Journal of Biomedical Informatics 45 (2012) 674–682

Contents lists available at SciVerse ScienceDirect

Journal of Biomedical Informatics

journal homepage: www.elsevier.com/locate/yjbin

Using LOINC to link 10 terminology standards to one unified standard in a specialized domain

Philip J. Kroth [a,*], Shamsi Daneshvari [a,b], Edward F. Harris [c], Daniel J. Vreeman [d], Heather J.H. Edgar [a,b]

[a] University of New Mexico, Health Sciences Library and Informatics Center, MSC09 5010, Albuquerque, NM 87131-0001, United States
[b] University of New Mexico, Department of Anthropology, MSC01 1050, Albuquerque, NM 87131-0001, United States
[c] Department of Orthodontics, University of Tennessee, 875 Union Avenue, Memphis, TN 38163, United States
[d] Regenstrief Institute, Inc., 410 W. 10th Street, Suite 2000, Indianapolis, IN 46202, United States

ARTICLE INFO

Article history:
Received 21 July 2011
Accepted 11 October 2011
Available online 19 October 2011

Keywords:
Cephalometry
Logical observation identifiers names and codes
Terminology as topic
Systems integration
Information storage and retrieval

ABSTRACT

Despite the existence of multiple standards for the coding of biomedical data and the known benefits of doing so, there remain a myriad of biomedical information domain spaces that are essentially un-coded and unstandardized. Perhaps a worse situation is when the same or similar information in a given domain is coded to a variety of different standards. Such is the case with cephalometrics – standardized measurements of angles and distances between specified landmarks on X-ray film used for orthodontic treatment planning and a variety of research applications. We describe how we unified the existing cephalometric definitions from 10 existing cephalometric standards to one unifying terminology set using an existing standard (LOINC). Using our example of an open and web-based orthodontic case file system, we describe how this work benefited our project and discuss how adopting or expanding established standards can benefit other similar projects in specialized domains.

1. Introduction

The benefits of using standardized terminologies for the coding of biomedical information are well described [1]. Perhaps equally well described are the difficulties, costs, complexities, and limiting historical precedents that often conspire to make it difficult to do so [2,3]. Indeed, the existence of the perfectly coded biomedical data set may be more of a theoretical concept to aim for rather than an obtainable goal. The existence of a plethora of terminology standards illustrates the complexity of the task. The National Library of Medicine's Unified Medical Language System (UMLS)® is a set of tools that attempts to catalog and where possible, inter-relate the over 1 million biomedical concepts from over 100 source terminology standards. Despite the standards, powerful tools, and recognized benefits of doing so, the coding of biomedical data sets using recognized standard terminologies remains a daunting challenge with only limited success when compared with the size of and number of biomedical data sets that could benefit from such standardization.

As part of a project to create an open, de-identified, web-based case file system in support of orthodontic treatment, training, and research [4], we came upon a biomedical domain ripe for standardization; that of cephalometric measurements.

Cephalometrics are measurements (angles and linear distances) defined by anatomical points as measured on X-ray film of the human head taken in a standardized manner. Some authors credit the scientist and painter Petrus Camper (1722–1789) or even earlier artists for originating cephalometric studies [5]. Most trace modern cephalometry to the efforts of Herbert Hofrath (1899–1952) in Germany and B.H. Broadbent, Sr., (1894–1977) in the United States, who simultaneously introduced radiography to the standardized study of facial skeletons and teeth. Broadbent, working with T. Wingate Todd (1885–1938) developed the "roentgenographic craniometer" which produced X-rays films of living human heads in a standardized manner, allowing accurate and reproducible scientific measurements [6]. This device was used as part of the Bolton-Brush Growth Study, a large, prospective, longitudinal cohort study of children intended to establish a detailed cephalometric data set on normal human growth and development [7]. Over time, cephalometric techniques have been elaborated and refined, and have become commonplace tools in growth studies as well as orthodontic training and practice. Today, while some question the value of cephalometric data in orthodontic treatment planning [8] several computerized systems exist to allow orthodontic practitioners to compare their patients' measurements with the data collected in the aforementioned growth studies, e.g. Dolphin Imaging [9] and CephX [10]. Additionally, cephalometric data has been used in a variety of other applications, including studies of facial attractiveness [11] and forensic anthropology [12,13].

* Corresponding author. Fax: +1 505 272 8254.
E-mail address: pkroth@salud.unm.edu (P.J. Kroth).

1532-0464/$ - see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.jbi.2011.10.003

Because we understand impact of radiation on human growth and development, especially on young children, the data from these studies could not ethically be reproduced today. Despite the broad availability of modern three dimensional imaging technologies, cephalometrics still continue to be useful despite the fact they are measurements of the three-dimensional human head as projected onto a two-dimensional X-ray film. Since these data are the foundation of what is now known about normal human skeletal growth and development, much contemporary research continues to actively reference these valuable data.

The advent of cephalometry as a scientific tool spawned a number of studies and analyses of cephalometric data sets. As the number of studies increased and as time went on, the terminology that defined the cephalometric landmarks and measurements diverged until at present, the same measurement may be defined in up to five different cephalometric "atlases" or "analyses".

One of the design requirements of the orthodontic case files database we developed was to make it possible for additional cases to be donated from other institutions or orthodontics practices. It became clear that to accept a donated case that included cephalometric data, a major step in the accession process would be to determine what the cephalometric measurements actually are. If, for example, a case used a given standard, it would be necessary to record that fact and then convert certain measurements to whatever standard our project would choose as *the* standard. Simply choosing one of the existing standards for our project would be arbitrary because each has its own proponents and yet none comprehensively represents all existing cephalometric measurements.

We also see the case file system's database could be used in the future as a source for web-based cephalometric analysis tools or perhaps even linked to other similar databases. To achieve these functionalities the data will have to be structured in such a manner to support semantic interoperability between the case file database and any cephalometric analysis tools or other databases. The most efficient way to achieve semantic interoperability would require one grand unifying terminology standard. Like any good terminology standard, the grand unifying standard would also require an infrastructure for expansion, revision, and support. In summary our problem required a unified cephalometric standard that would satisfy the following requirements:

(1) All cephalometric measurements in the existing standards can be directly mapped or represented in the unified standard. This would help solve the problem of managing donated cases with cephalometric data coded to different standards. By asking the donor to which standard the cephalometric measurements were made, all measurements can be directly mapped to the unifying terminology, allowing for a greatly simplified data representation in the database.
(2) The unified standard must have a robust process for continuous improvement, maintenance, and expansion. As a starting point, our project used the cephalometric measurements used by one forward thinking orthodontist in Albuquerque, NM [14–16]. However, we anticipate expansion of the number of cephalometric measurements as new, and perhaps more exotic cases are donated to the collection over time.
(3) The unified standard must be authoritative, internationally recognizable, and as widely known as possible. We did not want to create yet another standard that would have little credibility.
(4) The unified standard must have tools and support to assist in the discovery of existing terms that are either already defined or come close to defining what specific cephalometric measurements we want to standardize. Such tools will assure that users of the standard can find the right terms and facilitate efficient maintenance of the existing corpus of terms.

Because of these requirements we chose to use the Logical Observations Identifiers Names and Codes (LOINC®) standard [17] as the unifying standard for the cephalometric measurements contained in our collection of cephalometric measurements. LOINC was first released by the Regenstrief Institute in April of 1996 [18]. Since its inception, LOINC has been developed as an open standard and freely distributed worldwide. LOINC is a terminology standard focused specifically on providing observation identifiers. The development of LOINC is divided into two divisions: the Laboratory division focuses on the observations and measurements that can be made on specimens withdrawn from the body, and the Clinical division focuses on the observations and measurements that can be made on patients. The most recent LOINC release (Version 2.36, June 2011) contains 65,003 terms, of which 45,428 are laboratory terms and 19,575 are clinical terms [19].

LOINC has been widely adopted in both the public and private sectors, within the United States and more than 140 other countries. Several countries (including Brazil, Canada, Germany, the Netherlands, Mexico, and Rwanda) have adopted LOINC as a national standard, and there are large health information exchanges using LOINC in Spain, Singapore, and Korea as well. Within the US, LOINC has been adopted by many health information exchanges, large national reference laboratories, healthcare organizations, insurance companies, research programs, and national standards. LOINC has long been a source vocabulary included in the National Library of Medicine's Unified Medical Language System. Furthermore, the Department of Health and Human Services adopted LOINC as the standard across federal agencies for laboratory result names, laboratory test order names, and federally required patient assessment instruments.

LOINC is distributed at no cost from its website (http://loinc.org) as a database that contains the LOINC codes, term names, and many other attributes such as synonyms, alternate display names and example units of measure. New versions of LOINC are published twice yearly (typically June and December). In addition, The Regenstrief Institute develops and distributes at no cost a software program called RELMA (the *Re*genstrief *L*OINC *M*apping *A*ssistant) that helps browse and search the LOINC database, review the detailed accessory content for each term, and map local terms to LOINC.

Because of its open development approach, new content additions to LOINC are welcomed, and added based on submissions from the worldwide LOINC user community. There is a well-defined mechanism for creating new term requests (documented in the LOINC User's Guide) and rigorous quality review process that ensures that LOINC follows best practices for terminology development [2]. Requests for new content in domains not previously modeled in LOINC are discussed at one of the public LOINC Committee meetings that occur regularly.

We previously published a detailed rationale for choosing LOINC over a number of existing terminology standards for representation of cephalometrics [20]. This analysis considered representing cephalometrics using MeSH (Medical Subject Headings) [21], SNOMED-CT (Systematized Nomenclature of Medicine–Clinical Terms) [22], ICD-10 (International Classification of Diseases, 10th Edition) [23], LOINC [19], CDT (Current Dental Terminology) [24], CPT (Current Procedural Terminology) [25], ICD-DA (Application of the International Classification of Diseases to Dentistry and Stomatology) [26], SNODENT (Systematized Nomenclature of Dentistry – a subset of SNOMED), UMLS [27], HL-7 (Health Level Seven) [28], and DICOM (Digital Imaging and Communications in Medicine) [29]. This analysis involved looking at content coverage, scope, hierarchical structure, and term composition of each standard. For example, with regard to hierarchical structure, most medical vocabularies are arranged in hierarchies [2,3]. Though cephalometric measurements can be categorized by dimensionality (angles vs.