Semantic Content in EHR Systems

Save to myBoK

By Dr. S.B. Bhattacharyya, MD, MBBS, MBA, FCGP, and Diana Warner, MS, RHIA, CHPS, FAHIMA

An electronic health record (EHR) can greatly help the timely delivery of efficient healthcare if the information it contains is adequate, accurate, and accessible in a timely manner. An EHR is a longitudinal record of an individual's health journey made up of electronic documents from various health information technology (HIT) systems. Consequently, the unified view of EHR content from various HIT systems needs to be available in a logical and meaningful manner to anyone using it.

The Need for Semantic Unity in Healthcare

This uniformity needs to be attained without any loss of meaning, which presents a significant challenge since the various health information systems that support EHRs are created by different vendors whose semantic content capabilities need to be harmonized in order to make the EHR appear as one. By providing a patient's medical history, EHRs help providers quickly and accurately decide on the best course of treatment. EHRs also assist in population health, disease surveillance, clinical research, and health services research, making effective and efficient care delivery a reality. Considering that the content of EHRs can be a matter of life or death—and that the information in them can be used and re-used for different purposes—the importance of effectively dealing with semantic content is immense.

The EHR system is a health information system that ensures the longitudinal collection of electronic health information about patients. It enables electronic capture, management, and access to person-level and population-level information. Each encounter generates the health information that represents a view of the clinical status of the individual at the episode of care. This information is aggregated from a variety of clinical documents that are created during the encounter, such as physician notes, provider orders for procedures and results of the procedures, medication administration records, nursing notes, public health reports, and more.

There are three general steps for bringing information together to generate the episode of care record. These include:

1. **Identify** all information belonging to one uniquely identifiable individual in various HIT systems—including data sources such as the EHR, laboratory, pharmacy, and other ancillary systems. This is achieved through the use of unique identifiers for all records' components belonging to a single patient EHR regardless of their origin or current location.
2. **Consolidate** the content from different data sources into one single episode of care record with a unified structure from semantic and syntactic perspectives. This is achieved by using standardized terminologies and data formats, as well as information models and structures across source systems.
3. **Assemble** the contents of records from different sources so that they appear as a single unambiguous record. This is achieved through using standardized clinical definitions templates developed from the clinical pathways documentation.

These three steps are critical to achieving semantic interoperability in healthcare.

Semantics Explained

People express themselves in words that they understand best. And the meaning behind those terms should be unambiguous for the listener or reader. Furthermore, in healthcare there is the need for content that can be interpreted in any language as health information is exchanged globally.

Semantic content is about the "inner" or inherent meaning of a statement. Due to the richness of languages, one can express the same thought in multiple ways. Moreover, adding a single word to a phrase can have a dramatic impact in different contexts. For example, a record entry stating that a patient "has tuberculosis" is completely different than "had tuberculosis." The case management for this patient is significantly impacted by this seemingly slight grammatical change. Electronically captured information has to retain its semantic meaning, unambiguously, throughout its lifetime.

Clinical records routinely contain:

- **Homonyms** like "cold," which can either mean the common cold or a particular sensory experience
- **Synonyms** like coryza and nasal catarrh that both mean "runny nose"
- **Eponyms** (names of individuals) like von Recklinghausen disease or Turner syndrome
- **Acronyms** like investigations Urine RE (routine examination) or USG (ultrasonography) or COPD (chronic obstructive pulmonary disease)
- **Abbreviations** like a Stat order or ObGyn department, with many such instances being purely local

It's important to remember that any of these may need to be translated into different languages without any loss in meaning. This adds to the complexity of ensuring the information is adequate, accurate, and accessible in a timely manner.

When making entries into a clinical record, clinicians rely on their clinical knowledge. They record their observations according to best practices and personal preferences. For some clinicians, where they were trained affects their documentation style. Anyone subsequently using the record must be able to interpret what the care provider meant. If the meaning is not clear, the interpretation may inadvertently cause the patient harm.

There are two ways to ensure the meaning of the information in clinical documentation. The first way is the "semasiological," or a definition-based approach, where the starting point is the word or term. The second way is the "onomasiological," or concept-based approach, where the starting point is the concept. For clinical information, the concept-based approach is more practical, particularly when considering the issues related to ambiguity and language. Onomasiological (concept-based) approach relates the terms to concepts that can be expressed unambiguously and be interpreted the same way in any language through its support for Unicode Transformation Format-8 (UTF-8) encoding. The Systematized Nomenclature of Medicine – Clinical Terms (SNOMED CT) takes the second approach. SNOMED CT text files are encoded using UTF-8 which allows for the terms to be linked to concepts and interpreted in other languages without losing the meaning.

Using the SNOMED CT Standard

Owned, maintained, and distributed by the International Health Terminology Standards Development Organization (IHTSDO), SNOMED CT supports the development of comprehensive high-quality clinical content in health records. It provides a standardized way to represent clinical phrases captured by the clinician and enables automatic interpretation of these phrases. SNOMED CT is a clinically validated, semantically rich, controlled vocabulary that facilitates evolutionary growth in expressing clinical concepts when every concept means something unambiguous and distinctive. For example, "fever" is considered to be a "body temperature above reference range" and "interprets body temperature."

With more than 300,000 concepts, SNOMED CT covers a vast array of healthcare-related terms from A to Z (from abscess to zygote). Over one million human readable terms are used to describe the concepts, at an average of three terms per concept, making it easy for users to find the particular term closest to the thought being expressed. Expressions that best describe the observations are recorded using either one term or by combining several terms together as per compositional grammar rules.

Through the use of its description logic and machine-readable codes, SNOMED CT enables all instances of clinical data to be logically organized and interpreted by automated clinical systems. By way of robust support for query/retrieval, health information systems that deploy SNOMED CT facilitate efficient harnessing of clinical information in real time to provide relevant information in the process of care.

References

Benson, Tim. *Principles of Health Interoperability HL7 and SNOMED*. New York: Springer Science & Business Media, 2010.

Bhattacharyya, S.B. *Introduction to SNOMED CT.* Singapore: Springer Singapore, 2016, p. 19-23.

International Health Terminology Standards Development Organisation. "Guidelines for Translations of SNOMED CT." December 11, 2012. www.ihtsdo.org/snomed-ct/snomed-ct-worldwide/translations-of-snomed-ct.

International Health Terminology Standards Development Organisation. “SNOMED CT Starter Guide.” December 2014. http://ihtsdo.org/fileadmin/user_upload/doc/download/doc_StarterGuide_Current-en-US_INT_20141202.pdf?ok.

International Health Terminology Standards Development Organisation. “SNOMED CT Technical Implementation Guide. January 2015 International Release (US English Version).” http://ihtsdo.org/fileadmin/user_upload/doc/download/doc_TechnicalImplementationGuide_Current-en-US_INT_20150131.pdf?ok.

NEHTA Education. “SNOMED CT Overview - Part 1.” YouTube. November 10, 2008. www.youtube.com/watch?v=CrZ46eXunzo&list=PL4DE2FA9DA297015F.

Dr. S.B. Bhattacharyya (sbbhattacharyya@gmail.com) is a practicing physician, head of health informatics at Tata Consultancy Services, and a member of the National EHR Standards Committee, Ministry of Health and Family Welfare, Government of India. Diana Warner (diana.warner@ahima.org) is a director of HIM practice excellence at AHIMA.

Article citation:
Bhattacharyya, S. B; Warner, Diana. "Semantic Content in EHR Systems" *Journal of AHIMA* 87, no.6 (June 2016): 38-39.