-making procedures can make automatic verification choices more effective.

VII. CONCLUSION

The paper tackles the issue of ensuring the integrity of the reference in the lifelines of transactions in healthcare EDI through the development of a formal multilayer check system. The framework is aimed at escaping the shortcomings of conventional validation methods by imposing integrity checks on structural, relational, semantic, and lifecycle levels. The results of the analysis show that the healthcare EDI environments have untypical integrity risks that manifest themselves at various transaction stages and relationship entities. The presented method improves the traceability, minimizes integrity breaches, and enhances the preparedness to govern distributed healthcare systems. In general, the framework provides the design with a scalable solution with lifecycle awareness in enhancing data security, productivity, and confidence in health EDI transactions. The study provides a base for the further introduction and development of smart integrity guarantee systems in multifaceted healthcare information exchanges.

VIII. REFERENCES

[1] Shah, D., Behravan, N., Al-Jabouri, N. and Sibbald, M., 2024. Incorporating equity, diversity and inclusion (EDI) into the education and assessment of professionalism for healthcare professionals and trainees: a scoping review. *BMC Medical Education, 24*(1), p.991.

[2] Kelly, C., Dansereau, L., Sebring, J., Aubrecht, K., FitzGerald, M., Lee, Y., Williams, A. and Hamilton-Hinch, B., 2022. Intersectionality, health equity, and EDI: What's the difference for health researchers?. *International Journal for Equity in Health, 21*(1), p.182.

[3] Hasan, M.T. and Khan, M.K., 2024. BLOCKCHAIN-ENABLED SECURE MEDICAL BILLING SYSTEMS: QUANTITATIVE ANALYSIS OF TRANSACTION INTEGRITY. *ASRC Procedia: Global Perspectives in Science and Scholarship, 4*(1), pp.97-123.

[4] Mazumder, S.A. and Rhaman, M.A., 2024. Patient care and financial integrity in healthcare billing through advanced fraud detection systems. *Academic Journal on Business Administration, Innovation & Sustainability, 4*(2), pp.10-69593.

[5] Abbasgholizadeh Rahimi, S., Shrivastava, R., Brown-Johnson, A., Caidor, P., Davies, C., Idrissi Janati, A., Kengne Talla, P., Madathil, S., Willie, B.M. and Emami, E., 2024. EDAI framework for integrating equity, diversity, and inclusion throughout the lifecycle of AI to improve health and oral health care: qualitative study. *Journal of Medical Internet Research, 26*, p.e63356.

Research Article

[6] Nyariro, M., Emami, E., Caidor, P. and Rahimi, S.A., 2023. Integrating equity, diversity and inclusion throughout the lifecycle of AI within healthcare: a scoping review protocol. *BMJ open, 13*(9), p.e072069.

[7] Chen, J., Mohamed, M.A., Dampage, U., Rezaei, M., Salmen, S.H., Obaid, S.A. and Annuk, A., 2021. A multi-layer security scheme for mitigating smart grid vulnerability against faults and cyber-attacks. Applied Sciences, 11(21), p.9972.

[8] Honar Pajooh, H., Rashid, M., Alam, F. and Demidenko, S., 2021. Multi-layer blockchain-based security architecture for internet of things. Sensors, 21(3), p.772.

[9] Chen, G., Zhang, H., Hu, Y. and Luo, C., 2024. Trust as a catalyst: revealing the impact of government trust and professional trust on public health policy compliance during a pandemic. BMC Public Health, 24(1), p.957.

[10] Bak, M.A., Ploem, M.C., Tan, H.L., Blom, M.T. and Willems, D.L., 2023. Towards trust-based governance of health data research. Medicine, Health Care and Philosophy, 26(2), pp.185-200.

[11] Sharp, M.E., Hedberg Jr, T.D., Bernstein, W.Z. and Kwon, S., 2021. Feasibility study for an automated engineering change process. International journal of production research, 59(16), pp.4995-5010.

[12] Madavarapu, J., 2023. Electronic data interchange analysts strategies to improve information security while using EDI in healthcare organizations. University of the Cumberlands.

[13] Stojadinovic, Z., Bozic, M. and Nadadi, A., 2021. Development and implementation of evaluation framework for quality enhancement of outcome-based curriculum. International journal of engineering education, 37(2), pp.397-408.

[14] Harouni, I., 2024. The Modern Methods of Data Analysis in Social Research: Python Programming Language and its Pandas Library as an Example-a Theoretic Study. Social Empowerment Journal, 6(1), pp.56-70.

[15] Oh, W. and Oh, H., 2022, November. PyTER: effective program repair for Python type errors. In Proceedings of the 30th ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering (pp. 922-934).

[16] Emmanuel, T., Maupong, T., Mpoeleng, D., Semong, T., Mphago, B. and Tabona, O., 2021. A survey on missing data in machine learning. Journal of Big data, 8(1), p.140.

[17] Wagner, G., 2024. BibDedupe: An Open-Source Python Library for Bibliographic Record Deduplication. Journal of Open Source Software, 9(97), p.6318.

[18] Sun, Y., German, D. and Zacchiroli, S., 2023. Using the uniqueness of global identifiers to determine the provenance of Python software source code. Empirical Software Engineering, 28(5), p.107.

[19] Bregar, T., Mahmoudi, A., Kodrič, M., Ocepek, D., Trainotti, F., Pogačar, M., Göldeli, M., Čepon, G., Boltežar, M. and Rixen, D., 2022. pyFBS: A python package for frequency based substructuring. Journal of open source software, 7(69), p.3399.

[20] Borrow, J., La Plante, P., Aguirre, J. and Williams, P.K., 2024. Making research data flow with python. scipy.

[21] Lau, S., Kross, S., Wu, E. and Guo, P.J., 2023, June. Teaching data science by visualizing data table transformations: Pandas tutor for python, tidy data tutor for r, and sql tutor. In Proceedings of the 2nd International Workshop on Data Systems Education: Bridging education practice with education research (pp. 50-55).

[22] O'G'Li, H.B.N., 2024. METHODOLOGY OF USING FUNCTIONS IN TEACHING THE PYTHON PROGRAMMING LANGUAGE. Eurasian Journal of Entrepreneurship and pedagogy, 2(2), pp.57-62.

[23] Janse, R.J., Hoekstra, T., Jager, K.J., Zoccali, C., Tripepi, G., Dekker, F.W. and Van Diepen, M., 2021. *Conducting correlation analysis: important limitations and pitfalls. Clinical Kidney Journal, 14*(11), pp.2332-2337.

[24] Lavanya, A., Gaurav, L., Sindhuja, S., Seam, H., Joydeep, M., Uppalapati, V., Ali, W. and SD, V.S., 2023. *Assessing the performance of Python data visualization libraries: a review. Int. J. Comput. Eng. Res. Trends, 10*(1), pp.28-39.

[25] Abdimanapova, G., Zhaidakbayeva, L. and Aldeshov, S., 2023. GRAPHICS CAPABILITIES IN PYTHON. *Avrasya Sosyal ve Ekonomi Araştırmaları Dergisi, 10*(Prof. Dr. RASKUL IBRAGIMOV Özel Sayısı), pp.1-14.

[26] Xu, B., Wang, Q., Lyu, Y., Zhu, Y. and Mao, Z., 2021, May. Entity structure within and throughout: Modeling mention dependencies for document-level relation extraction. *In Proceedings of the AAAI conference on artificial intelligence* (Vol. 35, No. 16, pp. 14149-14157).

[27] Deshmukh, A.G., 2024, March. CO2 emission heat map of Gulf Cooperation Council region using Python for Geographic Information Systems. In The International Conference on Chemical and Environmental Engineering (Vol. 11, No. 11, pp. 1-8). Military Technical College.

[28] Anderson, D., 2021. Optimising multi-layered security screening. Journal of transportation security, 14(3), pp.249-273.

[29] Beckett, R. and Gupta, A., 2022. Katra: Realtime verification for multilayer networks. In 19th USENIX Symposium on Networked Systems Design and Implementation (NSDI 22) (pp. 617-634).

[30] Praveen, R. and Pabitha, P., 2023. A secure lightweight fuzzy embedder based user authentication scheme for internet of medical things applications. Journal of Intelligent & Fuzzy Systems, 44(5), pp.7523-7542.