

Micro-Identities Improve Healthcare Interoperability with Blockchain: Deterministic Methods for Connecting Patient Data to Uniform Patient Identifiers

Full Paper

**Peter B. Nichol, PMP, CSSMBB, CMPI, CQM,
CSM, SA, SP**

Digital and Innovation Expert
Managing Director
Oroca Innovations
peter.nichol@orocainnovations.com

William R. Dailey, MD, MEng, MSMI

Chief of Medical Information
Post Graduate Biomedical Informatics
Golden Valley Memorial Healthcare
wdailey@gvmh.org
docdailey@gmail.com

Abstract — The aim of this paper is to present solutions to the ongoing concerns about healthcare interoperability within the United States. Technological changes threaten to improve healthcare. Provider to provider data transfers within a trustless ecosystem is possible by leveraging blockchain technologies. Blockchain has the potential to enable healthcare interoperability alignment to address identity, confidentiality, integrity of data and accessibility. This paper presents a hybrid model, integrating HL7 FHIR, interoperability standards describing data formats and elements, as well as an Application Programming Interface (API) for exchanging electronic health records using blockchain technologies for better patient access to health information. This research expands traditional identity matching strategies to formulate a new solution for healthcare entities to match patient identities. This paper contributes to the literature on the potential for blockchain technologies, as it relates to improving patient care continuum, thus, empowering patient self-sovereignty.

Keywords — *Blockchain, Interoperability, Identity, Patient-centric Care, Patient Matching, and Microidentity.*

I. INTRODUCTION

A revolution is occurring. Blockchain technologies have the potential to change the world, affecting the property you buy, the food you eat and eventually the healthcare you receive. In order for healthcare to benefit from this emerging technology, change is required. Healthcare spending must align with outcomes. Breaking the trend of U.S. healthcare spending must start by enabling clinicians with the tools to provide better, faster and more cost-efficient access to care.

Blockchain technologies will provide the catalyst for this change. Health information technology (HIT) and the emergence

of blockchain technologies will solve healthcare interoperability.

U.S. healthcare spending grew 5.3 percent in 2014, reaching USD \$3.0 trillion or USD \$9,523 per person. The National Health Expenditure Accounts (NHEA) office offers estimates of healthcare spending in the United States as far back as 1960. NHEA reported accelerated healthcare spending of 5.3 percent in 2014, compared to 2.9 percent in 2013, was attributed to coverage expansions under the Affordable Care Act, specifically for Medicaid and private health insurance. The increasing cost of healthcare is not isolated; these increasing healthcare costs span every category of healthcare spending. Healthcare spending growth for Hospital Care increased 4.1 percent to USD \$971.8 billion; Physician and Clinical Services increased 4.6 percent to USD \$603.7 billion; Other Professional Services increased 5.2 percent to USD \$84.4 billion; Dental Services increased 2.8 percent to USD \$113.5 billion; Home Healthcare increased by 4.8 percent to USD \$83.2 billion; Nursing Care Facilities increased 3.6 percent to USD \$155.6 billion and Prescription Drugs increased by 12.2 percent to USD \$297.7 billion. The CMS National Health Expenditures 2014 report highlights that healthcare spending has increased across the care continuum.¹

Shifting focus to major sources of funds in 2014, Medicare spending grew 5.5 percent to USD \$618.7 billion. Medicaid increased by 11 percent in 2014 to USD \$495.8 billion compared to only 5.9 percent growth in 2013. Private insurance increased 4.4 percent to USD \$991.0 billion in 2014, when compared to 1.6 percent in 2013.²

The rampant rise of healthcare costs is not sustainable. It is easy to rationalize increasing healthcare costs in the U.S. given a recent economic analysis, predicting that global health spending

¹ Centers for Medicare & Medicaid Services. "National Health Accounts Projected," 2016. <https://www.cms.gov/research-statistics-data-and-systems/statistics-trends-and-reports/nationalhealthexpenddata/nationalhealthaccountsprojected.html>.

² International Data Corp. "Worldwide IT Spending Will Reach \$2.8 Trillion in 2019 with the Strongest Growth Coming from the Healthcare Industry, According to IDC," 2016. <http://www.businesswire.com/news/home/20160204005807/en/Worldwide-Spending-Reach-2.8-Trillion-2019-Strongest>.

would be expected to increase by about 6 percent per year if the global gross domestic product (GDP) grows by 4.5 percent annually in U.S. dollars over the next decade. The U.S. (GDP) was USD \$16,663.15 trillion in 2013 and USD \$17,348.08 trillion in 2014, growing U.S. GDP by 3.95 percent. Nominalizing 4.5 percent down to 3.95 percent, resulted in a forecasted annual increase in healthcare spending of 5.26 percent for the U.S.³ The runaway healthcare costs of care in the U.S. must be contained.

The U.S. healthcare system is one of the greatest healthcare ecosystems in the world. A perpetual pillar of strength for the U.S. healthcare system continues to be the advanced state of technology; high-quality services and pressing advancement in clinical research attract medical students from all over the world. In 2015, the U.S. produced over 18,000 medical school graduates.⁴ With over 854,698 doctors in the U.S and 809,845 active doctors of medicine, there are 2.8 physicians per 1,000 in the civilian population.⁵ Interestingly, patients only visited their primary care physician 54.6 percent of the times they went to the doctor.⁶

When patients go outside their network or visit doctors other than their primary care doctor, the challenge of interoperability presents itself and cost of care increases. Healthcare interoperability refers to the interactions of a system or collection of systems in uniform orchestration, to offer an integrated patient experience without special effort.

Do patient management systems (PMS) between provider facilities talk to one another? Can a doctor who sent a patient to a specialist yesterday see the results of that visit today? Is access to clinical treatment and administrative information integrated across payers and providers? Today, regrettably, the answer is that access to care is not an integrated experience in the United States. Together we can change this.

1. TECHNOLOGICAL ADVANCEMENTS IMPROVE HEALTH OUTCOMES

The U.S. healthcare system has been going through dynamic times over the last fifty years. President Lyndon Johnson enacted legislation introducing Medicare in 1965. Shortly thereafter in 1985, the Consolidated Omnibus Budget Reconciliation Act of 1985 (COBRA) amended the Employee Retirement Income Security Act of 1974 (ERISA) to give employees the ability to continue health insurance coverage after leaving employment. A little more than ten years later in 1996, the Health Insurance Portability and Accountability Act (HIPAA) was signed to protect health insurance coverage for workers and their families when they change or lose their jobs, and made it a legal requirement for health insurance companies to cover pre-existing conditions. In 1997, the State Children's Health Insurance Program, or SCHIP, was established by the federal government to provide health insurance to children in families at or below 200 percent of the federal poverty line. Then on March 23, 2010, the Patient Protection and Affordable Care Act, also known as Obamacare,

was enacted, providing for the phased introduction over four years of a comprehensive system of mandated health insurance. The recent and dramatic healthcare reform represents the most significant regulatory overhaul of the U.S. healthcare system since the passage of Medicare and Medicaid in 1965.⁷

The accelerated pace of change within the provider and payer settings are monumental in healthcare settings. We live in exciting times with unlimited potential for technological advancements.

Strong Artificial Intelligence (AI). Recursive self-improvement. Exponential growth. Technology singularity is a hypothetical event where, by leveraging artificial general intelligence (known as 'strong AI') a computer could theoretically be capable of recursive self-improvement (redesigning itself) – building a computer better than itself. Applying recursive improvements to big data means that structures unknown to humans today could be created within a decade. Applying recursive improvements to analytics means that correlations that have to be linear today, could be non-linear tomorrow, and although appearing seemingly unrelated, in fact, have extreme distance connections. Applying recursive improvements to biometric sensors could create new unique identifying characteristics currently unknown and unmonitored. This opens possibilities that, through enabled smart devices, we can ascertain new ways of establishing identity such as gait analysis (someone's walking style, formed through wearable device data recorded in the last 30 seconds).

Will superintelligence improve societal health? Futurist Ray Kurzweil, the principal inventor of the first charge-coupled device flatbed scanner, the first omni-font optical character recognition, the first print-to-speech reading machine for the blind and the first commercial text-to-speech synthesizer, believes that singularity will occur around the year 2045. Vernor Vinge argues that artificial intelligence, human biological enhancement or brain-computer interfaces could be possible causes of the singularity and that singularity will occur sometime before 2030.

"Within thirty years, we will have the technological means to create superhuman intelligence. Shortly after, the human era will be ended," according to Vernor Vinge. In Vinge's 1993 article, "The Coming Technological Singularity," he explains that once true superhuman artificial intelligence is created, no current model of reality will be sufficient to predict beyond it. When will the era of the robots start? It will be shortly after the death of the recommendation engines. A recommendation engine (recommender system) is a tool that predicts likeness (may like, may not like) among a list of given items. These preference recommendations could be around books, software, travel and many other areas. This, however, is not artificial intelligence (AI); this is a recommendation engine. A recommendation engine uses two pieces of known information, typically leveraging either collaborative filtering (arrives at a recommendation that's based on a model of prior user behavior) or content-based filtering (recommendations based on a user's behavior, e.g., historical

³ Thomas, Jason M., and Stephen H. Wise. "2016 Global Health Care Outlook: Reconciling Rapid Growth & Cost Consciousness," 2016., 6.

⁴ Kaiser Family Foundation. (2016). Total Number of Medical School Graduates. Retrieved from <http://kff.org/other/state-indicator/total-medical-school-graduates/>

⁵ Statista. "U.S. Physicians - Statistics & Facts," 2016. <http://www.statista.com/topics/1244/physicians/>.

⁶ CDC National Center for Health Statistics. "Ambulatory Care Use and Physician Office Visits," 2014. <http://www.cdc.gov/nchs/fastats/physician-visits.htm>.

⁷ Wilson, Robert H., Norman J. Glickman, and Laurence E. Lynn Jr, eds. LBJ's Neglected Legacy: How Lyndon Johnson Reshaped Domestic Policy and Government. 1 edition. Austin: University of Texas Press, 2015.

browsing) to determine one's likes or dislikes.

In contrast, artificial intelligence takes something known and creates something unknown. Netflix uses a form of machine learning, a subfield of AI that produces results for learning, prediction and decision-making. Collaborative filtering drives the Netflix engine, commonly used for research in combination with the Pearson correlation. The Pearson correlation measures the linear dependence between two variables (or users in this case) as a function of their attributes.⁸ Many algorithms become less reliable as the population sample grows exceptionally.

The Pearson correlation sifts down the sampling population to neighborhoods based on similarity (reading the same books, traveling to the same locations). This approach produces targeted predictions that are accurate within a small population sample while leveraging the population data, and relevant for a subsection or neighborhood of users.

The Turing Test evolution. IBM Watson.

John McCarthy cut the term 'Artificial Intelligence' in his 1955 proposal for the 1956 Dartmouth Conference. He also invented the Lisp programming language. Until 1956, this space was referred to as machine intelligence. When a conversation moves to the topic of AI, it's not long before talk of the Turing Test arises. Alan Turing in his 1950 paper, "Computing Machinery and Intelligence," was first published in *Mind* (a British peer-reviewed academic journal currently published by Oxford University Press on behalf of the Mind Association).⁹ It was within this seminal paper that the concept of what is now considered the Turing Test (TT), was introduced. The TT involves three participants in isolated rooms: a computer (which is being tested), a human and a judge (also human). Typing through a terminal, the computer and the human both try to convince the judge that they are human. The computer is the winner when the judge can't consistently tell which is which. This is the de facto test of artificial intelligence.¹⁰

Stevan Harnad, a cognitive scientist, contends that the TT has evolved since 1965 and today's Turing Test asks the question: "Can machines do what we (as thinking entities) can do?" Harnad also suggests that this test is not designed to trick the judge that a computer is a human, but rather establish AI's empirical goal of generating human scale performance capacity. The Turing Test represents what the science of AI intends to do - until then, AI remains a machine. The term 'intelligence' will only be bestowed to a computer, after successfully passing the TT test.

II. PROVIDER TO PROVIDER DATA TRANSFER WITHIN A TRUSTLESS ECOSYSTEM

Identity, security and cryptography are baked into middleware; customized performance for blockchains. Technology teams just got more productive.

Microsoft recently published to GitHub an overview of project Bletchley. Project Bletchley is a set of tools for

supporting SmartContracts on the blockchain, enabling secure access to off-chain information. The project supports open standards for protocol-level implementations of peer-to-peer networking, consensus and database, and virtual machines are vital to establishing trust within a blockchain ecosystem. Bletchley is a middleware toolset for developers, and provides an ecosystem to enable implementing identity, security, cryptography, scale, tooling, management, monitoring, and reporting for both on and off the blockchain. This is the first step towards client-driven performance flexibility. What Bletchley offers is performance flexibility for core, kernel and universal protocols. For example, a banking application will have different requirements for transactional, processing and nonfunctional requirements for scale when compared to a nonprofit using a basic digital ledger to record donations.

Two-tier client-server architectures are multi-tier computing architectures in which an entire application is distributed as two distinct layers or tiers. In this case, the presentation layer and data layer run on the server. In contrast, three-tier or n-tier architecture is usually separated into three major sections: The presentation/front-end tier, the business/application tier and the data/back-end tier. The process of blockchain architectures experienced a similar evolution as tier-level client architectures.

- Blockchain 1.0, simple state machine, used logic (stored procedures) to record transactions in sequence, where referential integrity was implemented using primary keys (PK) and foreign keys (FK).
- Blockchain 2.0, state machine and code, added SmartContracts. The 2.0 version also leverages PK and FK; however, Blockchain 2.0 also contains logic (code like a stored procedure) that can be executed.
- Blockchain 3.0, state machine and code, as well as cryptlets, allow for improved interoperability and scale on and off the blockchain. As a general clarification, there are some additional differences regarding tokenization and instantiation of transactions that are beyond the scope of this article.

Microsoft's recent release of Bletchley introduces the idea of the enterprise consortium node. In these scenarios, the client system makes a request, and the request is given to a future node (block database, state, history) that is connected to the block database (state and history, signing, VM and consensus). In short, the modular framework can choose the best components to fulfill the client's request. We're within grasp of the ability to handle dynamic scalable, requests; there is a new solution to access off-chain patient data.

Cryptlets establish the foundation for Microsoft's security blockchain middleware and run as a cloud-based service, provider-agnostic. Previously, when making data requests outside a SmartContract, the authenticity was broken for dependent transactions. For example, if you're running an app on your phone, but you need your payment information stored in another

⁸ Jones, M. T. (2013). Recommender systems, Part 1: Introduction to approaches and algorithms. Retrieved November 29, 2015, from <http://www.ibm.com/developerworks/library/os-recommender1/>

⁹ Turing, Alan (1950), "Computing Machinery and Intelligence", *Mind* LIX (236): 433–460, doi:10.1093/mind/LIX.236.433, ISSN 0026-4423. P 460.

¹⁰ Nichol, Peter B. "CIO Perspectives: Impact of Technological Singularity on Analytics." LinkedIn Pulse, 2015. <https://www.linkedin.com/pulse/cio-perspectives-impact-technological-singularity-analytics-nichol>.

SmartContract to process an order. Another example would be while running a phone app, you make a request to view your medical information, but when you click on your lab results details, a call (to the lab's SmartContract) is required to access that information (outside of the existing SmartContract). Today, the effect is that the pure integrity of the transaction is broken.

Cryptlets live off the blockchain, executing within a secure trusted container and communicating using secure channels. Cryptlets can also be written in any programming language and are called or instantiated by a CryptoDelegate (with the SmartContract). Cryptlets come in two flavors: Utility (providing core infrastructure and middleware services, e.g., encryption, time and date events, external data access and authentication services) and Contract (providing all the execution logic and securely storing the data in the SmartContract). Contract Cryptlets also don't run on the blockchain and, therefore, can execute in parallel on vertically scaled systems.

III. TECHNOLOGIES PROMISE TO ADVANCE HEALTHCARE

If Thomas Edison had gone to business school, we would all be reading by larger candles.

— Mark McCormack

Two-hundred to five-hundred applications and technologies comprise the average provider environment. It's a wonder that quality healthcare can be delivered in an environment where collaboration is siloed. For the physician starting a career, dreaming about changing patient health and making a difference, it is exciting to peer into technologies with the potential to transform patient care. Our long-held problems within healthcare with seemingly simple, yet, unobtainable solutions (national patient identifier) lead practitioners and academics alike to explore new frontiers, in hope of finding solutions. Three technologies that fit within this framework are the Health Level Seven (HL7), Fast Healthcare Interoperability Resources (FHIR) and blockchain technology.

Many people look at blockchain technology as a one-off or perhaps a "fringe" technology only applicable to pseudo-currency or SmartContracts requiring massive "proof of work" computations. At best, this technology is cloaked as a solution looking for a problem, but we do know the problem in front of us.

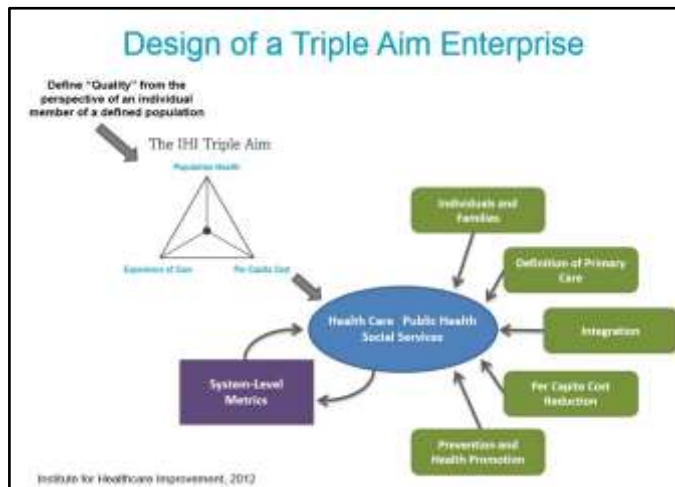
The Triple Aim is a framework developed by the Institute for Healthcare Improvement (IHI) to pursue three objectives for the U.S. Healthcare system:

1. Improving the patient experience of care (including quality and satisfaction).
2. Improving the health of populations.
3. Reducing the per capita cost of healthcare.

The IHI reports that according to the National Healthcare Expenditure Projections for 2010-2020, the U.S. healthcare

system is one of the costliest in the world, accounting for 17 percent of the gross domestic product with estimates that this percentage will grow to nearly 20 percent by 2020.¹¹

While the Triple Aim focused on fulfillment of its objective with a concentration on individuals and families, redesign of primary care services and structures, population health management, cost control platform and system integration and execution, it's hard to ignore the growing challenges associated with delivering healthcare in the United States.



Healthcare presents numerous problems, including obtaining low shared-cost services, improving the quality of care, strengthening patient outcomes and discovering solutions that provide distributed approaches for maintaining and protecting patient data. We all have a duty to ensure the exchange of patient information necessary for treatment is afforded to clinicians, while retaining the diligence to secure sensitive protected health information (PHI).

“Fast Healthcare Interoperability Resources (HL7 FHIR, pronounced “fire”) is a draft standard describing data formats and elements (known as “resources”) and an Application Programming Interface (API) for exchanging Electronic health records.”¹²

HL7 FHIR is a critical piece of the solution. HL7 FHIR allows for the RESTful exchange of patient information between trusted entities. “REST stands for Representational State Transfer, which is an architectural style for networked hypermedia applications. It is primarily used to build Web services that are lightweight, maintainable, and scalable. A service based on REST is called a RESTful service.”¹³ FHIR was developed from modern web technologies, RESTful services and familiar web specifications like XML, JSON, HTTP, Atom and OAuth. HL7 FHIR also supports leading specifications for Privacy & Security, including OAuth2 and OpenID.

Below defines web specifications that are collectively supported within the HL7 FHIR framework.

1. **XML:** Extensible Markup Language (XML) is a markup language that defines a set of rules for

¹¹ Office of the National Coordinator for Health Information Technology (ONC). Connecting Health and Care for the Nation A Shared Nationwide Interoperability Roadmap, 2015. <https://www.healthit.gov/sites/default/files/nationwide-interoperability-roadmap-draft-version-1.0.pdf>.

¹² Ibid.

¹³ Vaqqas, M. “RESTful Web Services: A Tutorial.” Dr. Dobb’s, 2016. <http://www.drdoobs.com/web-development/restful-web-services-a-tutorial/240169069>.

encoding documents in a format that is both human-readable and machine-readable.¹⁴

2. **JSON:** JavaScript Object Notation is an open standard format that uses human-readable text to transmit data objects consisting of attribute-value pairs. It is used primarily to transmit data between a server and web application, as an alternative to XML.¹⁵
3. **HTTP:** The Hypertext Transfer Protocol (HTTP) is an application protocol for distributed, collaborative, hypermedia information systems.¹⁶
4. **Atom:** The Atom Syndication Format is an XML language used for web feeds, while the Atom Publishing Protocol (AtomPub or APP) is a simple HTTP-based protocol for creating and updating web resources.¹⁷
5. **OAuth:** OAuth is an open standard for authorization, commonly used as a way for Internet users to log in to third-party websites using their Google, Facebook, Microsoft, Twitter, One Network, etc., accounts without exposing their password.¹⁸
6. **OAuth 2.0:** OAuth 2.0 is the next evolution of the OAuth protocol and is not backward compatible with OAuth 1.0. OAuth 2.0 focuses on client developer simplicity while providing specific authorization flows for web applications, desktop applications, mobile phones and living room devices.¹⁹
7. **OpenID:** OpenID is an open standard and decentralized authentication protocol.²⁰ [20]

HL7 FHIR's extensive and flexible framework can support mobility and mobile health, social media, personal health records, public health, payment systems and clinical research. This framework makes development and, therefore, integration straightforward. HL7 FHIR is faster to learn, faster to develop and faster to implement. This combination creates a framework that is not only flexible, but also easy to use. Unlike similar healthcare integration frameworks, HL7 FHIR is free. Only the HL7 FHIR name and logo are trademarked, but the specification is licensed without restriction or royalty.

To further the acceptance of the HL7 FHIR framework, the Argonaut Project was born. The Argonaut Project is a private sector initiative to advance industry adoption of modern, open interoperability standards.²¹ While founded in December 2014, the project didn't launch until February 2015 with thirteen core project member organizations, including Accenture, athenahealth, Beth Israel Deaconess Medical Center, Cerner, Epic, Mayo Clinic, MEDITECH, McKesson, Partners

HealthCare System, SMART at the Boston Children's Hospital Computational Health Informatics Program, The Advisory Board Company and Surescripts. Together these organizations accelerate the development of the HL7 FHIR specification.

The HL7 FHIR specification allows secure information exchange and is a positive step towards interoperability. HL7 FHIR separates the data structure from the wider problem of entity-to-entity trust, centralized server broadcasting and patient identification and matching spanning entities. HL7 FHIR, despite being revolutionary and a piece of the interoperability puzzle, has hurdles to overcome relating to connectivity and matching.

The HL7 FHIR specification will soon be complete and implemented by major electronic medical record (EMR) vendors supported by the Argonaut Project. At that time, each healthcare entity will have RESTful interface(s) to the information they choose to make available to trusted partners, accessible by way of their HL7 FHIR server(s); however, specific gaps exist in locating the appropriate server, obtaining entity trust and requesting collocated patient information at the next point of care (POC). Patient demographics such as date of birth, gender, country, postal code, ethnicity and blood type may be slightly different due to multiple factors and ensuring acquisition of the correct data. Identifying the patient correctly is essential to care.

Blockchain technology has the potential to address these technology gaps. Novel methods for identity verification and sharing healthcare "event" transactions between entities, will advance healthcare stakeholders towards complete solutions.

The requirement to link verification, patient identity and the healthcare entity is achieved through a subsequent proof-of-concept. This step leverages information that only the patient knows, e.g., soft multi-factor authentication utilizing the patient's knowledge of previous visits. HL7 FHIR's extensibility offers many solutions within a single framework.

IV. INTEROPERABILITY STATE OF THE UNION

You can accomplish anything in life if you don't mind who gets the credit.

— Harry Truman

We started with faxes, migrated to health information exchanges and made advancements with the Direct Project. Today, HL7 v2, FHIR highlights the future direction to achieve interoperability.

The invention of the fax machine was a great leap forward, facilitating the exchange of information between providers and patients. No longer were records required to be mailed or carried between provider offices. Incredibly, faxes are still in use today nationwide.

The Scottish inventor Alexander Bain, who worked on

¹⁴ Harold, Elliotte Rusty. XML 1.1 Bible. 3rd edition. Indianapolis, IN: Wiley, 2004.

¹⁵ Bassett, Lindsay. Introduction to JavaScript Object Notation: A To-the-Point Guide to JSON. 1 edition. Beijing; Sebastopol, CA: O'Reilly Media, 2015.

¹⁶ Wong, Clinton. HTTP Pocket Reference: Hypertext Transfer Protocol. 1 edition. Sebastopol, CA: O'Reilly Media, 2000.

¹⁷ Wittenbrink, Heinz. RSS and Atom: Understanding and Implementing Content Feeds and Syndication: A Clear and Concise Guide to Strategy, Structure, Selection with in Depth ... Coverage of Feed Formats and XML Vocabularies. Birmingham England: Packt Publishing, 2005.

¹⁸ Boyd. Getting Started with OAuth 2.0. 1 edition. Beijing; Sebastopol, Calif: O'Reilly Media, 2012.

¹⁹ Ibid.

²⁰ Siriwardena, Prabath. Advanced API Security: Securing APIs with OAuth 2.0, OpenID Connect, JWS, and JWE. 1st ed. edition. Berkeley, California: Apress, 2004.

²¹ HL7 Argonaut Project Wiki. "Main Page/Background - HL7 Argonaut Project Wiki," 2016. http://argonautwiki.hl7.org/index.php?title=Main_Page/Background.

chemical mechanical fax type devices, received British patent 9745 on May 27, 1843, for his “Electric Printing Telegraph”. Why would a technology invented over 170-years ago still be in use? The answer is simple: it is easy to obtain the fax number for a known entity, and fax machines are pervasive. Additionally, fax numbers are relatively permanent. Once two entities have established fax machine addresses, e.g., fax numbers, the entities can exchange data.

A fax cover sheet and header include the return address. The fax machine is largely non-linear, e.g., the user can send it and usually forget the fax was sent after they received confirmation the transmission was sent successfully. Faxes operate as part of a push or pull system. A push system sends data similar to a fax machine and a pull system asks for and retrieves it. This would be akin to the receiver of the faxed information requested and then responding by sending a fax back with the information. A pull-based production system explicitly limits the volume of work in the process of the system and a push production system does not have explicit limits on the system’s capacity for work volume.²²

There are obvious downsides when transmitting images, including no granular data, eventually being printed on paper, and difficulty with EMR integration. Additionally, images consume enormous amounts of storage for relatively few bytes of information.

Health Information Exchanges (HIE) have failed due to insufficient financial sustainability models, limited shared incentives, low utilization, complex connectivity adoption, as well as regionally specific data silos of unused information and the obvious patient matching issues. HIE had a troubled past and struggled to reach financial sustainability with a few localized exceptions. To-date, HIEs are regional, centralized silos that store much more information than they dole out.

The Direct Project is a part of the Nationwide Health Information Network. The Direct Project was created to specify a simple, secure, scalable, standards-based way for participants to send authenticated, encrypted health information directly to known, trusted recipients over the Internet. This project had great potential for advancing the point-to-point exchange of data. However, the Direct Project was limited to simple XML formatted Continuity of Care Documents (CCD) stored in the HIE because of the requirement for verifying sender and receiver identities defined in the HISP-in-a-Box architecture. HISP-in-a-Box is a software implementation model that might be offered by HIT Regional Extension Centers, multiple software vendors or even consultants.²³

Pure HL7 v2 is effective for well-established interfaces between closed-off, well-defined systems when implemented as one-offs. Often, these implementations contain variations and consequently require programmatic rework on the interfaces to ensure normal functionality. The HL7 specification is very complex; Therefore, difficult for programmers to implement while staying within compliance of published specifications. Implementations are plagued with challenges, e.g., recurrent patient matching and address issues. Comically, these issues are easily resolved by leveraging 170-year-old fax technology. Moreover, HL7 also does not have built-in encryption. Limited

encryption protocols increase security risk; as a result, data cannot be sent over open networks.

HL7 FHIR will be the RESTful interface standard for HL7. This interface offers specifications and encryption built into the standard from the “ground-up”. HL7 FHIR interfaces enable simple and easily understood encrypted queries. Due to the encryption, these queries can be executed safely over open networks. This integration approach for connecting provider-to-provider to share health information is easier to code, as it is very similar to modern SSL-based web technologies. Even ownership of the fastest and “highest performing” sports car is not particularly useful if there is no defined road to travel upon. Speed is only relevant when moving in the desired direction. There is no need to discard something valuable along with other things that are inessential or undesirable.

HL7 FHIR is heading in the right direction, akin to how HTML gained adoption during the establishment of the Internet. Roads are required for a fast sports car enthusiast; similarly, healthcare providers need to know the patient’s identity and the destination of the patient data.

A National Patient Identifier (NPI) is the utopian future state for patient and provider interoperability. However, it’s apparent this approach is not viable today with the current level of legislative resistance and Orwellian paranoia.

It is easy to re-identify a percentage of patients, whose records contain relatively incomplete demographic data (DOB, gender and Zip Code). The combination of HL7 FHIR, blockchain technologies and intelligent algorithms can tackle this challenge. Hope, as a strategy for the adoption of a national patient identifier, is not likely to yield successful outcomes. The NPI is unlikely to be the national strategy for healthcare interoperability.

A white-paper by the Sequoia Project sparked curiosity, raising the issue of “identity matching and addresses” that have been a systemic problem throughout healthcare systems for years, inhibiting interoperability.²⁴ This paper described a gold-standard dataset of 10,000 identities at Intermountain Healthcare and explained the pitfalls affiliated with the process of patient matching. Human factors affiliated associated with the imprecise entry of various demographic characteristics for patients were a critical cause of variability. The social security number (SSN) was presented as the most sensitive and specific identifier. The SSN is also unique and invariant over time; it is not recommended to use SSN for general identification. Moreover, the use of SSNs is irresponsible when applied in a healthcare setting for patient matching, due to widespread identity theft issues.

Less is more. This adage is true of patient matching: the more discreet items you try to match, the less likely they are to match. Keying errors, pseudonyms, misspellings and the like, impact the conformability of patient data. Also, poor data quality results in searches with no data matches. Successful matching strategies are a proprietary mix of deterministic and stochastic or probabilistic methods. They are proprietary for numerous reasons, which are beyond the scope of this paper. Stochastic methods are commonly used to overcome poor data quality.

²² Roser, Christopher. “The (True) Difference between Push and Pull.” AllAboutLean.com, 2015. <http://www.allaboutlean.com/push-pull/>.

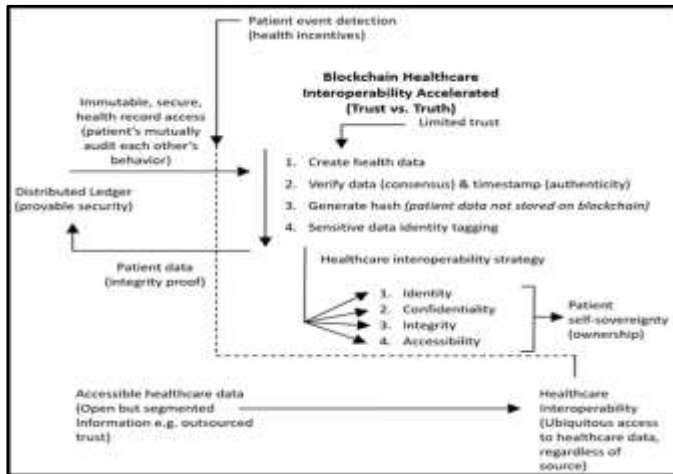
²³ Direct Project. “Direct Project - HISP in a Box,” 2016. <http://wiki.directproject.org/HISP+in+a+Box>.

²⁴ The Sequoia Project. “Framework for Cross-Organizational Patient Identity Matching.” The Sequoia Project, 2016. <http://sequoiaproject.org/framework-for-cross-organizational-patient-identity-matching/>.

V. BLOCKCHAIN HEALTHCARE INTEROPERABILITY FRAMEWORK

The below interoperability strategic healthcare framework establishes a model for incorporating blockchain technologies into healthcare.

Accessible healthcare data that is open-sourced but siloed will transition to healthcare interoperability where ubiquitous access to data will be secure and ever-present. In this new healthcare environment, patients will be able to mutually audit each other's behavior, obtaining provable security of patient health information.



VI. MATCHING PATIENT IDENTITIES AT GOLDEN VALLEY MEMORIAL HEALTHCARE

Let's do an experiment with a patient at the point of care and work our way backward to a solution. What information could be easily obtained from the patient that would inch us towards a solution? Two closely linked pieces of information are necessary from the patient: first, a specific date of care, and second, the location of that care (the care entity). All that is required to establish a unique identifier at that entity is the date of care and various non-specific identifying data. Non-specific data is easily matched, but independently, it does not enable the construction of the patient identity. In clinical parlance, this condition is referred to as high sensitivity and low specificity.

Creating a high sensitivity and low specificity condition using a patient's name is challenging, but solvable, as we will explain in the following steps.

1. **Normalize:** Normalize the name, by removing inconsistently utilized nuances such as non-alphabetic characters, capital letters, prefixes, suffixes, etc.
2. **Use wildcards:** Shorten each component of the name to a certain length, making it less specific and adding wildcards filters, if required.
3. **Statistical methods:** Use the "like" operator in a novel way (bidirectional) to maximize the effectiveness of the wildcards within a relational database system, instead of using an "=" operator that would otherwise require statistical methods.

This thought experiment inspired a data test at a small hospital system in Missouri, using real patient data. Golden

Valley Memorial Healthcare is a small 50 bed, rural healthcare facility. Interestingly, this facility has two separate, yet highly patient-concordant electronic medical records systems (inpatient and ambulatory), making this facility ripe for analysis. Each EMR contained approximately 70,000 to 90,000 distinct SSNs. Social security numbers were used as the "gold standard" for "true matches" of patient identities between entities. Subsequently, distinct SSNs were queried from each EMR and were then divided into two data sets. A set intersection was performed on the two sets, yielding the true matches between systems. All distinct medical records were then used to populate two separate tables: table one for EMR A and table two for EMR B.

Each table comprised of four fields:

1. Trimmed identifiers (1)
(Wildcard fillers " _ " and
ffffmmlllgmmddyyyy or ff_m_ll_gmmddyyyy)
2. Local medical record number
3. True match Boolean
4. My match Boolean

Table structure:

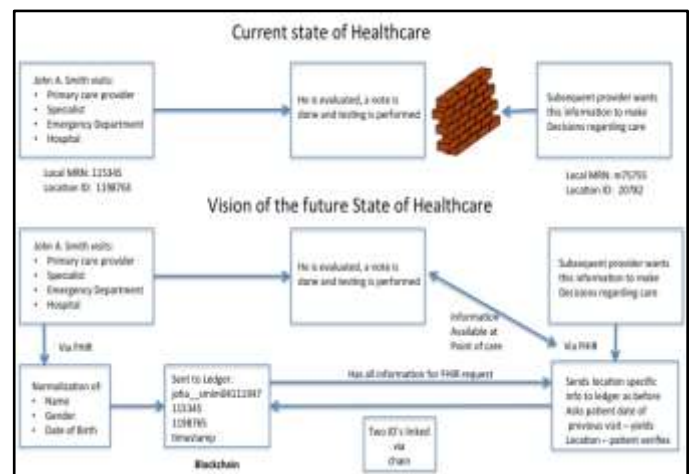
microid | localid | true_match | my_match

The true matches were marked as 'True' using the SSNs from the set intersection. Then each entry in EMR A was used via a novel "like" operation to EMR B:

*Select * from EMRB where* (2)

EMRA.microid like EMRB.microid or

EMRB.microid like EMRA.microid



This query maximized the "sensitivity value" of the wildcard operators used as fillers. All results were marked as 'True' in the my_match field. This same process was completed by iterating through microids in EMR B as compared to EMR A.

These tables allowed easy calculation of true positive (11), false negative (10), false positive (01) and true negative (00) integers. Therefore, we can calculate the sensitivity and specificity of the matching algorithm (below).

EMR A to EMR B

(3)

Sensitivity Specificity
 93.61 percent 94.93 percent

EMR B to EMR A (4)

Sensitivity Specificity
 96.65 percent 90.52 percent

This approach was deemed “pretty good matching” and the resulting microids can scarcely be considered protected health information, because these identifiers are not unique. In addition, these identifiers are personally identifiable only in the setting of that particular entity. They do contribute to solving the matching problem with a three step matching process.

1. Place the combined localid, locationid, microid and timestamp into a trusted ledger (aggregate event key).
2. Request the patient, at the next point of care, to provide a specific date and location of a previous healthcare encounter.
3. Pull the specific aggregate event key from step one, from the EMR using HL7 FHIR connectivity to the ledger. These identities are linked with another ledger entry. At this point, there is nearly a 100 percent match specificity.

This is the “golden ticket” when it comes to patient matching. Patient matching today requires providing ample patient demographics in clear text, in order to match patient identities. It’s true, “orphan” microids (marriages, misattributed genders, miskeyed, misspelled names or dates of birth) will be difficult to hard link. In these cases, a quick response code (QR code for short) can be generated; this code can be communicated by fax, mail delivery or hand-carried to the prior facility; clearly, this is not ideal. An alternative is to obtain or share the local identifier, microid and location identifier providing linkages to the orphaned identities.

VII. PATIENT CHECK IN WORKFLOW EXAMPLE

John Smith comes in for a visit today and the receptionist checks him in as John Smith, DOB: 04/11/1947 and Gender: Male. John states he was seen by the cardiologist May 5, 2016. The cardiologist had entered this into the hook within the EMR to link the blockchain ledger via HL7 FHIR workflow. Behind the scenes, it pulls from the following sample ledger:

microid	localid	locationid	timestamp
cark_coxf08261972	k94782	57583	2016-01-03 08:50:48
joha_smim04111947	k090912	57583	2016-02-05 09:42:33
cark_coxf08261972	s656754	98976	2016-02-06 08:50:51
joha_smim04111947	k08947	57583	2016-05-03 08:50:53
wilrobdaim04131969	m690984	56987	2016-08-03 08:50:59
anc_ricf09211944	5455	16909	2016-08-03 08:51:19

Looking at the sample log entries, the above search is executed with John’s normalized identifier. Without John’s middle name, that identifier becomes “joh__smim04111947”, the like operator yields the following results:

Combining the identifier with the date of the visit verifies that John Smith was seen at location 57583. By mapping to Research

Search: LIKE

microid	localid	locationid	timestamp
joha_smim04111947	k090912	57583	2016-02-05 09:42:33
joha_smim04111947	k08947	57583	2016-05-03 09:42:33

Medical Center and John’s name, John is verified as the patient. The only data viewable by the receptionist is an entry box labeled “date of remote visit” and once the date is entered, a list of matching locations appears in a dropdown box. The receptionist clicks “verify and link” and now all available data can be immediately pulled from all encounters at that facility, and localid, as well as all other ledger entries linked to the same micro. microid, localid, locationid and timestamp as a cascade of FHIR queries to all the respective locations.

VIII. RESEARCH PROPOSITIONS

The purpose of these propositions is not to pontificate on conceptual outcomes based largely on predictions, but rather to identify logical areas of future work, worthy of exploration. We developed these propositions based on our research with blockchain technologies and a strong understanding of the healthcare system within the United States. We offer three propositions (P1, P2, and P3) that have the potential to enable healthcare interoperability for the alignment of patient identity, confidentiality, integrity and accessibility. Expanding on the integration of HL7 FHIR and modern RESTful architectures, security becomes a paramount challenge within the healthcare landscape.

A. Proposition 1: Trust Verse Truth and The Role of Security in Patient Record Protection

Until 2009, trust was a belief and truth was unmeasurable. With blockchain technologies, truth can also be measured. Modern enterprises’ ability to secure patient data has become questionable of late. The Identity Theft Resources Center reported 572 data breaches, exposing 13,491,597 records as of August 2, 2016. Across banking, education, government and healthcare, healthcare was attributed to 206 of the breaches or 36.8 percent of all breaches in the first half of 2016. These healthcare breaches resulted in the loss of 4,962,136 health-related records. With almost 50 billion new devices scheduled for connectivity to the internet by 2020, this problem will amplify.

At some point, virtually every health system will be compromised. Healthcare leaders have a duty to independently verify the integrity of their healthcare systems. Today, this is done by adding new security components into the environment e.g. virus protection software, hard or soft firewalls, virtual private networks, etc. The fundamental assumption in the decision to path security gaps with hardware or software is that components will not be compromised. It is troubling that when transmitting data, it’s not possible to determine if new or old components have been compromised. Now, with blockchain, healthcare system administrators can prove the healthcare data has not been compromised. This is accomplished by establishing data authenticity with the chain of custody utilizing blockchain technologies.

The Keyless Signature Infrastructure (KSI) is designed to provide scalable digital signature based authentication for electronic data, machines and humans. Every health care data transfer can be captured and timestamped, creating proof of authenticity and restoring truth into our healthcare system. This paradigm shift offers data integrity and visibility, previously unheard of – moving healthcare towards transparent truth, not trust. KSIs can resolve the lack of consistent methods for conducting patient matching, and decrease the occurrence of out of date and incorrect patient matching errors. Leveraging KSI can prevent man-in-the-middle (MiM) attacks. MiM is a process

where a user gains unauthorized access to communication between two parties who believe they are directly communicating with each other. MiM attacks can alter valid matches, resulting in unauthorized users consuming the data for unknown and potential nefarious purposes, manufacturing “no matches found” despite the availability of valid matches. KSI helps to ensure data integrity and authenticity, protecting the patient.

B. Proposition 2: Cognitive Intelligence Improves Patient Health

Physicians are busy. IBM estimated that 160 hours of weekly reading is necessary to keep pace with the changes in the relevant medical information required for care treatment. As any physician would respond to this request, this level of weekly reading while a practicing physician is impractical.

There is, however, an alternative. With data acting as a witness to events, the reliability and integrity of data is possible. This digital ledger of health records could be anonymously utilized to provide a daily assessment of the national impact of new regulations, policies, procedures and drugs on the existing population. More specifically, the local physician could use cognitive intelligence to anonymously mine patients’ health records to determine the areas they may be affected, ranging from new regulations to drug trials.

This unique perspective empowers physicians to wade through 1 petabyte (5-years of NASA's Earth Observing System data), 5 exabytes (all words ever spoken by human beings), 1.9 zettabyte (the informational equivalent to every person on earth receiving 174 newspapers per day) or even yottabytes (the combined space of all the computer hard drives in the entire world does not amount to even one yottabyte) to provide better care.

We have experienced the birth of cognitive decision systems focused on patient health, blockchain fueled, driven by physicians. This new environment allows physicians to practice with the full knowledge of every policy, procedure and drug trial that could harm or enable their patients. Society just got healthier.

C. Proposition 3: Consumptive Collaboration for Population Health

Collaborative consumption is reshaping the economy and consumers are taking notice. Soon patients will also respond. Owning a product is being replaced by the sharing economy. Companies across the globe are buying into the models with new distribution channels bucking prevailing trends and consumers are taking action. Product services systems (paying for usage and access over ownership), redistribution markets (bartering, trading and sharing or swapping goods) and collaborative lifestyles (sharing intangibles) have triggered a new business philosophy: the introduction of the shared economy.

Let’s empower the patients to lead the charge to find their cures together. Who else would be more motivated? Welcome to ONC Health 2.0, where based on DNA sequencing you can locate individuals with similar health conditions, and leverage big data analytics. Communities working together, use precision medicine to discover better clinical outcomes.

Blockchain technology could provide conditional access to medical records and health related information. Context-driven experiences expanded to contextually driven access to medical data. Patients can locate other individuals in similar health situations and anonymously pool information for better health outcomes.

IX. THEORETICAL AND REGULATORY IMPLICATIONS

There are limitations to this conceptual analysis. First, the healthcare ecosystem is a complex beast with a vast number of stakeholders. Second, there are technical challenges with integration and uniform patient identities. Namely, providers and payers alike need to reach consensus on general standards for interoperability. These standards include policy, procedure, and technical guidelines for implementation. The following four challenges remain as the primary obstacles for adoption.

1. Securing stakeholder buy-in, trust and collaboration.
 - a. Key development stakeholders: HL7 (especially FHIR, EMR vendor consortia, Payer consortia and major players (CMS), Policy and strategic interoperability (ONC), pilot community/region entities.
 - b. Key use stakeholders: patients, healthcare providers, healthcare entities (hospitals, clinics, payers, data aggregators).
2. Orphan Identities – a kind of manual process to link miskeyed or married-name identities.
3. Unchaining mislinked identities – will need some central authority for unlinking or previous link (may be as simple as adding a Boolean for ignoring any links that represent errant links).
4. Optimal trimming of identifiers for microids (see 6-level full factorial analysis in direction for future work).

Lastly, excitement for solutions runs high. Inevitably, after society agrees that healthcare interoperability is vital for the economy, further exploration will be required to articulate practical technology solutions that can be implemented at reasonable costs.

X. LIMITATIONS AND DIRECTIONS FOR FUTURE WORK

First, this paper contributes to the advancement of healthcare interoperability. National healthcare interoperability is a foundational element for better patient outcomes by empowering clinicians with better contextual and relevant information to provide care. HL7 FHIR has made promising advancements over the last 2 years. Similarly, since 2009, the emergence of blockchains as a base technology has shown strong potential to improve the immutability of historical information, e.g., electronic health records. However, the intersection of HL7 FHIR and blockchain remains an area of continued research. The following limitations require further exploration to solidify the previously introduced concepts for identity matching and the construction of a technological framework amenable to payers and providers.

1. Integration: Develop an underlying blockchain to HL7 FHIR architecture, workflow and API hooks.
2. APIs: Develop EMR API hooks for linking, via HL7 FHIR to the blockchain.
3. Specifications: Establish blockchain specifications (vendor specific. open-source).
4. Development: Expand blockchain development for HL7 FHIR identities, certificates and trust.
5. Fine-Tune Data: Parametric analysis of optimal

trimming of name components for high sensitivity (allowing providers to “dial in” sensitivity; minimizing orphan microid, localid, locationid, timestamp composite keys). This process will be performed on the above dataset as a starting point. This process will also need to be performed on larger entities for verification and fine-tuning of data.

6. Align incentives: Cost analysis at every step to ensure the solution has shared costs and aligned incentives.
7. Verify Viability: Proof of Prototyping particular pairs or groups of entities for proof of viability.
8. Pilot Projects: Community or regional groups of entities. These can be multiple or singular entities but must have some level of patient concordance and scalability testing.
9. Rollout: Add additional “sandboxed” communities or regions with separate but identical specifications.
10. Interconnection and Adoption: Linking blockchains from all regions.
11. Use: The sky is the limit here with appropriately attributed data for all patients aggregated as if there was a national patient identifier.

There are additional challenges that must be solved to create a healthcare ecosystem that supports a mobile and patient-centric framework for healthcare information exchange.

XI. CONCLUSION

The healthcare blockchain stands to provide a nationwide-shared resource. This resource will enable patient identity matching, identity linking, redundant connectivity, location and the retrieval of granular patient data to and from any EMR. The emergence of blockchain into the healthcare ecosystem will be accelerated by leveraging the current and future HL7 FHIR interfaces and APIs. Patient data continues to reside within locally owned and operated EMRs without taking advantage of the public, distributed, immutable, timestamped and persistent capabilities – utilizing blockchain for healthcare.

Alternatively, present state EMRs store data on centralized repositories or federated providers moving data into siloed EMR systems. Blockchain technologies, when applied to healthcare, create concurrent, distributed, redundant and secure shared-cost national healthcare resource with a web-like structure, limiting or even removing intermediaries traditionally required for data exchange. Using blockchain technologies in the healthcare setting will represent a significant accelerator for healthcare interoperability.

We recommend creating a cross-functional consortium to explore the potential for blockchain technologies to accelerate healthcare outcomes.

XII. REFERENCES

[1] Centers for Medicare & Medicaid Services. “National Health Accounts Projected,” 2016. <https://www.cms.gov/research-statistics-data-and-systems/statistics-trends-and-reports/nationalhealthaccountspending/nationalhealthaccountspending.html>.

[2] International Data Corp. “Worldwide IT Spending Will Reach \$2.8 Trillion in 2019 with the Strongest Growth Coming from the Healthcare Industry, According to IDC,” 2016. <http://www.businesswire.com/news/home/20160204005807/en/Worldwide-Spending-Reach-2.8-Trillion-2019-Strongest>.

[3] Thomas, Jason M., and Stephen H. Wise. “2016 Global Health Care Outlook: Reconciling Rapid Growth & Cost Consciousness,” 2016. P 6.

[4] Kaiser Family Foundation. (2016). Total Number of Medical School Graduates. Retrieved from <http://kff.org/other/state-indicator/total-medical-school-graduates/>

[5] CDC National Center for Health Statistics. “Ambulatory Care Use and Physician Office Visits,” 2014. <http://www.cdc.gov/nchs/fastats/physician-visits.htm>.

[6] Wilson, Robert H., Norman J. Glickman, and Laurence E. Lynn Jr, eds. LBJ’s Neglected Legacy: How Lyndon Johnson Reshaped Domestic Policy and Government. 1 edition. Austin: University of Texas Press, 2015.

[7] Jones, M. T. (2013). Recommender systems, Part 1: Introduction to approaches and algorithms. Retrieved November 29, 2015, from <http://www.ibm.com/developerworks/library/os-recommender1/>

[8] Turing, Alan (1950), “Computing Machinery and Intelligence,” *Mind* LIX (236): 433–460, doi:10.1093/mind/LIX.236.433, ISSN 0026-4423. P 460.

[9] Nichol, Peter B. “CIO Perspectives: Impact of Technological Singularity on Analytics.” LinkedIn Pulse, 2015. <https://www.linkedin.com/pulse/cio-perspectives-impact-technological-singularity-analytics-nichol>.

[10] Office of the National Coordinator for Health Information Technology (ONC). Connecting Health and Care for the Nation A Shared Nationwide Interoperability Roadmap, 2015. <https://www.healthit.gov/sites/default/files/nationwide-interoperability-roadmap-draft-version-1.0.pdf>.

[11] Ibid.

[12] Vaqqas, M. “RESTful Web Services: A Tutorial.” Dr. Dobb’s, 2016. <http://www.drdoobbs.com/web-development/restful-web-services-a-tutorial/240169069>.

[13] Harold, Elliotte Rusty. XML 1.1 Bible. 3rd edition. Indianapolis, IN: Wiley, 2004.

[14] Bassett, Lindsay. Introduction to JavaScript Object Notation: A To-the-Point Guide to JSON. 1 edition. Beijing; Sebastopol, CA: O’Reilly Media, 2015.

[15] Wong, Clinton. HTTP Pocket Reference: Hypertext Transfer Protocol. 1 edition. Sebastopol, CA: O’Reilly Media, 2000.

[16] Wittenbrink, Heinz. RSS and Atom: Understanding and Implementing Content Feeds and Syndication: A Clear and Concise Guide to Strategy, Structure, Selection with in Depth. Birmingham England: Packt Publishing, 2005.

[17] Boyd. Getting Started with OAuth 2.0. 1 edition. Beijing; Sebastopol, Calif: O’Reilly Media, 2012.

[18] Ibid.

[19] Siriwardena, Prabath. Advanced API Security: Securing APIs with OAuth 2.0, OpenID Connect, JWS, and JWE. 1st ed. edition. Berkeley, California: Apress, 2004.

[20] HL7 Argonaut Project Wiki. “Main Page/Background - HL7 Argonaut Project Wiki,” 2016. http://argonautwiki.hl7.org/index.php?title=Main_Page/Background.

[21] Roser, Christopher. “The (True) Difference between Push and Pull.” AllAboutLean.com, 2015. <http://www.allaboutlean.com/push-pull/>.