Review Article Open Access

Marcia Lei Zeng,* Yi Hong, Julaine Clunis, Shaoyi He, L.P. Coladangelo

Implications of Knowledge Organization Systems for Health Information Exchange and Communication during the COVID-19 Pandemic

https://doi.org/10.2478/dim-2020-0009
received May 21, 2020; accepted May 23, 2020.

Abstract: This article aims to review the important roles of health knowledge organization systems (KOSs) during the COVID-19 pandemic. Different types of knowledge organization systems, including term lists, synonym rings, thesauri, subject heading systems, taxonomies, classification schemes, and ontologies are widely recognized and applied in both modern and traditional information systems. Apart from their usage in the management of data, information, and knowledge, KOSs are seen as valuable components for large information architecture, content management, findability improvement, and many other applications. After introducing the challenges of information overload and semantic conflicts, the article reviews the efforts of major health KOSs, illustrates various health coding schemes, explains their usages and implementations, and reveals their implications for health information exchange and communication during the COVID-19 pandemic. Some general examples of the applications, services, and analysis powered by KOSs are presented at the end. As revealed in this article, they have become even more critical to aid the frontline endeavors to overcome the obstacles due to information overload and semantic conflicts that can occur during devastating historic and worldwide events like the COVID-19 pandemic.

Keywords: knowledge organization systems, health terminologies, health information exchange, COVID-19 pandemic

***Corresponding author: Marcia Lei Zeng,** School of Information, Kent State University, Kent, OH, USA. E-mail: mzeng@kent.edu
Yi Hong, Department of Product Development, DeepThink Health, Inc., Richmond, CA, USA
Julaine Clunis, L.P. Coladangelo, School of Information, Kent State University, Kent, OH, USA
Shaoyi He, Department of Information and Decision Sciences, Jack H. Brown College of Business and Public Administration, California State University, San Bernardino, CA, USA

1 Introduction: Dealing with Information Overload and Semantic Conflicts

During a pandemic period, one common challenge is information overload. “Information overload” is a term popularized by Alvin Toffler in his bestselling book *Future Shock* (1970). It refers to the difficulty a person can have in understanding an issue and making decisions that can be caused by the presence of too much information. During times of crisis such as a global pandemic, users need access to immediate information. This is compounded by the fact that news reports from around the world are constantly being updated, providing users with yet more new information. This information may not be clearly delivered if the language or keywords referenced carry different meanings in different contexts. Moreover, one may also wish to trace the methods or criteria for collecting data used in data-driven messages and decision-making. There may even be doubts about or perceived discrepancies between the communicated messages and the data originally collected and analyzed. The challenge could grow dramatically when a report needs to be communicated across languages, regions, and cultures, especially if statistics were generated based on different contextual information and gathered from diverse resources.

Therefore, while struggling to obtain information on the new developments regarding the pandemic within one’s own community or as they are occurring in other places, one would inevitably face information overload, or even misinformation, that would lead to confusion, uncertainty, worries, fear, stress, or anxiety. When people are in a new and difficult situation, especially in a medical emergency, the availability of information that can be useful would be very crucial for decision making. On one hand, people may be easily overwhelmed by “the virtually unlimited amount of information available, information that is often poorly organized and of highly variable quality and relevance" (Jadad, Haynes, Hunt, & Browman, 2000, p. 262). On the other hand, for some people, such information overload could lead to cyberchondria, an excessive anxiety caused by obsessively searching the Web for medical information, when they are frightened by the wide-spread misinformation of severity and fatality of COVID-19 (Laato, Islam, A., Islam, M., & Whelan, 2020). In a recent paper published in the *Asian Journal of Psychiatry*, Sahoo et al. (2020) reported and discussed two cases in which two people separately attempted self-harming due to their apprehension of developing COVID-19 and were brought to the emergency room. The paper pointed out that such tragedies were the outcomes of information overload that caused normal people to develop severe anxiety and depression leading to self-harming. When discussing the important aspects to be considered for dealing with the impact of COVID-19 pandemic on mental health, Fiorillo and Gorwood (2020) defined information overload as an "infodemic" that was based on their observation of the spreading of a large amount of fake news and/or uncontrolled information about coronaviruses that moved faster than the coronavirus itself. Information overload as an infodemic has been given attention by more and more organizations as well as researchers. For instance, Shaw, Kim & Hua (2020) have discussed the impact of governance, technology and citizen behavior on coping with the COVID-19 infodemic in East Asia. The World Health Organization (WHO) launched WHO Information Network for Epidemics (EPI-WIN), a new information platform to overcome the infodemic around COVID-19 (Zarocostas, 2020).

Another major challenge is called "semantic conflict," which can occur within any data and information communication process. A relevant example is the naming of a new disease, including the reuse of previous names that may share some similarities, the adoption of a known accepted name which may carry different meanings at different times of history, and the inclusion of names of particular groups of people, places, or animals based on the cases reported earliest. During the 2009 outbreak of H1N1 influenza in humans, controversy started early on regarding the usage of various of terms by journalists, academics, and officials, such as "swine flu," "pig flu," "Mexican flu," "North American influenza," etc. This time, before getting its official name, "COVID-19," the disease was called "Wuhan SARS," "Wuhan Flu," and "Wuhan coronavirus." The virus was also referred as "China virus" and "Wuhan virus" repeatedly even after the official name was announced. A search for three of these labels—"Chinese Coronavirus," "China Coronavirus," and "Wuhan Coronavirus"—on Google Scholar in mid-April retrieved more than 1,280 items. Another search of these three names on Google showed millions of related hits. This situation demonstrates that disease-naming-based semantic conflicts are not only an issue of the past, but also an issue of the present, and the future, thus necessitating further investigation. These semantic conflicts undoubtedly added more confusion to an environment already overwhelmingly characterized by heavy information overload.

Furthermore, the challenges of information overload and semantic conflict directly impact the whole domain of healthcare, which deals with the maintenance or improvement of health via the prevention, diagnosis, treatment, recovery, or cure of diseases, illnesses, injuries, and other physical and mental impairments in people. Healthcare extends beyond the delivery of services to patients, encompassing many related sectors, and is set within a larger picture of financing and governance structures (Wikipedia, n.d.). At the frontlines of healthcare, the need to make instantaneous decisions at various levels is crucially based on the available and truthful data and information.

To deal with information overload and eliminate semantic conflicts, controlled, standardized, and shared vocabularies have become critical to information exchange and communication during the COVID-19 pandemic. Under the guidance of international and national health organizations and government agencies, KOSs, such as the *International Classification of Diseases* (ICD) maintained by the WHO, *SNOMED Clinical Terms* (SMOMED CT) of the International Health Terminology Standards Development Organization (IHTSDO), *Logical Observation Identifiers Names and Codes* (LOINC) of the LOINC.org, and *Medical Subject Headings* (MeSH) of the U.S. National Library of Medicine (NLM), responded promptly to the WHO's announcement of an official name for the disease caused by the 2019 novel coronavirus, "COVID-19," as evidenced by their actions discussed in Section 2.1 of this article, followed by the background about the purposes and usages of the major health KOSs in Section 2.2.

Even when a new concept is identified, and the related terminologies or naming conventions are controlled by national and international level institutions, semantic conflicts can still occur through the way concepts are classified and defined. Incorrect diagnoses and cause of death is a well-known problem with international morbidity and mortality statistics (O'Malley et al., 2005). This will be a particular problem with COVID-19 where cause of death is either being over- or under-attributed to COVID. That concern is the rationale behind the use of a classification system, which is considered to be a fundamental approach to organizing knowledge. Because of this issue that many communities, professions, and subject disciplines have developed different ways to classify things and organize their knowledge, the high-level effort to producing a unified classification system such as *International Classification of Diseases* (ICD) is critical, which provides a common language for reporting and monitoring diseases and has been used worldwide for morbidity and mortality statistics. By providing classification notations to represent concepts (refer to Figure 1 as an example), ICD greatly enhances the consistency of coding across languages, cultures, and healthcare systems (refer to Table 1 in Section 2.1.2). In thesauri and other types of KOSs which are not primarily in a hierarchical structure, the broader and narrower hierarchical relationships provide contextual information about a particular concept. Non-hierarchical but related concepts are presented in associated relationships. The contextual information provided by KOSs ensures consistency and helps to eliminate semantic conflicts.

In KOSs that do not use notations to represent concepts, the dedicated record for a concept (which has an assigned unique identifier) presents a standardized preferred lexical label for each of the languages included in the vocabulary. *UNESCO Thesaurus*,[1] and the *AGROVOC* thesaurus[2] of the Food and Agricultural Organization (FAO), both under United Nations (UN), are notable examples (refer to Section 5.4). Each record also provides terms and codes considered non-preferred but can lead to a standardized preferred label (refer to the MeSH record in Figure 2). The clarity provided by listing and linking synonyms becomes important as naming is more problematic early on and for selecting an official name consensually. Synonym rings with alternate names could then lead to the standardized terms, once established. Preferred and non-preferred terms linked in synonym rings have been widely used by search engines and information retrieval systems to provide access to authoritative public information such as news articles and reports that can explain various factors for the causes, effects, risks, and treatments, including recommendations for managing personal decisions regarding a virulent infectious disease, with terms and languages that are readily understood by normal people. KOSs play a role in aiding the provision of such accurate public information.

1 http://vocabularies.unesco.org/thesaurus/concept3505

2 http://aims.fao.org/aos/agrovoc/c_4ad07701

All KOSs' fundamental structures and functional requirements are defined by national and international standards, with the main functions of eliminating ambiguity, controlling synonyms or equivalents, establishing explicit semantic relationships such as hierarchical and associative relationships, and presenting both relationships and properties of concepts in the knowledge models (NISO Z39.19-2005, 2005; Zeng, 2008; ISO 25964-1:2011, 2011; ISO 25964-2:2013, 2013). In addition to providing a common framework and language for healthcare domain practitioners (discussed in Section 2), various types of KOSs provide a structured way to communicate complex concepts to the general public, as evidenced by the selected cases of timely useful information services (refer to Section 4 and Section 5). Similar to the case of COVID-19 (2019 – present), such efforts can be seen during other recent epidemic and pandemic periods including SARS ((Severe Acute Respiratory Syndrome), 2002–2003),[3] 2009 H1N1 Influenza, (2009–2010),[4] MERS ((Middle East Respiratory Syndrome), 2012–present),[5] and Ebola ((Ebola Virus Disease), 2014–2016).[6] In the next section, the selected health KOSs demonstrate semantic structures that lead to trustworthy and effective efforts to eliminate semantic conflicts and make an impact on the global fight against information overload during the COVID-19 pandemic.

2 Standardized Health KOSs and Coding Guidance

2.1 Efforts of the Major Health KOSs during the COVID-19 Pandemic

For the outbreak of a new viral disease, three most important names have to be decided: (1) the disease, (2) the virus, and (3) the species. The WHO is responsible for the first (disease), expert virologists for the second (virus), and the International Committee on Taxonomy of Viruses (ICTV) for the third (species) (International Committee on Taxonomy of Viruses, 2020). Establishing a name for a new disease provides a shared understanding for researchers and developers to discuss disease prevention, spread, transmissibility, severity, and treatment. Viruses are named based on their genetic structure to facilitate the

3 https://www.cdc.gov/sars/

4 https://www.cdc.gov/h1n1flu/

5 https://www.cdc.gov/coronavirus/mers/

6 https://www.cdc.gov/vhf/ebola/

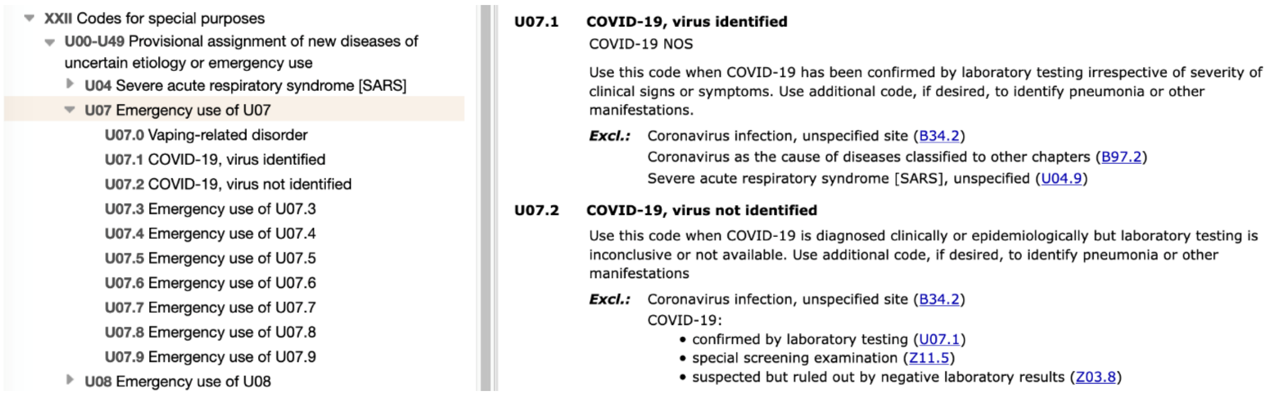


Figure 1. ICD-10 emergency codes for COVID-19.
Image captured 2020-04-26. Source: https://icd.who.int/browse10/2019/en#/U07

development of diagnostic tests, vaccines, and medicines (WHO, 2020a).

The following timeline reflects the recent efforts to eliminate ambiguities and semantic conflicts through naming of the disease led by the WHO and the actions of the *International Classification of Diseases, Tenth Revision* (ICD-10), as documented by the WHO's press conferences videos (WHO, 2020b) and the Centers for Disease Control and Prevention (CDC, 2020). The next sub-section provides a list of new codes and coding guidance from standardized health KOSs, which are used in healthcare workflows, in response to the WHO announcement of an official name for the disease caused by the 2019 novel coronavirus. ICD and other examples of standardized KOS described below allow the world to compare and share data in a consistent and standard way— between institutions, across regions and countries, and over a period of time. They facilitate the collection and storage of data for analysis and evidence-based decision-making. Together they are contributing to the actions of eliminating semantic conflicts and avoiding information overload in real-world healthcare systems.

2.1.1 Naming and Classifying by WHO and ICD-10

2020-01-30. The WHO declared the 2019 Novel Coronavirus (**2019-nCoV**) disease outbreak a public health emergency of international concern.

- The term "2019-nCoV" was used instantly, e.g., in a *Science* Jan. 31 article.[8]

2020-01-31. As a result of the declaration, the WHO Family of International Classifications (WHO-FIC) network's Classification and Statistics Advisory Committee (CSAC) convened an emergency meeting to discuss the creation of a specific code for this new type of coronavirus.

- The ICD-10 established a new emergency code ("**U07.1, 2019-nCoV, acute respiratory disease**"). At that time, the WHO Classification Team had noted that the virus name "2019-nCoV" was temporary and was likely to change (to be independent of date and virus family, and for consistency with international virus taxonomy).

2020-02-11. The WHO officially announced the name of the disease, **COVID-19**, an acronym for "coronavirus disease 2019."

- "Having a name matters to prevent the use of other names that can be inaccurate or stigmatizing," said Director-General of the WHO, Tedros Adhanom Ghebreyesus. "It also gives us a standard format to use for any future coronavirus outbreaks." The WHO referenced guidelines set in 2015 ensure that the name does not refer to a geographical location, an animal, an individual or group of people, while still being pronounceable and related to the disease (WHO, 2015). Public health experts agree with the choice not to name the disease after a geographic region in China. Wendy Parmet, a law professor at Northeastern University and public health expert, told the *TIME* magazine that if the new name had included a reference to Wuhan, it would have caused a "tremendous stigmatization on the people of Wuhan who are the victims" of the disease, (Mansoor, 2020; WHO, 2020b).

7 DOI: 10.1126/science.367.6479.727
8 DOI:10.1126/science.367.6477.492

- The ICD-10 was updated with two emergency codes: "**U07.1 COVID-19, virus identified**" was assigned to a disease diagnosis of COVID-19 confirmed by laboratory testing, and "**U07.2 COVID-19, virus not identified**" was assigned to a clinical or epidemiological diagnosis of COVID-19 where laboratory confirmation would be inconclusive or not available (Figure 1).
- The term "COVID-19" was used immediately worldwide. For example, in an article published in the journal *Science* on Feb. 14, the term used was "COVID-19".[7]
- The same day, a study group of the International Committee on Taxonomy of Viruses (ICTV) christened the novel virus as "severe acute respiratory syndrome coronavirus 2," or **SARS-CoV-2** (ICTV, 2020). ICTV is the official body of the Virology Division of the International Union of Microbiological Societies responsible for naming and classifying viruses. **SARS-CoV-2** then became the official name used by WHO. Meanwhile, WHO also began to refer the virus as "the virus responsible for COVID-19" or "the COVID-19 virus" when communicating with the public (WHO, 2020b).

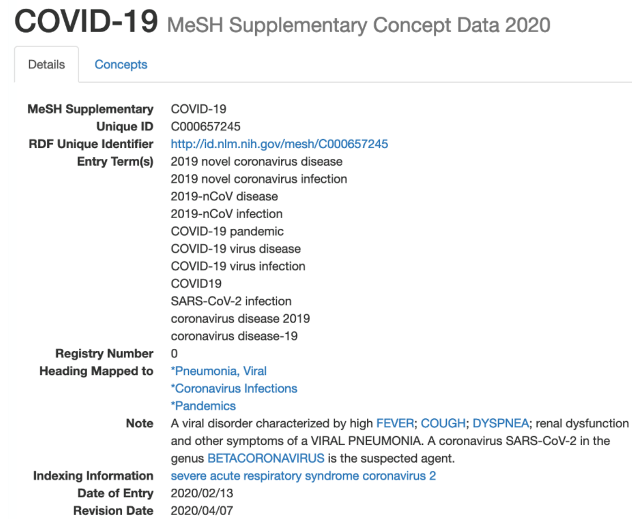
2.1.2 Coding Guidance Provided by the Standardized Health KOSs

In response to the WHO's naming of the disease COVID-19 and the virus SARS-CoV-2, and the declaration of this Novel Coronavirus Disease (COVID-19) as a "pandemic" on March 11 (WHO, 2020b), the standardized health KOSs immediately released their new codes and coding guidance. The coding guidelines have been updated since the first release (Table 1).

2.1.3 Supplementary Concept Record Added by the Medical Subject Headings (MeSH)

In addition to the KOSs listed in Table 1 which are directly used in the management of clinical practices and healthcare delivery systems, the U.S. National Library of Medicine (NLM)'s *Medical Subject Headings* (MeSH) added a Supplementary Concept Record (SCR) for COVID-19 to the 2020 MeSH Browser on February 13, 2020.[9] MeSH is a comprehensive controlled vocabulary for the biomedical life science bibliographical databases (such as MEDLINE/PubMed). It is also used by ClinicalTrials.gov registry to classify the diseases that are registered in ClinicalTrials. In this new Supplementary Concept Record, in addition to the preferred lexical label "COVID-19" and a Unique ID, more than 10 entry terms are provided (see Figure 2) to facilitate subject access to the related information resources. It has mapped to other related MeSH descriptors including "Coronavirus Infections," "Pneumonia, Viral," and "Pandemics." The Resource Description Framework (RDF) Unique Identifier facilitates the creation of an RDF record, downloadable in RDF/XML, RDF/N3, and JSON-LD serialization formats (Figure 2).

Since the MeSH subject headings' Unique IDs have been used as one of the three major controlled identifiers (together with that from ICD-10 and SNOMED CT) in Wikipedia, searching with MeSH ID on the Web brings dozens of results of multilingual Wikipedia entries aligned with this identifier. There will be more discussions later on in Section 5.5.


COVID-19 MeSH Supplementary Concept Data 2020

Details | Concepts

MeSH Supplementary	COVID-19
Unique ID	C000657245
RDF Unique Identifier	http://id.nlm.nih.gov/mesh/C000657245
Entry Term(s)	2019 novel coronavirus disease 2019 novel coronavirus infection 2019-nCoV disease 2019-nCoV infection COVID-19 pandemic COVID-19 virus disease COVID-19 virus infection COVID19 SARS-CoV-2 infection coronavirus disease 2019 coronavirus disease-19
Registry Number	0
Heading Mapped to	*Pneumonia, Viral *Coronavirus Infections *Pandemics
Note	A viral disorder characterized by high FEVER; COUGH; DYSPNEA; renal dysfunction and other symptoms of a VIRAL PNEUMONIA. A coronavirus SARS-CoV-2 in the genus BETACORONAVIRUS is the suspected agent.
Indexing Information	severe acute respiratory syndrome coronavirus 2
Date of Entry	2020/02/13
Revision Date	2020/04/07

Figure 2. MeSH Supplementary Concept Record for COVID-19. Image capture 2020-04-21. Source: https://meshb.nlm.nih.gov/record/ui?ui=C000657245

2.2 Purposes and Usages of the Major Health KOSs

The efforts listed in the previous section necessitate a discussion about the need for various types of KOS. In healthcare systems, there are many parties, including medical professionals, medical technicians and paramedics, clinical laboratory technologists and technicians, government agencies, international organizations, academic researchers, medical center

9 https://www.nlm.nih.gov/pubs/techbull/jf20/brief/jf20_mesh_novel_coronavirus_disease.html

Table 1
COVID-19 Coding Guidance

KOS	Code	Code Description	Coding guidance
ICD-10 *International Classification of Diseases – Version 10.* (Guidance released on 2020-03-25: https://www.who.int/classifications/icd/COVID-19-coding-icd10.pdf)	U07.1*	COVID-19, virus identified	Positive test result; COVID-19 documented as cause of death *Use intervention/procedure codes to capture any mechanical ventilation or extracorporeal membrane oxygenation and identify any admission to intensive care unit
	U07.1	COVID-19, virus not identified o Clinically epidemiologically diagnosed COVID-19 o Probable COVID-19 o Suspected COVID-19	Positive test result only, patient showing no symptoms
	U07.1 + codes for symptoms	COVID-19, virus identified	Use additional code(s) for respiratory disease (e.g. viral pneumonia J12.8) or signs or symptoms of respiratory disease (e.g. shortness of breath R06.0, cough R05) as documented
	U07.2; Z20.8 + codes for symptoms	Contact or suspected exposure	Suspected/probable cases. No other etiology; history of travel
	U07.2; Z20.8 + codes for symptoms	Contact or suspected exposure	Suspected/probable cases. Contact with confirmed or probable case
	U07.2 + codes for symptoms		Suspected/probable cases. No other etiology: hospitalization required
	U07.2 + codes for any symptoms		Suspected/probable cases. COVID-19 documented without any further information regarding testing
CPT *Current Procedural Terminology* (Guidance released on 2020-03-13; https://www.ama-assn.org/practice-management/cpt/covid-19-coding-and-guidance)	87635	SARS-COV-2 COVID-19 AMP PRB	Effective March 13, 2020, for novel coronavirus tests through infectious agent detection by nucleic acid
	86318	IMMUNOASSAY INFECTIOUS AGENT ANTIBODY	Effective April 10, 2020, for novel coronavirus tests through infectious agent detection by nucleic acid
	86328	IA NFCT AB SARSCOV2 COVID19	Effective April 10, 2020, for antibody tests using a single step method immunoassay. This testing method typically includes a strip with all of the critical components for the assay and is appropriate for a point of care platform
	86769	SARS-COV-2 COVID-19 ANTIBODY	Effective April 10, 2020, for antibody tests using a multiple step method. For severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) (Coronavirus disease (COVID-19)) antibody testing using single step method, use 86328

(continued) Table 1

COVID-19 Coding Guidance

KOS	Code	Code Description	Coding guidance
SNOMED CT *Systematized Nomenclature of Medicine – Clinical Terms* (Guidance released on 2020-03-09; https://confluence.ihtsdotools.org/display/snomed/SNOMED+CT+COVID-19+Related+Content)	840539006	COVID-19	Fully specified name (FSN) = Disease caused by severe acute respiratory syndrome coronavirus 2 (disorder)
	840544004	Suspected COVID-19	FSN = Suspected disease caused by severe acute respiratory coronavirus 2 (situation)
	840534001	SARS-CoV-2 vaccination	FSN = Severe acute respiratory syndrome coronavirus 2 vaccination (procedure)
	840536004	Antigen of SARS-CoV-2	FSN = Antigen of severe acute respiratory syndrome coronavirus 2 (substance)
	840535000	Antibody to SARS-CoV-2	FSN = Antibody to severe acute respiratory syndrome coronavirus 2 (substance)
	840546002	Exposure to SARS-CoV-2	FSN = Exposure to severe acute respiratory syndrome coronavirus 2 (event)
	840533007	SARS-CoV-2	FSN = Severe acute respiratory syndrome coronavirus 2 (organism)
LOINC *Logical Observation Identifiers Names and Codes* (Special use codes and terms pre-released in mid-March; https://clinicalarchitecture.com/covid-19-updates/)	94721-8	COVID-19 Evaluation note	These pre-released terms are not yet part of an official LOINC release and therefore not available as a direct download from LOINC website
	94723-4	Emergency department COVID-19 Initial Evaluation form	For a complete list of COVID-19 related LOINC codes, check https://loinc.org/prerelease/
	94722-6	COVID-19 Initial Evaluation form	

administrators, patients and families, and insurance providers. During the COVID-19 or any pandemic, users who want to find out the status, impacts, prevention, and control of the disease need to find accurate information while gaining a complete picture of its significance around the world. Data and information have to be exchanged among all of the above-mentioned parties. The ability of KOS to reduce potential semantic conflicts assists in connecting verified data, compatible data structures, and explicit scientific information, which ensures the quality and consistency of data and information communicated in scholarly publications as well as to the general public. As everything is increasingly digital, information and data exchange become even more critical. For many years, efforts have been given to produce clinical data standards and health information exchange protocols. Each of these KOSs has unique features and purposes, being used in different workflows and at different decision levels, appearing in electronic health records, clinical quality reports, subject indexing, and in other forms. "Different families of knowledge organization systems, including thesauri, classification schemes, subject heading systems, and taxonomies are widely recognized and applied in both modern and traditional information systems" (W3C, 2009). The main functions of the standard KOSs, also known as health terminology, as demonstrated above, include eliminating ambiguity, controlling synonyms or equivalents, and establishing explicit semantic relationships. Embodied as (Web) services, they can facilitate resource discovery and retrieval. They act as semantic road maps and make possible a common orientation for information professionals and future users (whether human or machine) (Hodge, 2000; Tudhope & Koch, 2006).

Electronic health records (EHRs) have become the central feature in the healthcare domain of the 21st century. In addition to the formal systems used, there are an increasing number of open-source EHR systems (Syzdykova, Malta, Zolfo, Diro, & Oliveira, 2017). In a healthcare information system, semantic interoperability is

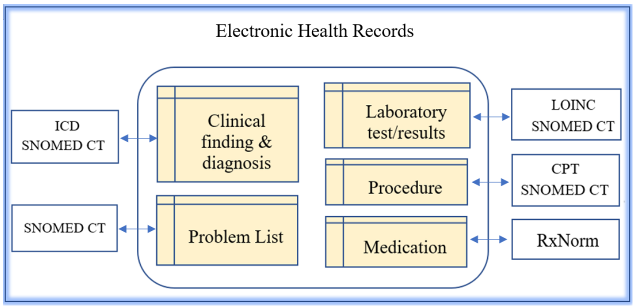


Figure 3. Variety of KOSs used in the world of EHR.

the ability to use digital health information across diverse settings and clinical software as the increasing amounts of health data from different locations make unique challenges in connecting and analyzing these data as a unified set. It is obvious that different systems that are modeled differently will have difficulties in data transferring, cross-dataset searching, and data reusing. Very often, this will also impact data integrity and could cause serious problems. For example, when two health information systems share e-records, they will need to be very careful about data transfer. For the data to be usable and exchangeable, those systems have to communicate using standard language and terminologies. This means that the codes from one system must be mapped to the codes from the other system. Though the data from multiple systems can be stored in one place, if those codes cannot be mapped to one another, the data would not be unlocked.

Standardized KOSs used in EHRs are illustrated in Figure 3, based on the practices in the United States. These include ICD, SNOMED CT (*Systematized Nomenclature of Medicine Clinical Terms*), LOINC (*Logical Observation Identifiers Names and Codes*), CPT (*Current Procedural Terminology*), RxNorm, etc. (Figure 3).

Given the fact that multiple and diverse KOSs have been widely used, it is necessary to describe them based on their main purposes and usages. Since it is the vocabulary that is used in the data exchange and information communication, they will be referred to as "terminology" in the following sub-sections. For most of these terminologies, the full titles are also provided in this section, in order to facilitate a better understanding of their contents. A list of major health KOSs is provided in the Appendix of this article. Sources of additional KOSs mentioned below are provided in footnotes. All of them are consolidated at the Unified Medical Language System (UMLS)[10] of the National Library of Medicine. Their full ontology deliveries are available at the BioPortal[11] terminology service, an authoritative, comprehensive repository of biomedical ontologies. There exist specific information systems in the world which may use different terminologies or self-defined ones. Yet, in general, those types of terminologies will also fall into these categories according to the purposes and usages.

2.2.1 WHY — Their Purposes and Responsibilities

First, health KOS may be categorized into administrative, clinical, reference, and interface terminologies according to their major responsibilities in application (Bronnert, Masarie, Naeymi-Rad, Rose, & Aldin, 2012).

An **administrative terminology** is developed to help administrative functions of healthcare, such as medical billing, reimbursement, classification of information, and other secondary data aggregation. Common standardized ones are ICD, *Current Procedural Terminology* (CPT), and *Healthcare Common Procedure Coding System* (HCPCS). WHO's ICD has been translated into 43 languages and it is used by all member States. Most countries (117) use it to report mortality data, a primary indicator of health status, according to ICD.[12]

A **clinical terminology** is developed for electronic exchange and aggregation of clinical data through encoding specific clinical entities involved in clinical workflow. Among the common standards is RxNorm, a normalized naming system for generic and branded drugs. It is a tool for supporting semantic interoperation between drug terminologies and pharmacy knowledge base systems. The purpose of RxNorm is to address the communication problems when the systems used in hospitals, pharmacies, and other organizations record and process drug information with different sets of drug names. Another universal standard is *Logical Observation Identifiers Names and Codes* (LOINC) developed for identifying laboratory observations.

A **reference terminology** is concept-based. It identifies semantic relationships between concepts, maintains a common reference point in the health information systems, and enables organization and aggregation of clinical information. This kind of terminology, e.g., *Systematized Nomenclature of Medicine Clinical Terms* (SNOWMED CT), represents a large number and range of possible concepts in a consistent

10 https://www.nlm.nih.gov/research/umls/index.html

11 http://bioportal.bioontology.org/

12 https://www.who.int/classifications/icd/revision/icd11faq/en/

manner and allows health information systems to get values from clinical data coded at the point of care. In general, reference terms are useful for decision support and aggregate reporting and are therefore more general than the highly detailed descriptions of actual patient conditions.

An **interface terminology** is a code set to ensure interoperability. Also known as colloquial terminologies, application terminologies, or entry terminologies, they are collections of healthcare related terms that support documentation of patient information. Since administrative codes and descriptors do not use the same language as different clinical, administrative, and colloquial terms that clinicians are using daily in healthcare, it makes it difficult for clinicians, information management professionals, and patients to find the terms they need when performing simple text searches. An interface terminology, e.g., Intelligent Medical Objects (IMO)'s *Clinical Interface Terminology* (CIT), [13] enables the mapping of the clinically relevant terms to the standard administrative and clinical terminologies and bridges the gaps. With interface terminology in place within an EHR, clinicians are able to find the right diagnosis and procedure terms to document and code, generating more comprehensive and accurate coding within their normal workflow.

2.2.2 WHERE — Their Major Usages

Secondly, from the perspective of usage, these common health KOSs may also be categorized into medical billing standards, clinically specific standards, pharmacy terminology standards, nursing terminology standards, and messaging standards. They play a big role in healthcare information systems to facilitate data normalization.

Medical billing terminology standards are used by all healthcare organizations to support medical billing, including ICD for diagnosis and procedure reimbursement, *Current Procedural Terminology* (CPT) for billing and reimbursement of outpatient procedures and interventions, and Diagnosis Related Group (DRG) classification[14] for billing a patient's hospital stay in the inpatient setting. They are mandated by the Health Insurance Portability and Accountability Act (HIPAA)[15] to code a patient's medical history.

Clinical terminology standards are used to describe health conditions and problems, supporting meaningful electronic exchange and aggregation of clinical data for better patient care. The most common examples include *Systematized Nomenclature of Medicine Clinical Terms* (SNOMED CT) and *Logical Observation Identifiers Names and Codes* (LOINC), while RadLex radiology lexicon and other clinically specific standards tend to have a specific clinical or workflow emphasis.

Pharmacy terminology domain is well-represented in many commonly used solutions and databases in pharmacy management and drug interaction software such as *National Drug File - Reference Terminology* (NDF-RT),[16] *National Drug Code* (NDC),[17] *First Databank* (FDB),[18] *Multum MediSource Lexicon* (Multum),[19] *Medi-Span Drug Data* (Medi-Span),[20] and United States Pharmacopeia (USP) Compendial Nomenclature.[21] The interoperability standard recommended for pharmacy terminology is the open-source RxNorm of the U.S. National Library of Medicine that provides normalized names for clinical drugs, which has links to many existing drug vocabularies to facilitate interoperability of drug information.

Nursing terminology standards have been developed to describe the nursing process, document nursing care with interventions and outcomes, and facilitate aggregation of data for comparisons at local, regional, national, and international levels. The *Nursing Outcomes Classification* (NOC)[22] and *Nursing Intervention classifications* (NIC)[23] are used to describe nursing practice in both home and acute care. The *International Classification for Nursing Practice* (ICNP)[24] and *NANDA International* (NANDA-I)[25] are being used internationally to describe nursing diagnoses, interventions, and outcomes.

Messaging standards enable health information systems to communicate using the industry standards of health information exchange. Digital Imaging

13 https://www.imohealth.com/imo-precision-sets
14 https://hmsa.com/portal/PROVIDER/zav_pel.fh.DIA.650.htm
15 https://aspe.hhs.gov/report/health-insurance-portability-and-accountability-act-1996

16 https://www.nlm.nih.gov/research/umls/sourcereleasedocs/current/NDFRT/
17 https://www.fda.gov/drugs/drug-approvals-and-databases/national-drug-code-directory
18 https://www.fdbhealth.com/
19 https://www.cerner.com/solutions/drug-database
20 http://www.wolterskluwercdi.com/drug-data/
21 https://www.usp.org/health-quality-safety/compendial-nomenclature
22 https://www.nlm.nih.gov/research/umls/sourcereleasedocs/current/NOC/
23 https://www.nlm.nih.gov/research/umls/sourcereleasedocs/current/NIC/
24 http://www.who.int/classifications/icd/adaptations/icnp/en
25 http://www.nanda.org/

and Communications in Medicine (DICOM)[26] is the international standard to transmit, store, retrieve, print, process, and display medical imaging information. The Health Level Seven (HL7)[27] standard is a set of international standards dedicated to providing a comprehensive framework and related standards for the exchange, integration, sharing and retrieval of electronic health information that supports clinical practice and the management, delivery and evaluation of health services. HL7 requires using standardized terminologies to represent health data. Besides developing its own standardized code sets to identify administrative data such as gender code, data type, and status codes, HL7 has employed existing standardized health KOSs to support unambiguously health information exchange.

In general, health KOS standards are essential parts of the framework supporting unique applications and services for healthcare research and practice. Initiatives to understand the evolution and spread of viruses are necessary for public health measures as well as for surveillance and tracking. As viruses evolve rapidly, it becomes necessary to make inferences about their epidemic history from genomic data, and this is possible if the right applications can be created to quickly harness the genomic data. No matter whether it is during a pandemic or not, data integration will be a vital component in the delivery of quality services as data are ingested, captured, and collected from multiple sources. The integration and interoperability of these resources are key to enabling applications that will answer questions currently elude us.

3 Schemas for Web-based Information Management and Communication

When dealing with millions of published and unpublished resources from various sources, it is necessary that search engines, authors, users, databases, and programs are all able to reference conceptual terms from the same language. Search engines, for example, rely on the content markup embedded in the websites to improve the display of search results, making it easier for people to find the right web pages. Again, to eliminate semantic conflicts and handle information overload, shared vocabularies are necessary for webmasters and website developers to provide semantically enriched structured data regarding the contents of their web resources.

One such common and extensive vocabulary is Schema.org. Schema.org is an open, community-based, collaboratively developed vocabulary to promote the structuring of data on the Internet, originally founded by Google, Microsoft, Yahoo, and Yandex (Schema.org Community Group, n.d.-a.). The extensible vocabulary covers different types of entities, relationships between entities, and is used by over ten million sites to markup web pages and email messages. This markup ability is especially important for the purposes of web indexing and exposing web-based information to search engines, as the vocabulary can be used to structure the contents of websites as HTML Microdata.

In terms of improving the dissemination of health-related information on the Web, the vocabulary also specifically supports the embedding and structuring of health and medical information through a schema type *MedicalEntity* and its subtypes (Schema.org Community Group, n.d.-b.). Citing both the existence of high-quality health information on the Web and the difficulty users experience in finding and navigating such information, developers looked for a way to allow webmasters and publishers to mark up the medical-related content on their websites. Developers worked with experts from the U.S. National Center for Biotechnology Information (NCBI), physicians from Harvard, Duke, and other institutions and from health websites, as well as contributors from the W3C Healthcare and Lifesciences group to address the markup of the implicit structure in health-related information. The goal was to bridge the gap between medical knowledge found in high-quality information on the Web and the keywords to be used in web searches through search engines. The resulting schema models medical entities such as conditions, signs and symptoms, risk factors, therapies, studies and trials, and guidelines. Although not intended for clinical data exchange or to supplant existing medical vocabularies, the schema covers information meant for consumers and professionals, and provides ways to annotate entities with codes from existing controlled medical vocabularies (e.g., MeSH, SNOMED, ICD, RxNorm, UMLS). The Health and Lifesciences Section[28] within Schema.org enumerates an extension of defined classes and properties for structuring medical information.

Most recently, developers and community leaders of Schema.org quickly responded to the global pandemic

26 https://www.dicomstandard.org
27 http://www.hl7.org/
28 https://health-lifesci.schema.org/

by releasing the versions 7.0 through 7.04 within a month (Schema.org Community Group, 2020). Some of the following examples of changes to the vocabulary demonstrate the vocabulary's flexibility in addressing medical and health related properties and the urgent need for structuring public and official information regarding COVID-19. For example, in the creation of new types such as *SpecialAnnouncement* (Thing > CreativeWork > SpecialAnnouncement),[29] the motivating scenario is the coronavirus pandemic, and the initial vocabulary is oriented to this emergency situation. Note that terms in Schema.org for classes/types are named with capital letters with no spaces, or upper camel case (e.g., *MedicalEntity*) and properties are named in lower camel case (e.g., *medicineSystem*).

- 2020-03-17, Version 7.0
  - Un-used medical health properties with "inappropriately general" names were removed.
  - Motivated by the need to structure urgent, crisis-related information during the COVID-19 pandemic, a new type *SpecialAnnouncement* and its properties were created.
  - A subtype of *MedicalClinic*, labelled *CovidTestingFacility*, was implemented to describe locations where testing for COVID-19 is available. This also allowed reuse of existing Schema.org structured data, such as contact information, address, and operating hours, for existing described instances.
  - A property, *hasDriveThroughService*, was created to indicate the presence of drive-through services, like drive-through virus testing.
  - Due to work at home policies and the migration to online work settings, a *VirtualLocation* type was developed, including a way to indicate that an event has moved from a physical location to an online or virtual location.
- 2020-03-22, Version 7.01
  - *SchoolDistrict* was added as a subtype of *AdministrativeArea*.
  - A property, *datePosted,* was included in *SpecialAnnouncement*, which has the dates showing changes or revisions done by official announcements as they are developed in real time.
- 2020-03-31, Version 7.02
  - Another property, *announcementLocation*, was added in *SpecialAnnouncement*, to clarify the spatial coverage for an announcement.
- 2020-04-06, Version 7.03
  - Due to the requirements of medical and government authorities to aggregate data from different medical facilities and the U.S. CDC definition of data fields to facilitate exchange of COVID-19-related data, a 1:1 Schema.org representation of the required CDC CSV data reporting format, as well as schema documentation for its implementation, was developed (Brickley, 2020).
  - Developers issued a clarification that the property *webFeed* can be used to describe a *SpecialAnnouncement*.
- 2020-04-16, Version 7.04
  - Developers clarified that an *Educational Organization* was also a subtype of *Civic Structure*, in order to be a value for an *announcementLocation* of a *SpecialAnnouncement*.

In addition to structuring information that will be published on the Web, indexed by search engines, and disseminated to the public, one of the notable developments to the Schema.org vocabulary mentioned in the timeline above is the creation of a type of *StructuredValue* which encodes CDC data fields as properties of a *CDCPMDRecord*. A *CDCPMDRecord* (which stands for CDC Patient Module Denominator Record) is defined as "a data structure representing a record in a CDC tabular data format used for hospital data reporting" (Bradley, 2020). The properties of this class ensure consistency with CDC records by corresponding with required CDC data fields, and have been prefixed with "cvd" to differentiate these properties with other properties that may occur elsewhere in the Schema.org vocabulary. For instance, where a CDC record has a data field for a number of available inpatient hospital beds ("numbeds"), a corresponding property *cvdNumBeds* is used in Schema.org to structure such data, and to distinguish it from the property *numberOfBeds* that already exists for describing hotel rooms, apartments, etc. The prefixing of "cvd" keeps data regarding the *CDCPMDRecord* type (14 properties as of the beginning of April)[30] relatively self-contained (Brickley, 2020). Although developers of Schema.org note that data encoded in this way may not be published on the public Web or made available to search engines, it does mitigate future semantic conflicts by ensuring that consistency between data reported to the CDC under its reporting requirements and Schema.org encoded records if they were made available later on the Web as linked datasets.

29 https://schema.org/SpecialAnnouncement

30 https://schema.org/docs/cdc-covid.html

Based on the swift and evolving response from its development community, Schema.org's actions during the COVID-19 pandemic filled gaps caused by the dramatic increase in the amount of information on the Web and satisfied the need for managing and communicating that information through effectively structured data. These efforts to organize and classify relevant information and resources directly contribute to addressing the challenges presented by information overload and semantic conflict.

4 Ontologies for COVID-19 Research

Another significant set of KOS products are the ontologies being created or experimented with regarding to COVID-19 research and development. In the fields of computer science and knowledge organization, an ontology has been defined as "an explicit specification of a conceptualization" (Gruber, 1995, p. 908), in which the basic terms for concepts and their relationships within a particular domain are defined, as well as the rules for combining and associating terms and relationships to represent more complex concepts (Neches et al., 1991). In contrast to the health standards introduced in the previous sections in this article, ontologies are used for specific-use scenarios and situations. In the case of COVID-19, ontologies are being used to present interoperability solutions by linking equivalencies and associations between terms in different vocabularies and data encoded from different standards as well as semantic solutions to make data processible by both humans and machines. These innovative usages can be seen in the creation of knowledge graphs, automatic reasoning, and using artificial intelligence to help tackle the information regarding coronaviruses. The following selections provide a handful of examples that demonstrate the ways in which ontologies can power the organization and exchange of massive amount of information over Web-based applications and collected through automated systems.

4.1 Infectious Disease Ontology

Infectious Disease Ontology (IDO) Core is a set of interoperable ontology modules covering entities and relations generally relevant to the infectious disease domain (Babcock, Beverley, Cowell, & Smith, 2020). This core is then extended to focus on specific diseases and pathogens. IDO is itself related to and extended from the *Ontology for General Medical Science* (OGMS), where IDO Core describes relationships among and distinguishes between entities of an infectious disease, sign or symptom of the infectious disease, infectious disease diagnosis, infectious disease course, and infection. Extensions of IDO Core relevant to COVID-19 include *IDO Virus* (VIDO), the *Coronavirus Infectious Disease Ontology* (CIDO), and *IDO-COVID-19 Infectious Disease Ontology*, with VIDO and IDO-COVID-19 being currently in development. IDO Core serves as an important backbone in reducing semantic conflicts and promoting interoperability among heterogeneous bodies of research data as well as among ontologies designed for specific kinds of infectious diseases.

4.2 Coronavirus Infectious Disease Ontology

One important extension of IDO is CIDO, the *Coronavirus Infectious Disease Ontology* (He, 2020), an expansive, open-source, community-developed biomedical ontology aimed at providing human- and machine-readable annotation and representing numerous aspects of coronavirus infectious diseases in general. With over 3,700 classes of entities and 82 properties, the ontology covers terms for disease causes, transmission, diagnosis, prevention, and treatment, with properties for integrating other forms of medical vocabulary and relevant coding systems such as matching terms with their Chinese labels and ways of linking terms from MeSH, SNOMED, RxNorm, UMLS, and Unique Ingredient Identifiers (UNII) from the U.S. Food and Drug Administration (FDA) Substance Registration System.

4.3 COVID-19 Surveillance Ontology (COVID19)

COVID-19 Surveillance Ontology is a small application ontology (Liyanage, de Lusignan, & Williams, 2020) intended to support surveillance of COVID-19 in primary care settings. Classes of entities include terms for classifying means of exposure, methods of testing and diagnosis, and subsequent courses of action such as isolation and contact tracing. The ontology is envisioned to facilitate COVID-19 case monitoring, including related respiratory conditions and data gathered from computerized medical systems.

4.4 WHO COVID-19 Rapid Version CRF Semantic Data Model (COVIDCRFRAPID)

This data model (Bonino, 2020) allows semantic referencing of data gathered from the "Global COVID-19: Clinical platform: Novel coronavirus (COVID-19): Rapid version form," a clinical characterization case record form (clinical CRF) developed by the WHO to standardize the collection of clinical data regarding characteristics of the disease and its treatment (WHO, 2020c). The model enumerates the medical and biological terminology that appears on the WHO CRF and maps those terms to their semantic equivalents in other KOSs.

4.5 Cochrane PICO Ontology and Linked Data Vocabulary

Not-for-profit organization Cochrane originally developed its *PICO Ontology* to characterize four areas addressed by clinical studies (Population, Intervention, Comparison, and Outcome), which is supported by the approximately 400,000 terms of the *Cochrane Linked Data Vocabulary* (Cochrane, 2020). This vocabulary is linked to existing health vocabularies—including MeSH, SNOMED-CT, *Medical Dictionary for Regulatory Activities* (MedDRA), and RxNorm—to support semantic standardization, and now includes a term for a Condition, COVID-19. The *PICO Ontology* and the *Cochrane Linked Data Vocabulary* provide part of the data architecture on the back-end to enable front-end applications like the recent Cochrane COVID-19 Study Register,[31] which provides up-to-date information on clinical trials and published research (Wilton, 2020).

This small subset of ontologies discussed above reflects some of the immediate efforts to directly respond to the organization and classification of the overwhelmingly large amount of information that has proliferated as a result of both clinical data collection efforts and research efforts to combat the virus. Many other ontology development projects exist worldwide, and some may be extended or revised to integrate new entities and relationships regarding COVID-19. For instance, the *Gene Ontology* has compiled and included in their knowledge base functions of human proteins used by the SARS-CoV-2 virus to enter a human cell and those possibly targeted by the virus after cell entry, as well as semantic annotations of relevant human and viral genes (Gene Ontology Consortium, 2020). Moreover, experimental application of a COVID-19 ontology with tools for descriptive logic reasoning has been proposed in the context of detecting "fake news" to identify inconsistencies between trusted medical sources and sources of unknown veracity (Groza, 2020). In addition to the open source ontologies and novel approaches involving ontological modeling introduced here, there may be many other proprietary or local ontologies or ontology-related studies being used and developed in research labs and beyond.

Additionally, other KOSs can be used as the basis for, or converted into, ontological models. Such thesauri, classifications, and taxonomies converted into OWL ontologies can be found in the BioPortal ontology repository. A search conducted on April 25, 2020 for the ontological class "Coronavirus" led to 19 ontologies containing this term, which included various entries on the family of Coronaviruses of which COVID-19 is a part, as well as other human and animal viruses. A search for "COVID-19" brought up ontologies such as the *Experimental Factor Ontology* (EFO) and the *Vaccine Ontology* (VO), in addition to the *COVID-19 Surveillance Ontology* and WHO's Rapid version CRF data model described earlier. As information continues to augment regarding COVID-19 disease and the SARS-CoV-2 virus, more ontologies and semantic applications will adapt and expand their classes, entities, and related properties to structure emerging biological and medical data about the pandemic.

5 Applications, Services, and Data-based Analysis Empowered by KOS

Data-driven decision-making has become increasingly important in our daily lives as well as in business, research, and clinical settings. The ever-increasing data volumes and complexities will become even more overwhelming for users already overburdened by information overload. Interesting studies on information overload and decision making in various situations have been continuously conducted. When examining the scope of research on health information overload in consumers, Khaleel et al. identified predictors of health information overload, such as low education level, health literacy, and socioeconomic status; and found that videotaped consultations and written materials can be used as interventions for information overload (Khaleel et al., 2020). Bawden and Robinson (2020) came up with some ideas for avoiding information overload, such as filtering, withdrawing, queuing, and satisficing, as well as

31 https://covid-19.cochrane.org/

designing better information systems, managing personal information effectively, and promoting digital and media literacy. Consequently, it is therefore crucial to develop strategies that avoid or negate the harmful impacts of information overload at this very difficult time of the COVID-19 pandemic. Ortutay and Klepper (2020, April 24) suggested the following strategies for individuals to avoid information overload: (1) look for the source; (2) check the websites of the CDC and the WHO; (3) act like a journalist; (4) pause before reposting the news; and (5) do not believe everything you see.

The challenges of information overload could increase dramatically when a report or a message needs to be communicated across languages, regions, and cultures, especially when the statistics are generated based on different types of contextual information from diverse resources. To illustrate, Nielsen, Fletcher, Newman, Brennen, and Howard (2020) investigated and compared the information overload in terms of accessing and rating news and information about coronavirus by people in six countries based on the data from news media, blogs, and social media. Baniamin, Rahman, and Hasan (2020) further tried to answer the question: "Why are some countries more successful than others in the COVID-19 pandemic?" Their findings show that the better performing countries have done well with the following aspects: consideration of people's attitudes, demographic profile, citizen trust, culture, magnitude of policy learning, state structure, and technological and administrative readiness.

It should be noted in all the examples above that well-designed and purposeful technology plays a key role in reducing the negative impacts of information overload. Data can allow for assessment of trends and patterns and be used to make inferences from contextual information that humans may lose sight of. Studies have shown that various applications and services such as knowledge graphs, recommendation engines, and decision support systems can provide remarkable support as aids in decision-making and positively impact decision quality. The field of healthcare—in particular the clinical processes of diagnosis treatment, monitoring and prognosis—is largely driven and shaped by knowledge. Thus, there is a critical need to integrate vast amounts of structured, semi-structured, and weakly structured data and a tremendous volume of unstructured information in addition to enhancing the functioning of workflows, processes and guidelines, thereby minimizing costs and inefficiencies to improve research and practice (Holzinger & Jurisica, 2014). New methods and applications that support knowledge discovery and interpretation of complex data in integrated and interactive ways are needed and are being developed. Keselman et al. (2010) noted that a proliferation of disaster health information results in information overload among responsible professionals, requiring efforts be made to manage and organize information in ways that support decision making using semantic technology, data mining and machine learning approaches, graphs, and integration of systems. As such clinical health and public health has seen the development of multiple tools to harness voluminous data and perform surveillance, tracking, analysis, prediction and modeling for various healthcare crises. The next sections provide a review of some tools being used to support the efforts of eliminating information overload during the COVID-19 pandemic.

5.1 Health Dashboards

5.1.1 HealthMap

HealthMap (https://www.healthmap.org/en/) is a health surveillance system supporting public health by focusing on event-based monitoring of infectious diseases. The tool continually aggregates reports on new and ongoing infectious disease outbreaks which it then extracts, categorizes, filters, and integrates in an effort to facilitate knowledge management and early detection (Brownstein, Freifeld, Reis, & Mandl, 2008). The system serves as a source for libraries, local governments, governments and multinational agencies, as well as normal users. HealthMap uses text mining algorithms to perform characterization of the sources from which it draws data. This includes identifying disease and location, ascertaining the relevance, grouping reports, and removing duplicates (Freifeld, Mandl, Reis, & Brownstein, 2008). To perform extraction of locations and disease names it relies on various KOSs dealing with pathogens and geographic names to classify and tag the data. Because multiple sources are being integrated, HealthMap often ascribes a reliability score to the information shown by giving increased weight to official sources of information and less to media reports. Currently HealthMap is being used to track the spread of COVID-19 and support public health official efforts to fight the disease. The website provides multiple faceted filters, allowing a user to look for information about the outbreaks and alerts by location, time, diseases, species, and sources. Each of them is supported by pre-defined vocabularies, name authorities, and taxonomies; hence the information is exchanged without possible semantic conflict.

5.1.2 Nextstrain

Nextstrain (https://nextstrain.org/) is an open-source initiative that uses pathogen genome data in its provision of real-time views of the spread and evolution of viruses. It consists of data curation analysis and visualization components (Hadfield et al., 2018) and is facilitated by collaborations with subject matter experts. This system relies on taxonomies from the National Center for Biotechnology Information (NCBI) of the United States National Library of Medicine and data from the Virus Pathogen Resource (a searchable database of virus genomes) which itself stores genomic, proteomic, and annotation data from various ontologies and protein databases. Nextstrain is unique in that it tracks and is able to recreate mutations as well as convey the geographical spread of the virus and the underlying genomic data that is supporting that. Nextstrain is being used by researchers to trace the virus back to its origin point and date in time and space and to perform analysis of its genomic variations in order to facilitate treatment, containment, and development of vaccines.

5.1.3 Johns Hopkins COVID-19 Dashboard

In response to the COVID-19 pandemic, the Johns Hopkins University developed an interactive web-based dashboard that could visualize and track cases in real time (https://coronavirus.jhu.edu/us-map). It aggregates local government and media reports around the world and confirms the data with local health authorities before publishing. Researchers, public health authorities, and the general public are using the tool to track the outbreak as it happens. In the US it reports data at the city level and in China at province level. It uses a semi-automated living data stream strategy to aggregate data from multiple sources. This tool relies strongly on federal information processing standards and geographic identifiers from KOSs created by The American National Standards Institute (ANSI), US Census Bureau, US Department of Education, and the US Geological Survey (USGS) to organize the information presented (Dong, Du & Gardner, 2020).

5.1.4 FAPDA— Food and Agriculture Policy Decision Analysis Tool

The FAPDA is a FAO initiative that aims to promote evidence-based decision making by providing up-to-date information on national policy decisions and policy frameworks, through the FAPDA policy database. The FAPDA database (http://www.fao.org/in-action/fapda/fapda-home/en/) is a global and comprehensive database which contains over 10,000 national policy decisions and 2,000 national policy frameworks for 100 countries around the world. In the current exceptional situation, a timely tool is provided for discovering policy measures and urgent actions directly related to COVID-19. The search interface provides faceted filters, which enable assessment of the data from different points of views, or from different aspects. Similar to HealthMap, facets of What, Where, When, Who, and How enable access points for browsing or searching of the policies with different traits. Behind each of the facets, there is a KOS which supports browsing and searching with standardized terms and semantic relationships. The *Policy Decision Classification* allows for the presentation of consumer-oriented, producer-oriented, and trade-oriented policies. Combined with schemes which support the food security dimension (*access_to*, *availability_of*, and *utilization_of* food), as well as the commodities, the policy decisions found can match with carefully crafted queries.

As exhibited above, in recent weeks the mobilization of the research community in response to the novel coronavirus has led to the development of multiple tools and applications to organize and harness knowledge. Without these efforts and the support of the KOSs much of this knowledge would remain hidden. Integrating these data sources has allowed for reasoning and analysis of the information on a scale that supports clinical decision making and local and international policy making.

5.2 Knowledge Graphs

Another type of application that benefits from the existence of KOS is knowledge graphs. On the Web knowledge graphs are essential components of information systems that need to access semantically structured data and support knowledge discovery. They are highly curated representations of knowledge and their function is most often seen to support semantic knowledge in web searches. Knowledge graphs describe real world entities and the relationships between them, defining classes and relations of entities in a schema. They enable the creation of relationships between arbitrary entities and can cover a wide variety of domains (Paulheim, 2017). Knowledge graphs drawn from multiple data points and present connections between data points to enable multiple perspectives in one glance and infer meaning, tailoring the information returned to the unique context of the search. Some of the most well-known implementations of knowledge graphs are those which can be seen when running Web searches. As an example, in a search for H1N1 which is an influenza virus, a number of data points are made available to users in the graph, including a general description of the virus as well as the virus family, class, higher classification and other pieces of data pulled from multiple sources. Currently a quick search for COVID-19 in a web browser will have the knowledge graph returning multiple data points including a quick statistic summary starting with your location and then extending it to worldwide.

Of note, the Cochrane COVID-19 Study Register (https://covid-19.cochrane.org/) is a knowledge graph technology built in direct response to COVID-19. In their efforts to produce evidence-based systematic reviews they utilize a flexible knowledge graph based linked data architecture. Systematic reviews are a highly knowledge-bound, domain-specific, time-intensive, and complicated task requiring oversight by subject matter experts to complete. A knowledge graph can support and simplify this process by capturing evidential statements about the contents of documents and describing the evidence at the right place within the content using structured linked data. A model for the knowledge graph implementation includes (1) a model for describing clinical questions, (2) a linked data vocabulary to describe and construct the questions, (3) a content model for description, (4) an annotation model to capture provenance, (5) a knowledge graph implementation to store the linked data vocabulary, the content metadata, and the evidence, and (6) a tool to enable all this (Wilton, 2020). At multiple steps of this process one can see its reliance on various types of KOS (from the simple to the complex) and the process itself is an integrated KOS.

5.3 Electronic Health Records

KOSs are often used for organizing the contents of electronic health records (EHRs), as discussed in Section 2.2. This is necessary because of heterogeneity of resources and the need to enable semantic interoperability (Park, Kim, & Min, 2012). Their use can improve population and health outcomes through the collection of data that can be shared across multiple organizations. The KOSs supporting these allow for the collection of standardized, systematic data that can be used for surveillance data submission, and electronic laboratory reporting. Information from these records is what providers use to transfer public and population health data to public health officials and organizations which they in turn use to monitor, prevent, and manage disease. In the current pandemic, tools built into the EHRs are supporting the work of clinicians through screening, patient education, updates for clinicians, decision support, ambulatory orders, reporting and analytics, secure communication, and telemedicine (Reeves et al., 2020).

Information found in EHRs is also being used in new ways to facilitate clinical research through secondary use of the data collected. The quality and context of secondary data is often questioned. KOSs can be used to improve the consistency and completeness of the data as well as to enable archiving and sharing the data after clinical studies are completed. Published studies have shown that they are helping clinicians assess risk factors for mortality (Zhou et al., 2020), transmission potential in pregnant women and asymptomatic carriers (Chen et al., 2020; Bai et al., 2020), and decipher clinical, epidemiological and diagnostic features of both adult and child patients (Qiu et al., 2020; Ng et al., 2020; Wu & McGoogan, 2020).

5.4 Multilingual Vocabulary and Concept Hubs

Information overload and semantic conflict challenges are multiplied when communication involves many languages, let alone more than one. Similar to other KOSs for information organization and retrieval use, multilingual thesauri are structured collections of concepts, terms, definitions and relationships. They differ from monolingual thesauri in that they cover many world languages and are often core resources for many countries and regions. Another feature of multilingual thesauri is that their development and maintenance are dependent on collaborative editing and community contributions.

Centralized vocabulary and concept hubs are typically provided by international organizations. The FAO of the UN, for example, has a well-established multilingual thesaurus, AGROVOC, since early 1980s. Each concept in AGROVOC has corresponding terms used to express those concepts in various languages. As of April 2020, AGROVOC consists of over 37,000 concepts and more than 750,000 terms in up to 38 languages.[32] The copyright for AGROVOC content that occurs in official FAO languages — Arabic, Chinese, English, French, Russian and Spanish — is retained by the FAO and licensed under the international

32 http://aims.fao.org/standards/agrovoc/concept-scheme

Table 2
Wikipedia and Wikidata Entries Related to COVID-19 (Data Collected on May 20, 2020)

Wikipedia			Wikidata			
Wikipedia entry	# of entries (languages)	Matching KOS IDs	Wikidata English Label and ID	scope notes	# of "Also Known as" in English	# of mapped "Identifier"
Coronavirus disease 2019	128	•ICD-10: U07.1 •ICD-10: U07.2 •MeSH: C000657245 •SNOMED CT: 840539006	COVID-19 (Q84263196)	zoonotic respiratory syndrome and infectious disease in humans, caused by SARS coronavirus 2	19	21
Coronavirus	69	•ICD-10:B97.2 •MeSH:D017934	Coronavirus (Q89469904)	group of related viruses that cause diseases in mammals and birds	1	6
COVID-19 pandemic	125	none	COVID-19 pandemic (Q81068910)	ongoing pandemic of COVID-19	15	23
Severe acute respiratory syndrome coronavirus 2	102	•ICD-10: U07.1 •MeSH: C000656484 •SNOMED CT: 840533007	SARS-CoV-2 (Q82069695)	strain of virus causing the ongoing pandemic of coronavirus disease 2019 (COVID-19)	16	14

Creative Commons Attribution License (CC-BY), while content in other languages rests with the institutions that authored it.[33] Following a request for more specific terminology on COVID-19, AGROVOC added new concepts related to the current world health crisis in an additional release on April 10 (the concepts coronavirus and pandemics were already present in the thesaurus). These include entries for "coronavirus disease" (alternative term COVID-19) and "severe acute respiratory syndrome coronavirus 2" (alternative SARS-Cov-2). Other concepts were added for MERS, SARS, movement restrictions, and supply chain disruptions. As of April 10, the lexical labels for these concepts were available in 16 languages.[34]

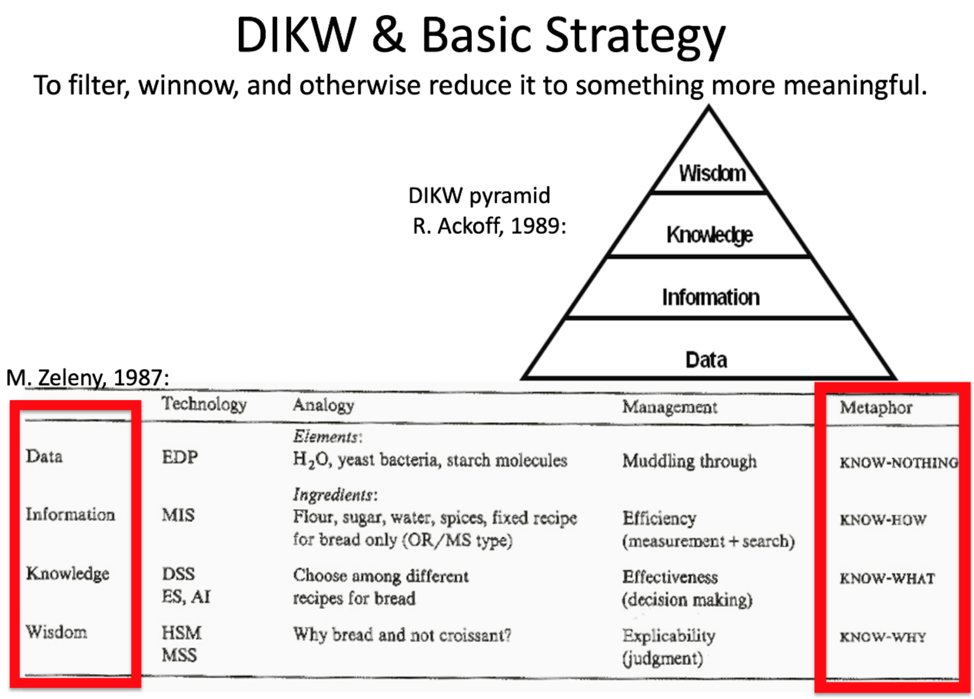
De-centralized vocabularies and concept hubs are also growing with the advancement of Wikimedia. As the COVID-19 pandemic became a dominant topic in news and social media, Wikipedia and Wikidata have had entries established and updated continuously. For example, the Wikipedia entry "Coronavirus disease 2019" had entries in 128 languages (as of the middle of May); the concept is matched to ICD-10, MeSH, and SNOMED CT. This concept's entry in Wikidata (as a name authority) has the label "COVID-19" with the identifier Q84263196, includes scope notes, has many synonyms or equivalents (e.g., 19 in English), and is mapped to 21 identifiers found in other controlled terminologies, news topic IDs, Google Knowledge Graph ID, etc. The data is updated frequently, while requests for changes in entry titles were noted in March and April. For example, Wikipedia changed the article's entry title from "2019–20 coronavirus pandemic" to "COVID-19 pandemic" in May 2020.

These multilingual vocabulary and concept hubs have become central resources for ordinary users during the COVID-19 pandemic, particularly for the purpose of communicating across different languages and cultures worldwide. By using a unique identifier for each concept, the collaborative controlled vocabularies provide a significant number of synonyms and equivalent terms used by different populations and agents. The mapped identifiers greatly enhance semantic interoperability among different commonly used KOSs as well as special vocabularies used by institutions and branches of the media.

6 Conclusion

Classification systems, taxonomies, controlled glossaries, thesauri, subject headings, ontologies, and other types of KOS have existed for many years. The health and medical fields, as well as other disciplines, have used KOSs to organize, represent, and conceptualize the specialized information that is vital to understanding their respective domains. KOSs provide a common framework and language for domain practitioners and a structured way to communicate complex concepts to the general public. The Digital Age has only brought more attention to KOSs in the overall global information infrastructure, as they increasingly form the backbone of automated and Web-based systems and services. As

33 http://aims.fao.org/vest-registry/vocabularies/agrovoc
34 http://aims.fao.org/activity/blog/AGROVOC_Covid-19



	Technology	Analogy	Management	Metaphor
Data	EDP	*Elements*: H_2O, yeast bacteria, starch molecules	Muddling through	KNOW-NOTHING
Information	MIS	*Ingredients*: Flour, sugar, water, spices, fixed recipe for bread only (OR/MS type)	Efficiency (measurement + search)	KNOW-HOW
Knowledge	DSS ES, AI	Choose among different recipes for bread	Effectiveness (decision making)	KNOW-WHAT
Wisdom	HSM MSS	Why bread and not croissant?	Explicability (judgment)	KNOW-WHY

Figure 4. Data-Information-Knowledge-Wisdom (DIKW) and basic strategy
Source: Image generated based on Ackoff (1989) and Zeleny (1987).

shown earlier, they have become even more critical to aid the frontline endeavors to overcome the obstacles of information overload and semantic conflicts that can occur during special historic and worldwide events like the COVID-19 pandemic.

This review of the efforts, usages, and functions of the significant KOSs discussed here reveals their essential roles in supporting health data exchange and information management, as well as ensuring consistency and interoperability of data collection and reuse among various providers and care settings. They also support many other services and applications, as they have been developed and routinely updated for administrative, reference, and interface purposes. This attention to consistency, responsiveness, and systematic maintenance also demonstrates their flexibility and swiftness to respond to a flood of new information, such as that experienced during the COVID-19 pandemic. In a digital information environment, these standardized health KOSs increasingly play a larger and more important role in healthcare information systems to facilitate data normalization, which is a fundamental requirement for any subsequent data analysis and information management. Additionally, both health KOSs and non-health- oriented KOSs which have a medical and health component vocabulary facilitate the dissemination of important and accurate information, in multiple languages, to audiences as diverse as healthcare providers, medical researchers, public health and government officials, the news media, and the general public. They also enable search and retrieval systems, interactive maps and charts, repositories and databases supporting research efforts, policy and decision making by governments, businesses, and medical facilities to have far reaching applications on patients and healthcare consumers.

Ultimately, the networking environment in which we live also means that we are in a big-data world. The use cases of KOSs presented here have been examined in context of their abilities to solve two major problems occurred as a result of this big-data world: first, dealing with information overload; second, resolving semantic conflicts. The fundamental approaches of KOS, including eliminating ambiguity, managing semantic relationships, and modeling ontological classes and their properties, enable effective and efficient management of information and knowledge. In categorizing and contextualizing vast amounts of developing information and by controlling or eliminating the pitfalls of semantic conflicts, KOSs ultimately underpin how experts and ordinary citizens understand the complexities of rapidly unfolding, global situations like those occurred during the COVID-19 pandemic.

The authors would like to conclude this article by revisiting the Data-Information-Knowledge-Wisdom (DIKW) pyramid (Figure 4) used in knowledge management (Ackoff 1989). The DIKW pyramid represents the most basic strategy for understanding a world that is far beyond the capacity of our brain: the ability to filter, winnow, and otherwise reduce the chaos and confusion of the world to something more meaningful, building from a state of knowing "nothing," to understanding "how," "what," and "why" (Zeleny 1987). As supported by the examples discussed in this article, this sense-making hierarchy is bolstered by the foundation provided by KOSs. Especially in challenging times, KOSs provide the crucial infrastructure to transform disparate data and overloaded information into real, actionable knowledge as the basis of a wise decision-making.

Acknowledgments: The authors would like to thank the three anonymous reviewers for providing their valuable feedback and guidance.

References

Ackoff, R. L. (1989). From data to wisdom. *Journal of Applied Systems Analysis*, *16*(1), 3–9.

Babcock, S., Beverley, J., Cowell, L. G., & Smith, B. (2020). The Infectious Disease Ontology in the Age of COVID-19. doi:10.31219/osf.io/az6u5.

Bai, Y. , Yao, L., Wei, T., Tian, F., Jin, D.Y., Chen, L., & Wang, M. (2020). Presumed asymptomatic carrier transmission of COVID-19. *Journal of the American Medical Association, 323*(14), 1406–1407.

Baniamin, H.M., Rahman, M., & Hasan, M.T. (2020). The COVID-19 Pandemic: Why Are Some Countries More Successful Than Others? doi:10.2139/ssrn.3575251.

Bawden, D. & Robinson, L. (2020). Information overload: An overview. In *Oxford Encyclopedia of Political Decision Making*. Oxford: Oxford University Press. Retrieved from https://openaccess.city.ac.uk/id/eprint/23544/1/

Bonino, L. (2020). *WHO COVID-19 rapid version CRF semantic data model*. Retrieved from https://bioportal.bioontology.org/ontologies/COVIDCRFRAPID

Bradley, A. (2020). *New schema. Org vocabulary in response to COVID-19 and its use by search engines* [Google slides presentation]. Retrieved from https://docs.google.com/presentation/d/1OaXkItfzolwISZv4Io-f7P5kT0tMLSoY-2VPQF8oJXE

Brickley, D. (2020). *Schema.org COVID-19: US CDC data table fields*. Retrieved from https://schema.org/docs/cdc-covid.html

Bronnert, J., Masarie, C., Naeymi-Rad, F., Rose, E., & Aldin, G. (2012). Problem-centered care delivery: How interface terminology makes standardized health information possible. *Journal of American Health Information Management Association, 83*(7), 30–35. Retrieved from http://bok.ahima.org/doc?oid=105588#.XqxKG_ZFzKe

Brownstein, J. S., Freifeld, C. C., Reis, B. Y., & Mandl, K. D. (2008). Surveillance Sans Frontières: Internet-based emerging infectious disease intelligence and the HealthMap project. *PLoS medicine, 5*(7), e151. doi:10.1371/journal.pmed.0050151

CDC. (2020, March 18). *New ICD-10-CM code for the 2019 novel coronavirus (COVID-19)*. Retrieved from https://www.cdc.gov/nchs/data/icd/Announcement-New-ICD-code-for-coronavirus-3-18-2020.pdf

Chen, H., Guo, J., Wang, C., Luo, F., Yu, X., Zhang, W. , . . . Zhang, Y. (2020). Clinical characteristics and intrauterine vertical transmission potential of COVID-19 infection in nine pregnant women: A retrospective review of medical records. *Lancet*, *395*(10226), 809–815. doi: 10.1016/S0140-6736(20)30360-3

Cochrane. (2020). *The cochrane linked data vocabulary*. Retrieved from https://data.cochrane.org/concepts/

Dong, E., Du, H., & Gardner, L. (2020). An interactive web-based dashboard to track COVID-19 in real time. *The Lancet: Infectious Diseases, 20*(5), 533–534. doi: 10.1016/S1473-3099(20)30120-1

Fiorillo, A. & Gorwood, P. (2020). The consequences of the COVID-19 pandemic on mental health and implications for clinical practice. *European Psychiatry, 63*(1), 1-2. doi:10.1192%2Fj.eurpsy.2020.35

Freifeld, C. C., Mandl, K. D., Reis, B. Y., & Brownstein, J. S. (2008). HealthMap: Global infectious disease monitoring through automated classification and visualization of Internet media reports. *Journal of the American Medical Informatics Association*, *15*(2), 150–157. doi:10.1197/jamia.M2544.

Gene Ontology Consortium. (2020). *SARS-CoV-2 – Coronavirus*. Retrieved from http://geneontology.org/covid- 19.html

Groza, A. (2020). Detecting fake news for the new coronavirus by reasoning on the Covid-19 ontology. arXiv preprint arXiv:2004.12330. Retrieved from https://arxiv.org/abs/2004.12330

Gruber, T. R. (1995). Toward principles for the design of ontologies used for knowledge sharing. *International Journal of Human-Computer Studies, 43*(5-6), 907–928.

Hadfield, J., Megill, C., Bell, S. M., Huddleston, J., Potter, B., Callender, C., . . . Neher, R. A. (2018). Nextstrain: Real-time tracking of pathogen evolution. *Bioinformatics, 34*(23), 4121–4123. doi:10.1093/bioinformatics/bty407

He, Y. O. (2020). *Coronavirus infectious disease ontology*. Retrieved from https://bioportal.bioontology.org/ontologies/CIDO

Hodge, G. (2000). *Systems of knowledge organization for digital libraries: Beyond traditional authority files*. Retrieved from http://www.clir.org/pubs/reports/pub91/contents.html

Holzinger, A. & Jurisica, I. (2014). Knowledge discovery and data mining in biomedical informatics: The future is in integrative, interactive machine learning solutions. In A. Holzinger & I. Jurisica (Eds.), *Interactive Knowledge Discovery and Data Mining in Biomedical Informatics, 8401*. 1-18. Berlin: Springer. doi: 10.1007/978-3-662-43968-5_1

International Committee on Taxonomy of Viruses. (2020). *Naming the 2019 Coronavirus*. Retrieved from https://talk.ictvonline.org/

ISO 25964-1:2011. (2011). *Thesauri and Interoperability with Other Vocabularies, Part 1: Thesauri for Information Retrieval*. ISO/TC 46/SC 9.

ISO 25964-2:2013. (2013). *Thesauri and Interoperability with Other Vocabularies, Part 2: Interoperability with Other Vocabularies*. ISO/TC 46/SC 9.

Jadad, A. R., Haynes, R. B., Hunt, D., & Browman, G. P. (2000). The Internet and evidence-based decision-making: A needed synergy for efficient knowledge management in health care. *Canadian Medical Association Journal, 162*(3): 362–365. PMID:10693595

Keselman, A., Rosemblat, G., Kilicoglu, H., Fiszman, M., Jin, H., Shin, D., & Rindflesch, T.C. (2010). Adapting semantic natural language processing technology to address information overload in influenza epidemic management. *Journal of the American Society for Information Science and Technology, 61*(12), 2531-2543. doi:10.1002%2Fasi.21414

Khaleel, I., Wimmer, B. C., Peterson, G. M., Zaidi, S. T. R., Roehrer, E., Cummings, E., & Lee, K. (2020). Health information overload among health consumers: A scoping review. *Patient Education and Counseling, 103*(1), 15–32. doi:10.1016/j.pec.2019.08.008

Laato, S., Islam, A. K. M. N., Islam, M. N., & Whelan, E. (2020). Why do people share misinformation during the COVID-19 pandemic? *arXiv preprint arXiv:2004.09600*. Retrieved from https://arxiv.org/ftp/arxiv/papers/2004/2004.09600.pdf

Liyanage, H., de Lusignan, S., & Williams, J. (2020). *COVID-19 surveillance ontology*. Retrieved from https://bioportal.bioontology.org/ontologies/COVID19/

Mansoor, S. (2020). What's in a name? Why WHO's formal name for the new coronavirus disease matters. *Time*, February 11, 2020 5:55 PM EST. Retrieved from https://time.com/5782284/who-name-coronavirus- covid-19/

Neches, R., Fikes, R. E., Finin, T., Gruber, T. R., Senator, T., & Swartout, W. R. (1991). Enabling technology for knowledge sharing. *AI Magazine, 12*(3), 36–56.

Ng, M. Y., Lee, E. Y., Yang, J., Yang, F., Li, X., Wang, H., . . . Hui, C. K. M. (2020). Imaging profile of the COVID-19 infection: Radiologic findings and literature review. *Radiology: Cardiothoracic Imaging*, *2*(1). doi:10.1148/ryct.2020200034

Nielsen, R. K., Fletcher, R., Newman, N., Brennen, J. S., & Howard, P. N. (2020). *Navigating the 'Infodemic': How people in six countries access and rate news and information about Coronavirus.* Retrieved from https://www.politico.eu/wp-content/uploads/2020/04/Navigating-the-Coronavirus-infodemic.pdf

NISO Z39.19-2005. (2005) *Guidelines for the construction, format, and management of monolingual controlled vocabularies*. Retrieved from https://ils.unc.edu/courses/2015_fall/inls151_002/Readings/NISO.pdf

O'Malley, K. J., Cook, K. F., Price, M. D., Wildes, K. R., Hurdle, J. F., & Ashton, C. M. (2005). Measuring diagnoses: ICD code accuracy. *Health Services research*, *40*(5 Pt 2), 1620–1639. doi:10.1111/j.1475-6773.2005.00444.x

Ortutay, B. & Klepper, D. (2020, April 24). Virus outbreak means (mis)information overload: How to cope. *ABC News.* Retrieved from https://abcnews.go.com/Business/wireStory/virus-outbreak-means-misinformation-overload-cope-69734267

Park, H.A., Kim, H.Y., & Min, Y.H. (2012). Use of clinical terminology for semantic interoperability of electronic health records. *Journal of the Korean Medical Association, 55*(8), 720-728. doi:10.5124/jkma.2012.55.8.720

Paulheim, H. (2017). Knowledge graph refinement: A survey of approaches and evaluation methods. *Semantic Web*, *8*(3), 489–508. doi:10.3233/SW-160218

Qiu, H., Wu, J., Hong, L., Luo, Y., Song, Q., & Chen, D. (2020). Clinical and epidemiological features of 36 children with coronavirus disease 2019 (COVID-19) in Zhejiang, China: An observational cohort study. *The Lance: Infectious Diseases, 20*(6), 689–696. doi:10.1016/S1473-3099(20)30198-5

Reeves, J. J., Hollandsworth, H. M., Torriani, F. J., Taplitz, R., Abeles, S., Tai-Seale, M., . . . Longhurst, C. A. (2020). Rapid response to COVID-19: Health informatics support for outbreak management in an academic health system. *Journal of the American Medical Informatics Associatio,* ocaa037. doi:10.1093/jamia/ocaa037

Sahoo, S., Bharadwaj, S., Parveen, S., Singh, A. P., Tandup, C., Mehra, A., ... Grover, S. (2020). Self-harm and COVID-19 Pandemic: An emerging concern–A report of 2 cases from India. *Asian Journal of Psychiatry*. doi:10.1016%2Fj.ajp.2020.102104

Schema.org Community Group. (n.d.-a). *About Schema.org.* Retrieved from https://schema.org/docs/about.html

Schema.org Community Group. (n.d.-b). *Documentation for health/medical types*. Retrieved from https://schema.org/docs/meddocs.html

Schema.org Community Group. (2020). *Releases*. Retrieved from https://schema.org/docs/releases.html

Shaw, R., Kim, Y. K., & Hua, J. (2020). Governance, technology and citizen behavior in pandemic: Lessons from COVID-19 in East Asia. *Progress in Disaster Science, 6*. doi:10.1016/j.pdisas.2020.100090

Syzdykova, A., Malta, A., Zolfo, M., Diro, E., & Oliveira, J. L. (2017). Open-Source Electronic Health Record Systems for Low-Resource Settings: Systematic Review. *JMIR Medical Informatics*, *5*(4), e44. doi:10.2196/medinform.8131

Toffler, A. (1970). *Future shock*. New York, NY: Bantam Books.

Tudhope, D. & Koch, T. (2004). New applications of knowledge organization systems: Introduction to a special issue. *Journal of Digital Information, 4*(4). Retrieved from https://journals.tdl.org/jodi/index.php/jodi/article/view/109/108

W3C. (2009). *SKOS simple knowledge organization system reference*. W3C Recommendation 18 August 2009. Retrieved from www.w3.org.https://www.w3.org/TR/2009/REC-skos-reference-20090818/

WHO. (2015). *WHO best practices for naming of new human infectious diseases*. Retrieved from https://www.who.int/topics/infectious_diseases/naming-new-diseases/en/

WHO. (2020a). *Naming the coronavirus disease (COVID-19) and the virus that causes it.* Retrieved from https://www.who.int/emergencies/diseases/novel-coronavirus-2019/technical-guidance/naming-the-coronavirus-disease-(covid-2019)-and-the-virus-that-causes-it

WHO. (2020b). *Coronavirus disease (COVID-19) Pandemic.* Retrieved from https://www.who.int/emergencies/diseases/novel-coronavirus-2019

WHO. (2020c). *Global COVID-19: Clinical platform: Novel coronavirus (COVID-19): Rapid version*. Retrieved from https://www.who.int/publications-detail/global-covid-19-clinical-platform-novel-coronavius-(-covid-19)-rapid-version

Wikipedia. (n.d.). Health care. Retrieved from https://en.wikipedia.org/wiki/Health_care
Wilton, P. (2020, April 7). *How knowledge graph technology is helping Cochrane respond to COVID-19*. Retrieved from https://datalanguage.com/blog/how-knowledge-graph-technology-is-helping-cochrane-respond-to-covid-19
Wu, Z. & McGoogan, J. M. (2020). Characteristics of and important lessons from the Coronavirus disease 2019 (COVID-19) outbreak in China: Summary of a report of 72 314 cases from the Chinese Center for Disease Control and Prevention. *Journal of the American Medical Association, 323*(13), 1239–1242. doi:10.1001/jama.2020.2648
Zarocostas, J. (2020). How to fight an infodemic. *Lancet*, *395*(10225), 676. doi:10.1016/S0140-6736(20)30461-X
Zeleny, M. (1987). Management support systems: Towards integrated knowledge management. *Human Systems Management*, *7*(1), 59–70.
Zeng, M. L. (2008). Knowledge organization systems (KOS). *Knowledge Organization*, *35*(2-3), 160–182.
Zhou, F., Yu, T., Du, R., Fan, G., Liu, Y., Liu, Z., . . . Cao, B. (2020). Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: A retrospective cohort study. *Lancet, 395*(10229), 1054–1062. doi:10.1016/S0140-6736(20)30566-3

Appendix

List of Common Health KOS Standards (in alphabetical order)

KOS	Full title	Description	Provider	URL
CPT	Current Procedural Terminology	Offer doctors and health care professionals a uniform language for coding medical services and procedures to streamline reporting, increase accuracy and efficiency. CPT codes are also used for administrative management purposes such as claims processing and developing guidelines for medical care review.	American Medical Association	https://www.ama-assn.org/amaone/cpt-current-procedural-terminology
FDB	First Databank Drug Database	Provide clinical and descriptive drug knowledge that's integrated into healthcare information systems around the world.	First Databank	https://www.fdbhealth.com/
ICD	International Classification of Diseases	A diagnostic classification standard for all clinical and research purposes. ICD defines the universe of diseases, disorders, injuries and other related health conditions. (Previous name: International Statistical Classification of Diseases and Related Health Problems)	World Health Organization (WHO)	https://www.who.int/classifications/icd/en/
ICNP	International Classification for Nursing Practice	Provide an agreed set of terms that can be used to record the observations and interventions of nurses across the world.	International Council of Nurses (ICN)	https://www.icn.ch/what-we-doprojectsehealthicnp-download/icnp-download
LOINC	Logical Observation Identifiers Names and Codes	The international standard for identifying health measurements, observations, and documents, which may help the receiving facility to better understand the results and make appropriate treatment choices based upon the laboratory results.	Regenstrief Institute, Inc.	https://loinc.org
Medi-Span	Medi-Span drug databases	Used across the healthcare continuum to help inform medication-related decisions.	Wolters Kluwer	https://www.wolterskluwercdi.com/drug-data/why-medispan/
MeSH	Medical Subject Headings	A controlled and hierarchically organized vocabulary, which is used for indexing, cataloging, and searching of biomedical and health-related information.	National Library of Medicine (NLM)	https://www.nlm.nih.gov/mesh/meshhome.html
Multum	Multum MediSource Lexicon	A foundational database with comprehensive drug product and disease nomenclature information. It includes drug names, drug product information, disease names, coding systems such as ICD-9-CM and NDC, generic names, brand names, and common abbreviations.	Cerner Corporation	https://www.cerner.com/solutions/drug-database
NANDA-I	NANDA International	Facilitate the development, refinement, dissemination, and use of standardized nursing diagnostic terminology	International Nursing Knowledge Association	https://www.nanda.org/
NDC	National Drug Code	Serves as a universal product identifier for drugs. FDA publishes the listed NDC numbers and the information submitted as part of the listing information in the NDC Directory which is updated daily.	Food and Drug Administration (FDA)	https://www.fda.gov/drugs/drug-approvals-and-databases/national-drug-code-directory

KOS	Full title	Description	Provider	URL
NDF-RT	National Drug File - Reference Terminology (NDF-RT)	An extension of the VHA National Drug File (NDF). It organizes the drug list into a formal representation, which NDF-RT combines the NDF hierarchical drug classification with a multi-category reference model. Used for modeling drug characteristics including ingredients, chemical structure, dose form, physiologic effect, mechanism of action, pharmacokinetics, and related diseases.	U.S. Department of Veterans Affairs	https://www.nlm.nih.gov/research/umls/sourcereleasedocs/current/NDFRT/index.html
NIC	Nursing Intervention Classification	A comprehensive, research-based, standardized classification of interventions that nurses perform.	University of Iowa College of Nursing	http://www.nursing.uiowa.edu/cncce/nursing-interventions-classification-overview
NOC	Nursing Outcomes Classification	A standardized classification system of patient outcomes for evaluating the effects of nursing interventions.	University of Iowa College of Nursing	http://www.nursing.uiowa.edu/cncce/nursing-outcomes-classification-overview
RadLex	RadLex radiology lexicon	Provide medical imaging terms which are not found in other medical terminologies. It may unify terms used in radiology reports, bridge the terminology gap among radiologists, and promote radiological knowledge sharing.	RSNA	http://radlex.org/
RxNorm		Provide normalized names for clinical drugs and links its names to many of the drug vocabularies commonly used in pharmacy management and drug interaction software, including those in First Databank, NDF-RT, Micromedex, Gold Standard Drug Database, Multum, and U.S. Pharmacopeia (USP) Compendial Nomenclature.	National Library of Medicine (NLM)	https://www.nlm.nih.gov/research/umls/rxnorm/index.html
SNOMED CT	Systematized Nomenclature of Medicine – Clinical Terms	A global common language for clinical terms. The most comprehensive clinical terminology in use around the world.	SNOMED International	http://www.snomed.org/