Data Mapping of Clinical Terminologies, Classifications, and Ontologies

By James R. Campbell, MD, FACP, FACMI

Editor's Note: This article is an excerpt from Chapter 17 of Healthcare Code Sets, Terminologies, and Classification Systems, Fourth Edition, published by AHIMA Press.

In 2007, the Institute of Medicine set a vision for the evolution of the healthcare record in the digital age.[1] That vision, Learning Healthcare System, described a future in which data recorded in an electronic health record (EHR) would support clinical care and decision-making. That data would then be securely re-used to improve the process of healthcare delivery, clinical research, and the needs of public health for population health management in a continuous cycle of process improvement. To achieve that vision, data in the EHR should be managed to support semantic interoperability, expecting that a data item can be transmitted to a different computer system, recognized for its meaning or semantics, and re-used by that computer system without human intervention. Although the Office of the National Coordinator for Health Information Technology (ONC) has specified a plan for achieving interoperability in the United States, data re-use still faces major challenges due to inconsistent and incomplete terminologies, fragmented information architectures, and proprietary information management schemes.[2]

Data Maps

One common approach to the re-use of data from one computer system to another is to employ a data map, sometimes called a cross map. The International Standards Organization (ISO) defines a data map as the creation of links from concepts or terms in one classification, terminology, or ontology to another for some purpose of data re-use.[3] Maps are built to meet specific business requirements and serve one or more use cases, activities, or step-wise descriptions of events that describe real-world employment and expected outcomes resulting from a piece of software or data map, for re-use. The links that are used in constructing the map are driven by the use case and the editorial features of the two terminologies. For semantic (the meaning of a concept) interoperation, an equivalence between the two concepts would be desirable, but this is not always achievable. Theoretically, identifying computable semantic equivalence is only reliable between two well-constructed ontologies with overlapping semantic domains—they apply to the same types of concepts such as diagnoses—and also share a concept model, or "the set of rules that determines the permitted sets of relationships between particular types of concept."[4] Since this is seldom true, the majority of maps have a directionality to the link, mapping from a concept usually in the source terminology from which the data originates to a concept in the target terminology, and the terminology to which the concept links. This helps employ a type of link or relationship which serves the pragmatic purpose of the use case. Maps can be developed, published, and shared as simple tables identifying pairs of concepts from source and target, "triples" in Resource Description Framework format constructed as records linking source, relationship (link type), and target, or as a knowledge resource which might consist of a network of IF-THEN rules supporting context-based mapping.

Mapping Purposes

Maps may be employed throughout the healthcare domain for a variety of purposes and to meet many different use cases. ISO suggests that there are three general goals or purposes for data mapping:

1. Converting legacy data to newer coding schemes
2. Re-using data for purposes other than originally intended
3. Semantic interoperation of data where computer syste are using different coding schemes. (See Table 17.1.)

Table 17.1 Purposes of data map creation and interoperation of clinical terms

Table B	Use Case	Example
Data conversion	Map EHR data from a legacy coding system to new standards	Convert coded diagnosis records from ICD-9-CM to ICD-10-CM: General Equivalence Mappings (GEMs)
Re-use	Choose medication records from the EHR corresponding to orders with a specific treatment reason	RXCLASS mapping of RxNorm medications to VA drug classes to identify cancer therapies
Re-use	Map SNOMED CT conditions to ICD-10-CM for reporting morbidity and mortality	NLM release of SNOMED CT – ICD-10-CM rules-based map
Interoperability	Extend and integrate terminology standards across domains for interoperation	The SOLOR project

Types of Maps

Historically, data maps have been constructed employing a variety of approaches and can serve many different use cases. There is no standard for classifying data maps, although the ISO technical document does specify a set of quality indicators as noted in the following section. A data map may be described minimally based upon the protocols involved in creating map records, the nature of the link used in the map, or by the characteristics of the terminology resources that are linked. Therefore, a descriptive map type may be assigned as a list of one of each of the following descriptive sets.

Mapping Protocol:

- Lexical mapping: Identifies sets of terms to be linked by comparing words or character strings and is very dependent upon the language and dialect employed. The National Library of Medicine's (NLM) MetaMap program supports mapping of biomedical text to concepts in the Unified Medical Language System (UMLS) and is an example of a lexical approach.
- Semantic mapping: Employs the full defined meaning of a concept in a source reference terminology or ontology to identify equivalent or related concepts in the target. A reference terminology is a controlled

terminology employing a set of terms and relationships which capture and define the meaning of each concept. The map of SNOMED CT to ICD-10-CM discussed in the use cases later in the chapter is an example of a map that was constructed semantically.

Nature of the Link Relationship:

- **Equivalence maps** are designed to identify pairs of identical concepts between the source and target. Semantic equivalence and lexical equivalence can lead to very different types of map sets that may not serve the same use cases well.
- **Hierarchical (Taxonomic) maps** order sets of concepts as having more or less specificity when traversing from source to target. They employ link relationships such as "less-than," "child-of," "more-specific than," or "is-a." A classification such as ICD-10-CM includes such implicit relationships which specifies its hierarchical structure.
- **Knowledge-based (Rules-based) maps** employ features of coding context, business case, and workflow assumptions, which apply to the interpretation and application of map records, in order to create maps for terminologies where the standards developer imposes constraints on the use of the terminology scheme. The best example of a knowledge-based map is the map of SNOMED CT to ICD-10-CM in which World Health Organization (WHO) classification guidelines and additional context from the patient demographics and problem list are employed by a set of rules to map each SNOMED CT concept recorded as a problem (condition) in the EHR.

Characteristics of the Source and Target:

Maps may be typed based upon the nature of the terminologies involved in the map records, which may include classifications, controlled terminologies, reference terminologies, or ontologies. A map such as the SNOMED CT to ICD-10-CM map may be described as "reference terminology to classification."

Notes

1. Olsen, LeighAnne, Dara Aisner, and J. Michael McGinnis. *Institute of Medicine Roundtable on Evidence-Based Medicine. The Learning Healthcare System Workshop Summary*. Washington, DC: The National Academies Press, 2007. http://www.ncbi.nlm.nih.gov/books/NBK53494/.
2. Office of the National Coordinator for Health Information Technology (ONC). "Connecting Healthcare for the Nation: A Shared Nationwide Interoperability Roadmap." 2017. https://www.healthit.gov/sites/default/files/hie-interoperability/nationwide-interoperability-roadmap-final-version-1.0.pdf.
3. International Standards Organization (ISO). "ISO/TR 12300:2014 Health Informatics – Principles of Mapping between Terminological Systems." 2014. https://www.iso.org/standard/51344.html.
4. SNOMED International. "SNOMED CT Editorial Guide." 2019. https://confluence.ihtsdotools.org/display/DOCEG/SNOMED+CT+Editorial+Guide.

James R. Campbell is a board-certified physician and a professor of medicine at the University of Nebraska Medical Center.

Article citation:
Campbell, James R.. "Data Mapping of Clinical Terminologies, Classifications, and Ontologies" *Journal of AHIMA* 90, no.8 (August 2019): 46-47.