Integrated Knowledge Management (IKM) Volume 9

Version 1 - Last Updated 1/25/2024

Table of Contents

List of Figures

Part I. Open-Source Overview

Table of Contents

1. Open-Source

1.1. How to Contribute

For community members and users who would like to collaborate and contribute, please visit the 'Community' tab on the *IKM.dev* website or click on **Community** in the top header for additional options.

1.2. What is Open-Source?

Open-source software makes its source code, blueprints, documentation, and other products publicly available for users to modify, contribute to, and distribute. Open-source software uses a decentralized approach to software development that supports open collaboration, testing, and improvements from the community of users. In open-source, no one group is solely responsible for maintaining or improving software because peer-to-peer production of code and documentation is encouraged. While a governing body may exist in practice to maintain the software and ensure contributions meet community guidelines, they may not hold special rights to the software. Reasonable guidelines, such as extensive documentation procedures, enhance the benefits and trust in open-source solutions, while simultaneously preventing an “anything goes” mentality.

1.3. Why Does Open-Source Matter?

The current restrictive and proprietary nature of software licensing within the healthcare ecosystem significantly hinders interoperability, particularly by preventing or restricting community engagement and contributions that could improve the software or code to meet emerging needs. To achieve lossless data exchanges between and within systems, the Health Information Technology (IT) community must consider open-source software and open-source licensing models. While these licenses enable use, regardless of intent, and the ability to modify, adapt, and extend software with minimal restrictions, they also establish rules of the road and codify best practices to ensure adoptability and ease of use for both the initial user and the broader downstream community. Open-source licensing will support the public-private consortium needed to address gaps created by varying rates of adoption and implementation of proprietary software.

Due to the decentralization and wide-spread availability of open-source software, errors and bugs are found much quicker because the code is utilized in a variety of contexts and stretched much farther. Instead of a single development team being tasked with identifying and repairing errors, users can serve as developers and contribute to addressing errors much faster. As a result, open-source software employs a high degree of modularity where it can be adaptive and flexible to the diverse needs of its users.

Allowing for open licensing encourages improvements to terminology and standards by ancillary bodies. The open licensing also creates a more robust and flexible system that can adapt to the varying needs of organizations.

1.4. Preference for Open-Source

Open-source software is preferred and is managed through open-source licensing. A permissive license will provide source code to users, carries minimal restriction over adaption, modification, and distribution of intellectual property of organizations, and allows the issuing body to impose some restrictions, such as monetization. However, the permissive license does not exert special rights to derivative works. Permissive licensing throughout the foundational architecture, terminology knowledge, and statement model layers removes barriers imposed on modifications and extensions and improves the compatibility of different terminologies to support a common understanding of knowledge.

2. Summary of Findings - Managing the Best Practice for Open-Source Communities

2.1. The Current State of Healthcare Knowledge Management

Over the last decade, billion dollars have been poured into achieving interoperability in the healthcare industry. Patients, providers, and healthcare agencies have collectively identified the need for a more harmonized, integrated health system to solve the challenges surrounding clinical data, such as its interpretability and accuracy. The current structure of the healthcare ecosystem continues to fail in providing adequate structures conducive to collaborative, community-based solutioning. A review in 2021 of American Hospital Association Information Technology survey data found that by 2018, 98.3% of hospitals have adopted Electronic Health Record (EHR) systems – the primary creators and curators of digital health data. Yet, across the board, health systems have struggled to move beyond simple adoption and implement advanced use in patient engagement and clinical data analytics. [1] Notably, critical access hospitals in rural areas were less likely to demonstrate advanced usage with an expanding deficit to their counterparts in clinical data analytics since 2015, leading to gaps in patient access of quality healthcare.

Today, many hospital systems have the goal of improving data interoperability within and beyond their system. To be truly useful, clinical data must move through numerous systems without loss of meaning. In an idealized situation, data transfers would complete a single round-trip with total integrity intact. In reality, only 22% - 68% of data integrity is successfully preserved through its transfer journey. [2] A patient's record may correctly showcase the correct numerical result from a test but could lose the context and meaning surrounding the test. Ever-increasing system complexity and resulting low-quality data, however, has ultimately impacted the quality of care across the healthcare ecosystem, such as inefficiencies of redundant testing and procedures being ordered, expanding the time frame of care for patients. Additionally, the very systems that have been designed to integrate EHRs into the health space can create barriers to interoperability through closed, proprietary systems. Multiple, sometimes competing, encoding standards have been developed over the last few decades to address the varying formats clinical data embodies. Some clinical standardization forms, however, are very broad in its acceptance criteria resulting in large amounts of variation between encodement while others are too restrictive, limiting the situations in which they can be utilized. As a result, the need for a consistent form to map standards to and between other standards in creation of a common model is required for improvement of interoperability within the healthcare landscape where clinical decisions can be captured regardless of what point they originate in patient care.

Defining a Key Issue: Restrictive Healthcare Licenses

The root cause of insufficiencies within Health IT is not limited solely to a group of vendors or organizations but is rather a foundational issue that permeates throughout the ecosystem. Restrictive licenses are just one of these causes and are used in varying forms across the landscape. However, this spurs unintended consequences of limiting harmonization. Instead of just treating the symptomatic results of poor data quality, attention should also be spent understanding the reasons beneath recurring pain points. Various standards have been developed to compete and improve upon the limitations of other encoding standards. As multiple organizations, such as by clinicians, laboratorians, and providers integrate patient data together, the need of various standards to transfer openly and completely becomes even more apparent. A piecemeal approach has occurred where different standards carry licensing restrictions to protect against modifications and extensions of their code for proprietary oversight, creating confusion on how and to what extent standards can be shared. Healthcare data standard organizations seek to protect their software and standards, by design, from unauthorized subversion, modification, and duplication of their work through their licensing agreements. These restrictions often allow for direct oversight of derivative works from the issuing body as well as monetization from commercial usage of their product. In a highly profitable landscape, restrictive licenses allow organizations to prevent their work from being duplicated by rivals and seek longevity in the marketplace. Proprietary licensing leads to siloed operations by providers and developers, as well as higher implementation costs, limiting stronger adoption. These restrictive licenses can pose challenges on the ease of transfers within and between systems and are creators of insufficiencies within Health IT. For example, issuing standards do not always fit into the exact context practitioners need and require adaption into localized knowledge. Licenses often prevent against local knowledge being shared between system and require a mapping to an existing code for meaningful data usage. Data integrity can be impacted as there may be no exact match that encompasses all the information from local codes and be transferred incompletely. Tools to aid in standardization may be restricted due to modification and extension restrictions within proprietary licenses.

The current restrictive and proprietary nature of licensing within the healthcare ecosystem significantly hinders interoperability. To achieve lossless data exchanges between systems, the Health IT community must consider movement towards open-source software and open-source licensing models. Gaps created by varying rates of adoption and implementation requires various avenues of collaboration within the Integrated Knowledge Management (IKM) community to address these issues within the entire Health IT ecosystem.

2.2. The Case for a More Open Future in Healthcare

The rapid expansion of technology and connectivity has created ill-defined guidelines on management and ownership of data. As a result, researchers and users find themselves navigating different and competing guidance on utilizing and adapting software for their own needs. One of the main movements on expanding access and resolving conflict between proprietary systems is that of granting wide, sweeping rights through the open-source movement. Notably, open-source software like Mozilla Firefox has enabled web access utilized by millions every year. Additionally, open software systems, like Apache 2.0, have had broad adoptive use because it allows software to be used for any purpose, modified or otherwise.

The following sections aim to define the open-source software and licensing more clearly:

Open-Source Software (OSS)

There are two branches of software utilized by individuals and companies: open-source software and proprietary, or closed-source, software. At the most theoretical level, in an open-source environment, the source code is freely available and accessible to its end-users regardless of their intent, be it for personal or commercial use. The term “free” is under debate regarding the cost of accessing the software and under contention on if the source code can be monetized for access. Here, free is used as in libre, or freedom, and not necessarily in cost. The discussion of free licensing extends to freedoms guaranteed through the licensing for the software, and not in the context of the monetary cost to access the software. Many software systems are considered to be open-source but are not free from implementation cost. There are third-party costs associated with integrating and maintaining open-source.

The Free Software Foundation (FSF) defines free as in:

- The freedom to run the program as you wish, for any purpose (freedom 0).
- The freedom to study how the program works and alter it so it computes as you wish (freedom 1). Access to the source code is a precondition for this.
- The freedom to redistribute copies so you can help others (freedom 2).

- The freedom to distribute copies of your modified versions to others (freedom 3). [3]

This is controlled in the forms of licenses and is discussed further below. Closed source systems protect the source code and modifications cannot be made where the end-user can only execute the code within the confines of the original intent of the creator.

In open-source, no one group is solely responsible for the software because it allows for a peer-to-peer production of codes and documentation. A governing body may exist to oversee and maintain the software as long as no one group holds special rights to the software. Due to the decentralization of software, errors and bugs are found much quicker because the code is utilized in a variety of contexts and stretched much farther. Instead of a single development team being tasked with identifying and repairing errors as they are found, users can also be developers and contribute to identifying errors much faster. As a result, open-source software utilizes a high degree of modularity where it can be adaptive and flexible to the diverse needs of its users.

In its purest form, open-source software lowers the burden of buy-in required by consumers where they can build on the work of others through a collaborative community without reinvention. Lower implementation cost and maintenance of open-source software helps expand the possible consumer base. It is not restricted to those who can devote the time and resources to stand up the required infrastructure and knowledge base. Instead, end-users can guide and shape the software to their niche rather than be limited to the specific intended usage at its inception. The creation of open-source leads to a symbiotic ecosystem where software is more robust and applicable than its closed counterparts.

Many software systems today utilize qualities of open-source software but lie in a gray area where they maintain proprietary techniques. These limits are imposed by their licensing models through their end-user agreements.

Real World Example: General Data Protection Regulation (GDPR) and Data Privacy in Healthcare

The General Data Protection Regulation (GDPR) is an European Union (EU) privacy and security law that was put into effect May 25, 2018. This EU law imposes obligations onto organizations anywhere in the world if they target or collect data related to EU citizens or residents. The law outlines seven protection and accountability principles, listed below [4]:

1. Lawfulness, fairness and transparency
2. Purpose limitation
3. Data minimization
4. Accuracy
5. Storage limitation
6. Integrity and confidentiality
7. Accountability

The GDPR has influenced data management and privacy practices in the healthcare sector by compelling healthcare organizations to adopt stricter data protection measures and transparency in data handling. According to GDPR law, health data is defined as personal data related to the physical or mental health of a natural person, including the provision of health care services, which reveal information about their health. This health data also includes genetic and biometric data. [5]

Healthcare organizations must adapt to these strict new regulations, and are charged with managing patient consent, data portability, and the consequences of non-compliance, which could be fines up to €22M+. For U.S. healthcare organization this means complying with GDPR as well as Health Insurance Portability and Accountability Act (HIPPA), which is a separate level of consent and information the provider has to ensure the patient receives. Additionally, GDPR places the burden of proof of consent on the provider, and they must implement an audit trail for electronic consents. This example demonstrates how regulatory changes are shaping the future of data management in healthcare and aligns with the need for clarity in data management guidelines.

Open-Source Licensing (OSL)

The implementation of open-source software is managed in-hand through open-source licensing (OSL). The license of the software sets out the guidelines on the usage and implementation of the specific version of software being used. The license acts as a copyright agreement where organizations and creators can control the distribution, duplication, and modification of their work.

There are many different types of licenses, as seen in Figure 2.1 below, and are often created to be intentionally vague to apply to a multitude of situations and preserve the rights of the licensor:

Figure 2.1. Restrictions Imposed by License Type

Public-Domain: Public domain refers to any information, software, or intellectual property that carries no restrictions in regard to copyright law. It is free to use by any individual. No one group can exercise control over it. In a software setting, source code must be provided.

Permissive License: Also known as an open-source license. This type of license carries minimal restriction over adaption, modification, and distribution of intellectual property of organizations. An issuing body may impose some restrictions, such as the monetization, but it does not exert special rights to derivative works. In a software setting, source code must be provided.

Protective License: This type of license, also called a copyleft license, grants rights on modification and redistribution but also requires derivative works to include and grants the same rights of the original license. Similar to public domain and permissive licenses, source code must also be accessible.

Noncommercial: The license only grants special rights to be used in a noncommercial setting where an external entity may exercise proprietary control or monetization of the applicable intellectual property. This can also decrease the cost of access, especially for organizations for academic or research purposes.

Proprietary: The most traditional form of licensing where no rights must be granted to other entities. There are varying significant restrictions on modification and distribution, where it is often disallowed completely.

Trade Secret: Often left unpublished, information is not open to the public or non-entity members in any form. It is the most restrictive, private type of licensing. In software, no source code is available.

Open-source licensing is seen as a method to resolve the dispute between unauthorized copying and distribution of content as well as the copyright laws that have been created to protect individual and businesses of their own work. [6] Instead of users finding ways to circumvent barriers of access to content, the license serves as a pathway on how to freely access and modify software to the needs of users while protecting the proprietary interests of the originator. In this context, discussion surrounding the term of “free” applies here as well as in open-source software. Many software systems are considered to be open-source and utilize an open-source license but are not free in cost to the licensee.

Open-source licensing agreements give access to software that may have been prevented under copyright law. These types of licenses enable use regardless of intent and the ability to modify, adapt, and extend software with minimal restrictions. The license establishes rules of the road and codifies best practices to enable adoptability and ease of use for both the initial user and the broader downstream community. For example, many licenses require modifications to be disclosed and thoroughly documented for third-party individuals to avoid endorsements and confusion regarding what has been published directly from organizations themselves.

In practice, many licenses are not considered open-source due to certain conditions like distribution, access, and modification being prohibited. It is important to note that the quality of software and source code is not inherently affected by the structure of the license. A balance, however, is required ensuring that the licenses do not cause unintended consequences surrounding the proper development and evolution of the code by the user. [7]

2.3. Examining License Agreements in Action

For many organizations, it is unrealistic to fully adopt all parts of open-source software and licensing. Organization's structure licenses dependent on the intended usage of software and codes and aim to embody qualities of open-source. For example, an individual may be able to access software and source code for no-cost or at a greatly reduced cost if it is meant for an academic usage. The moment where the user would like to adapt, transmute, and distribute software, they will often need a different license such as a commercial one. Rather than being purely open or closed source, in application, many organizations lie somewhere in the middle where commercialization of parts of their software support the ongoing research, development, and maintenance of itself. Organizations argue that the monetization of their software also allows them to create detailed guidance for implementation as well as more direct support to its customers while also protecting the proprietary source and implementation of its product. This can contribute to unintended downstream effects that require a multistep approach to solving. Although there are understandable reasons for organizations to license some or all of their software, limiting it's use inhibits contributors who bring diverse and valuable perspectives and solutions. Every part of an organization's work in relation to healthcare ultimately has an effect on patient welfare, quality, and care. Ideas that are locked behind licenses cannot be part of the larger solution to interoperability. Advocacy towards open-source, permissive licenses is one aspect to solutioning.

Users are often left to determine their obligations to different licenses, especially when wanting to combine related software. How do researchers, laboratorians, and health providers navigate the intentional vague complexity of licenses when attempting to solve interoperability? Instead, they experience restrictions and lock-in from standards that act as a hindrance to solving patient problems. Two related but distinguishable implication of licenses are [8]:

1. the explicit terms and rights maintained on content derived from software, and

2. the incompatibility of licensing in competing and related standards.

The current siloed approach of licensing models prevents interoperability within Health IT and results in lower quality data with unattended, downstream patient effects.

Real-world Example: The Linux Kernel and Android Operating System

The Linux Kernel General Public License (GPL) is a real-world example of an open-source licensing agreement that can be referenced and learned from as we continue our IKM efforts. In the Linux Kernel contribution model, contributors need to complete the following process to submit and merge code:

1. **Patch Development**– External contributors create a new patch branch to author and finalize their code. The patch is then developed and details the differences between the main branch and the newly authored branch.

2. **Review**– Once the new code is finalized in its separate branch, it is submitted into a standardized and well-documented code review process. The Module Maintainer Team then conducts an initial, manual review of the code.

3. **Testing**– If the Module Maintainer Team deems that the patch passes the initial review process, it is entered into the Continuous Integration/Continuous Delivery (CI/CD) system that performs acceptance, integration, functionality, scaling, and performance tests.

4. **Merging**– If the patch passes all tests, the Module Maintainer Team notifies the contributor that their patch has been entered into the Merge Process queue before being merged into the main branch. However, if a patch fails a test, the Module Maintainer Team notifies the contributor that their patch has been rejected or needs revisions before it can be resubmitted.

Some defining characteristics of the Linux Kernel contribution model include:

- **Mailing List-Based Review**– Patch submissions are submitted to a centralized Module Maintainer Team then initiates the review process. This supports transparent discussion and visibility among the developer and contributor community but provides a less structured process that can be confusing for newer contributors.

- **Maintainer-Centric Approach**– The Module Maintainer Team have an incredibly important role reviewing and approving all submitted branches. While this ensures a standardized and consistent process, it heavily relies on the maintainers.

- **Scalability**– The Linux Kernel contribution model remains relatively unchanged as the size of the project increases, supporting effective scaling and continual content development.

The Linux Kernel GPL allows any stakeholder to use, modify, or distribute its code freely. Android is a large stakeholder that participates in the Linux Kernel GPL and uses Linux Kernel as the base and foundation for the Android operating system (OS). Android then employs its own, more permissive, open-source licensing agreement through Apache License Version 2.0 as part of the Android Open-Source Project (AOSP). The AOSP licensing agreement allows manufacturers to build custom versions of the operative system, which facilitates the use of the Android OS in healthcare technology. The prevalence of the Android OS, or a modified version through the AOSP, throughout healthcare highlights the impact that licensing has on the development, adoption, and implementation of code, software, and products in the healthcare industry and the importance of having a well-curated licensing agreement. [9-10]

Case-Study: LOINC® and SNOMED International Licensing Agreements

Logical Observation Identifiers Names and Code® (LOINC®) and Systemized Nomenclature of Medicine – Clinical Terms® (SNOMED CT®) are two of the most widely used health data standards in the U.S. for clinical terminology. LOINC® is overseen by the Regenstrief Institute and utilizes a global team for maintaining, developing, and updating guidance on its terminology with the vision of promoting open terminology standards across every clinical information system [11]. LOINC® advocates for open terminology standards but does not issue an open-source license. Instead, all users are required to hold a license to utilize LOINC®, with some licenses requiring a fee to implement and utilize the full functionality of LOINC®. Similarly, SNOMED CT® is overseen by SNOMED International with the intent to support, maintain, and implement SNOMED CT®. All users are required to accept SNOMED International's licensing agreement through an end-user agreement and are issued a proprietary license with some requiring a fee depending on the usage and intent of the license. Both organizations issue guidance on new and existing codes to reflect current events and needs at least biannually by engaging in a rigorous feedback function where users can engage. SNOMED International is moving towards more timely standard updates, not limited to the biannual timeline, especially for rapid developing arenas, such as COVID-19.

Where Licenses Break Interoperability: Redistribution and Modification

The Regenstrief institute highlights guiding principles such as openness and agility in improving and widening LOINC®'s implementation. Regenstrief does not allow the distribution of modified codes under their license. Often, health systems will adapt an existing code to better represent their localized environment where there is no one-for-one match. Users must create their own solution limited to the licensing agreement, engage in a feedback process with the Regenstrief Institute, and wait until a new code is issued directly from Regenstrief to be able to share between systems.

SNOMED International does not allow for derivative work or extensions to be created if it has not already been issued a Namespace Identifier. Modifications are disallowed under their SNOMED CT® Affiliate license and must go through a feedback process where users can petition for added codes in a future release cycle. This is to avoid confusion and ambiguity of SNOMED International codes and prevent against inconsistencies between systems. [12] As a result, there will be many forms of similarly derived codes in a localized format.

This waiting cycle creates gaps that directly impact the quality and interoperability of data. While organizations petition to have new codes created, they must continue to standardize their codes before their data is transferred to another system. During this step, data can be lost during the transfer from a local code to a standard LOINC® or SNOMED CT® code. Doctors may lose the much-needed context surrounding patient care, leading to redundancy in additional ordered tests and referrals. Moreover, the issuance of new codes does not circumvent the usage of localized knowledge. New codes can be used in place of local knowledge, but it can also be further adapted into knowledge for the specific needs of practitioners. This does not solve the transfer errors that can occur when local codes are standardized.

2.4. Putting a More Open Healthcare Ecosystem Into Practice

The IKM community has been longing to address proper standardization encodement of healthcare information across ever changing system. Many avenues of collaboration would help solve the data quality and interoperability issue that the healthcare IT ecosystem is facing. The meaningful semantic difference within terminologies requires preservation to prevent against varying clinical interpretations and treatment. An **IKM approach** can act as a method to standardize clinical statements and terminologies by providing abstraction and elimination of unnecessary complexity. Instead of a local code being incorrectly mapped to a standard code, the IKM platform would provide guidance on the most applicable codes and reduce data loss by ensuring a consistent, reproducible format allowing for the same test to be represented the same way, every time. Proper achievement requires collaboration across government regulatory bodies, public health laboratories, private healthcare systems, and In-Vitro-Diagnostic (IVD) device manufacturers. This also extends to the bodies that oversee terminology standards and the licensing agreements governing them.

For successful implementation of an integrated knowledge ecosystem, licensing bodies should implement components of open-source licensing, specifically on modification and distribution of copyrighted standards. Currently, the process that exists on how licenses work together in common spaces and the subsequent requirements are ambiguous. Would the implementation of a third-party software trigger licensing restrictions and inhibit the willingness of organizations to adopt them? Is the usage of an IKM platform by an organization to reconcile local codes considered to be an extension of a modified code, disallowed under the existing licensing structure? Some licenses consider any work to be referential to be a “derivative” without fully specifying the scope of the reference. This is seen, for example, in the form of the GNU's Not Unix (GNU) General Public License (GPL) and in references to GPL licensed libraries. [13] The navigation of differing licenses and determining the varying responsibilities imposed by them can be a costly, time-consuming burden that often acts as a barrier for meaningful progress. It can quickly become a convoluted process that inhibits success of implementation, unintended by the licensor and their vision of interoperability as seen in LOINC®. Licenses must walk a tight line on preventing extension of standards, thereby slowing down the evolution of these standards, and allowing full extensions that can subvert the standard in and of itself, creating confusion where modifications look to be endorsements by the issuing body. [5]

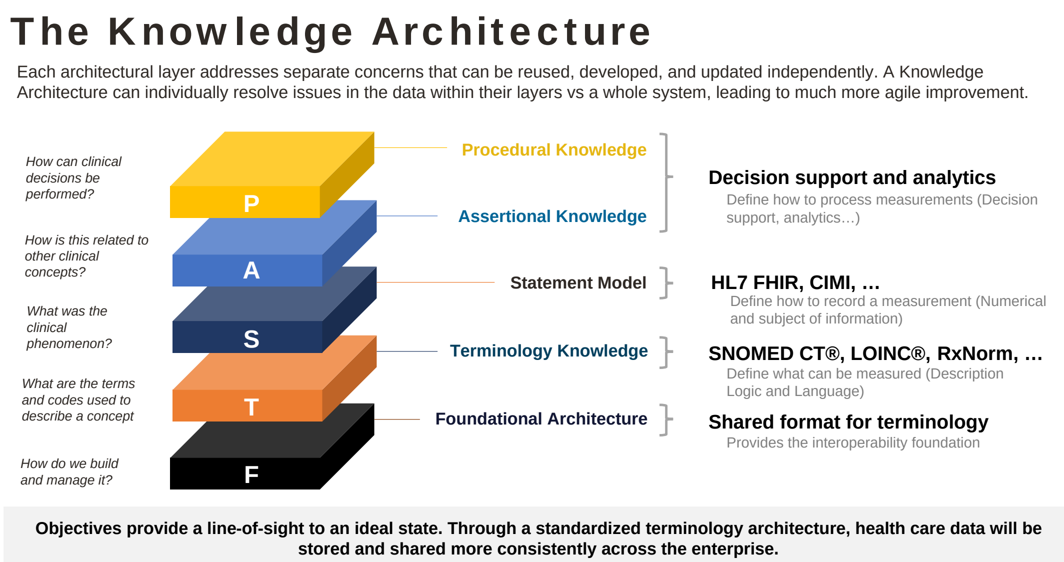
In practice, strong open-source licenses can encourage the development and evolution of standards without subversion of standards through guardrails on community contributions. Instead of a few chosen contributors, solutions can be developed openly with the collaboration of multiple stakeholders where solutioning becomes more robust and agile to the various needs of its users. Reasonable guidelines, such as extensive documentation procedures especially in extensions, enhances the trust within open-source solutions and allow for the benefits of open-source without allowing an "anything goes" mentality where confusion ensues. Additionally, open-source licenses can continue to protect against impersonations and false endorsements through their agreements.

Open-Source Licensing Within a Knowledge Architecture

A Knowledge Architecture is a framework for clinical information that is organized into distinct layers such that each higher layer relies upon artifacts from the lower layer. It aims to define a standardized form of clinical statements and harmonize existing terminologies together into a single system.

A knowledge architecture intended for use in the current Health IT ecosystem is listed below:

Figure 2.2. Layers Within Knowledge Architecture



Rather than a broad transformation of existing licenses in a purely open-source model, not all components of software and standard licensing require a pivot away from proprietary oversight to lead to successful implementation of an IKM ecosystem. The commercial usage of aspects of licenses allows for ongoing support and maintenance of software and standards. The move towards a permissive license does not require a seismic shift in organization structure or its ability monetize its commercial license, especially for organizations already promoting open-standards. *Allowing for open-licensing encourages improvements of terminology and standards by ancillary bodies and creates a more robust, flexible system that can further adapt into the varying needs of organizations.*

Data harmonization improves when foundational layers of Knowledge Architecture are integrated into common, collective sources of data. If utilized correctly under a Knowledge Architecture, a single-data journey can be completed with integrity intact every time. Permissive licensing throughout the foundational architecture, terminology knowledge, and statement model layers removes barriers imposed on modifications and extensions and instead aids in the creation of common understandings of knowledge. The Terminology Knowledge layer within Knowledge Architecture oversees the structure of medical terminologies, such as the language and semantic hierarchy within data. It is in this layer where the IKM platform arbitrates the valid codes and expressions to be used in higher layers. Currently, the usage of terminology standards like SNOMED CT®, LOINC®, and additional standards pose questions regarding compatibility due to modification and extensions limitations. The complexity widens when further standards based on other standards are integrated further in the terminology layer, like the Laboratory Interoperability Data Repository (LIDR) and the TermINology Knowledge Architecture (Tinkar). The loss of data often occurs in this layer and breaks the functionality of standards where local codes are mapped to the incorrect standard codes. The IKM platform aims to solve this issue by harmonizing standards but cannot be successfully implemented unless the underlying licenses allow an IKM platform to reconcile variances. In the Statement Model, the artifacts as defined below are reused and demonstrate how the data elements should be packaged into clinical statements through clinical formatting, such as Health Level 7 Fast Healthcare Interoperability Resources (HL7 FHIR) or Clinical Informational Modeling Initiative (CIMI). The shift towards a permissive license in regard to terminologies and standards improves the compatibility of how different terminologies work together and thereby is integral to the functionality of higher layers and interoperability overall.

2.5. Conclusion

Achievement of interoperability within healthcare systems requires a multi-faceted approach through the extensive collaboration of various stakeholders involved. In its current form many licenses serve as unintended barriers to interoperability where downstream patient harm can occur. The increasing interconnectedness of industries and partners requires a reevaluation of existing licensing structures that impose barriers often antithetical to goals set out by issuing organizations themselves. For solutions like an Integrated Knowledge Platform to succeed, licensing must be structured to support and encourage growth of ancillary bodies. The adoption of open-source licensing encourages diversity of thought where stronger software often results to the benefit of organizations. Advocacy towards open-source licensing must also occur internally by managing teams. Together, common solutioning reduces siloed company operations, moving one step closer to harmonization across the Health IT landscape.

2.6. Appendix

	LOINC®	*SNOMED CT®*
What is the type of license utilized?	LOINC® utilizes proprietary licenses throughout. It is not considered to be open-source, especially under the commercial agreement. Notably, all users are required to hold a license to utilize LOINC®.	SNOMED CT® is not considered open-source and instead issues proprietary licenses. SNOMED CT® has a variety of licenses dependent on the usage and intent, with some requiring a fee. SNOMED CT® requires all users to have a license.
Who owns and maintains the standard?	Regenstrief Institute oversees LOINC® and utilizes a global team to contribute to LOINC®. They oversee and are responsible for the maintenance of LOINC®, as well as issuing up-to-date guidance.	SNOMED International oversees SNOMED CT® and is supported in part by the International Health Terminology Standards Development Organization (IHTSDO) to maintain and implement SNOMED CT®. Additional licensing for non-IHTSDO countries is available.
What is the cost associated with usage?	LOINC® is not completely free to use and is specified by its intent of the license holder. Nonprofit,	SNOMED International, which includes SNOMED CT®, has fees associated with its licensee

	academic, and individual licenses are available under a free license; however, a commercial license requires a fee to implement. There may be third-party costs associated with implementing and supporting LOINC® external to the licensing.	depending on the usage by the holder. For non-commercial use, such as by academic bodies, there is no fee for SNOMED International licensing, but are limited on redistribution. For commercial use, however, like adding into existing healthcare systems for usage, then a fee is incurred for the license. There are multiple licenses to best support the varying needs of users. There are also fees associated with non-IHTSDO member countries.
What does redistribution look like?	LOINC® allows for distribution of their original code without restriction as long as their original license is included in it. Limitations exist for modified codes and cannot be redistributed without the consent of LOINC® copyright holders. If the distribution is intended for commercial usage, a commercial license may be required with associated fees.	SNOMED International redistribution is only allowed under a commercial license as part of their license and cannot be done as a standalone extension. Additionally, the licensee must still protect the proprietary requirements as listed in their license to distribute. Redistribution cannot be done in non-commercial licenses. Modifications of SNOMED International codes are allowed under certain distribution licenses but must maintain compatibility with the original SNOMED CT® as well as be clearly documented and labeled.
How often do releases occur? Who contributes to what is updated?	Releases occur twice a year, typically every 6 months. LOINC® utilizes a feedback function where users can request new codes as well as gathering feedback from stakeholders, regardless of license status. They additionally test these codes and update guidance on existing codes to reflect current events (ex. COVID-19).	Releases occur in January and July every year and contain updated guidance for their codes to reflect current events, similar to LOINC®.

2.7. References

1. Apathy, N. C., Holmgren, A. J., and Adler-Milstein, J. (2021, September). A decade post-HITECH: Critical access hospitals have electronic health records but struggle to keep up with other advanced functions. Journal of the American Medical Informatics Association, 1947-19554. Retrieved from https://doi.org/10.1093/jamia/ocab102

2. Bernstam, E. V., Warner, J. L., Krauss, J. C., Ambinder, E., Rubinstein, W. S., Komastsoulis, G., and Miller, R. S. (2022, May). Quantitating and assessing interoperability between electronic health records. Journal of the American Medical Informatics Association, 29(5), 753-760. doi:https://doi.org/10.1093/jamia/ocab289

3. Free Software Foundation. (n.d.). What is Free Software? Retrieved from GNU Operating System: https://www.gnu.org/philosophy/free-sw.en.html#copyleft

4. Wolford B. What is GDPR, the EU's new data protection law? [Internet]. GDPR.eu. 2018. Available from: https://gdpr.eu/what-is-gdpr/#:~:text=The%20General%20Data%20Protection%20Regulation%20%28GDPR%29%20is%20the

5. GDPR and healthcare: Understanding health data and consent [Internet]. www.pega.com. 2018. Available from: https://www.pega.com/insights/articles/gdpr-and-healthcare-understanding-health-data-and-consent

6. Wilbanks, J. (2008). Public domain, copyright licenses and the freedom to integrate science. Journal of Science Communication, 7(2). doi:https://doi.org/10.22323/2.07020304

7. Reynolds, C. J., and Wyatt, J. C. (2011, February). Open Source, Open Standards, and Health Care Information Systems. Journal of Medical Internet Research. doi:https://doi.org/10.2196/jmir.1521

8. Hosking, R., and Gahegan, M. (2013). The Effects of Licensing on Open Data: Computing a Measure of Health for Our Scholarly Record. International Semantic Web Conference, 432-439. doi:https://doi.org/10.1007/978-3-642-41338-4_28

9. Wikiwand - Android (operating system) [Internet]. Wikiwand. [cited 2024 Jan 17]. Available from: https://www.wikiwand.com/en/Android_(operating_system)

10. Licenses [Internet]. Android Open Source Project. Available from: https://source.android.com/docs/setup/about/licenses

11. Mission, Vision, and Principles for Open Terminology Development. (n.d.). Retrieved from LOINC®: https://loinc.org/principles/

12. SNOMED CT Affiliate License Agreement. (n.d.). Retrieved from SNOMED CT®: https://www.snomed.org/_files/ugd/900274_689013e9e0c74d23892abe9caee02612.pdf

13. Scheibner, J. (2017). What price freedom (of software)? A guide for Australian legal practitioners on open source licensing. Retrieved from Australian Lawyers Alliance: http://classic.austlii.edu.au/au/journals/PrecedentAULA/2017/23.html#fnB3

Part II. Ecosystem Engagement

Table of Contents

3. Ecosystem

3.1. Purpose of a Retrospective and Industry Evaluation

In today's healthcare ecosystem, there is a need for high-quality and accurate data as healthcare organizations becoming increasingly reliant on EHRs, laboratory information systems (LISs), and other electronic information systems. Terminology standards are a primary tool that many organizations use to work towards this goal and improve the quality and accuracy of their data. Terminology standards aim to ensure information is captured and represented in a repeatable and standardized format so that other organizations and stakeholders using said terminology standard have a common understanding of the data. While terminology standards are useful, extensive knowledge management (KM) is needed to identify, manage, and share the standardized data that these systems and standards capture. However, as data flows within and between healthcare systems, the use of different terminology standards, KM processes, and systems themselves can pose a threat to the integrity of data.

IKM is an approach to data management that aims to maintain the meaning, integrity, and quality of data as it travels within and between healthcare systems. IKM is an evolving approach to manage disparate terminology standards using a common, centralized, and repeatable representation of data. The use of a common model supports the interoperability of data between disparate terminology standards, well documented version control even when different terminology standards employ their own versioning practices, and the incorporation of new or emerging terminology standards. IKM will use an iterative approach to build upon previous work on Resource Description Frameworks (RDFs) and property graphs with novel ideas and concepts. There is no reference or best practice guide for how an IKM solution can or should operate and will therefore require extensive research, prototyping, and development.

While IKM may pose as a daunting task for many organizations due to the large barrier to entry and resource commitment, our team is not the only stakeholder working towards IKM solutions. SNOMED CT® and LOINC® have announced a collaborative partnership to increase the interoperability of their data and standards by facilitating the identification of synonymous concepts and reducing duplication in the standards. Our work prototyping and trialing an IKM reference implementation has given us insight into various challenges, work arounds, and best practices. Our work has identified functional and non-functional areas of IKM that require further exploration and refinement to support more widespread adoption, which can be shared with other organizations and stakeholders to shape their work and support a unified approach.

3.2. Summary of Findings – Functional and Non-Functional Issues Related to Knowledge Management Proof of Concept

Over the past year, our team has been working to develop an IKM reference implementation, Komet, to demonstrate IKM capabilities. Our work with Komet builds upon data property graph and RDF foundations and incorporates new concepts like versioned concept-oriented property graph representation. Due to the relatively novel nature of Komet and IKM, our team did not have guidelines or best practices to follow. Our initial work serves as a prototype that identifies functional and non-functional areas related to IKM that should be further explored. Throughout our experience, we have documented the following lessons learned, best practices, and challenges to improve the adoptability, functionality, and development of IKM as our team and other stakeholders work towards future iterations.

3.2.1. Knowledge Management

The team rapidly developed an IKM prototype leveraging the latest Tinkar data model standard to showcase IKM capabilities at the SNOMED CT® conference in October of 2023. Through the effort over the past year and leading up to that event, we've identified a few findings that have been resolved or will need to be resolved in the future.

3.2.1.1. Requirements

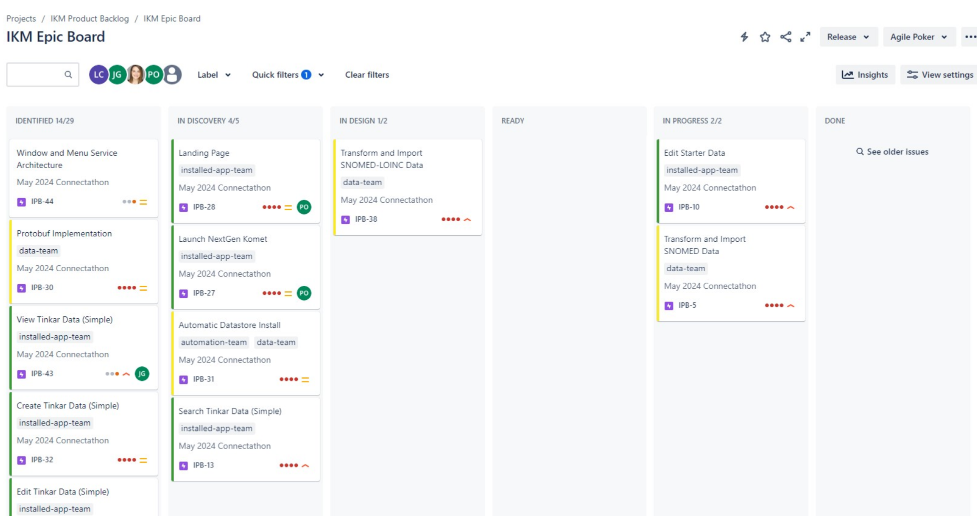
As each clinical terminology standard has its own proprietary data representation and tooling to both view and author data, the ability to represent multiple standards with a common underlying data model and view or author across standards in a single environment is an unprecedented idea. IKM would influence a large community of users, such as hospitals, laboratories and clinical informaticists so there are many thoughts and opinions about how IKM should work and what features would be important. Though the potential stakeholder base is large, the access to the community to illicit requirements is very limited. While we were able to conduct user interviews during our User Interface and User Experience (UI/UX) analysis to gain some understanding of necessary system functionalities, we do have not regular access to those participants.

To solve for this, we're leveraging experts to identify high priority capabilities that will be prioritized for upcoming milestones. These milestones will take shape in the form of conferences, community events, and other demonstration opportunities. For example, we plan to organize a Health Level Seven (HL7) Connect-a-thon event to conduct real-world testing scenarios with interested stakeholders.

Leading up to the SNOMED CT® conference, we found that while we wanted to support rapid development, we still needed a clear understanding of what system capabilities were important to showcase for the demonstration and detailed documentation in the form of user stories for the developers to reference.

As seen in the figure below, we established a new process to describe the system capabilities in the form of Agile epics and plot them against the appropriate milestones to generate our product feature backlog. As per agile development best practices, these epics are revisited weekly to be defined, refined, and prioritized. These epics are then decomposed into user stories with defined acceptance criteria for the development teams to develop against in three week increments or sprints. [1]

Figure 3.1. IKM Product Backlog Epic Board



As the IKM community grows, we will leverage the open-source nature of the project to track feature requests through submitted GitHub issues. [2] This external issue tracker will be integrated more fully in the future but will allow us to track feature requests against other priorities and communicated releases back to the IKM community. [3]

3.2.1.2. User Interface and User Experience (UI/UX) Design

In early 2023, a UI/UX assessment for IKM was conducted to better understand the needs and pain points of healthcare professionals in their daily operations working with healthcare knowledge. The goal was to use findings from the assessment to determine user profiles, requirements, and priority, which were then used to strategize a human centered UI/UX design for a Next Gen IKM platform that promotes data interoperability.

The assessment was conducted in five phases:

1. A landscape analysis to evaluate the competitive landscape of similar products.
2. A heuristic evaluation of the old Komet version to identify features to keep in the Next Gen version.
3. User interviews with multiple professionals in the healthcare space—such as clinicians, health informaticists, public health scientist—to uncover specific needs and pain points of potential users.
4. A UI/UX workshop to discuss findings from interviews and align on future UI/UX strategy.
5. Creation and prioritization of user stories for future sprints based on workshop results.

Some of the challenges faced by the UI/UX team include representing different knowledge standards with different requirements for viewing/editing in a standardized way, abstracting complexities behind-the-scenes for an easier user experience and designing in a way that supports iterative development while still meeting future requirements.

The team plans to tackle these challenges by identifying commonalities and differences between user interfaces and data schema of existing knowledge standards. The aim of these planned activities is to better understand how to design a standard UI and conduct periodic user evaluation sessions for voluntary user feedback within the healthcare industry. Healthcare stakeholders can help the UI community explore the latest design and share insightful information in the user experience. They can also influence the process of building modular and re-usable UI components that allow flexible and adaptable development in an agile environment with changing requirements.

3.2.1.3. Development and Operations (DevOps) Processes

In the current state of Health IT, healthcare data standard organizations have implemented restrictive licensing around their proprietary data and technology which creates silos and the inability to collaborate across standards. IKM provides an open-source solution that could be used across the community to collaborate and further reduce these data silos. Komet is an open-source solution in the form of an installed Java application that provides users the ability to visualize, manage and extend disparate medical terminologies. [4]

Early in the development phase, we identified that we would need to make an investment in Development and Operations (DevOps) technologies and tooling to allow our open-source solutions to be rapidly developed and released to reach the largest audiences possible. We began by hosting our DevOps tools, as it was the quickest way to begin rapid prototyping. However, through this process, we've found that the hosting of these tools requires more overhead and manhours than we would like. Our larger goals and priorities require developers' time to contribute to our solutions rather than supporting our DevOps tools. We are currently investigating and analyzing different alternative approaches to move to Software-as-a-Service (SaaS) solutions for DevOps.

Additionally, due to the licensing issues around proprietary data, as well as other open-source technologies, we leverage both an internal and external code repository. This arrangement allows us to continue development while we work through licensing with the goal of becoming fully open-source. [5] Currently, this complication keeps us from leveraging open-source supported tooling and requires us to support complex DevOps processes to restrict the publishing of code. Once these license issues are resolved, we'll be able to leverage additional tools to streamline and simplify our processes.

However, our DevOps processes are largely successful as we're able to collaboratively develop across a large team and release both code and artifacts to our public repository often and with little manual effort. Additionally, every promotion to GitHub releases Windows, Mac and Linux versions of the Komet installer to be available to the community for use.

3.2.1.4. Data Layer Development

As new knowledge standards are transformed and introduced to the Tinkar Core, the more the need for ease of code troubleshooting and management becomes apparent. [6] Reinforcing the code to also handle different types of data and types of programming languages will also keep the data ingestion and export free of conflicts. Additionally, adding in a layer of abstraction is needed to allow all types of users to set up and use Komet. For large scale development and automation to be possible these improvements to code and architecture will be crucial.

3.2.1.5. Bindings

Currently, Tinkar Core uses Tinkar Terms to call certain Java variables that represent retrievable concepts from the Tinkar datastore. As the codebase develops and changes, these can be difficult to maintain, especially when more complexity is added. Each Tinkar Term affected must be manually identified and changed if needed. Not only is it time consuming, but it can be susceptible to human error. By creating an automated solution like a bindings library, updates will be made efficiently and can be integrated into a dependency management system.

3.2.1.6. Protocol Buffers

With the requirement of being able to support many types of information systems, the ingestion of data requires a programming language neutral format. Protocol buffers solve this issue, by serializing the structure data with a format that supports many programming languages. Additionally, it is open-source code and does not require a licensing to implement. This would allow the data to be exported in easy to share formats that can be simply shared and imported.

3.2.1.7. Data Retrieval and Curation

As the Komet users increase, adoption of the tool will need to accommodate a variety of experienced users. Currently, the Tinkar database requires a significant amount of coding and or prior knowledge of lower-level data structures to successfully retrieve and manipulate Tinkar data elements. This inevitably limits the immediate number of users who can implement meaningful solutions using Komet and the Tinkar data. Creating an abstraction layer and or user experience that allows users to have an easier and streamlined approach working with and updating Tinkar data will result in users utilizing the Komet to its full potential.

3.2.2. Solution Adoption

To date, the IKM Solution Adoption team has worked to identify key stakeholders and tailor content to their understanding of the goals of IKM. In this phase of increasing awareness, the team has developed and disseminated informational presentations and trainings both internally and externally to federal and industry partners. Through this and additional presentations and demos of Komet at industry conferences (i.e., SNOMED International Expo, American Medical Informatics Association), the team has been able to identify champions for the IKM movement across the ecosystem. As awareness has increased, the team has been able to maintain stakeholder relationships through personalized communications to individual stakeholders at organizations of significance (i.e. Food and Drug Administration, Office of the National Coordinator for Health Information Technology, Graphite Health, Royal College of Pathologists of Australasia). [8] Through such engagements, the team has recognized the following key findings of developing successful communications that will propel the efforts up the communications curve trajectory.

Through such engagements, the team has recognized the following key findings of developing successful communications that will propel the efforts up the communications curve trajectory.

3.2.2.1. Curating Technical Communications

There is a need in IKM communications to provide adequate technical guidance for stakeholders interested in collaboration and/or integration. Therefore, the development of proper processes and documentation of work has been critical to the IKM strategic communications efforts and will continue to be strengthened as we move towards driving engagement.

3.2.2.2. Defining IKM

IKM is an evolving and novel movement in the healthcare information technology field. Therefore, communicating to all potential stakeholders effectively on the mission to increase cohesion in health data standards and what problems that it may solve is paramount to advancing the movement. Based on feedback from stakeholder engagement sessions, the team has been iteratively refining strategic communications initiatives to develop concise and aimed messaging to key stakeholders.

3.2.2.3. Ensuring End User Adoption

A more nuanced challenge that the IKM movement faces as technologies and integration methods develop into end-user capabilities is the adoption and implementation of these tools by the end-users themselves, whether they are clinicians, healthcare organizations, laboratories, or other healthcare-related entities. Many of these end-users may not have the technical background or insight into the issues that IKM addresses and aims to solve and therefore may not fully be invested in adopting such a solution into their operations.

As future efforts in the Integrated Knowledge Management space are advanced, there will be an increased need for intentional strategies centered around communications to the various stakeholders that IKM impacts to ensure engagement and ultimate adoption. Technical professionals in the data standards field will need to have transparent and detailed information on the methodologies used in previous efforts in order to continuously improve upon the collaborative process. Further, end-user specific communication plans will be developed that articulate both the problems that IKM attempts to solve along with a clear explanation about the tools' capabilities.

3.2.3. Knowledge Representation

The advantages of utilizing a generalized, self-describing data structure are well-documented. This data structure comprises not just the database but also includes metadata that outlines and explains both the data and the connections between tables [15]. Also, descriptions of data structures and constraints are included, allowing storage of a description of its own structure within the database. While this self-describing approach is reference terminology-agnostic, making it capable of accommodating various data sources, challenges for determining the importance of terminology-specific data components are introduced [16].

Representing data properties such as concepts, patterns, or semantics becomes a complex task, especially when importing data from multiple terminology standards. As clinical reference terminologies, SNOMED CT® and LOINC® employ different data structures to represent their respective content. These data structures do not translate easily between terminologies. SNOMED CT® utilizes Concepts, Descriptions, and Relationships, organized hierarchically to convey clinical meanings to its terms. LOINC® codes are represented by five or six main parts. And hierarchical relationship between LOINC® terms is manually applied outside of the LOINC® SNOMED CT® Collaboration. The components from each of these terminologies may function as crucial data structure components for axioms and related data properties within the database. [9]

Once imported into a self-describing data structure, determining equivalence with another terminology standard or which parts of the data components are important may create unanticipated challenges [15]. Unlike more specific terminology models, Tinkar lacks pre-defined references to identify crucial components or properties within incoming data, necessitating a comprehensive understanding of the dataset. This challenge extends to determining the importance or equivalence of specific data components, prompting the creation of new properties and concepts before the data can be seamlessly imported into Tinkar. While the generic nature of Tinkar provides valuable flexibility for accommodating diverse data standards, it underscores the need for a proactive and nuanced approach to customization. This adaptability allows Tinkar to serve as a versatile platform capable of handling a variety of data sources, but it also highlights the importance of thoughtful customization to ensure effective integration. [6]

3.3. Evaluation of Current Systemized Nomenclature of Medicine – Clinical Terms® (SNOMED CT®) Logical Observation Identifiers Names and Code® (LOINC®) Collaboration

In October 2022, Regenstrief Institute (RII) and SNOMED International, also known as International Health Terminology Standards Development Organization (IHTSDO), signed a cooperative agreement to collaborate the work between LOINC® and SNOMED CT®. This will eliminate duplicative data, allowing the two terminologies to be used together by creating an extension of the SNOMED CT® with the LOINC® knowledge standard included. Using the SNOMED CT® formats and concepts, the LOINC® knowledge standard will be more defined and accessible. [10]

The SNOMED LOINC® Collaborative Agreement indicates a pivotal progression towards data interoperability as the two organizations are industry leaders in health data standards and terminology. [11] We conducted an analysis of the current state of the collaboration project to better understand the successes and pain points as well as assess how IKM may provide additional solutions.

3.3.1. Knowledge Management: SNOMED LOINC® Collaborative Agreement Demonstration Project Progress Reports

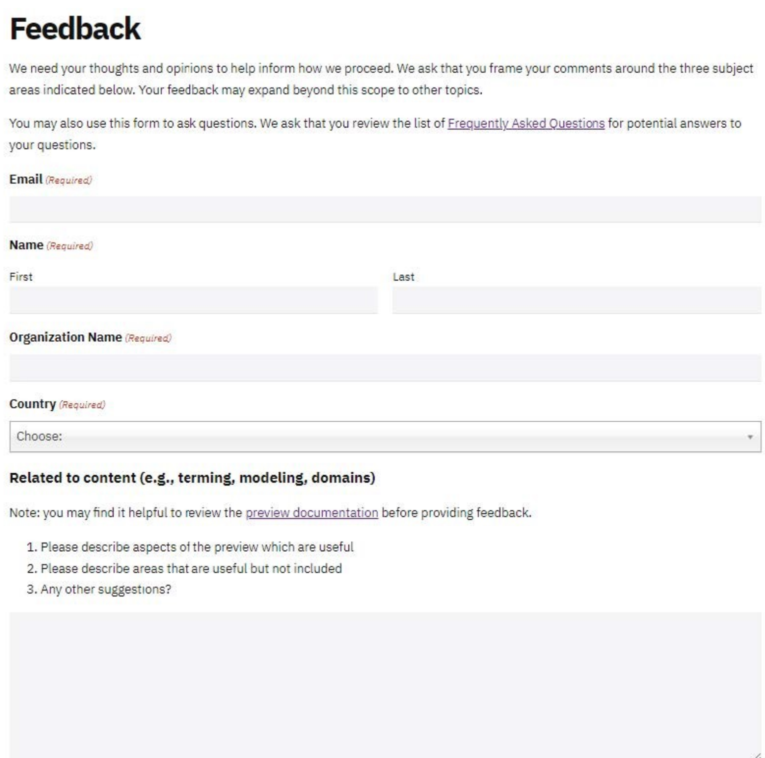
The SNOMED LOINC® Collaborative Agreement is a critical first step forward towards data interoperability and a testament towards the collective need for collaborative solutions. As described on the website, the LOINC® Ontology can be accessed in two ways: through a full data download of the LOINC® Ontology file or viewable through the LOINC® Ontology SNOMED CT® Browser. The browser provides the same capabilities of the SNOMED CT® browser but provides visibility to LOINC® data through the appropriate search or query. The website, while brief, provides additional information to the community such as release notes, frequently asked questions, and the ability to provide feedback and sign up to receive project updates. [11]

Regarding Knowledge Management, our analysis will be focused on the systems and services provided by the LOINC® and SNOMED CT® Interoperability Solution as described on their website. [12] The analysis of the user interface / user experience will be discussed in section **UI/UX: SNOMED LOINC® Collaborative Agreement Demonstration Project Progress Reports** and analysis of the data representation implications are expanded on in section**Knowledge Representation: SNOMED LOINC® Collaborative Agreement Demonstration Project Progress Reports**. Further details regarding the communications to the community is discussed in section **Ecosystem Engagement: SNOMED LOINC® Collaborative Agreement Demonstration Project Progress Reports** .

3.3.1.1. Hearing from the Community

The LOINC® and SNOMED CT® Interoperability Solution encourages the community to provide input to help inform how the solution will proceed. [11] Responders can provide their perspective related to content, releases and use case. Additionally, responders can leverage the feedback form to ask questions.

Figure 3.2. Screenshot of Feedback Page on loincsnomed.org



Feedback

We need your thoughts and opinions to help inform how we proceed. We ask that you frame your comments around the three subject areas indicated below. Your feedback may expand beyond this scope to other topics.

You may also use this form to ask questions. We ask that you review the list of Frequently Asked Questions for potential answers to your questions.

Email (Required)

Name (Required)

First

Last

Organization Name (Required)

Country (Required)

Choose:

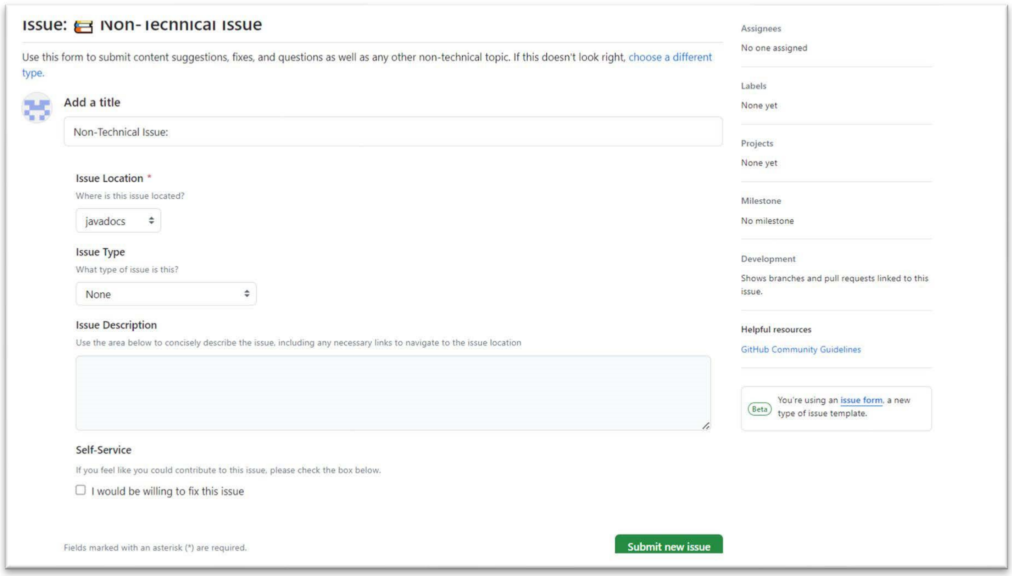
Related to content (e.g., terming, modeling, domains)

Note: you may find it helpful to review the preview documentation before providing feedback.

1. Please describe aspects of the preview which are useful
2. Please describe areas that are useful but not included
3. Any other suggestions?

Community involvement is a key to ensure the solutions that are being made available meet the needs of the users. Similarly, IKM provides the ability for users to provide feedback through GitHub issues. GitHub provides custom form layouts to give users the best experience for submitting technical and non-technical issues or questions. [12]

Figure 3.3. Screenshot of GitHub Non-Technical Issue Submission Page

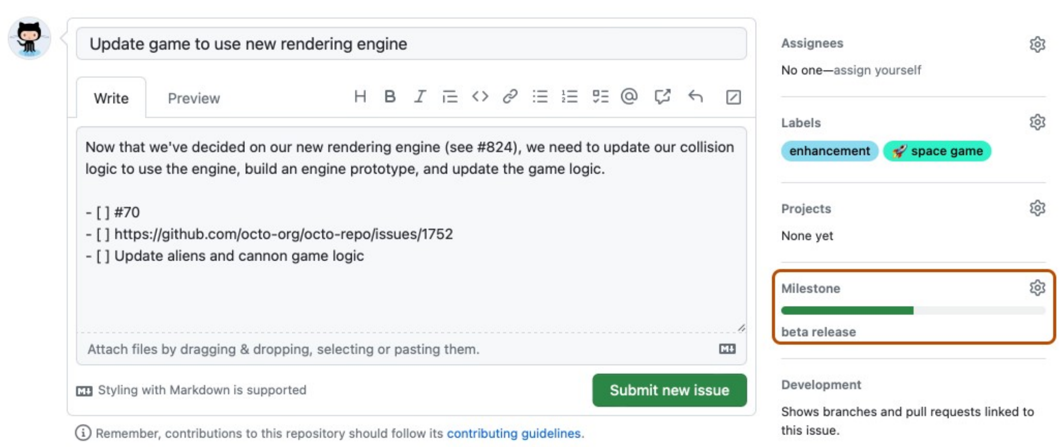


3.3.1.2. Complications in Community Visibility

Though the LOINC® and SNOMED CT® Interoperability Solution collects community feedback and issues, there is no visibility into that information. Community members have no insight into what ideas and questions have already been submitted. Furthermore, there is no visibility into what feedback will be acted upon or what questions have been answered. According to GitHub, a leader in open-source community best practices, keeping communication public is key so that all members have access to the same information. [11]

IKM provides the ability for all issues submitted to be visible. The team can provide transparent updates, such as commenting to answer questions, assigning the issue for work, or assigning the issue to a milestone for release. This will allow the issue creator and the rest of the community to have full insight into when the issue may be addressed.

Figure 3.4. Screenshot of GitHub Issue Screen



3.3.2. UI/UX: SNOMED LOINC® Collaborative Agreement Demonstration Project Progress Reports

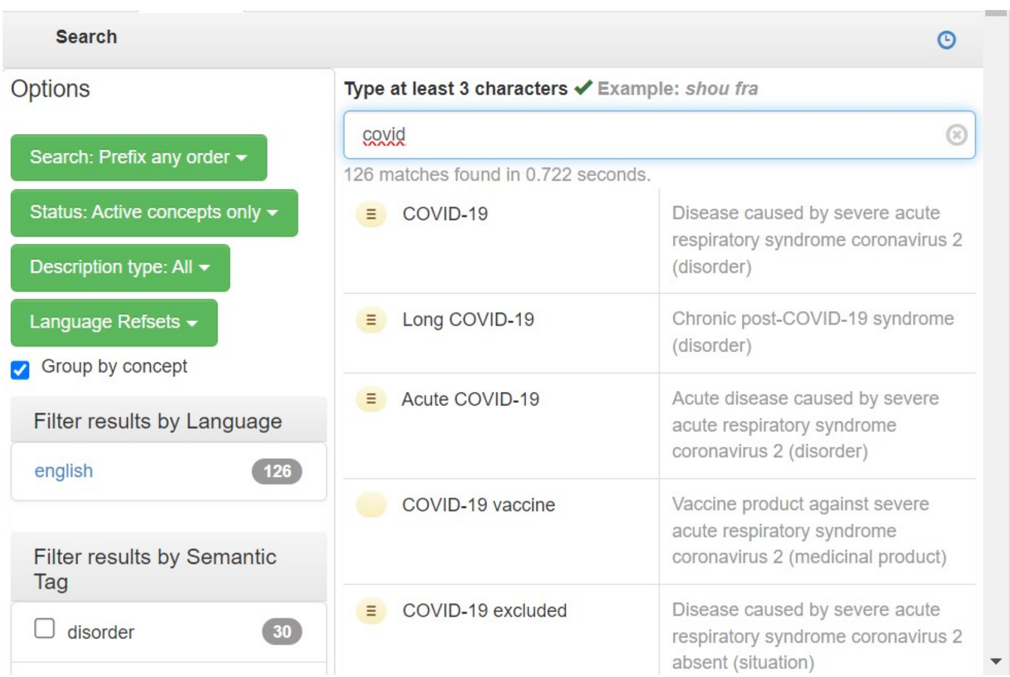
Through their collaborative agreement, LOINC® and SNOMED CT® have developed the LOINC® Ontology SNOMED CT® Browser, which allows users to navigate concepts, view concept details, and determine equivalency between LOINC® and SNOMED CT® concepts. [10] This is an important development for the medical terminology field because it represents a move toward achieving interoperability. The collaboration data set and supporting browser tool, while useful, is a work in progress and has some user challenges in its current state.

3.3.2.1. Utilizing Search in the LOINC® Ontology SNOMED CT® Browser

From a user perspective, navigating through the LOINC® Ontology SNOMED CT® browser to locate and view concepts is done in two primary ways: searching by typing or navigating the taxonomy. [10] Using the built-in search engine, the user can find results with as few as three letters. The search results automatically update as the user types, so there is no need to hit “Enter”, just edit the input field and wait for the results to refresh.

In addition to the search bar, the search option also includes a list of filters to narrow search results as shown on the left side of the image below. Prior to entering text in the search bar, only the four green filter options and the “Group by concept” checkbox appear. After entering text, additional filters appear, which are filtering by: language (e.g., English), semantic tag (e.g., observable event), module (SNOMED CT® core vs. LOINC® Extension module), and refset (e.g., Web Ontology Language [OWL] axiom reference set). [13]

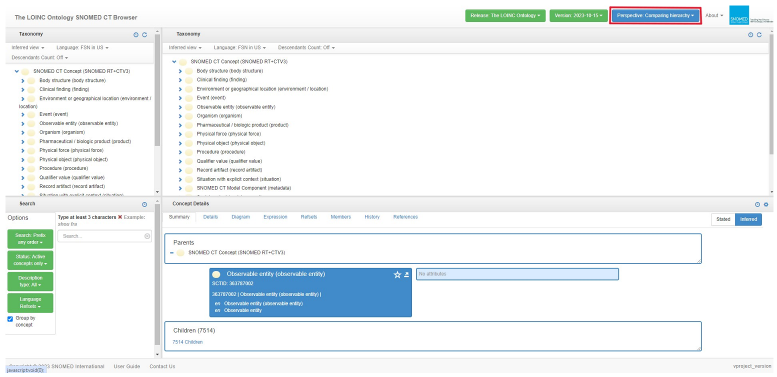
Figure 3.5. LOINC® Ontology SNOMED CT® Browser Search Filters



The user can also navigate through concepts using the taxonomy structure in the LOINC® Ontology SNOMED CT® Browser. In the default "Perspective”, the taxonomy structure lists all SNOMED CT®

observable entity concepts in alphabetical order. If the user changes the “Perspective” to “Comparing Hierarchy”, the system displays a more robust taxonomy of SNOMED CT® concepts, as shown below. The user can then navigate the hierarchy to find concepts. The user cannot currently view and navigate the equivalent LOINC® hierarchy.

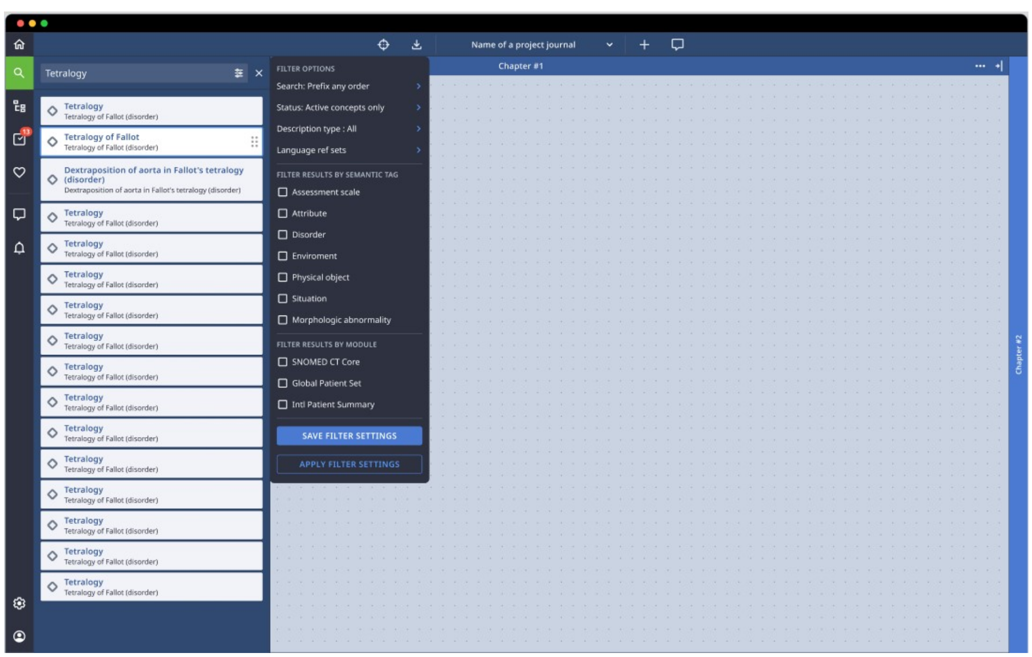
Figure 3.6. Taxonomy Structure in the LOINC® Ontology SNOMED CT® Browser



In the current state, the user experience in searching and navigating the LOINC® Ontology SNOMED CT® Browser can be improved by improving the default UI settings, allowing the user to edit the views to fit their needs, and adding the ability to search and navigate within LOINC® data specifically. In the default UI, some features are hidden from the user and may require time and effort to uncover them. Surfacing functional user settings will enhance user experience.

Noted improvements can be made to search to enhance the function and experience. Improving the search user interface, making search and navigation more intuitive (e.g., filters, taxonomy), and optimizing page space use will enhance the tool’s usability. The image below demonstrates what an optimized search function might look like for an Integrated Knowledge Management tool. The view below is simplified and clean but simultaneously more functional and user-friendly. [14]

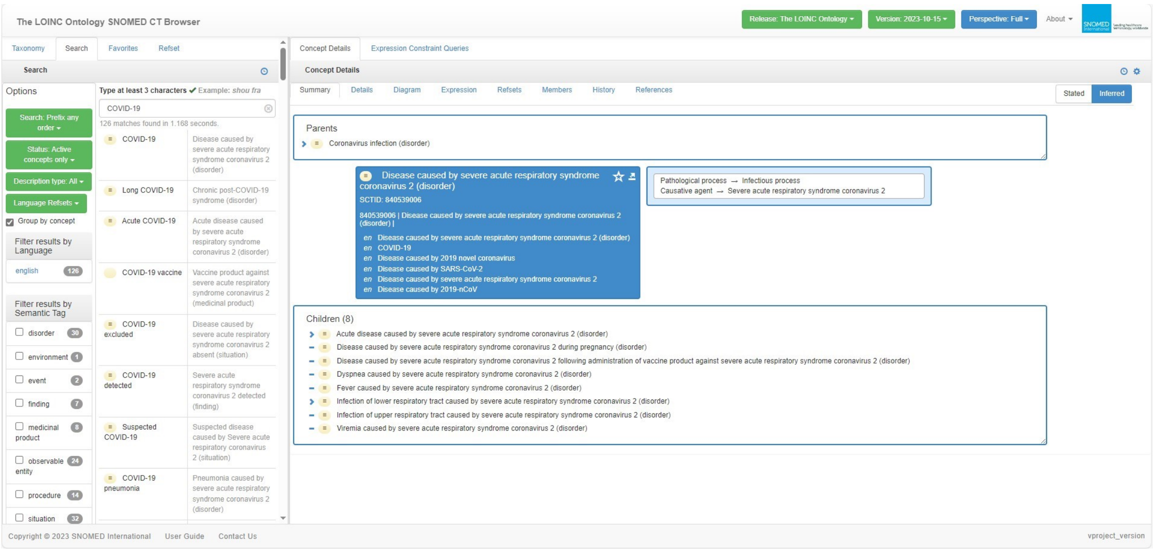
Figure 3.7. Hypothetical Optimized Search Function



3.3.2.2. Viewing a Concept and its details in the LOINC® Ontology SNOMED CT® Browser

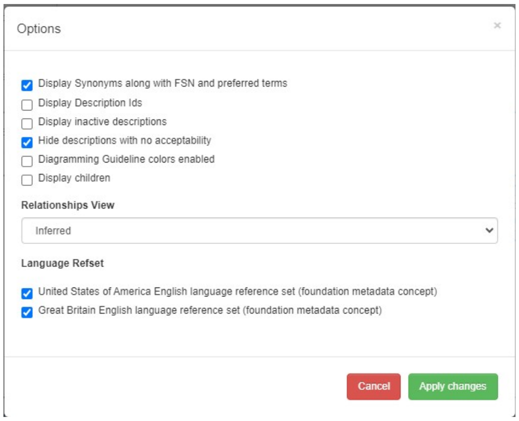
Upon selecting a concept to view within the LOINC® Ontology SNOMED CT® browser, the Concept Details tab on the right side of the screen displays information pertaining to that concept. The initial view is the Summary tab that displays parent/child information within the SNOMED CT® hierarchy as well as LOINC® equivalents (if applicable). The other tabs, Details, Diagram, Expression, Refsets, Members, History, and References provide further information on the concept. [15]

Figure 3.8. View of the Summary Tab



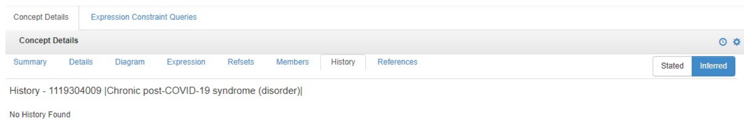
In addition, there is a settings button in the upper right corner of the Concept Details window that allows the user to change some settings to their preference. This type of customizable functionality is needed for a tool that will be so widely used across fields. [15]

Figure 3.9. Concept Details Window



These features provide the user a functional view of concepts in a straightforward manner, but there stands room for improvement. For example, most tabs of a given concept do not present useful or novel information, if they have information at all. Notably, after searching through dozens of concepts, this team was unable to locate one that had information in the History tab. Because the team was unable to locate a concept with a populated history, it was not clear whether this functionality was not yet available or simply empty. It would benefit the user to communicate more clearly which concepts are missing information. For future implementations, developers of similar tools can look to prioritize communication with the user, so that it is clear to the user what they are seeing (and not seeing). [13]

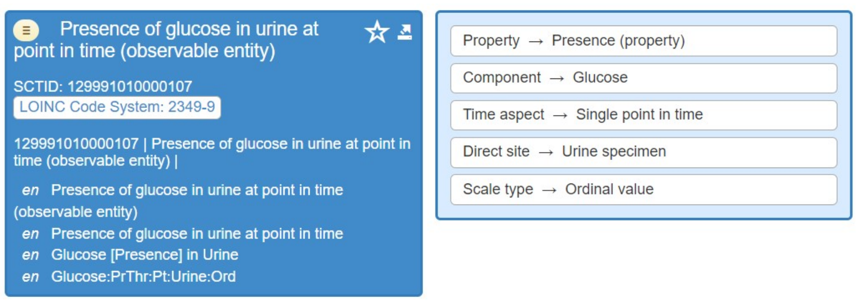
Figure 3.10. History Tab of Concept Details



3.3.2.3. Utilizing the LOINC® Ontology SNOMED CT® Browser for Purpose and Contributing

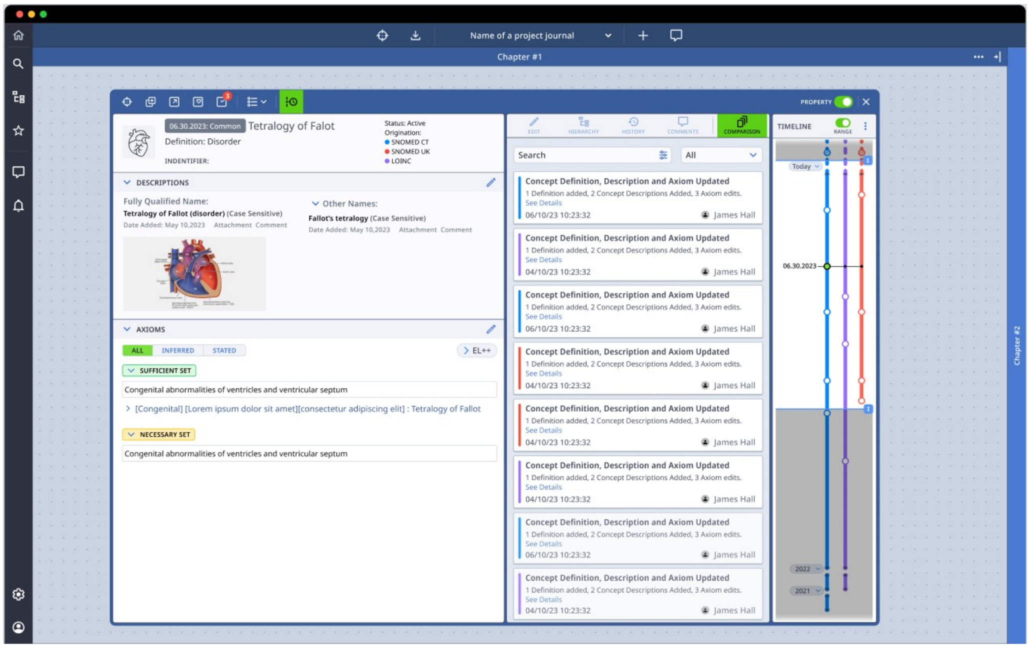
The LOINC® Ontology SNOMED CT® browser, a collaborative effort between LOINC® and SNOMED CT®, seemingly relies disproportionately on the SNOMED CT® dataset. In fact, finding equivalency between concepts may be best executed if the user already knows the existing LOINC® concept and needs to find the equivalent SNOMED CT® concept. Trials in finding LOINC®-equivalents using SNOMED® concepts often proved fruitless as many SNOMED® concepts did not produce a LOINC® equivalent. Upon finding a concept that has a LOINC® equivalent, the indication is rather small and does not provide a great deal of information, as shown below. [13]

Figure 3.11. LOINC® Ontology SNOMED CT® Browser Indication of Equivalence



To enhance the existing system, there should be a more complete record for concepts, including the ability to view how they have changed over time. An example of how this could be done in an integrated knowledge management system is shown below.

Figure 3.12. Example of Hypothetical Concepts Versioning Record

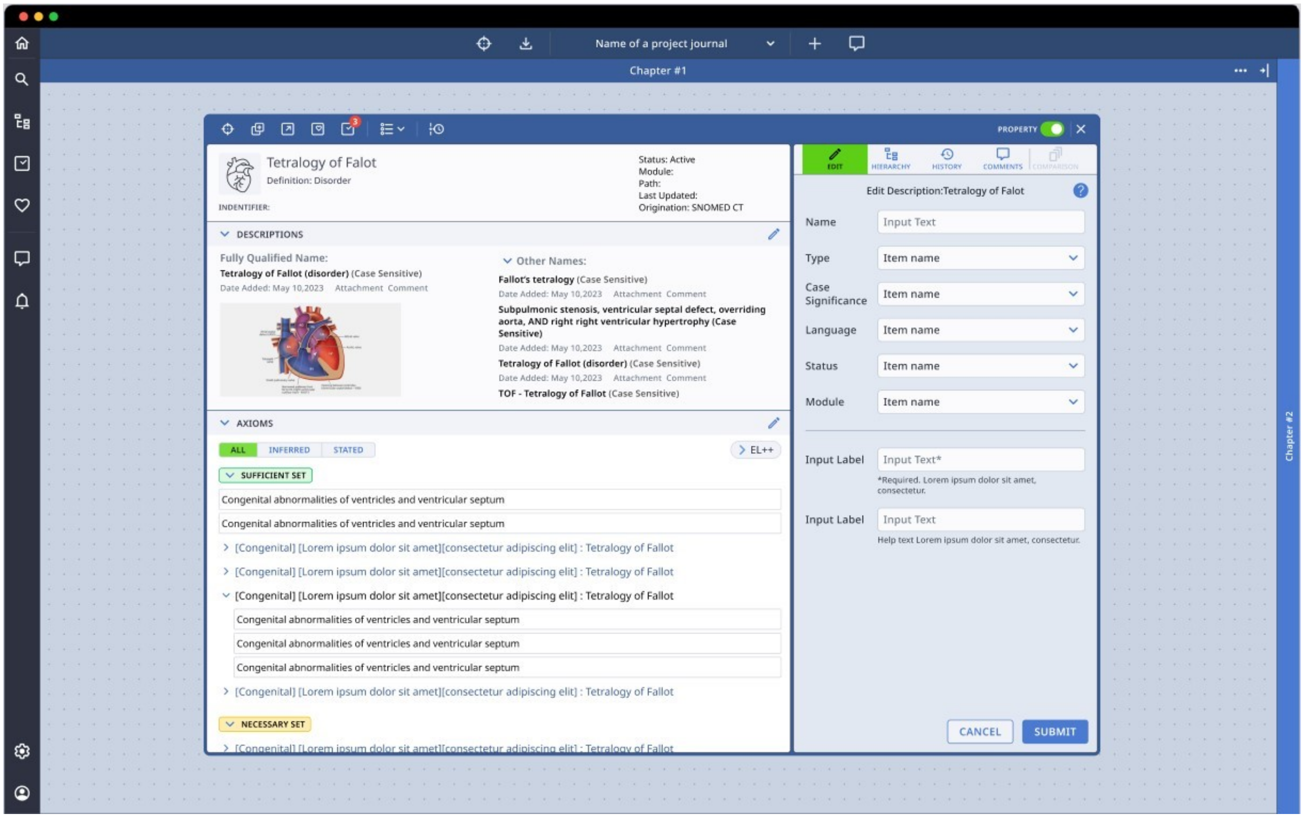


While the current system relies heavily on the SNOMED CT® hierarchy, the system must look ahead to not favor one terminology over another and allow the user to search all terminologies. LOINC® terminologies should have a larger presence within the tool, so that it is easier to navigate concepts and determine equivalence without needing significant background knowledge. [9]

Using and navigating the LOINC® Ontology SNOMED CT® browser requires the creation of an account. This account, however, does not allow the user to make contributions to the platform, only view data. Editing the tool and terminology requires access to a separate system, which is a barrier to potentially valuable contributions. Specialized tools like this one are better off when they capture the expertise of the individual user through open-source collaboration rather than restricting editing rights.

The image below captures what collaboration within an IKM tool might look like. Allowing users to edit and contribute increases the knowledge base and the momentum at which the tool itself improves.

Figure 3.13. Hypothetical IKM Tooling Collaboration



The LOINC® Ontology SNOMED CT® browser is a useful tool that can be improved by focusing on the end user to improve UI/UX. Improvements and enhancements to the tool's user experience will be a key component to its long-term viability.

3.3.3. SNOMED LOINC® Collaborative Agreement Demonstration Project Progress Reports

While the success of the technical integration is critical, of equivalent significance is the organization's ability to communicate about its efforts to enable awareness and adoption of the newly integrated model. Evaluating the partnership's solution adoption efforts against the communications strategy of the IKM movement can provide indicators of success.

3.3.3.1. Achievements in Technical Communications

Similar to the need for technical communications, the ontology browser homepage links to the documentation release, providing an introductory technical explanation of the processes utilized to link the two data standards. The release includes details on the content inclusions of the browser, modeling guidelines used to integrate standards, concept/attribute representation, and future steps that the technical partnership will embark on to improve the browser. This document is critical for the public to access the information regarding the collaborative processes and technical efforts made on the project and provides a background for non-technical individuals to understand its development and technical individuals to see where their own collaboration or implementation may align.

3.3.3.2. Challenges of Integrated Knowledge Communications

Regenstrief LOINC® and SNOMED International collaborated on a singular virtual location for the LOINC®-SNOMED CT® Ontology Browser to be housed and available to the public. This site has the entirety of their current progress integrating the two data standards up to their most recent version release (October 15, 2023). While the site provides some introductory background into the objectives of the partnership between the two organizations, there is little communication of what the overarching goal of that partnership is aimed to accomplish. Integrating knowledge requires an active community effort to advance healthcare information technology, and the challenge of engaging champions for this effort can be mitigated by increasing the community's understanding and support. Garnering the community's trust is vital in progressing the efforts towards an integrated knowledge management solution. [11]

3.3.3.3. Ensuring End User Adoption

As communication strategies are developed, crafting specific messaging for end user segmentations is critical to promoting adoption of tools and IKM advancements. The LOINC®-SNOMED CT® collaboration will benefit from a similar approach in identifying end user groups and then building compelling communications to each grouping based on key interests and pain points. The collaborative partnership's overall end user adoption strategy will be a duality of technical information that supports the operational collaboration and the non-technical communications positioning that engages users to learn and apply tools. The technical guidance will need to provide detailed support for technical experts to involve themselves effectively within their efforts and may include user guides, collaboration forums, and proper feedback channels. The non-technical communications will need to cater towards conveying benefits of the browser and what problems this collaboration addresses for a given end user segmentation.

3.3.4. Knowledge Representation: SNOMED LOINC® Collaborative Agreement Demonstration Project Progress Reports

Mapping LOINC® terms to SNOMED CT®, specifically for the laboratory component of LOINC® required extensive discussion among stakeholders and an extension of the SNOMED CT® concept model to accommodate LOINC® Axis properties. The table below shows the SNOMED CT® attribute crosswalk to the LOINC® Axis.

LOINC® Axis	SNOMED CT® Attribute	Note
Component	246093002 | Component (attribute) OR 704319004 | Inheres in (attribute)	The Property of the LOINC® Term determines if the LOINC® Component is modeled with 246093002 | Component (attribute) or 704319004 | Inheres in (attribute), e.g., a term with property of mass concentration is modeled using 246093002 | Component (attribute) whereas a term representing an inherent part of component such as type is modeled using 704319004 | Inheres in (attribute). -When the LOINC® Component includes an adjustment and/or a count,
	704326004 | Precondition (attribute) (included when the LOINC® Component includes a challenge)	

		this additional information is included in the terming of the LOINC® concept but not by an attribute in the model.
Property	370130000 | Property (attribute)	
Scale	370132008 | Scale type (attribute)	
System	704327008 |Direct site (attribute) OR 704319004 |Inheres in (attribute)	704319004|In heres in (attribute) Is used when the LOINC® Term has a property type of Prid, Type, Number of cells, ID
Time	370134009 | Time aspect (attribute)	
Method	246501002 | Technique (attribute)	

Release notes acknowledge some terms defined by LOINC® Parts require additional discussion before finalizing the modeling and terming. An example of this issue is LOINC® terms defined by LOINC® Part Antibiotic XXX (‘other antibiotic’). In addition, inclusion of the Great Britain (GB) English refset for LOINC® terms is still pending. [11]

LOINC® contains properties that define a Category Status of “Order Only,” “Observation Only”, or “Both Order and Observation”. SNOMED CT® hierarchies define whether such concepts are an order (Procedure hierarchy) or an observation (Observable Entity hierarchy). The SNOMED CT® equivalent of the LOINC® “Both Order and Observation” Category Status was not found. [11]

A goal is for the majority of the concepts created in the LOINC® Extension of SNOMED CT® is to be designated as sufficiently defined. Where this is not possible, the concept is designated as primitive. Primitive concepts within the SNOMED CT® terminology and its extensions interfere with use of a classification process where the formally stated definitions of each concept compute the subsumption hierarchies and defining properties of each concept. Some of the concepts will remain as primitive in future releases and some will change to fully defined as the work progresses. Some concepts are currently designated as primitive as a map from the needed LOINC® Part to SNOMED CT® concept does not yet exist. However, this is expected to be resolved in future releases. [11]

Across many domains, the meaning of a particular observation can be best understood in the context of the set of possible answers (result values). For example, qualitative laboratory tests often have fixed answer lists for expected lab values, as seen in the screenshot below. Answers are not linked to Observable Entity concepts in SNOMED CT®. [14]

Figure 3.14. Example of Qualitative Laboratory Tests with Fixed Answer Lists

91065-3

Long Common Name

14-3-3 protein [Presence] in Tissue by NAA with probe detection

Excerpt from LoincAnswerListLink.csv

LoincNumber	LongCommonName	AnswerListId	AnswerListName	AnswerList LinkType	Applicable Context
91065-3	14-3-3 protein [Presence] in Tissue by NAA with probe detection	LL1937-3	Present|Absent	EXAMPLE	

Excerpt from AnswerList.csv

AnswerListId	AnswerListName	Ext Defined YN	AnswerStringId	Sequence Number	DisplayText	ExtCodeId	ExtCodeDisplayName	ExtCodeSystem
LL1937-3	Present|Absent	N	LA9633-4	1	Present	52101004	Present (qualifier value)	http://snomed.info/sct
LL1937-3	Present|Absent	N	LA9634-2	2	Absent	2667000	Absent (qualifier value)	http://snomed.info/sct

AnswerListOID: 1.3.6.1.4.1.12009.10.1.1112
ExtCodeSystemVersion: http://snomed.info/sct/900000000000207008/version/20170731
ExtCodeSystemCopyrightNotice: This material includes SNOMED Clinical Terms® (SNOMED CT®) which is used by permission of the International Health Terminology Standards Development Organisation (IHTSDO) under license. All rights reserved. SNOMED CT® was originally created by The College of American Pathologists. "SNOMED" and "SNOMED CT" are registered trademarks of the IHTSDO. This material includes content from the US Edition to SNOMED CT, which is developed and maintained by the U.S. National Library of Medicine and is available to authorized UMLS Metathesaurus Licensees from the UTS Downloads site at https://uts.nlm.nih.gov. Use of SNOMED CT content is subject to the terms and conditions set forth in the SNOMED CT Affiliate License Agreement. It is the responsibility of those implementing this product to ensure they are appropriately licensed and for more information on the license, including how to register as an Affiliate Licensee, please refer to http://www.snomed.org/snomed-ct/get-snomed-ct or info@snomed.org. This may incur a fee in SNOMED International non-Member countries.

A Clinical LOINC® crosswalk for standardized assessment instruments and document ontology has not been addressed yet in the LOINC® SNOMED CT® Collaboration.

3.4. Conclusion

This deliverable serves as an initial report to document the challenges our team faced during the research and initial development of an IKM solution and environment. Our summary of findings for various components of the SNOMED CT® LOINC® collaborative agreement that aims to develop an IKM environment and the identification of functional and non-functional issues our team experienced with IKM have identified critical areas that will require additional and continued efforts to improve the implementation, functionality, and adoption of IKM solutions.

While this deliverable will inform the direction and focus areas for our future work, this report can also function as a resource and reference implementation for SNOMED CT®, LOINC®, and other stakeholders working to develop IKM solutions. As our team works to develop an open-source environment for IKM, we look forward to and encourage collaboration with other stakeholders to gain a deeper understanding of the challenges they faced, best practices, and IKM as a whole.

3.5. References

1. IKM Epic Board - Agile board - Jira [Internet]. [cited 2024 Jan 8]. Available from: https://ikmdev.atlassian.net/jira/software/c/projects/IPB/boards/40

2. A library for manipulating with proofs that use inference rules. [Internet]. GitHub. [cited 2023 September 15]. Available from: https://github.com/liveontologies/puliz

3. Bug and Issue Reporting Process Framework. [Deloitte Deliverable]. December 20, 2023 [Cited 2024 Jan 9] Available Internally at Deloitte Consulting LLP. DeloitteBAA_Task_2.3.4_Bugs and Issue Reporting Process Framework.pdf

4. The Integrated Knowledge Management Application | Komet [Internet]. [cited 2024 Jan 19]. Available from: https://www.ikm.dev/application

5. Reynolds CJ, Wyatt JC. Open source, open standards, and health care information systems. J Med Internet Res. 2011 Feb 17; 13(1):e24.

6. HL7 International [Internet]. 2021. HL7 Logical Model: Standardized Terminology Knowledgebase, Release 1. Available from:https://www.hl7.org/documentcenter/private/standards/HL7_LM_TERM_KB_R1_INFORM_2021AUG_FINAL.pdf

7. Systemic Harmonization and Interoperability Enhancement for Laboratory Data (SHIELD) | FDA [Internet]. [cited 2024 Jan 9]. Available from: https://www.fda.gov/medical-devices/diagnostic-data-program/systemic-harmonization-and-interoperability-enhancement-laboratory-data-shield

8. SHIELD - Standardization of Lab Data to Enhance Patient-Centered Outcomes Research and Value-Based Care | ASPE [Internet]. [cited 2024 Jan 9]. Available from: https://aspe.hhs.gov/shield-standardization-lab-data-enhance-patient-centered-outcomes-research-value-based-care

9. LOINC — The freely available standard for identifying health measurements, observations, and documents. [Internet]. Loinc.org. 2019. Available from: https://loinc.org/

10. LOINC and SNOMED CT: LOINC SNOMED CT Extension - Why, What and How | LinkedIn [Internet]. [cited 2023 Jan 8]. Available from: https://www.linkedin.com/pulse/loinc-snomed-ct-extension-why-what-how-termlex/?trk=pulse-article

11. Regenstrief and SNOMED International broaden collaboration to facilitate interoperability nationally and globally by linking LOINC and SNOMED CT [Internet]. [cited 2024 Jan 9]. Available from: https://www.snomed.org/news/regenstrief-and-snomed-international-broaden-collaboration-to-facilitate-interoperability-nationally-and-globally-by-linking-loinc-and-snomed-ct

12. Changing the layout of a view - GitHub Docs [Internet]. [cited 2024 Jan 9]. Available from: https://docs.github.com/en/issues/planning-and-tracking-with-projects/customizing-views-in-your-project/changing-the-layout-of-a-view

13. SNOMED CT - Observable entity (observable entity) [Internet]. [cited 2024 Jan 10]. Available from: https://browser.loincsnomed.org/?perspective=full&conceptId1=363787002&edition=MAIN/LOINC/2023-10-15&release=&languages=en

14. Answer File – LOINC [Internet]. [cited 2024 Jan 10]. Available from: https://loinc.org/answer-file/

15. 1.3: Chapter 3 Characteristics and Benefits of a Database - Engineering LibreTexts [Internet]. [cited 2024 Jan 9]. Available from: https://eng.libretexts.org/Bookshelves/Computer_Science/Databases_and_Data_Structures/Database_Design_(Watt)/01%3A_Chapters/1.03%3A_Chapter_3_Characteristics_and_Benefits_of_a_Database

16. Abad-Navarro F, Martínez-Costa C. A knowledge graph-based data harmonization framework for secondary data reuse. Computer Methods and Programs in Biomedicine [Internet]. 2024 Jan 1 [cited 2024 Jan 9];243:107918. Available from: https://www.sciencedirect.com/science/article/pii/S0169260723005849