

1. Introduction to the Integrated Knowledge Management (IKM) Volumes

Table of Contents

1.1. IKM Volumes Motivation	1
1.2. Overview of IKM Volumes	2
1.2.1. Purpose of This Document	2
1.2.2. IKM Guiding Principles	2
1.2.3. IKM Value Proposition	3
1.3. The Knowledge Architecture	4
1.4. Preface from Keith Campbell	6
1.5. IKM Volumes Introduction	6
1.5.1. The Promise of Health Information Technology	6
1.5.2. The Problems with Electronic Health Data	7
1.5.3. A Brief Overview of Health Data Standards	8
1.6. References	11

1.1. IKM Volumes Motivation

In today's healthcare ecosystem, providers continually rely on data from an increasing number of sources to make informed clinical decisions. While this data is a critical component of providing patients with the best care possible, major challenges exist with the accuracy, interoperability, and quality of data that is transmitted within and between healthcare systems. Clinical terminology standards are calling for one strategy to standardize data collection and improve the quality of patient care with standardized and repeatable guidelines for encoding health data. However, these standards often lose critical aspects of the original data capture when they are shared with other providers and healthcare systems. To enhance the quality, interoperability and usability of patient data, the health care industry requires a solution that harmonizes the different clinical terminology standards and generates a common model that supports patient care, data querying and analysis, public health, and policy making. Interoperable data in an ideal world would eliminate patient harm and other inefficiencies stemmed from data misrepresentations and errors in the health care ecosystem. The motivation for this work is to explore the pathway to unifying a patient's care experience through reliable and interoperable data that supports accurate and timely clinical decision support that is solely prompted and driven by real-world data and provides real-time assistance to clinicians.

The essential challenge of informatics practice within the healthcare enterprise is to quickly deliver a high fidelity reasoned interpretation of principles and facts to the point of care—and then to quickly aggregate these point of care experiences for analytic analysis so that new principles and facts can be formulated and validated as part of a continuous optimization of healthcare knowledge and delivery. To effectively answer this challenge, we must focus on simplification and integration of knowledge assets, and on build, test, deploy, and release processes for delivering these assets to the points of care and analysis. This focus on perhaps mundane topics is not because we think that novelty has no place in our work; rather, that without a focus on aspects of our delivery challenge that are often treated as peripheral to the overall problem, we cannot achieve reliable, rapid, low-risk knowledge-asset development and delivery in an efficient manner.

1.2. Overview of IKM Volumes

These Integrated Knowledge Management (IKM) Volumes serve as an educational resource to allow stakeholders to perform a deep dive into the IKM Architecture this team has worked to develop.

1.2.1. Purpose of This Document

These volumes will act as 1) a reference for understanding the complex, multi-layered, and highly inter-dependent Health IT landscape, and 2) a framework for the implementation of a layered, architecture-based approach to integrating and managing knowledge in the healthcare ecosystem that will seek to solve the previously identified issues.

This framework, referred to throughout the volumes as the Knowledge Architecture, is a methodology seeking to enable the concept of Integrated Knowledge Management throughout the healthcare ecosystem.

Through examples and illustrations, these volumes seek to demonstrate that such an approach for managing knowledge will allow for:

- *The ability to recognize equivalence between different standards that are representing the same clinical data,*
- *The ability to represent clinically significant concepts that are needed in a system of record, such as an electronic health record (EHR) or laboratory information system (LIS), unambiguously,*
- *The ability to prevent the inclusion and/or perpetuation of errors in clinical data represented within healthcare systems of record, and*
- *The ability to ensure the safe and reliable evolution of terminology standards, solutions, and processes as new technologies and policies are implemented.*

The intent of the IKM Volumes is to have a direct connection to real world events and circumstances. For a granular and detailed breakdown of real world examples, please refer to the forthcoming Appendix.

1.2.2. IKM Guiding Principles

The following key principles are critical in our summarization and discussion of Integrated Knowledge Management.

Functional decomposition—often referred to as a Separation of Concerns (SoC)—across components or sections with a specific purpose is a foundational design principle for complex system architecture. SoC allows a complete system to be subdivided into distinct sections or components with well-defined functionality and dependencies. If successful, this approach allows individual sections to be reused, as well as designed, implemented, and updated independently to address emerging requirements. This is especially useful and important in a medical context given how many different health information and clinical terminology projects are ongoing at any given time. Efforts are often uncoordinated and led by disparate and unrelated standards development organizations. In these cases, SoC allows teams to simultaneously work independently and in coordination with each other and reuse the resulting artifacts.

The architecture presented in these volumes is inspired by core evolutionary design principles called “Understandable, Reproducible, and Useful,” upon which Systematized Nomenclature of Medicine Clinical Terms (SNOMED CT®) development is still based. [1, 2] These criteria describe an approach for improving data quality and increasing data integrity and agility:

- **Understandable:** The content can be processed by health IT systems and understood by most healthcare providers without reference to private or inaccessible information.

- **Reproducible:** Multiple users or systems apply the data to the same situations and source data with an equivalent result.
- **Useful:** Data is fit-for-purpose – it has practical value for data analysis in support of health information exchange, research, and public health that requires information aggregation across health IT systems.

In order to produce highly reliable software, the following tenets must be applied within a tooling's software development methodology: Use-case driven, architecture-centric, iterative and incremental.

- **Use-case Driven:** Defining, developing, and refining software design(s) to meet known and emerging use cases enable software solutions to meet critical real-world needs.
- **Architecture-centric:** Formally documenting the “layers” of capabilities that a software system may have enables developers to identify dependencies between layers, allowing for shared responsibility of reducing functional complexity across the entire software solution.
- **Iterative and Incremental:** Similar to the Agile mindset, developing software solutions should in small, iterative, and incremental ways thus affording time to perform adequate Independent Validation and Verification (IV&V) methodologies. IV&V methodologies include general-purpose validation approaches conducted by an organization that is not involved in development to generate metrics for data and software quality. From a health information exchange perspective, IV&V can help compare and test source data against requirements of a destination system through manual, semi-automated, and automated inspection of source-to-destination data reviews, thereby helping ensure a sound data strategy intended to reduce patient safety risk and improve data quality for analysis and tracking. [3]

1.2.3. IKM Value Proposition

Issues with the interoperability of health data could cost the U.S. healthcare system upwards of \$30 billion per year. The introduction of industry-wide interoperability could improve upon these expenditures by reducing adverse events, redundant testing, clinician's time spent on manual data entry, and reducing length of stay, resulting in \$1.9 billion, \$1.5 billion, \$12 billion, and \$18 billion in savings, respectively.[4-7] Implementing Integrated Knowledge Management will reduce the challenges associated with unpredictable, poorly controlled, and reactive health IT data. This lack of interoperable data hinders advancements and innovation in the healthcare system, including use of Artificial Intelligence (AI) and Big Data, medical communications, research, and international coordination, as new technologies must rely on suboptimal data and cannot achieve their full potential. [8] The Knowledge Architecture described above will provide the foundation for robust configuration management and quality assurance of health IT data while allowing for a standardized update and extension process. This approach will support increased clinical and laboratory data quality and enable healthcare systems across the industry to conduct robust and meaningful data analysis and increase overall interoperability, which ultimately enhances the quality of care across all medical facilities.

The purpose of employing the Knowledge Architecture and enabling Integrated Knowledge Management can be succinctly described through its benefits: increased interoperability, improved patient outcomes, and meaningful data analysis. These benefits cut across important aspects of healthcare including regulatory adherence, patient-centric care, operational excellence, data management, innovation and collaboration, and cost and fraud control. Additionally, incorporating aspects of Highly Reliable Organization (HRO) data governance and data quality, HRO principles can minimize errors, increase efficiency, and promote a culture of continuous learning and improvement, all of which are highly relevant to this work. [9]

1.2.3.1. Increased Interoperability

Improved interoperability will reduce the need for clinicians to access important health records using dated methods, such as manual, phone, and fax transmission of health data and break down information silos.

Medical facilities will have more interoperable capabilities, opening the door for future opportunities to integrate patient data with other agencies, community partners, and the private sector. [10]

1.2.3.2. Improved Patient Outcomes

Patient care will be optimized and efficient, enabling clinicians to provide high-quality care to patients across medical facilities that improves patient outcomes and provider productivity and reduces adverse events and redundant testing. The potential of digital innovation in health care delivery can also contribute to advancing diagnosis and treatment options, facilitating introduction of telemedicine, promoting and supporting self-management, and reducing error. [11]

1.2.3.3. Meaningful Data Analysis

Medical facilities will be able to easily query specific health factors to gain necessary insights, draw conclusions, and make recommendations based on accurate patient data. Improved data interoperability will also support policy makers and public health experts make informed decisions that more appropriately address evolving needs and priorities. [12]

1.3. The Knowledge Architecture

Increased reliance on computerized health records, including EHR Systems, requires standardized medical terminology to encode health information consistently across systems and enterprises. Clinicians require not only objective quantitative measurements (e.g., 90 beats per minute for a patient's pulse), but also procedural context (e.g., pulse oximetry, manual) about past observations or requests for future interventions. While two quantitative measurements may be the same, the procedural information could indicate meaningful semantic differences and lead to different clinical interpretation and treatment. As information is exchanged across systems, the solution requires a common understanding of data, a method to support knowledge representation, and clinical decision rules based on common terminology and statements. Each component must address an aspect and together need to address the requirements of clinicians. Current HL7 standard implementations rely on profiles and templates to disambiguate statement and terminology and provide sufficient precision for transactions, documents, and standards-based Application Programming Interfaces (APIs). Therefore, the architectural approach described here is applicable to standards organizations developing interoperability for enterprise and project-specific implementations in equal measure.

The Knowledge Architecture, by definition, is a framework for clinical information that is organized into distinct layers such that each higher layer relies upon artifacts from the lower layer. The purpose of the Knowledge Architecture is to standardize clinical statements and terminologies by providing abstraction & eliminating unnecessary complexity. It helps to do this by:

Integrating essential clinical terminologies into a coherent integrated system:

- Localized terminology specific to one organization or system
- EHR Vendor Terminologies (such as Cerner's Code Sets, Event Codes, Event Set Hierarchy)
- Interoperability standards: SNOMED CT®, Logical Observation Identifiers Names and Codes® (LOINC®), RxNorm, and similar

Defining a Standardized Form for clinical statements:

- Reduce unnecessary complexity and improve logical structure
- Create a coherent integrated healthcare data system

- Create an analysis normal form for clinical statements

The proposed architecture, shown visually below, serves as an implementation of SoC, and helps to break out and organize inter-dependent types of knowledge in a way that allows for reuse, ongoing development, and independent updates. If employed, this model can help the healthcare ecosystem to individually resolve issues in their data within their layers, rather than the entire system, leading to much more agile improvement

Figure 1.1. PASTF Model

Detailed descriptions of each layer will be spelled out in the chapters that follow, but brief descriptions of each can be found below:

Foundational Architecture – The foundational layer of the Knowledge Architecture provides the common elements of interoperability, such as: object identity, versioning, modularity, and knowledge representation. It includes (a) the foundation and building blocks of the common model; (b) how the repeatable transformation process of disparate standards into the common model promotes interoperability with other environments; and (c) how the modules of the architecture are tightly version controlled over time.

Terminology Knowledge – The terminology knowledge layer is responsible for structured sets of medical terms and codes that define concepts of interest, including descriptions, dialects, language, and semantic hierarchy. SNOMED CT®, LOINC®, and RxNorm are part of this layer. It defines what valid codes or expressions may be used by higher level layers.

Statement Model – The statement model layer is responsible for defining how data elements are combined to create a statement. This layer reuses the artifacts defined in the terminology knowledge layer.

Assertional Knowledge – The assertional knowledge layer makes use of the Terminology Knowledge layer concepts to specify non-defining facts that may be used by procedural knowledge algorithms. An example fact might be that “thiazide diuretics treat hypertension.” Assertional knowledge may also indicate what symptoms may be associated with a disorder.

Procedural Knowledge – The procedural knowledge layer, also known as imperative knowledge, is the knowledge exercised in the performance of some task. An example would be determining a hypertension treatment plan by analyzing a combination of a patient’s clinical statements and the available assertional knowledge. The procedural knowledge is responsible for information about standard ways to carry out specific procedures, as well as other procedural guidelines, (e.g., treatment protocols for diseases and order sets focused on certain patient situations). Procedural knowledge, together with assertional knowledge, enables clinical decision support, quality measurement, and patient safety. This layer relies on the architectural foundation and terminology layers, incorporates the statement model for information retrieval, and uses assertional knowledge. Procedural knowledge artifacts may include clinical alert rules, reminders, etc., that trigger actions or recommend interventions.

Examining a clinical procedure for controlling hypertension illustrates each of the layers of the informatics architectural separation of concerns.

- At the terminology knowledge layer, there may be various codes and terms from disparate source terminologies to define a concept (e.g., hypertension). Ideally, these overlapping codes and terms would be oriented to the same parent concept during the transformation and integration process at the foundational architecture layer.
- The statement model layer enables representation of blood pressure measurement values (e.g., systolic BP = 140 mmHg), or the categorical data (e.g., pregnancy induced hypertension vs. renal hypertension)

within a standard data structure to facilitate information exchange or retrieval, such as within a standards-based clinical statement (i.e., Fast Healthcare Interoperability Resources, or FHIR).

- The assertional knowledge layer represents non-procedural statements, or facts, such as “Stage 2 high blood pressure is over 140 systolic or 90 diastolic,” or “beta-blockers and ACE inhibitors may be used to treat hypertension”, or “beta-blockers are contraindicated in patients with a diagnosis of reactive airway disease.”
- Finally, the procedural knowledge layer provides algorithms to analyze clinical statements about a patient, in combination with the assertional knowledge, to recommend a treatment protocol for different kinds of hypertension, including the considerations of, (e.g., patient age, co-morbidities etc., which can be generated by an electronic clinical decision support system [statement + assertional layers]). This layer adds support for workflow and conditional logic (i.e., if-then-else).

1.4. Preface from Keith Campbell

This section is currently a placeholder.

1.5. IKM Volumes Introduction

1.5.1. The Promise of Health Information Technology

Health Information Technology (IT), or the use of computer hardware, software, or infrastructure to record, store, protect, and retrieve clinical, administrative, or financial information, has become a critically important form of organizing, managing, analyzing, and exchanging patient health information. [13] Though first employed in the 1960s, Health IT entered the scene and obtained widespread adoption in the early-1990s promising:

- The enhancement of quality and safety of healthcare,
- The ability to more easily measure and reduce the cost of care,
- The ability to share information with multiple providers across organizations in a continuum of care,
- The ability to implement and automate high-quality decision support into clinical workflows, and
- The ability to improve clinician/provider satisfaction and reduce burden and burnout. [14]

The most commonly known forms of Health IT systems, EHRs, or a systematized collection of electronically stored patient and population health information, have undoubtedly enabled dramatic advancements in patient care, allowing for breakthrough technologies with advanced analytical capabilities (e.g., Machine Learning(ML), AI) to more easily identify new diseases and develop treatments and reduce the need for patients to maintain and carry paper records from provider to provider.

Despite advancement in technologies like these, a Johns Hopkins study conducted in 2016 suggested that preventable medical errors, including those due to EHR errors, accounted for approximately 250,000 patient deaths every year – making it the third-leading cause of death in the United States. [15] When Controlled Risk Insurance Company (CRICO), a medical professional liability organization, identified 147 cases of patient harm over five years in which an EHR was identified as a contributing factor, a staggering 80 percent of these cases involved moderate or severe harm to the patient. [16] In total, an estimated \$61 million in direct payments and legal expenses resulted from these cases, at an average cost of \$415,000 per case.

On the surface, exciting new technology has continued to advance in the healthcare ecosystem, but cases of patient harm, redundant testing, and over-expenditure persists. This begs the question: **is Health IT truly fulfilling its promises?**

Technology has the potential to revolutionize patient care – concepts like AI and ML promise better patient outcomes and reduced clinician burden. However, these systems are only as good as the data from which they operate. EHRs, for example, were created to provide healthcare automation, patient care and patient access, and reduce clinician workload. The advancements in EHRs have had a great impact for patients, but users continue to face complications due to the lack of standardized data exchange. Unstandardized data stacking has made it difficult to interpret a patient’s chart with accurate data reporting. [17] As far as healthcare technology has come, the quality of existing technology platforms is limited through the quality of encoded data – which is where the problem begins.

1.5.2. The Problems with Electronic Health Data

While healthcare information originates across a variety of disparate systems and organizations, the United States healthcare environment requires interoperable, coordinated, and accurate information to support high quality patient care, clinical and laboratory research, public health monitoring, and regulatory decision making. Many needs must be met in the existing, fragmented healthcare setting which includes a network of interconnected information systems.

Historically, the United States healthcare system has been composed of relatively autonomous organizations that maintain loose relationships in support of patient care. Patients typically visit one or more primary care clinicians and are frequently sent, along with orders and patient data, to specialists, laboratories, and other care teams. The observation and interpretation of patients’ healthcare data across these settings is crucial to the processes of making safe clinical decisions, tracking patient outcomes, anticipating future health problems, identifying deviations from expected trends, informing preventive and public health measures, supporting clinical research, and providing a legal and historical record. [18]

For example, when a primary care provider records a diabetes diagnosis in the provider's local system, it is critical for that system to share the data accurately and seamlessly to the pertinent pharmacy and laboratory systems. These systems must then in turn communicate with one another and with the primary care system to alert the provider that the patient has previously been prescribed gentamicin and has recently been given an order for a radiologic examination that requires intravenous iodine dye – both of which are known to affect kidney function. If this information is not accurately communicated across healthcare settings in a way that can be readily accessed, the patient may face unnecessary harm or risks that could have been avoided.

The following sections aim to spell out some of the current problems with electronic health data more clearly.

1.5.2.1. The Need for Unmatched Precision

The degree to which data must maintain precision in healthcare is unmatched in other industries. Health data is unique in its direct association with a person’s health outcomes. Nuanced concepts exist in patient care settings creating challenges for the representation thereof in data capture and transmission. Laboratory testing on a specimen involves a plethora of biological concepts which must be represented by data in a meaningful way while maintaining precision as well as consumption for other entities or systems in the ecosystem. When an association to a patient’s health outcomes is combined with such nuanced concepts, exceptional rigor must be taken in the maintenance of data precision and representation.

1.5.2.2. The Size, Scale, and Complexity of the System

The ecosystem of healthcare entities, which share responsibility in the delivery of care, is vast. Various stakeholders include government health authorities, EHR companies, medical device manufacturers,

health insurance companies, medical laboratories, hospitals and healthcare systems, and clinical providers among various site types. A healthcare entity's data requirements differ in context from other industries regarding the facilitation of patient care and handling of health data in methods that are specific to their context. Industries such as transportation and logistics are more confined to a common context for the data required to successfully facilitate operations. The array of contexts for healthcare entities creates an unsolved challenge for interpreting and exchanging data without compromising its intended meaning.

1.5.2.3. The Lack of Adherence and Enforcement of Standards

Standards in healthcare exist to promote successful interpretation and exchange of data; however, they are not enforced and can experience poor adherence in practice, undercutting their effectiveness and purpose. Despite a significant history of emphasis on healthcare data standards, entities continue to leverage proprietary, or localized, methods for data representation creating inherent and avoidable challenges in care delivery, reporting, and the meaningful application of available technology. Motivations for proprietary or localized data methods cannot be confirmed, however, considerations include protection of market share from competitors and avoidance of public scrutiny for misalignment to standard processes. Maintaining alignment with standards development organizations (SDOs) is resource intensive and invites SDOs to measure and scrutinize an organization against the standard processes to which an entity claims conformity.

1.5.2.4. The Inability to Share Between Systems

Numerous distinct systems and organizations across the healthcare ecosystem lead to a large variation in how data are handled, e.g., data collection, data manipulation, data transmission, data storage, and real-world application. Each system has a proprietary, standard, or ad-hoc information model and is typically configured to satisfy organization-specific needs, resulting in differences in data handling within systems. While this variation is not an innate issue, the difficulty exchanging data between systems and the lack of interoperability compromises data quality and ultimately patient safety.

1.5.2.5. The Significant Variation in Types of Data Needed

Medical data range in format from narrative and textual data to numerical measurements, images, drawings, and more. It is common practice for several different observations of a singular patient to be made concurrently, the observation of the same patient parameter to be made over the course of several points in time, or both. [18] It is also important to keep a record of the circumstances under which data are obtained, further adding to the complexity and volume of data. For example, a simple blood pressure reading may factor if a blood pressure reading is taken on the arm or the leg, whether the patient recently exercised, and what kind of device was used.

Current approaches for representing data from various, disparate health information systems introduce additional opportunities for inefficiencies, redundancies, and inconsistencies. Unpredictable and denormalized data formats reduce the quality of data processing and the ability to conduct safe and reliable analytics. When incongruent data is repeatedly collated, it requires extensive expertise and resources to synthesize basic trends, information, and insights that should be readily available. A history of numerous scholarly and practical efforts illustrates the challenges of using observational data for safe patient care, comparative effectiveness research, and analyses for biases, workflow differences, and issues with variations in data collection, such as invalid, inconsistent, and missing data. [19 – 34]

1.5.3. A Brief Overview of Health Data Standards

Standards are generally required when excessive variation creates inefficiencies or impedes effectiveness. [18] There have been many valuable standards applied throughout history such as railroad tracks, telephones, WiFi, and financial transactions. The broad goal of the application of standards to electronic health data is unambiguous data exchange, or “interoperability” – the ability of two parties, either human or

machine, to exchange data or information so that it is transportable, organizable, and interpretable in an expected format.

Standards for electronic health data have existed since the creation of the EHR in the 1960s. Computer engineers responsible for EHR development were tasked with representing health concepts in a digital format that could be shared with users and connected systems for common interpretation. In 1965, the Systematized Nomenclature of Pathology (SNOP) was developed to provide a common way to describe morphology and anatomy. A decade later, SNOP expanded into its known name today, Systematized Nomenclature of Medicine Clinical Terminology (SNOMED CT®). As the adoption of EHRs increased during the 1980s, Health Level Seven International (HL7) was formed in 1987 as a non-profit, American National Standards Institute (ANSI)-accredited standards development organization to develop standards that provide global health data interoperability. Since then, several SDOs in the healthcare space have appeared and, along with HL7 and SNOMED CT®, continued to commit resources to the development and adoption of standardized health data processes. In 2004, under the United States Department of Health and Human Services (HHS), the Office of the National Coordinator for Health Information Technology (ONC) was born to promote a national health information technology infrastructure and oversee its development. While ONC is not a SDO, it works in close coordination with SDOs to support the adoption and application of health data standards in the U.S.

While interoperable health data is a complex and nuanced topic, a primary component is maintaining the meaning and intent of clinical data from the point of consumption throughout the transmission and use of that data across various systems. As data is exchanged within and between multiple healthcare systems, the original provider's meaning and intent must be electronically communicated in a way that, upon receipt, receiving systems and providers can understand in an identical manner. Data exchange standards provide a formal process for defining specific constructs to be exchanged between machines and must be used to support this seamless transfer of data.

1.5.3.1. Structure and Meaning

In simple terms, *Structure* and *Meaning* are two considerations to ensure data maintains its full integrity and can be communicated in a way that supports interoperability.

Structure is the ability of systems to agree on how to format the data. Structural interoperability ensures the seamless communication of the data's structure between systems but does not guarantee that the receiving system will interpret the data with the original intent. Web pages, for instance, are developed using HTML or XML to ensure structural interoperability. While this guarantees a web page can be read by any machine with a web browser, it does not ensure the meaning of that page is identical across machines. There are a variety of healthcare structural standards, including FHIR, HL7 Version 2 Messaging (HL7 v2), and Consolidated Clinical Document Architecture (C-CDA).

To understand structural interoperability, consider the two sentences below. While both sentences have the same sentence *structure*, their *meanings* are completely different and only one makes logical sense.

Figure 1.2. Structural Interoperability

Meaning is the ability of systems to share the same interpretation of the data. Meaning interoperability ensures the original meaning and intent of data is interpreted identically across all parties but allows for structural variance. When people and systems use different words and codes to describe synonymous or related things, healthcare meaning standards translate the data to a common meaning. Healthcare terminology and meaning standards include SNOMED CT®, LOINC® and RxNorm (a terminology defining medications and ingredients from the National Library of Medicine).

In the below example, the two sentences share a common *meaning*, however the sentence *structure* varies drastically.

Figure 1.3. Meaning Interoperability

While Figure 1.2, “Structural Interoperability” and Figure 1.3, “Meaning Interoperability” intend to illustrate *structure* and *meaning* simply, a more real-world example like, “the patient was given pain medication” and “the patient was given medication for pain,” highlights potential issues when determining equivalence. Objectively, the two sentences do not follow the same *structure*; however, they may or may not have the same *meaning* depending on your perspective or interpretation. The first sentence could be read to mean “the provider prescribed the patient pain medication to address pain” or that “the patient acquired pain medication from another source other than the provider.” The complex relationship between these two concepts can lead to confusion and will need to be addressed to accurately characterize equivalence.

1.5.3.2. Example of Overlapping and Intersecting Nature of Standards

Hospitals and independent reference labs are two primary entities in the health data ecosystem which share direct relationships and interfaces in the exchange of clinical data. Healthcare interfaces must leverage various existing standards for *structure* and *meaning* and use them together in integrated ways. An electronic interface provides an avenue, or connection point, for sharing data in a meaningful, coherent way between two disparate systems. Without an interface, two disparate systems are unable to communicate in a common language. In the U.S. and elsewhere, the most common and accepted messaging structural standard for EHR-to-Reference Lab interfacing is the Health Level Seven (HL7) Version 2.x (V2.x) standard. While other messaging standards exist and can be used, the HL7 V2.x messaging standard, first released in 1987, is used in 95% of U.S. healthcare organizations and in more than 35 countries. [35]

To illustrate how standards for *structure* and *meaning* may be currently implemented, we will explore a scenario of a typical hospital-to-reference lab order message generated by a hospital EHR system sent electronically to a Reference Lab. This laboratory order message is received, processed, and acted upon by which the laboratory will perform the test specified in the order message received. After a laboratory test is performed and a test result is available, the Reference Laboratory in this scenario will return a Laboratory Result message to the hospital EHR system.

Figure 1.4, “Erroneous LOINC® Mapping” below shows an example of an HL7 v2 Observation (OBX) segment from two different mock lab result messages. In this scenario, a laboratory has conducted two distinct blood glucose tests on two distinct specimen sources. For the first test, a specimen from serum or plasma was used. For the second test, the specimen was from whole blood. The figure below illustrates how the HL7 v2 *structure* standard is populated using existing terminology standards for *meaning*, in this case LOINC®. In the first HL7 segment for test 1, the LOINC® code is for a “Glucose [Mass/volume] in Serum or Plasma”, whereas for test 2, the LOINC® code is for a “Glucose [Mass/volume] in Blood”. Different tests were conducted. Therefore, different LOINC® codes were used in the HL7 v2 messages. The local laboratory information system and EHR also define and use their own code systems to encode the tests; these are called “local codes.” Although not universal, local codes are another way for health information systems to represent *meaning* for clinical concepts. The HL7 v2 observation data in Figure 1.4, “Erroneous LOINC® Mapping” shows blue text representing a local code, and the observation data denoted in red represents a corresponding LOINC® code.

Figure 1.4. Erroneous LOINC® Mapping

The figure shows how the same local code can be erroneously mapped to different LOINC® codes and transmitted in a Lab Result message. In this case example, we see the same laboratory conducting two distinct tests and encoding each test with a different LOINC® code; however, both are mapped to the same

local code. This mock scenario highlights one example of the myriad of interoperability challenges faced in healthcare today. [35] Successful interoperability is dependent on either using universal code systems for data capture or through mapping local content to a universal code system. Local code systems serve the same function as universal coding systems such as SNOMED CT® and LOINC®, however are internal identification schemas that consist of a set of local codes and their associated definitions.

Because the receiving system will parse the identical local codes from the two mock HL7 v2 segments above, it will interpret the lab result as having been produced from identical lab tests. In simpler terms, the receiving system will interpret these messages as the same glucose test with a “count of 2 observations” rather than two different glucose tests, each with a “count of 1 observation.” Since the source specimen and lab tests are different, but represented identically in the data transmission, semantic meaning is lost compromising patient safety and accurate reporting. The receiving system will interpret these two HL7 v2 segments as meaning the “same thing” or representing the “same glucose test”; however, any terminologist or laboratory professional would interpret this as “different glucose tests.”

To improve the quality and safety of patient care, there is a need to refine data exchange within the healthcare ecosystem. We often hear comments as, “Why is the problem of data exchange in healthcare so difficult? After all, other industries—banking, for example— have solved the problem. Healthcare transactions can’t be that much more complicated than those managed by a geographically distributed, multiple line-of-business banking system.”

While other industries, like banking, have resolved issues with their transmission and exchange of complex banking data across geographical locations and between numerous unaffiliated businesses, the complexity with healthcare data arises, in part, because neither the standards for structure (i.e., statement model) nor standards for meaning (i.e., terminology model) can represent clinical data alone in the current data structures. Both types of models can be arbitrarily complex, and, at some point, semantics not represented in one must be represented in the other if one is to achieve interoperability. [18]

1.6. References

1. Spackman K, Guillermo R. Examining SNOMED from the perspective of formal ontological principles: Some preliminary analysis and observations. KR-MED 2004 Proceedings 2004; 72-80.
2. Campbell KE. Distributed development of a logic based controlled medical Terminology. PhD Dissertation, Stanford University, June 1997.
3. Analysis Normal Form Informative Ballot. HL7 CIMI Work Group. Sept 2019. http://www.hl7.org/documentcenter/public/ballots/2019SEP/downloads/HL7_CIMI_LM_ANF_R1_I1_2019SEP.pdf.
4. Middleton B. The Value of Healthcare Information Exchange and Interoperability [Internet]. 2014 [cited 2023 Jul 31]. Available from: https://www.researchgate.net/profile/Blackford-Middleton/publication/228874733_The_value_of_healthcare_information_exchange_and_interoperability/links/0fcfd51476ee58b1e0000000/The-value-of-healthcare-information-exchange-and-interoperability.pdf.
5. Houlne D. Technology and the Cost-of-Care Convergence [Internet]. HIMSS; 2021 [cited 2023 Jul 31]. Available from: <https://www.himss.org/news/technology-and-cost-care-convergence#:~:text=Better%2C%20cost-effective%20care%20starts%20with%20interoperability.%20One%20study,U.S.%20health%20system%20over%20%2430%20billion%20a%20year.>
6. Analysis by West Health Institute finds medical device interoperability could save more than \$30 billion a year. West Health; 2013 [cited 2023 Jul 31]. Available from: <https://www.westhealth.org/press-release/new-analysis-by-west-health-institute-finds-medical-device-interoperability-could-save-more-than-30-billion-a-year/>.

7. Benson T, Grieve G. Why Interoperability Is Hard. In: Principles of Health Interoperability: FHIR, HL7 and SNOMED CT. Cham: Springer; 2021. p. 21–40.
8. Lehne M, Sass J, Essenwanger A, Schepers J, Thun S. Why Digital Medicine Depends on Interoperability [Internet]. U.S. National Library of Medicine; 2019 [cited 2023 Jul 31]. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6702215/?report=classic>.
9. Frankel A, Leonard M, Brooks A, Peach A, Proulx J, Cramer D, et al. The Framework for High Reliability Healthcare [Internet]. [cited 2023 Jul 31]. Available from: <https://www.safeandreliablecare.com/the-framework-for-high-reliability-healthcare2>.
10. Patzer A. Council Post: Why a Lack of Interoperability in Healthcare is a Detriment to Patients [Internet]. Forbes Magazine; 2020 [cited 2023 Jul 31]. Available from: <https://www.forbes.com/sites/forbestechcouncil/2020/03/02/why-a-lack-of-interoperability-in-healthcare-is-a-detriment-to-patients/?sh=1dc13d761120>.
11. Abernethy A, Adams L, Barrett M, Bechtel C, Brennan P, Butte A, et al. The promise of Digital Health: Then, now, and the future [Internet]. U.S. National Library of Medicine; 2022 [cited 2023 Jul 31]. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9499383/>.
12. Ricker-Kiefert K, Block L. State Strategies to Advance Health Data Interoperability [Internet]. NGA; 2021 [cited 2023 Jul 31]. Available from: <https://www.nga.org/wp-content/uploads/2021/03/State-Strategies-to-Advance-Health-Data-Interoperability.pdf>.
13. What is Health IT? [Internet]. HealthIT.gov. [cited 2023Jan13]. Available from: <https://www.healthit.gov/faq/what-health-it>.
14. Net Health. The History of Electronic Health Records (ehrs) - updated [Internet]. Net Health. 2022 [cited 2023Jan13]. Available from: <https://www.nethealth.com/the-history-of-electronic-health-records-ehrs/>.
15. McMains V. Johns Hopkins study suggests medical errors are third-leading cause of death in U.S. [Internet]. The Hub. 2016 [cited 2023Jan13]. Available from: <https://hub.jhu.edu/2016/05/03/medical-errors-third-leading-cause-of-death/>.
16. McGinley P. Malpractice claims analysis confirms risks in ehRs [Internet]. Patient Safety & Quality Healthcare. 2014 [cited 2023Jan13]. Available from: <https://www.psqh.com/analysis/malpractice-claims-analysis-confirms-risks-in-ehrs/>.
17. Shortliffe EH, Cimino JJ, editors. Biomedical Informatics: Computer Applications in Health Care and Biomedicine. 5th ed. Cham, Switzerland: Springer Nature; 2021.
18. How Doctors Feel about Electronic Health Records [presentation]. Stanford Medicine. National Physician Poll by The Harris Poll. [cited 2023Jan13].
19. Kahn MG, Brown JS, Chun AT, Davidson BN, Meeker D, Ryan PB, et al. Transparent reporting of data quality in Distributed Data Networks. eGEMs (Generating Evidence & Methods to improve patient outcomes) [Internet]. 2015;3(1):7. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4434997/>.
20. Kahn MG, Callahan TJ, Barnard J, Bauck AE, Brown J, Davidson BN, Estiri H, Goerg C, Holve E, Johnson SG, Liaw ST, Hamilton-Lopez M, Meeker D, Ong TC, Ryan P, Shang N, Weiskopf NG, Weng C, Zozus MN, Schilling L. A Harmonized Data Quality Assessment Terminology and Framework for the Secondary Use of Electronic Health Record Data. EGEMS (Wash DC). 2016 Sep 11;4(1):1244. doi: 10.13063/2327-9214.1244. PMID: 27713905; PMCID: PMC5051581.

- 21.Cholan RA, Weiskopf NG, Rhoton DL, Colin NV, Ross RL, Marzullo MN, Sachdeva B, Dorr DA. Specifications of Clinical Quality Measures and Value Set Vocabularies Shift Over Time: A Study of Change through Implementation Differences. *AMIA Annu Symp Proc.* 2018 Apr 16;2017:575-584. PMID: 29854122; PMCID: PMC5977609.
- 22.Cholan RA, Weiskopf NG, Rhoton D, Sachdeva B, Colin NV, Martin SJ, Dorr DA. From Concepts and Codes to Healthcare Quality Measurement: Understanding Variations in Value Set Vocabularies for a Statin Therapy Clinical Quality Measure. *EGEMS (Wash DC).* 2017 Sep 4;5(1):19. doi: 10.5334/egems.212. PMID: 29881739; PMCID: PMC5983064.
- 23.Hemler JR, Hall JD, Cholan RA, Crabtree BF, Damschroder LJ, Solberg LI, Ono SS, Cohen DJ. Practice Facilitator Strategies for Addressing Electronic Health Record Data Challenges for Quality Improvement: EvidenceNOW. *J Am Board Fam Med.* 2018 May-Jun;31(3):398-409. doi: 10.3122/jabfm.2018.03.170274. PMID: 29743223; PMCID: PMC5972525.
- 24.Colin NV, Cholan RA, Sachdeva B, Nealy BE, Parchman ML, Dorr DA. Understanding the Impact of Variations in Measurement Period Reporting for Electronic Clinical Quality Measures. *EGEMS (Wash DC).* 2018 Jul 19;6(1):17. doi: 10.5334/egems.235. PMID: 30094289; PMCID: PMC6078150.
- 25.Hogan WR, Wagner MM. Accuracy of data in computer-based patient records. *J Am Med Inform Assoc.* 1997 Oct;4(5):342–55.
- 26.Aronsky D, Haug PJ. Assessing the quality of clinical data in a computer-based record for calculating the pneumonia severity index. *J Am Med Inform Assoc.* 2000 Feb;7(1):55–65.
- 27.Arts D, de Keizer N, Scheffer G-J, de Jonge E. Quality of data collected for severity of illness scores in the Dutch National Intensive Care Evaluation (NICE) registry. *Intensive Care Med.* 2002 May;28(5):656–9.
- 28.Thiru K, Hassey A, Sullivan F. Systematic review of scope and quality of electronic patient record data in primary care. *BMJ.* 2003 May 15;326(7398):1070.
- 29.Hasan S, Padman R. Analyzing the effect of data quality on the accuracy of clinical decision support systems: a computer simulation approach. *AMIA Annu Symp Proc.* 2006:324–8.
- 30.Cruz-Correia RJ, Rodrigues P, Freitas A, Almeida FC, Chen R, Costa-Pereira A. Data quality and integration issues in electronic health records. In: Hristidis V, editor. *Information discovery on electronic health records.* Chapman and Hall/CRC; 2009:55–95.
- 31.Botsis T, Hartvigsen G, Chen F, Weng C. Secondary use of EHR: Data quality issues and informatics opportunities. *AMIA Summits on Translational Science proceedings AMIA Summit on Translational Science.* 2010;2010:1–5.
- 32.Hripesak G, Knirsch C, Zhou L, Wilcox A, Melton G. Bias associated with mining electronic health records. *J Biomed Discov Collab.* 2011;6:48–52.
- 33.Brown JS, Kahn M, Toh S. Data quality assessment for comparative effectiveness research in distributed data networks. *Medical care.* 2013;51:S22–S29.
- 34.Rusanov A, Weiskopf NG, Wang S, Weng C. Hidden in plain sight: bias towards sick patients when sampling patients with sufficient electronic health record data for research. *BMC Med Inform Decis Mak.* 2014;14:51.
- 35.Health Level Seven international [Internet]. HL7 Standards Product Brief - HL7 Version 2 Product Suite | HL7 International. [cited 2023Jan12]. Available from: https://www.hl7.org/implement/standards/product_brief.cfm?product_id=185