*In proceedings of IEEE International Conference on Digital Health (ICDH) 2023 at IEEE World Congress on Services 2023*

# Policy Integrated Blockchain to Automate HIPAA Part 2 Compliance

James Clavin and Karuna P. Joshi
*Department of Information Systems*
*University of Maryland, Baltimore County*
Baltimore, MD, US
{jclavin, karuna.joshi}@umbc.edu

***Abstract*—Healthcare organizations exchange sensitive health records, including behavioral health data, across peer-to-peer networks, and it is challenging to find and fix compliance issues proactively. The Healthcare industry anticipates a growing need to audit substance use disorder patient data, commonly referred to as Part 2 data, having been shared without a release of information signed by the patient. To address this need, we developed and evaluated a novel methodology to detect Part 2 data exchanged between organizations that integrates Blockchain technologies with knowledge graphs. We detect substance use disorder data in patient encounters exchanged using clinical terminology based upon the value sets provided by the National Institutes of Health for the Substance Abuse and Mental Health Services Administration. Generally, we consider sharing Part 2 data without consent as Byzantine medical faults, as they represent data shared between known and trusted network participants, that is valid, but is not relevant, and sharing it causes a breach. In this paper, we present our methodology in detail along with the experiment results. We model a medical network of hospitals based upon the most recent healthcare legislation, TEFCA, and generate synthetic patient encounter data dynamically in HL7 format. We convert exchanged encounter data into a knowledge graph data model so that we can use SNOMED-CT for identifying Part 2 data. For cohorts of 1,000 patients, we detect Part 2 data in a subset of their encounter data shared between organizations and log that securely on an Ethereum-based blockchain.**

***Index Terms*—HIPAA, Substance Use Disorder, Part 2, Byzantine Medical Faults, TEFCA, Automated Compliance**

## I. Introduction

Medical organizations share protected health information by following the guidelines set forth in Health Insurance Portability and Accountability Act (HIPAA) and by adhering to whatever data use agreement governs the variables and cohort of data that can be shared between them. Depending upon the use case, the organization may send data sets via API endpoint or through secure file exchange. Regardless of channel, when a recipient organization receives a data set, some quality checks are typically done to check for syntactic accuracy before assuming the data conforms to the expected format. Most costly and done less frequently is a deeper analysis of the data set, evaluating clinical terms against the intention behind whichever data use agreement governs what data can be shared between organizations. We refer to sharing medical data that is not relevant between trusted partners as Byzantine Medical Faults. Identifying Byzantine Medical Faults is challenging, and these types of issues will likely surface more often in the near future.

The need to share data has grown, driven by socioeconomic forces, including new technology, the pandemic, and legislation. Key to the future of medical data exchange in the U.S. is legislation passed in 2016 - the 21st Century Cures Act. The Cures Act encourages the interoperability of health data through new regulations and organizational structures [1]. The Department of Health and Human Service's Office of the National Coordinator (ONC) is leading the implementation in partnership with a designated Recognized Coordinating Entity (RCE), a non-profit called the Sequoia Project [2]. The RCE is creating a meta-network of qualified health information networks (QHINs). A QHIN must qualify in order to participate, but can be any organization of health care providers, institutions and health information exchanges (HIE) with data sharing agreements in place and IT platforms to exchange data. The Sequoia Project is responsible for holding stakeholder meetings and workgroups to facillitate TEFCA implementation, approves QHIN applications, and in future will monitor the entire network to ensure information blocking does not occur [3]. Sequoia worked with stakeholders to form the draft versions of the Trusted Exchange Framework and Common Agreement (TEFCA) that form the technical and legal foundation for the meta-network of QHINs [2], [4]. The TEF is a list of seven technical principles for QHINs to adhere to, and concentrates around existing standards, quality measures, and information security best practices [4]. The CA is a binding contract for QHINs to complete and submit to the RCE. The Cures Act's Common Agreement stipulates that qualified health information networks must:

> ...maintain, throughout the term of this Common Agreement...a policy or policies of insurance for cyber risk and technology errors and omissions..." [2].

We assume a future state in which TEFCA has been adopted by several QHINs throughout the United States.

The potential benefits of TEFCA for better care, better health, and reduced costs could be substantial moving into the 2020's, particularly in the COVID-19 era, and its design fulfills many of the recommendations identified by the American Medical Informatics Association (AMIA) in 2007 for widespread, secure health data exchange [5], [6]. Using current technologies, handling a high-volume of secure transactions as envisioned in TEFCA will be a remarkable achievement, and in this future state those data exchange issues that typically occur within networks will surface, but with new intermediaries in place. Moving health data within complex networks in an unblocked manner to support TEFCA requires the development and adoption of many new applications and systems. We focus upon the need to insure that Part 2 data is not shared without the patient's consent using blockchain and knowledge graphs.

### A. Ethereum and Health Information Exchange

We migrate the TEFCA meta-network to Ethereum, and create a smart contract to log when any network participant shares Part 2 data out of compliance, while protecting the privacy of the patient and the security of the meta-network. We visualize the number of compliance errors so that observers, such as the RCE, can review the quality and accuracy of data being shared between organizations, the cost incurred, and develop plans for penalizing participants who incorrectly share data. Our purpose is to enable effective oversight of data exchange between QHINs without having to see the data. Breaches requiring notification to the Office of the National Coordinator (ONC) can either be stopped before reaching the minimum threshold of 500 patients per the HIPAA Breach Notification Rule, 45 CFR §§ 164.400-414, or can provide a full audit trail for use in reporting to the ONC [7]. In addition breach notification rules in CA Section 12 state that network participants, including other QHINs and the RCE, must be notified within five days of a breach, including the "natural and likely scope" of the breach [2].

## II. Related Work

The current state of health information exchange builds upon decades of federal government regulation intended to secure patient health data while at the same time making it more portable.

### A. Regulatory Frameworks and the Current State of National HIE

Beginning with 1996's Health Insurance Portability and Accountability Act (HIPAA) there has been regulatory momentum to support digitizing protected health information. The objectives have been to enable patient data exchange between healthcare organizations to improve patient outcomes, enhance quality, and reduce costs. Since then the 2009 HITECH Act encouraged the adoption of electronic health records through financial incentives to providers. And in 2016 the 21st Century Cures Act made patient information blocking - where a patient requests their health data and is not provided it - by healthcare providers illegal.

During this timeframe, EHRs developed as the glue between different sub-systems, connecting previously silo-ed applicatoins designed to support various clinical workflows, e.g., pharmacy order fulfillment or radiology imaging. As EHRs matured, they brought the data generated by sub-systems into one unified patient record with one interface. Initial focus was therefore primarily internal to the healthcare organization, with less incentive provided to create externally facing data exchanges - with the exception of those states that required health information exchange and set up the infrastructure and incentive to share data. In the 2010's large EHR vendors began creating APIs for third party applications to interface with and to integrate with HIEs. While progress was being made, there was not a federal regulation in place to enable efficient interstate health data exchange; most such exchange was done either point-to-point through data use agreement between organizations, or through the patient themselves attempting to coordinate the data exchange. Prior to the Cures Act EHR vendors Cerner and Epic felt they were driving towards effective national health information exchange through market forces, and that the Act was unnecessary and costly [8], [9]. Both companies take up most of the market share for EHR vendors, and as TEFCA advanced both have embraced it, with their respective subsidiaries Epic TIS and CommonWell selected as founding members of the inaugural qualified health information network cohort of six organizations [10]–[13].

*1) HIPAA and 42 CFR Part 2:* Included in HIPAA is 42 Code of Federal Regualtions (CFR) Part 2, which provides protections for sensitive patient information, such as substance use disorder and mental health data [14]. Part 2 data is subject to patient consent; such patient data can only be shared under certain circumstances, with the general usage being a signed release of information from the patient allowing the transfer of data. This regulation is critical for protecting opioid addiction information, an extremely sensitive set of patient data. To encourage patient management of releasing this information, the ONC funded the development of a pilot application called Consent2Share, designed to enable a patient to set up data segmentation of their behavioral and mental health data [15]. Consent2Share was found to have discrepancies between the value sets used for defining sensitive information and what practitioners considered sensitive [16]. Recommendations to improve Consent2Share included using knowledge-based value sets based upon clinical terminologies such as SNOMED-CT for enabling algorithmic approaches to detect sensitive information and using NLP techniques to automate

detecting sensitive information in the medical record [17].

*2) TEFCA:* The network that TEFCA creates will include only approved qualified health information networks (QHINs); as such it is a permissioned (each participant is known and has established their identity) and the threat model is one where participants are generally honest but curious about what data is provided. A recipient will evaluate all data provided by another participant - even if the data provided includes patient information that should not have been shared, such as substance use disorders that require a patient's release of information.

Under TEFCA, the minimum required terms and conditions specify that treatment, utilization, quality assessment and improvement, business planning and development, public health, individual access services, and benefits determination are the defined set of exchange purposes; Qualified Health Information Networks (QHINs) must be able to exchange this minimal data set [18].

The TEF principles determine how data should be exchanged between network participants. It states that data should be exchanged using existing EDI formats, or using standards set by recognized standards defining organizations [4]. The TEF also refers the reader to existing mechanisms for testing the syntactic accuracy of data, such as the ONC C-CDA scorecard and the ONC Inferno program for FHIR [4], [19], [20]. In Principle #2B the TEF requires that any agreement between information networks must include an addendum, or separate data use agreement, that specifies the minimum amount of data to be shared [4]. Under Principle #4A "Privacy, Security, and Safety," the ONC focuses on data accuracy through the application of the C-CDA and FHIR standards [4]. The TEF also states that clinicians should ensure the accuracy of patient data before it is exchanged with external qualified health information networks [4].

The Common Agreement (CA) is a legally binding contract between participants and the RCE which details what the required information types are for exchange between Qualified Health Information Networks (QHIN). Section #10B Individual Access Services Metadata and Data Segmentation addresses requirements for sharing data that is relevant to the intended recipient and for protecting sensitive health information such as substance abuse data [18].

### *B. Technologies for HIE*

*1) Semantic Web:* While exchanging sensitive datasets, organizations need to exchange information, queries, and requests with some assurance that they share a common meaning. This is critical not only for the data but also for the data protection policies followed by the organizations. The interoperability requirement is not just for the data itself but even for describing policies for data protection. One possible approach to this issue is employing Semantic Web techniques to model and reason about standards, such as the semantically rich clinical terminology of SNOMED-CT. The Semantic Web deals primarily with data instead of documents. It allows data to be annotated with machine-understandable meta-data, permitting the automation of their retrieval and their usage in incorrect contexts. Semantic Web technologies include languages such as Resource Description Framework (RDF) and Web Ontology Language (OWL) for defining ontologies and describing meta-data using these ontologies as well as tools for reasoning over these descriptions [21], [22]. These technologies can be used to provide common semantics of medical datasets and policies enabling all software agents who understand basic Semantic Web technologies to communicate and use each other's data and services effectively. Knowledge graphs have been used to extract knowledge about the security and privacy provisions included in HIPAA [23]–[27].

*2) Blockchain:* Blockchain is a promising concept to build upon because it can provide much higher levels of security over existing technologies [28]–[32]. Most blockchain projects in the healthcare sector focus on utilizing blockchain as a reliable platform for secure data sharing, and attempt to address the security and compliance challenges inherent in large-scale health information exchange systems [33]. There has been research into using Ethereum based smart contracts for remote patient monitoring [34]. However, the applicable blockchain-based approaches proposed so far have not directly addressed TEFCA requirements because they are either too specialized to convey to medical data exchange, or too generic to overlay into health systems data [31], [32]. One blockchain-based proposal considers TEFCA alignment, but is not tethered tightly to it, and leaves the technical implementation open-ended [35].

Blockchains can be categorized into two types: permissionless blockchains and permissioned blockchains [36]–[42]. Permissioned blockchains restrict network participants to those who have proven their identity, and run code that is provably secure through Byzantine Fault Tolerance (BFT). Generally speaking BFT protocols state that one must have $3f + 1$ nodes to overcome $f$ failures, in order to guarantee a system will function correctly. This is intuitive, but the underlying principle has proven difficult to reason about [43], [44]. BFT is the core primitive behind blockchains like Ethereum for ensuring system safety (nothing bad happens) and liveness (something good will happen) [45].

## III. Methodology

We develop an Ethereum test network (testnet) of the recently approved inaugural class of qualified health information networks as depicted in Fig. 1 [13]. As shown in Table I all hospitals and QHINs are assigned network addresses for sending and receiving data. The

recognized coordinating entity, the Sequoia Project, is responsible for only the data shared between QHINs. Managing the testnet is a smart contract to handle adding and removing QHINs, and other network participants, and for recording out of compliance transactions to the blockchain. Out of compliance transactions include the sender address, the recipient address, and the compliance error message; excluded is patient data.

For validation data we generate cohorts of synthetic patient data for each hospital in HL7 comma separated value (csv) format, taking into account each hospital's region for demographics and mortality, including incidences of encounters with patients for reasons of opioid use disorder [46]. The data sets are associated with a sender hospital and receiving hospital, which minimally traverses across at least two QHINs. We transfer encounter data, as if in response to a query from another hospital. The outbound data is converted from HL7 into Terse RDF Triple Language (ttl) files so that our knowledge graph can be used to check for compliance errors [47]. The National Institutes of Health maintains value sets of clinical terminology that can be used to identify part 2 data [16].

Using the NIH value set of SNOMED-CT terms that identify encounters that require patient consent, we evaluate the data that was sent to see if it contained substance use disorder encounter data [48]. In the event Part 2 data was shared a log entry is made by the smart contract indicating the sender organization, the recipient organization, and a message stating Part 2 data was shared. This causes a record to be added to the blockchain indicating possible compliance failure.

TABLE I: Ethereum Testnet Accounts

| Type | Name | Ethereum Addr (hex) | QHIN |
|---|---|---|---|
| RCE | The Sequoia Project | 0x00...EA72 | — |
| QHIN | CommonWell | 0x17...5CE | — |
| QHIN | eHealth Exchange | 0x59...a52 | — |
| QHIN | Epic TIS | 0xB3...F90 | — |
| QHIN | Health Gorilla | 0xcC...E4a | — |
| QHIN | Kno2 | 0xb...E1D | — |
| QHIN | KONZA | 0xb6...0bd | — |
| Hospital | Johns Hopkins | 0x4F...6EB | CommonWell |
| Hospital | University Maryland | 0x90...33B | eHealth |
| Hospital | Wake Forest | 0xC1...917 | Epic |
| Hospital | Duke | 0x16...a377 | Health Gorilla |
| Hospital | UCSF | 0x91...5e8 | Kno2 |
| Hospital | K State | 0xA8...458 | KONZA |

We utilized a server running the Ubuntu 20.04.3 LTS operating system allocated with 376GB memory and 72 Intel(R) Xeon(R) Gold 6140 @ 2.30GHz CPU's. We selected Synthea to produce patient encounter data sets in HL7 format and converted those into RDF using a Python library [46], [47], [49]. For querying the RDF data set of encounters and determining if Part 2 data is included, we use RDFox. RDFox is selected because its queries are optimized by keeping the knowledge graph in memory, and given its ability to adapt to new information in real time through forward chaining - features well-suited to big data healthcare [50], [51]. RDFox provides an http endpoint for encounter RDF data to be imported, queried, and deleted in real time through API calls. We develop an API in the Elixir language using the Phoenix web framework to interface with the endpoint as Elixir is a functional language with proven fault tolerance attributes and the Phoenix framework has a rich community of open source solutions, including libraries to interact with Ethereum [52], [53]. The Ethereum blockchain for recording the Part 2 data compliance checks runs locally using OpenEthereum, and is visualized using an Ethereum blockchain browser called BlockScout [54].

## IV. Results

The process for generating patient data in HL7 csv format, then converting into RDF, was measured for single hospitals with varying patient cohort sizes. As shown in Table II, the largest data set generated was 1,172 (1,000 living and 172 deceased) patients and took the longest to generate (125.4*s*). The average time to generate a cohort of 1,000 patients data in HL7 and convert to RDF is approximately 109*s*. We optimize the data generation by creating the data sets in parallel, thereby constraining the data generation to the longest running process.

TABLE II: Data Generation Runtimes

| Patients | Synthetic data runtime (*s*) | TTL data runtime (*s*) | Total (*s*) |
|---|---|---|---|
| 11 | 18 | 0.3 | 18.3 |
| 115 | 23 | 3.5 | 26.5 |
| 116 | 32 | 3.6 | 35.6 |
| 116 | 27 | 4.3 | 31.3 |
| 117 | 25 | 3.3 | 28.3 |
| 118 | 32 | 4.3 | 36.3 |
| 125 | 33 | 4.9 | 37.9 |
| 130 | 29 | 5.3 | 34.3 |
| 1117 | 58 | 41.2 | 99.2 |
| 1130 | 59 | 41.5 | 100.5 |
| 1130 | 59 | 42.5 | 101.5 |
| 1148 | 60 | 40 | 100 |
| 1152 | 69 | 43.1 | 112.1 |
| 1154 | 65 | 49.8 | 114.8 |
| 1154 | 67 | 51.2 | 118.2 |
| 1172 | 74 | 51.4 | 125.4 |
| 1172 | 66 | 45.9 | 111.9 |

Once in RDF format, the system must explore the graph efficiently within that timebound to determine if any encounter data contained one of the SNOMED-CT codes for opioid dependence, before a new data set is generated. We run benchmark tests on our API and

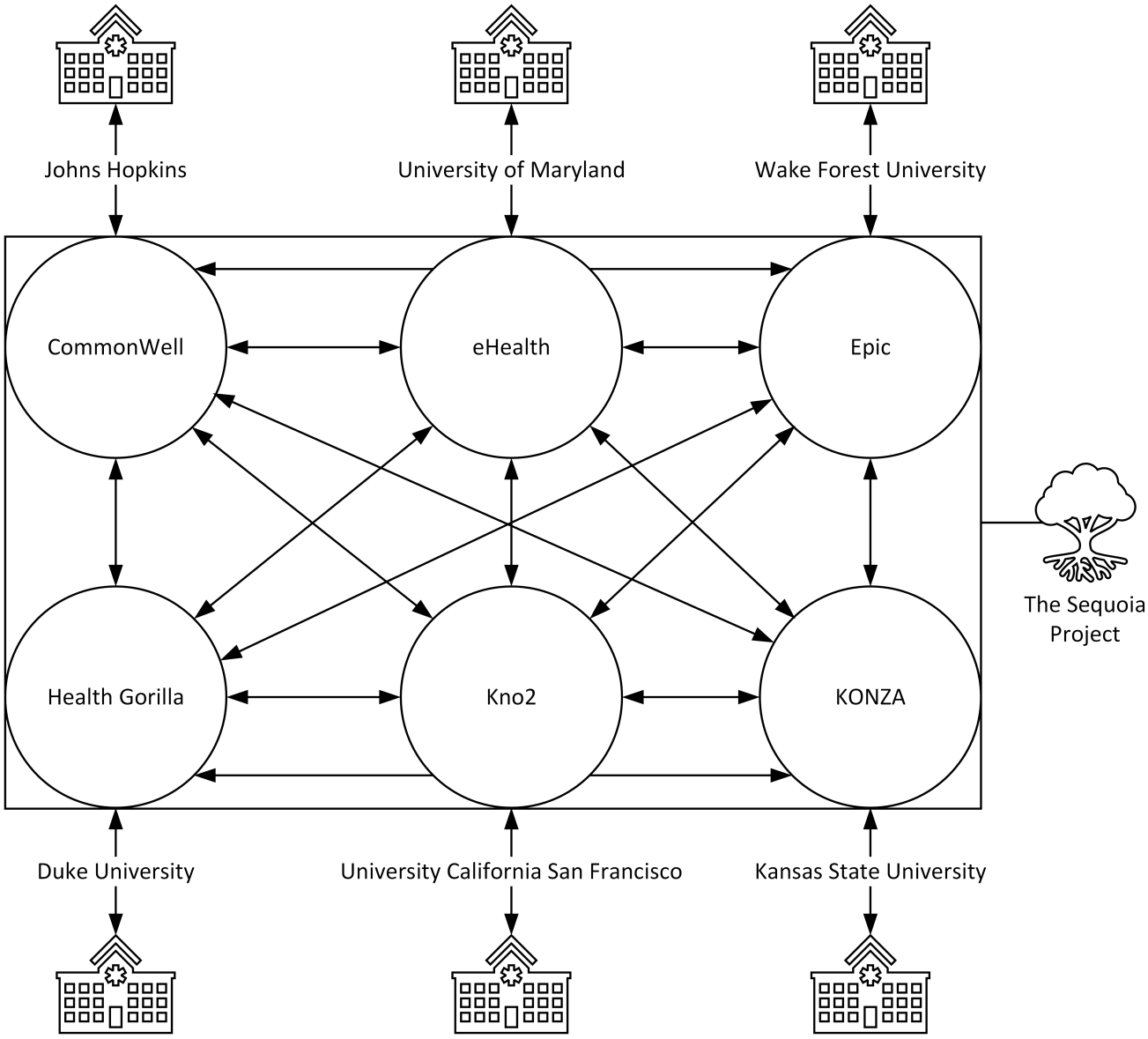


Fig. 1: TEFCA Network

gather the results shown in Table III. Each compliance check includes minimally 1,000 patients and their associated encounters. The longest running compliance check was 11.16$s$ on a file size of 110$MB$.

As Part 2 data is detected, the Elixir-based API invokes the system smart contract to record the sender, receiver, and a log message stating Part 2 data was shared to the testnet. Fig. 2 shows blocks from the testnet, including the addition of a network participant (blue block) and the compliance validation contract calls (green blocks). When Part 2 data has been shared, the transaction details show the sender and the receiver, along with a message, in hex as depicted in Fig. 3. To review the compliance message, the application allows switching to UTF-8, as shown in Fig. 4. The average time between blocks is 10.8$s$ for each 1,000 patient cohort, and the block size is consistently 0.69$KiB$.

TABLE III: Transfer times to check compliance

| **Sender** | **Receiver** | **Encounter file size (*MB*)** | **Compliance Check runtime (*s*)** |
|---|---|---|---|
| Hopkins | University Maryland | 106 | 10.88 |
| University Maryland | Hopkins | 110 | 11.16 |
| Wake | Duke | 108 | 9.58 |
| Duke | Wake | 94 | 10.52 |
| UCSF | Kansas State | 104 | 9.16 |
| Kansas State | UCSF | 95 | 8.85 |

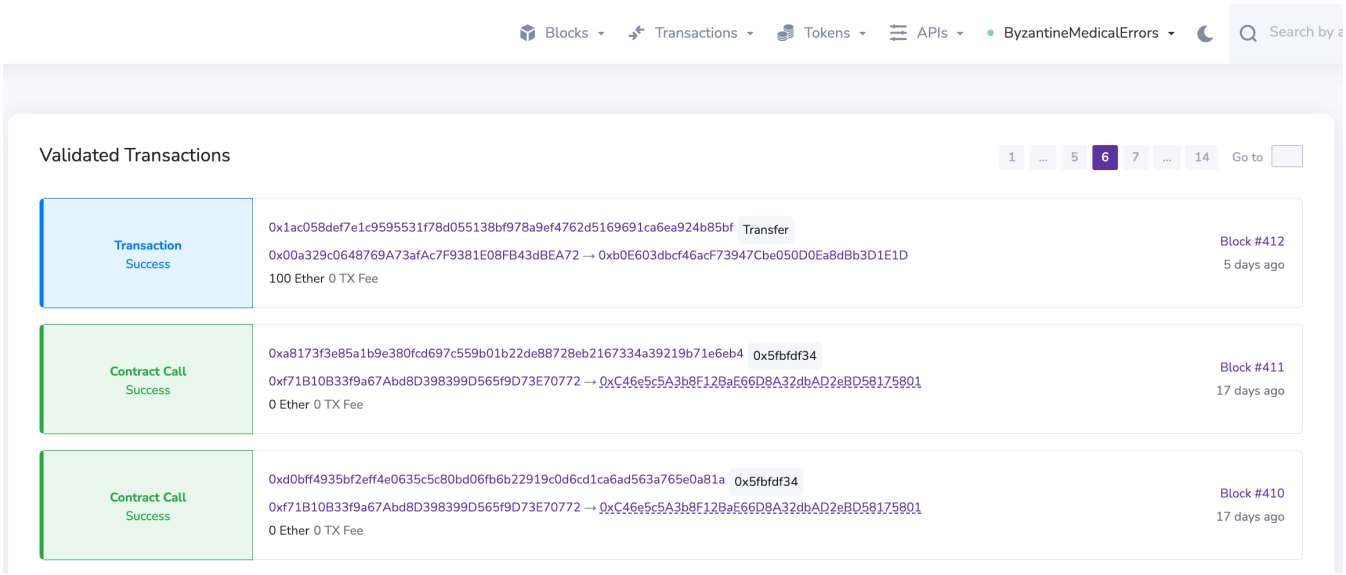


Fig. 2: TestNet Transactions

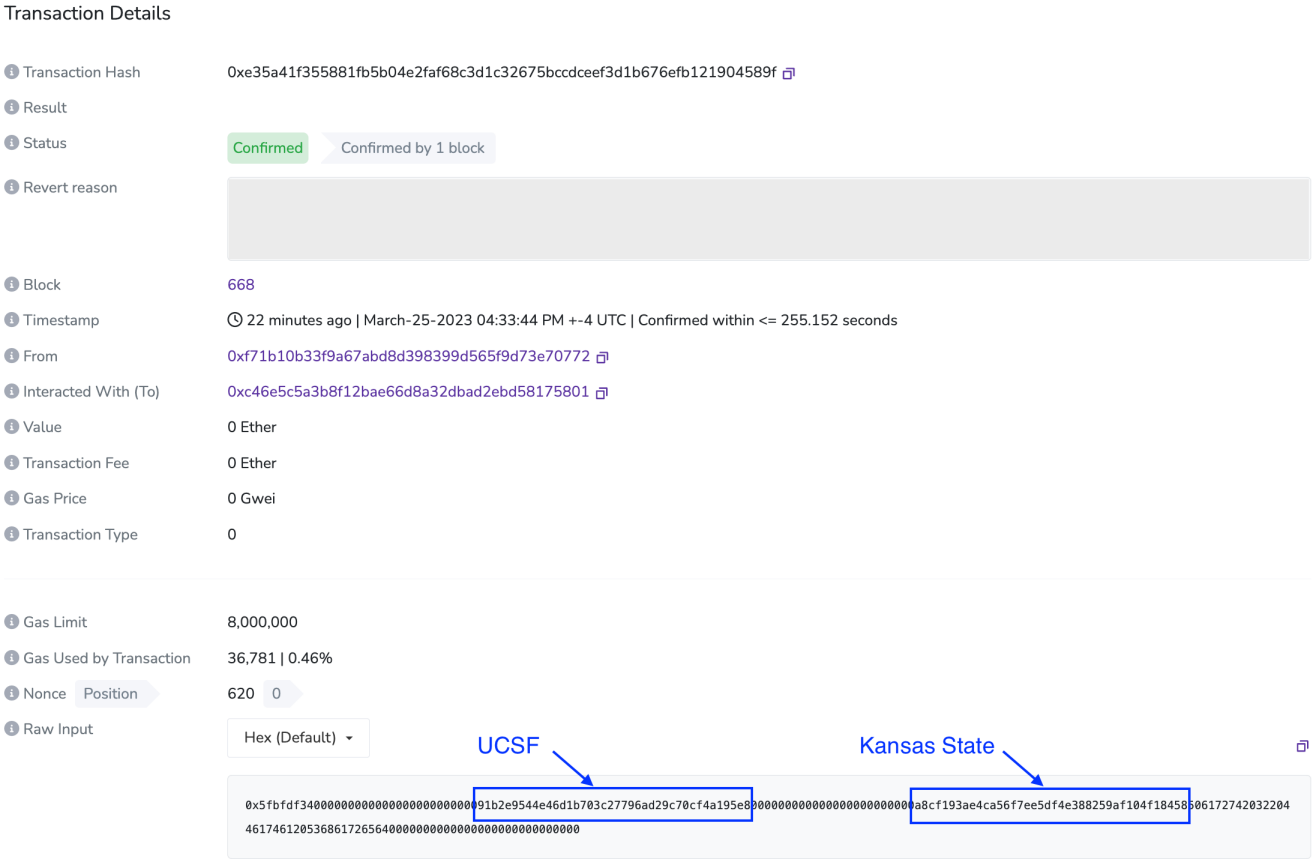


Fig. 3: Compliance Transaction Details - Raw Data Shows Sender and Receiver in Hex

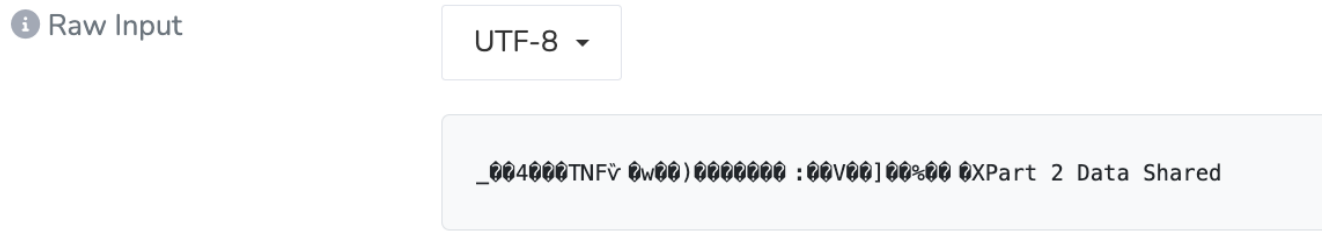


Fig. 4: Compliance Transaction Details - Raw Data Shows Compliance Message in UTF8

## V. Conclusion

In this study, we demonstrated how Knowledge graphs can be integrated with permissioned Blockchain technology to enable the automatic identification of compliance issues within the near future TEFCA network. We present in comprehensive detail the performance metrics for each phase of the data generation, exchange, and compliance checking. As part of our ongoing work, we are augmenting our system and testing it with greater data volumes - more hospitals and more patients - to determine how well it scales. Additionally, we will expand the value set of SNOMED-CT codes to incorporate other clinical coding schemes, such as ICD-9CM, ICD-10CM, and RxNorm. The ability of the system to reason about encounter data based upon HL7 and determine if Part 2 data has been shared, and record that for later inspection, may be of value to QHINs in the future.

## Acknowledgment

This research was partially supported by a DoD supplement to the NSF award 1747724, Phase I IUCRC UMBC: Center for Accelerated Real time Analytics (CARTA), and Office of Naval Research grant # N00014-18-1-2452 and N00014-18-1-2453.

## References

[1] US Department of Health and Human Services. (2020) U.s. department of health and human services. hhs finalizes historic rules to provide patients more control of their health data. https://www.hhs.gov/about/news/2020/03/09/hhs-finalizes-historic-rules-to-provide-patients-more-control-of-their-health-data.html, Last accessed on 2020-03-21.

[2] The Sequoia Project. https://rce.sequoiaproject.org/wp-content/uploads/2022/01/common-agreement-for-nationwide-health-information-interoperability-version-1.pdf.

[3] US Department of Health and Human Services Press Office. (2019) Onc awards the sequoia project a cooperative agreement for the trusted exchange framework and common agreement to support advancing nationwide interoperability of electronic health information. https://www.hhs.gov/about/news/2019/09/03/onc-awards-the-sequoia-project-cooperative-agreement.html, Last accessed on 2020-08-16.

[4] Office of the National Coordinator for Health Information Technology. https://www.healthit.gov/sites/default/files/page/2022-01/trusted_exchange_framework_0122.pdf.

[5] D. F. Sittig and H. Singh, "Covid-19 and the need for a national health information technology infrastructure," *Jama*, vol. 323, no. 23, pp. 2373–2374, 2020.

[6] C. Safran, M. Bloomrosen, W. E. Hammond, S. Labkoff, S. Markel-Fox, P. C. Tang, and D. E. Detmer, "Toward a national framework for the secondary use of health data: an american medical informatics association white paper," *Journal of the American Medical Informatics Association*, vol. 14, no. 1, pp. 1–9, 2007.

[7] U.S. Department of Health Human Services, "Adverse events, near misses, and errors," 1996, https://www.hhs.gov/hipaa/for-professionals/breach-notification, Last accessed on 2022-11-19.

[8] Farr C., "Epic's ceo is urging hospital customers to oppose rules that would make it easier to share medical info," 2020, https://www.cnbc.com/2020/01/22/epic-ceo-sends-letter-urging-hospitals-to-oppose-hhs-data-sharing-rule.html, Last accessed on 2020-03-20.

[9] SEC EDGAR Database, "Form 10-k cerner corporation," 2019, https://www.sec.gov/ix?doc=%2FArchives%2Fedgar%2Fdata%2F804753%2F000080475320000007%2Fa201910-k.htm, Last accessed on 2020-03-21.

[10] J. Sorace, H.-H. Wong, T. DeLeire, D. Xu, S. Handler, B. Garcia, and T. MaCurdy, "Quantifying the competitiveness of the electronic health record market and its implications for interoperability," *International Journal of Medical Informatics*, vol. 136, p. 104037, 2020.

[11] Epic, "A leap toward universal interoperability: Epic gets approval to join tefca," https://www.epic.com/epic/post/a-leap-toward-universal-interoperability-epic-gets-approval-to-join-tefca, 2021, [Accessed: March 19, 2023].

[12] Cerner, "TEFCA: A leap toward achieving nationwide interoperability," https://www.cerner.com/newsroom/tefca-a-leap-toward-achieving-nationwide-interoperability, 2021, [Accessed: March 19, 2023].

[13] Sequoia Project, "Meet the prospective QHINs," https://rce.sequoiaproject.org/meet-the-prospective-qhins/, n.d., [Accessed: March 19, 2023].

[14] "42 cfr part 2," Electronic Code of Federal Regulations, accessed: March 25, 2023. [Online]. Available: https://www.ecfr.gov/current/title-42/chapter-I/subchapter-A/part-2

[15] F. Systems, "Patient-driven information consent management," https://staging.feisystems.com/solutions/behavioral-health/consent2share/, 2022, accessed: 2023-03-18.

[16] A. Grando, D. Sottara, R. Singh, A. Murcko, H. Soni, T. Tang, N. Idouraine, M. Todd, M. Mote, D. Chern *et al.*, "Pilot evaluation of sensitive data segmentation technology for privacy," *International journal of medical informatics*, vol. 138, p. 104121, 2020.

[17] L. Chu, V. Kannan, M. A. Basit, D. J. Schaeflein, A. R. Ortuzar, J. F. Glorioso, J. R. Buchanan, D. L. Willett *et al.*, "Snomed ct concept hierarchies for computable clinical phenotypes from electronic health record data: comparison of intensional versus extensional value sets," *JMIR Medical Informatics*, vol. 7, no. 1, p. e11487, 2019.

[18] Office of the National Coordinator for Health Information Technology, "Trusted exchange framework and common agreement," Interoperability Standards Advisory, accessed: March 25, 2023. [Online]. Available: https://www.healthit.gov/isa/trusted-exchange-framework-and-common-agreement

[19] "Office of the national coordinator for health information technology," U.S. Department of Health and Human Services, accessed: March 25, 2023. [Online]. Available: https://site.healthit.gov/home

[20] Office of the National Coordinator. https://github.com/onc-healthit/onc-certification-g10-test-kit.

[21] O. Lassila, R. R. Swick *et al.*, "Resource description framework (rdf) model and syntax specification," 1998.

[22] D. L. McGuinness, F. Van Harmelen *et al.*, "Owl web ontology language overview," *W3C recommendation*, vol. 10, no. 10, p. 2004, 2004.

[23] K. P. Joshi, Y. Yesha, and T. Finin, "An Ontology for a HIPAA compliant cloud services," *4th International IBM Cloud Academy Conference ICACON 2016*, 2016.

[24] M. Joshi, K. Joshi, and T. Finin, "Attribute based encryption for secure access to cloud based ehr systems," in *2018 IEEE 11th International Conference on Cloud Computing (CLOUD)*. IEEE, 2018, pp. 932–935.

[25] D.-y. Kim and K. P. Joshi, "A semantically rich knowledge graph to automate hipaa regulations for cloud health it services," in *2021 7th IEEE Intl Conference on Big Data Security on Cloud (BigDataSecurity), IEEE Intl Conference on High Performance and Smart Computing,(HPSC) and IEEE Intl Conference on Intelligent Data and Security (IDS)*. IEEE, 2021, pp. 7–12.

[26] M. Joshi, K. P. Joshi, and T. Finin, "Delegated authorization framework for ehr services using attribute based encryption," *IEEE Transactions on Services Computing*, 2019.

[27] L. Elluri, A. Piplai, A. Kotal, A. Joshi, and K. P. Joshi, "A policy-driven approach to secure extraction of covid-19 data from research papers," *Frontiers in Big Data*, vol. 4, 2021.

[28] T.-T. Kuo, H.-E. Kim, and L. Ohno-Machado, "Blockchain distributed ledger technologies for biomedical and health care applications," *Journal of the American Medical Informatics Association*, vol. 24, no. 6, pp. 1211–1220, 2017.

[29] F. Casino, T. K. Dasaklis, and C. Patsakis, "A systematic literature review of blockchain-based applications: current status, classification and open issues," *Telematics and Informatics*, vol. 36, pp. 55–81, 2019.

[30] A. H. Mayer, C. A. da Costa, and R. d. R. Righi, "Electronic health records in a blockchain: A systematic review," *Health informatics journal*, vol. 26, no. 2, pp. 1273–1288, 2020.

[31] G. Zyskind, O. Nathan, and A. Pentland, "Enigma: Decentralized computation platform with guaranteed privacy," *arXiv preprint arXiv:1506.03471*, 2015.

[32] M. S. Ali, M. Vecchio, G. D. Putra, S. S. Kanhere, and F. Antonelli, "A decentralized peer-to-peer remote health monitoring system," *Sensors*, vol. 20, no. 6, p. 1656, 2020.

[33] J. Clavin, S. Duan, H. Zhang, V. P. Janeja, K. P. Joshi, Y. Yesha, L. C. Erickson, and J. D. Li, "Blockchains for government: use cases and challenges," *Digital Government: Research and Practice*, vol. 1, no. 3, pp. 1–21, 2020.

[34] K. N. Griggs, O. Ossipova, C. P. Kohlios, A. N. Baccarini, E. A. Howson, and T. Hayajneh, "Healthcare blockchain system using smart contracts for secure automated remote patient monitoring," *Journal of medical systems*, vol. 42, pp. 1–7, 2018.

[35] M. L. Gagnon and G. Stephen, "A pragmatic solution to a major interoperability problem: Using blockchain for the nationwide patient index," *Blockchain in Healthcare Today*, 2018.

[36] J. Yli-Huumo, D. Ko, S. Choi, S. Park, and K. Smolander, "Where is current research on blockchain technology? a systematic review," *PloS one*, vol. 11, no. 10, p. e0163477, 2016.

[37] S. Nakamoto, "Bitcoin: A peer-to-peer electronic cash system," Bitcoin, 2008.

[38] V. Buterin *et al.*, "A next-generation smart contract and decentralized application platform," *white paper*, 2014.

[39] M. Vukolić, "Rethinking permissioned blockchains," in *BCC*. ACM, 2017, pp. 3–7.

[40] E. Buchman, "Tendermint: Byzantine fault tolerance in the age of blockchains," *Master thesis, The University of Guelph*, 2016.

[41] M. Vukolić, "The quest for scalable blockchain fabric: Proof-of-work vs. BFT replication," in *iNetSeC*, 2015, pp. 112–125.

[42] C. Cachin and M. Vukolić, "Blockchain consensus protocols in the wild," in *DISC*, 2017, pp. 1:1–1:16.

[43] T. D. Chandra, R. Griesemer, and J. Redstone, "Paxos made live: an engineering perspective," in *PODC*. ACM, 2007.

[44] D. Ongaro and J. Ousterhout, "In search of an understandable consensus algorithm," in *ATC*, 2014, pp. 305–319.

[45] L. Lamport, R. Shostak, and M. Pease, "The byzantine generals problem," *ACM TOPLAS*, vol. 4, no. 3, pp. 382–401, 1982.

[46] J. Walonoski, M. Kramer, J. Nichols, A. Quina, C. Moesel, D. Hall, C. Duffett, K. Dube, T. Gallagher, and S. McLachlan, "Synthea: An approach, method, and software mechanism for generating synthetic patients and the synthetic electronic health care record," *Journal of the American Medical Informatics Association*, vol. 25, no. 3, pp. 230–238, 2018.

[47] Leroy Kim and contributors, "Synthea rdf," https://github.com/leroykim/synthea-rdf, 2021, accessed on March 19, 2023.

[48] U.S. National Library of Medicine, "Value set authority center (vsac)," https://vsac.nlm.nih.gov/welcome, 2021, accessed on March 19, 2023.

[49] A. E. Johnson, T. J. Pollard, L. Shen, L.-w. H. Lehman, M. Feng, M. Ghassemi, B. Moody, P. Szolovits, L. Anthony Celi, and R. G. Mark, "Mimic-iii, a freely accessible critical care database," *Scientific data*, vol. 3, no. 1, pp. 1–9, 2016.

[50] Y. Nenov, R. Piro, B. Motik, I. Horrocks, Z. Wu, and J. Banerjee, "Rdfox: A highly-scalable rdf store," in *The Semantic Web-ISWC 2015: 14th International Semantic Web Conference, Bethlehem, PA, USA, October 11-15, 2015, Proceedings, Part II 14*. Springer, 2015, pp. 3–20.

[51] J. Finin, "Re: Rdfox tutorial," Email, December 2022, email correspondence.

[52] S. Juric, *Elixir in action*. Simon and Schuster, 2019.

[53] C. McCord, J. Valim, and B. Tate, "Programming phoenix 1.4: Productive |> reliable |> fast," *Programming Phoenix 1.4*, pp. 1–325, 2019.

[54] BlockScout, "Blockscout," 2022, https://github.com/blockscout/blockscout, Last accessed on 2022-11-20.