

We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900

Open access books available

186,000

International authors and editors

200M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



Terminology Services: Standard Terminologies to Control Medical Vocabulary. “Words are Not What they Say but What they Mean”

Daniel Luna, Carlos Otero, María L. Gambarte and Julia Frangella

Additional information is available at the end of the chapter

<http://dx.doi.org/10.5772/intechopen.75781>

Abstract

Data entry is an obstacle for the usability of electronic health records (EHR) applications and the acceptance of physicians, who prefer to document using “free text”. Natural language is huge and very rich in details but at the same time is ambiguous; it has great dependence on context and uses jargon and acronyms. Healthcare Information Systems should capture clinical data in a structured and preferably coded format. This is crucial for data exchange between health information systems, epidemiological analysis, quality and research, clinical decision support systems, administrative functions, etc. In order to address this point, numerous terminological systems for the systematic recording of clinical data have been developed. These systems interrelate concepts of a particular domain and provide reference to related terms and possible definitions and codes. The purpose of terminology services consists of representing facts that happen in the real world through database management. This process is named Semantic Interoperability. It implies that different systems understand the information they are processing through the use of codes of clinical terminologies. Standard terminologies allow controlling medical vocabulary. But how do we do this? What do we need? Terminology services are a fundamental piece for health data management in health environment.

Keywords: terminology server, interface vocabulary, controlled vocabularies, semantic interoperability, standard terminology

1. Introduction

Recently, major healthcare stakeholders around the world have emphasized on the importance of establishing electronic health records (EHR) for all health care institutions. Their goals for doing so include increasing patient safety, reducing medical errors, improving efficiency and reducing costs [1, 2]. Everyday practical data entry, presentation and document retrieval for clinical tasks must be taken into account, so that the differences between the needs of users and the needs of available software's are addressed. Data entry is an obstacle for the adoption of EHR with structured data method and the acceptance of healthcare providers, who prefer to document healthcare findings, processes and outcomes using unfettered "free text" or narrative text in natural language [3]. Natural language is huge and very rich in details but at the same time ambiguous, having great dependence on context, it uses jargon and acronyms and it lacks of rigorous definitions.

1.1. The importance of narrative

Free text narrative formats allow physicians to share complex ideas in an efficient and effortless manner. Its use in electronic health records allows them to synthesize facts and to point a full picture rich with meaning that it might be easily interpreted by other health care providers [3]. Between physicians' register motivation, the main one is their own use of the information. Many current systems that provide EHRs use template-based system in order to capture structured data elements in databases. Structure data entry does not support the expressiveness and flexibility to which clinicians are accustomed, and it can be difficult to interpret and reconstruct meaning from structure data due to loss of contextual information [4]. To represent medical knowledge, it is necessary to represent patient's data from different sources including: problem list and sometimes progress notes, procedures, medication list, labs and complementary tests results, social determinants of health environmental information, people's decisions about health and medical treatments, genomics and proteomics, etc. As a result, ambiguities must be resolved and vocabulary standardized.

1.2. The need of a standard codification system

To accomplish these, EHR should capture the clinical data in a structured and preferably coded format. Looking at the definition of codifying, we found "To reduce to a code" [5]. Codes are usually numeric or alphanumeric. In order to represent facts that happen in the real world to be managed in a database, the need of a standard codification system (SCS) arise. Evans et al. stated that the medical community required a "common, uniform, and comprehensive approach to the representation of medical information" [6].

This SCS should be able to capture clinical findings, index medical records, index medical literature and represent medical knowledge, etc. Provided that possible, the codification should be one-to-one: one term should only exist for a given object. Each term should describe only one object. The aim is to avoid ambiguity through polysemy or homonymy [5].

In fact, many SCS have been proposed but their adoption has been slow and incomplete. System developers generally indicate that, while they would like to make use of standards, they cannot

find one that meets all their needs. Each author who expresses a need for a controlled vocabulary does so with a particular purpose in mind, so there are also multiple characteristics that it should accomplish [7, 8]. Because of all the reasons mentioned before, for a long time, there has been a discussion regarding the use of free text versus structured text for data entry in EHR that later must be codified. Free text has the advantage of allowing health care providers to express themselves freely, but as disadvantage it has the need for an arduous codification process to allow further analysis. Structured text allows a quick codification process but has the disadvantage of being time consuming for the physician and contains expressions to the level of detail of the selected entry terminology [9]. It has been suggested that tension between clinical usability and meticulous knowledge representation may result from a fundamental conflict between the needs of humans and those of computer programs that use terminologies [3].

1.3. Primary coding versus secondary coding

Ideally, clinical data should be coded by the practitioner at the time of the consultation, which is known as primary coding, so they can utilize their knowledge of the patient situation while being aware of the limitations set down by the selected classification system [10]. However there are several practical difficulties in setting primary coding. It is time consuming for the practitioner and requires major efforts in their training, to ensure that the same code would be chosen in the same situation by different physicians [10]. It also limits the physician's expression in the registry and creates high levels of resistance in its use. Enforcing mandatory as opposed to optional modifier codes results in lower rates of incomplete coding [11]. One answer to this problem is centralized secondary coding, where a reduced number of trained persons codify the narrative text recorded by the physicians taking care of the patients. Centralized secondary coding by non medical coders had proved to be reliable and can be used for coding medical problems from an electronic problem-oriented medical record [10]. As regards as the coding tool, manual coding versus computerized coding, it has been demonstrated that the use of a computerized coding tool can save time and result in higher quality coding. A study that compares both of them had shown that manual coding takes 100% longer [11]. It is fundamental to contemplate that time spent on coding may be underestimated when we look at individual coding times instead of looking at the whole task of processing a clinical scenario [11]. The completeness of coding had also been demonstrated that can be improved using a computerized coding tool [11]. A step-forward option is to achieve text-autocoding, allowing free or narrative text entry together with dynamic interaction of the information system at the time of entering data.

As a result, the challenge consists on finding the complex balance between the freedom of use of free text and the benefits of structured text for data entry in EHR. In order to answer to this need, interface terminology and a terminology server were developed. It is crucial to highlight that communication is successful only if the sender and the receiver know both, the language (code) and the context. Notice the importance of the context, which must be read in an identical manner for both parties involved [5].

1.4. But again, why do we need to codify in electronic health records?

Some of the aims to coding in EHR are:

- to support health services research: this system can promote quality of care by providing a link to medical knowledge and current publications that can be used for outcome measurement.
- to enable decision support programs use at the point of clinical care: a computer-based EHR system might work with a diagnostic expert system to backing physicians’ decisions. In order to achieve optimal integration, the transference of patient information from EHR to the diagnostic expert system would need to be automated. The major barrier to do so, are the variance between the controlled vocabularies of the two systems [7, 8, 12].
- to exchange data between health information systems: the concept of Semantic Interoperability arises. We defined it as the possibility of different systems to understand the information they are processing through the use of codes of clinical terminologies.
- for epidemiological analysis: it can be used by patients, physicians, researchers, quality control and management personnel and other administrative functions like accounting, billing and coding personnel.
- for the process of codifying medical information systems actually count with vocabularies (artifacts that describe and systematize meanings of terms), with the common distinctions between terminologies (which provide standardized meanings), thesauri (which introduce semantic relations between groups of terms) and classifications (which introduce exhaustive partitions for statistical purposes). Some of them are used in an international level, while others have been defined according to local needs. (for more information, see **Table 1**)

Terminology	Collections of words or phrases, called terms, aggregated in a systematic fashion to represent the conceptual information that makes up a given knowledge domain.
Classification System	Intended for classification of clinical conditions and procedures to support statistical data analysis across the healthcare system.
Thesaurus	List of terms created from free text inputs extracted from the clinical data repository. The terms included in the thesaurus are divided into concepts (real clinical entities) and descriptions (different ways of naming these clinical entities). The thesaurus has capabilities to reject invalid terms already flagged as not appropriate for the intended use [9].
Clinical coding	Designating descriptions of diseases, injuries and procedures into numeric or alphanumeric designations. It involves the use of a EHR as the source for determining code assignment [19]. Primary coding: clinical data are coded by the practitioner at the time of the consultation [10]. Secondary coding: a reduced number of trained persons codify the narrative text recorded by the physicians [10].
Semantic Interoperability	It refers to human interpretation of the content. There is a common comprehension among people about the meaning of the information that is being exchanged (correct interpretation is guaranteed, for this reason formal definitions of each entity, attribute, relationship, restriction and exchanged term are needed [20].

Table 1. Introduction’s definition summary.

While many terminologies have been developed, no single terminology has been accepted as a universal standard for the representation of clinical concepts. By contrast, individual terminologies or components have been identified by standards organizations as candidates for specific uses [13]. The recommended terminologies include the Systematized Nomenclature of Medicine Clinical Terms (SNOMED CT); Logical Observation Identifiers Names and Codes (LOINC) and Unified Medical Language System (UMLS) [14], between many other. These will develop in the following sections [15, 16].

In the nineteenth century, the advancement of clinical pathology and technology changed the framework of classification, moving emphasis from the patient's experience to phenomena determined by physician using diagnostic procedures [17].

The diagnostic entities in medicine are changing as a consequence of expanded and revised knowledge of the functions of the human body. Techniques for differential diagnostic strategies contribute to new categories, while old labels are gradually abandoned. There is a need to acknowledge the potency of classification systems as dynamic tools for medical practice and research [17]. According to all these changes, during twentieth century, the importance of "concept orientation" in terminology construction arises. Concept orientation allows a terminology to be helpful in several situations, depicted in different languages and easily evaluated for quality [18]. This transition from the use of Classification Systems to Reference Terminology was not only a change in institution's choices, but also both of them defined their purposes, potential functions, strength and limitations.

In the following sections, we will briefly present Classifications Systems, Reference Terminology and Interface Terminology. Finally, we will present our experience developing and implementing our Terminology Server.

2. Classification systems

A classification is "a system that arranges or organizes like or related entities" [21] (for more information, see **Table 1**). Classifications provide a useful framework for a systematic representation and codification of medical concepts. Monoaxial classifications form a hierarchy of terms based on a common root. The most commonly used example of monoaxial hierarchical classifications is the International Classification of Diseases, Tenth Revision, Clinical Modification and International Classification of Diseases, Procedure Coding System (ICD-10-CM/PCS), published by the World Health Organization, represents an example of the clinical classification systems. It has been designed for providing outputs in terms of reports and statistics [5, 22, 23]. Multi-axial or multifaceted classifications combine terms belonging to different classes that themselves may be organized in a hierarchy. SNOMED is an example of this type of classification [5]. Classification systems are intended for classification of clinical conditions and procedures to support statistical data analysis across the healthcare system. They are mutually exclusive and exhaustive and they can provide standards for comparisons of health statistics at national and international levels. They have been used:

- to support other applications in healthcare including reimbursement,
- for public health reporting,
- to improve quality of care assessment,
- education,
- research,
- performance monitoring [21–23].

2.1. International classification of diseases background

Work on classification systems began in the middle of the seventeenth century with John Gaunt's refinement of the late sixteenth-century classification scheme for the London Bills of Mortality [24, 25]. International Classification of Diseases (ICD) was first adopted in Paris in 1900 [24, 25]. Architecturally, the ICD has not fundamentally changed from the sixteenth century model of the London Bills, in that each new code is added as a new row in a single list. United States did not choose to adopt ICD-10 until the end of our 25-year window, in 2015. Besides, at the time that ICD-10 was introduced, it stayed as paper book, ICD-10 was not published in electronic format [26].

In Australia, The National Center for Classifications in Health at the University of Sydney, was the first group to migrate ICD-10 into an electronic format [26].

By 2005, the WHO-FIC, an organization chartered by the World Health Organization (WHO) comprising national centers for classification around the world, created an international forum to advise on the content and evolution of WHO's Family of International Classifications (WHO-FIC) which includes ICD and the International Classification of Functioning (ICF). Currently, the WHO manages an electronic revision and update platform with WHO-FIC as a web page [26]. ICD-10-CM/PCS is an output system that was designed for general reporting purposes, public health surveillance, administrative performance monitoring, and reimbursement of healthcare services [22].

The ICD was developed to code death certificates but its use was extensive to include a large range of statistical reporting. ICD-10 has been used since the 1990s to collect mortality statistics around the world. The WHO defines coding as "the translation of diagnoses, procedures, co-morbidities and complications that occur over the course of a patient's encounter from medical terminology to an internationally coded syntax" [21]. According to this definition, ICD system has capability of being used for clinical coding and classification to enable international comparisons as regard to mortality and morbidity statistics [27].

Professional coders, who used to manually assign codes to patients' diagnoses and procedures, performed ICD-10-CM/PCS coding. Nowadays, coders use computer-assisted coding applications. These applications can facilitate accurate and efficient coding by automatically

suggesting codes based on the clinical documentation in the EHR system. Thus, ICD-10-CM/PCS coding is semi-automated at best and requires human intervention to either assign or validate selected codes [27].

2.2. Diagnosis-Related Groups (DRGs)

With the advent of capitated payments, the inevitable need of how to objectively determine severity of illness, in order to appropriately adjust capitated payments, or case mix. As outlined above, traditional disease classifications such as the ICD did not enjoy explicit severity of illness parameters; all that could be done was to infer disease severity on the basis of co-morbidity [26]. However, there may be no evidence demonstrating causality between the condition of interest and the co-morbidities. Case mix required some objective metrics and co-morbidity was it [26]. The set of measures for co-morbidity found everywhere has been the Diagnosis-Related Groups (DRGs) [28]. Since their beginning, multiple versions have continued to come out, changing architecturally combining demographic, diagnoses, and procedures into several hundred categories of care. These categories can in turn be considered to have, or not have, "complications" [26]. The 11th version, ICD-11, is now being developed through a continuous revision process, it will be finalized in 2018. For the first time, through advances in information technology, public health users, stakeholders and others interested can provide input to the beta version of ICD-11 using an online revision process. Peer-reviewed comments and input will be added through the revision period. When finalized, ICD-11 will be ready to use with EHR and information systems. WHO encourages broad participation in the 11th revision, so that the final classification meets the needs of health information users and is more comprehensive [27].

2.3. ICD: strengths and limitations

ICD's strengths include non-redundancy, meaning by this that each concept should only be expressed in one way. If two terms refer to the same concept, the sensitivity of the replies to database queries will be reduced. The ability to manage synonyms, this is important because allows the presence of authorized intermediate terms that refer to a unique term used to encode, index, and find the useful information. And finally, there are explicit relationships this refers to the types of relationships between terms in a nomenclature are clear [5].

As regard as SCS's limitations, we can name the completeness and the non-ambiguity. A full description of the medical vocabulary is very hard to achieve. About non-ambiguity, if two different types of data are stored under the same term, the specificity would be affects [5].

2.4. Others standard classification systems

Currently, many classification systems exist and are maintained by responsible agencies. Next, in **Table 2** we will briefly name and describe some of them.

ICD	The ICD is the global health information standard for mortality and morbidity statistics. ICD is increasingly used in clinical care and research to define diseases and study disease patterns, as well as manage health care, monitor outcomes and allocate resources [29].
DRG	Statistical system for classify all inpatient stay into groups for the aim of payment. The DRG classification system divides possible diagnoses into more than 20 major body systems and subdivides them into almost 500 groups. It was born for the purpose of Medicare reimbursement. In order to determine the payment, factors consider include the diagnosis involved and the resources necessary for treating the condition [30].
LOINC	Logical Observation Identifiers Names and Codes was developed to provide a definitive standard for identifying clinical information in electronic reports. Its database provides a set of universal names and ID codes for identifying laboratory and clinical test results [31]. It aims is providing a means of uniquely identifying the information elements in EHR. LOINC is remarkable for being the first completely open clinical terminology, making all content available without royalties or charges; this was driven by its creator Clem McDonald [26].
NANDA	Prior to the year 2002, "NANDA" was an acronym for the North American Nursing Diagnosis Association. In 2002, they officially became NANDA International Nursing Diagnoses Classification. They are in charge of definitions and classification of the guide to nursing diagnoses [32].

Table 2. Examples of standard classification systems.

3. Reference terminology

According to the International Standards Organization (ISO), terminologies should be formal aggregations of language-independent concepts, that concepts should be represented by one favored term and appropriate synonymous terms, and that relationships among concepts should be explicitly represented [33, 34]. The ISO specification also stated that terminologies must define their purpose and scope, quantify the extent of their domain coverage, and provide mappings to external terminologies designed for classification and to support administrative functions [33, 34]. The ISO also highlighted the value of mapping among separate terminologies designed to meet different needs. This would allow, for example, a physician to choose a concept from a clinically oriented terminology for constructing a patient’s problem list and a mapped concept in an administrative classification (like ICD-9-CM) could be selected in an automated fashion for billing purposes [33, 34].

3.1. Reference terminology: a new paradigm

In 1998, J. Cimino summarized several works groups’ toward defining the precise attributes of a multipurpose and shareable terminology [7, 8]. He stressed the value of “concept orientation” pending terminology construction. Concept orientation imply “...to use concepts as basic building blocks ahead words, terms, or phrases”. It allows a terminology to be useful in several situations, represented in different languages and easily assess for quality. For Cimino, the aim was to have a universal single clinical terminology that would cover a specialty domain’s concepts completely at multiple levels of detail. Nonspecific phrases such as “not elsewhere classified” must be avoid [7, 8]. It is important to point out the need for complete and comprehensive domain coverage using non-ambiguous, non-overlapping concepts. In the absence of complete domain coverage, terminologies should integrate with other terminologies. Terminologies need to support synonymy and compositionality [35]. “High-quality vocabulary” has been defined as the vocabulary approaches completeness, is well organized and has terms whose meanings are clear [7, 8].

After Cimino’s Desiderata, the difference between Terminology Systems like SNOMED CT and Classification Systems like ICD-10-CM/PCS became clearer. Both coding schemes provide the necessary data structure needed to support healthcare clinical and administrative processes. Clinical terminology systems as well as clinical classification systems were originally designed to serve different purposes and different users’ requirements [27]. ICD-10 is a classification system and it was designed as an output general reporting purposes like public health surveillance, administrative monitoring, and repayments of healthcare services. For all of these reasons, a classification system can be less detailed than a clinical terminology. Contrary, SNOMED CT (Table 3) is a clinical terminology, it was developed to attend as a standard data infrastructure for clinical application, for these reason it requires a higher degree of specificity [36].

3.2. SNOMED CT: background

SNOMED has been used successfully on an international basis in areas such as anatomy-pathology and radiology. It has been translated into several languages. The Systematized Nomenclature of Medicine (SNOMED) nomenclature is an example of a multi-axial classification, developed by North American pathologists and extended from the Systematic Nomenclature of Pathology (SNOP) [5].

In 1965, the Systematized Nomenclature of Pathology (SNOP) was published by the College of American Pathologists (CAP) to describe morphology and anatomy.

In 1975, CAP expanded SNOP to create the Systematized Nomenclature of Medicine (SNOMED). In 1979, the most extensively adopted version of SNOMED named as SNOMED II was published. In 2000, in collaboration with Kaiser Permanente, CAP developed a new logic-based version named SNOMED RT. In the UK during twentieth century, Dr. James Read developed the Read Codes. In the end, under the National Health Service, they evolved into Clinical Terms Version 3 (CTV3). The first version of SNOMED CT, was published in January 2002, after a merge the CTV3 and SNOMED RT, performed by CAP. The merged product was called SNOMED Clinical Terms, which was shortened to SNOMED CT. SNOMED International considers SNOMED CT to be a brand name, not an acronym [37].

<i>"Complete coverage of domain specific content"</i>
<i>"Use of concepts rather than terms, phrases, and words"</i> (concept orientation)
<i>"Concepts do not change with time, view, or use"</i> (concept consistency)
<i>"Concepts must evolve with change in knowledge"</i>
<i>"Concepts identified through nonsense identifiers"</i> (context free identifier)
<i>"Representation if concept context consistently from multiple hierarchies"</i>
<i>"Concepts have single explicit formal definitions"</i>
<i>"Support for multiple levels of concept detail"</i>
<i>"Methods, or absence of, to identify duplication, ambiguity, and synonymy"</i>
<i>"Synonyms uniquely identified and appropriately mapped to relevant concepts"</i>
<i>"Support for compositionality to create concepts at multiple levels of detail"</i>

From Cimino et al. [7].

Table 3. Desiderata for controlled medical vocabularies in the twenty-first century.

SNOMED has been translated into several languages and successfully implemented around the world, in specialties such as anatomy-pathology and radiology. Novel development concerns the use of SNOMED as a reference terminology for health care. The next version nominated as SNOMED RT, will include data related to the causes and symptoms of diseases, treatment of patients, and the outcome of health care process [5]. SNOMED RT has the possibility to represent multiple types of hierarchies and to make the types fully explicit, after the proposed changes.

3.3. SNOMED CT as clinical reference terminology

Reference terminology was defined as a set of concepts and relationships that provide a common reference point for comparisons and aggregation of data about the entire health care process, recorded by multiple different individuals, systems, or institutions [38]. Cornet et al. defined it as “...a system of concepts with assigned identifiers and human language terms, typically involving some kind of semantic hierarchy. Some systems may support the assignment of multiple terms, or synonyms, to a given concept...” [26] SNOMED CT was developed to serve as a standard data infrastructure for clinical application, which requires a greater degree of specificity. A classification system can be less detailed than a clinical terminology [22]. In fact, the systems complement each other and contribute to providing quality data for different domains of the healthcare system [39]. Accordingly, both systems may be use depending on which degree of specificity is required: SNOMED is a better election to recognize unusual illness, mind while ICD-10 is consider more efficient for statistical reporting, such as collecting the top reasons of mortality and morbidity [40].

In order to accomplish “domain coverage”, terminology developers have created new concepts by the utilization of two methods: pre-coordination and post-coordination. With pre-coordination, also named enumeration, is possible to model suitable levels of detail with distinct concepts, derived from real world, non-restricted usage by physicians. Generally, only clinically meaningful concepts are pre-coordinated [41]. By contrast post-coordination, also called compositionality, complex concepts can be composed from simple concepts [42]. Pre-coordination and post-coordination can complement each other, with pre-coordination providing logic and complexity and post-coordination, allowing expressivity and more complete domain coverage.

Existing terminologies that allow post-coordination are better capable to represent phrases and concepts extracted from clinical documents compare to pre-coordinated terminologies [43]. The reason is because users can both: access existing concepts and dynamically compose new concepts according to their needs, such terminologies may improve terminology domain coverage. However, even using post-coordination, it has not yet successfully modeled the entire scope of medical knowledge.

SNOMED CT provides a unified language, it may be used as a standard for communication among healthcare providers. It also highly promotes to semantic interoperability in healthcare information systems [44–46]. Its standardized logical structure and its wide acceptance make it more appropriate for high-level information exchange at national and also international levels [44–46].

SNOMED CT also includes several descriptions that can be used as an entry terminology. Finally, SNOMED CT has a standard cross mapping model; the official distribution includes data for mapping to ICD-9 (ICD—International Classification of Diseases). Additional ICD-10

cross map data has also been developed. These mappings provide the aggregate terminology features to SNOMED CT [47]. However, coding in SNOMED CT is different from conventional coding using ICD-10-CM/PCS. Coding using SNOMED CT is always automated: end users cannot view the codes assigned by the system. For this reason, software developers and EHR vendors are using SNOMED CT to help communication between different applications through a SCS. In fact, we can think of SNOMED CT as a programing language; users utilize applications that apply it without knowing what is at work in the backend [27].

3.4. SNOMED CT: strengths and limitations

SNOMED CT provides functionalities in three layers: entry terminology, reference terminology and aggregate Terminology. Between SNOMED CT's strengths, we can name the completeness, the non-ambiguity (terms must refer to only one concept), the ability to manage synonyms and finally the explicit relationships (this refers to the types of relationships between terms in a nomenclature are clear) [5].

About limitations, they include non-redundancy, meaning by this that each concept should only be expressed in one way. If two terms refer to the same concept, the sensitivity of the replies to database queries will be reduced [5]. It is also remarkable that for those cases when an institutional term cannot be represented with a standard SNOMED CT code, to create new concepts is not allow (for more information, see **Table 4**).

SNOMED	Nomenclature created by the CAP and is evolving into an international Standards Development Organization and currently regarded as the most advanced initiative in knowledge-based representations with clinical application. Each of the 300,000 terms included are defined using relationships with other terms, creating a powerful semantic network. SNOMED CT data model allows continuous extension of the nomenclature, adding new terms, always following the same compositional concept representation model, called Description Logics [48]
Terms	Collections of words or phrases, aggregated in a systematic fashion to represent the conceptual information that makes up a given knowledge domain. Terms in a terminology generally correspond to actual events or entities and to their cognitive representations in people's minds (called concepts) [14, 43]
Concept	Unit of symbolic processing in control vocabulary, a representation of a particular meaning. Concept orientation means that terms must correspond to at least one meaning and no more than one meaning. Meanings correspond to no more than one term [7, 8, 18]
Non-vagueness	Terms must correspond to at least one meaning [7, 8, 18]
Non-ambiguity	Terms must correspond to at least one unequivocal meaning and no more than one meaning, based on context. A distinction must be made between ambiguity of the meaning of a concept and ambiguity of its usage [7, 8, 18]
Non-redundancy	Meanings correspond to no more than one term [7, 8, 18]
Explicit relationships	The kind of relationships between terms in a nomenclature is not clear. Is-a, is-part-of, causes, associated-with, equivalent-to, is-in are the most usual relationships [5]
Concept orientation	Each concept in the vocabulary has a single, coherent meaning, although its meaning might vary, depending on its appearance in a context (such as a medical record). Terminologies also typically contain hierarchical organizations and other representations of linkages among concepts, such as the "is-a-type-of" relationship between "high blood pressure" and "disorder of cardiovascular system" [33, 34, 49]

Table 4. Definition summary.

4. Interface terminology

Interface terminology (IT), which has also been called colloquial terminologies, application terminologies and entry terminologies, has been defined as a systematic collection of healthcare-related phrases (terms) that supports clinicians' entry of patient-related information into computer programs [42]. But how does it happen? When health care providers type into EHR, IT links free text patient descriptors to structured, coded internal data elements used by specific clinical computer programs. Interface terminologies also facilitate display of computer stored patient information to clinician-users as simple human readable text [42]. These terminologies generally embody a rich set of flexible, user-friendly phrases displayed in the graphical or text interfaces of specific computer programs. The "entry" terminologies allow users to interact easily with concepts through common colloquial terms and synonyms. Entry terms can then map to explicitly defined concepts in a more formal terminology, such as a reference terminology, which can then define relationships among concepts [50]. EHR depend on interface terminologies for successful implementation in clinical settings because such terminologies provide the translation from clinicians' own natural language expressions into the more structured representations required by application programs [42]. Interface terminologies are crucial to foment direct categorical data entry by physicians in EHR. Historically, the efforts performed by terminology developers and the standards community, have been orientate to other kind of terminologies, like reference and administrative, instead of interface terminologies.

Between the aims of interface terminology, we can mention: to provide an institutional vocabulary for all user interfaces so they interact with known terms, including local jargon and preferences; to proportion concept lookup functions with loose lexical matches and options, to be employed for the time of data entry process of new items in a problems list or similar user interfaces. It is also important to provide short pick-lists definitions for more structured data entry in specific use templates, with a short list of valid entries and different preferred terms for the same concept in different settings. It should include the ability to accept new terms from the user, in case a concept or description is not represented and detect inappropriate terms for being too general or not valid in a subset [47].

The "usability" of an interface terminology refers to the ease with which its users can accomplish their intended tasks using the terminology. In addition, it has been demonstrated that interface terminology usability correlates with the presence of attributes that enhance efficiency of term selection and composition [51, 52]. The usability of a clinical interface terminology designed correlates with the presence of relevant insertional medical knowledge; adequacy of synonymy; a balance between pre-coordination and post-coordination; and mapping to terminologies having formal concept representations. IT enhances its usability by decreasing the number of steps required for users to find or compose the terms needed for a given task [41, 53].

Synonymy refers to the number of individual terms that can correctly represent a unique concept. Synonym types may include alternate phrases, acronyms, definitional phrases and eponyms [53]. Clinical interface terminologies are specifically designed to represent the variety

of common colloquial phrases in medical discourse; rich synonymy should improve the nuance with which users can express themselves when using the terminology [53].

A very frequently asked question is why to use TS instead of only SNOMED as interface terminology? Between the reasons why we chose it, we can name:

- It is simpler for end users.
- When a single concept is not enough to define the information is possible to build a new one using post-coordination, understood as the representation of a clinical meaning using a combination of two or more SNOMED concept identifiers.
- Thesaurus allows to manage: synonyms (different descriptions related to a concept), list of valid and not recognized terms (error typing, etc), validated jargon and acronyms, list of "Not Valid" terms, thesaurus with local extension in a continuous learning process and drug composition information (commercial products) [47].
- SNOMED has pharmaceutical information as a single entity, not represented independently: Quantity of drug in the pharmaceutical presentation, measurement unit or pharmaceutical form. ut for clinical use, we need to identify single data components.

According to all the limitations mentioned before, terminology services arise.

5. Terminology services

Many definitions for terminology service exist. In previous publications, we defined as complex system of conceptual representation of medical knowledge, with relationships between concepts, with external representations of concepts in lists of standard terms (classifications) and with lexical tools that facilitate the search for terms [54].

A terminology server (TS) is a software that is composed of (**Figure 1**): a thesaurus or local interface vocabulary. This is a list of terms created from free text inputs extracted from the clinical data repository. The terms restrained in the thesaurus are split into concepts (real clinical entities) and descriptions (different ways of naming clinical entities). Thesaurus has been mapped to a reference a vocabulary, for example to SNOMED CT [9, 54]. The TS also is able to reject invalid terms before pointed out as not appropriate for the intended use [9]. The TS should also provide interactive information for refining concepts. This feature of the TS is achieved using semantic information included on SNOMED CT, navigating the subtypes/super-types hierarchy [9]. On the desiderata for TS, Chute et al. [50] attempt to articulate the functional needs of a terminology server oriented toward the clinical needs of care providers using applications in an operational environment. Between the desirable characteristics for a terminology server they included: Word Normalization, Word Completion, Target Terminology Specification, Spelling Correction, Lexical Matching, Term Completion, Semantic Locality, Term Composition, Term Decomposition (**Figure 1**).

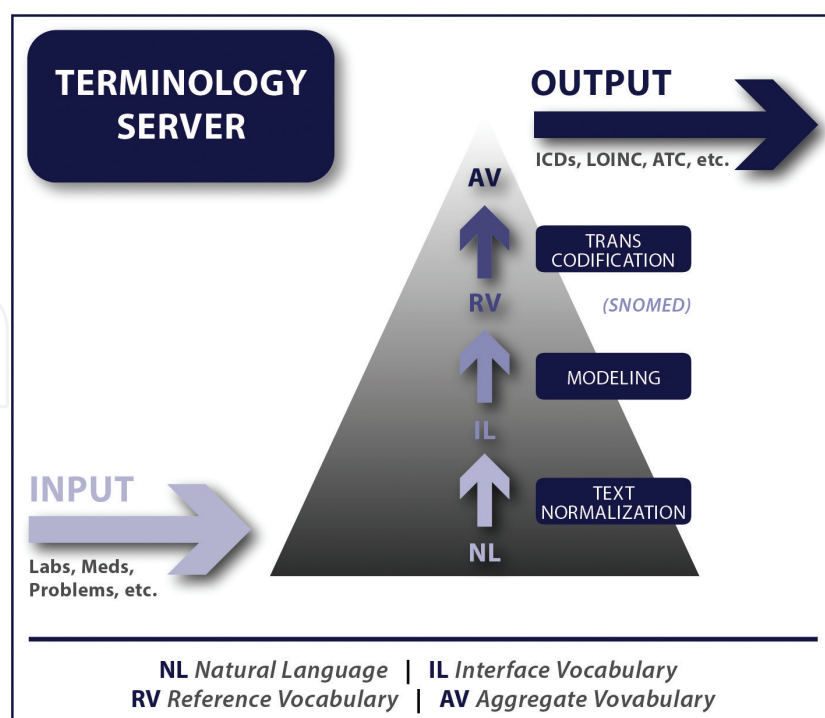


Figure 1. Schema of the functionality of the terminology server in reference to the pyramid of terminological systems.

6. Italians' hospital of Buenos Aires terminology services experience

6.1. Setting

The Hospital Italiano de Buenos Aires (HIBA) is a non-profit healthcare academic center founded in 1853, with over 2700 physicians, 2700 other health team members (including 1200 nurses) and 1800 administrative and support employees. Since 2015, it is a Joint Commission International (JCI) accredited institution. The HIBA has a network of two hospitals with 750 beds (200 for intensive care), 41 operating rooms, 800 home care beds, 25 outpatient clinics and 150 associated private practices located in Buenos Aires city and its suburban area. It has a Health Maintenance Organization (Plan de Salud) that covers more than 160,000 people and also provides health services to another 1,500,000 people who are covered by affiliated insurers. Annually, over 50,000 inpatients were admitted to its hospitals, there were 45,000 surgical procedures (50% ambulatory) and 3,000,000 outpatient visits. In addition, the HIBA is a teaching hospital, with over 30 medical residency-training programs and 34 fellowship programs. There are currently 400 residents and fellows in training. Since 1995, the HIBA runs an in-house developed health information system, which includes clinical and administrative data. Its EHR system called *Italica*, is an integrated, modular, problem-oriented and patient-centered system that works in different clinical settings (outpatient, inpatient, emergency and home care). *Italica* allows computer physician order entry for medications and medical tests, and storage and retrieval of tests results, including images through a picture archiving and communication system. In 2017, HIBA has been certified by the HIMSS as level 7 in the EHR Adoption Model, being the first hospital in Argentina and the second in Latin America reaching this stage [55]. Several health informatics standards had been implemented, including HL7, CDA Version 2, ICD-9, DRG, ICD-10, and ICPC.

6.2. Terminology server of HIBA

The terminology server of HIBA is composed of a local interface terminology (thesaurus) mapped to a reference terminology, SNOMED CT. Our main objective was to design a new terminology system, whose objectives can be related to the functions of the terminology system previously described (entry, reference and aggregate terminology) **Figure 2**.

The IT is updated every day by a team of experts, who audit, assign codes and link each new term to the SNOMED CT (reference terminology), and use the official mapping into SNOMED to another classification (like ICD 9). If SNOMED does not offer an official mapping, the team generates a manual cross-link through functionality on the terminology server [56].

6.3. Terminology server of HIBA: evolution

In 1998, the terminology work team started centralized secondary coding, where a reduced number of trained persons codify the narrative text recorded by the physicians taking care of the patients. The coding included problem list, diagnostics and procedures [10].

In 2004, we achieved 1 million of narrative text secondary coded. After this, we started an auto-codification process, through a thesaurus using interface terminology as a centralized service [56].

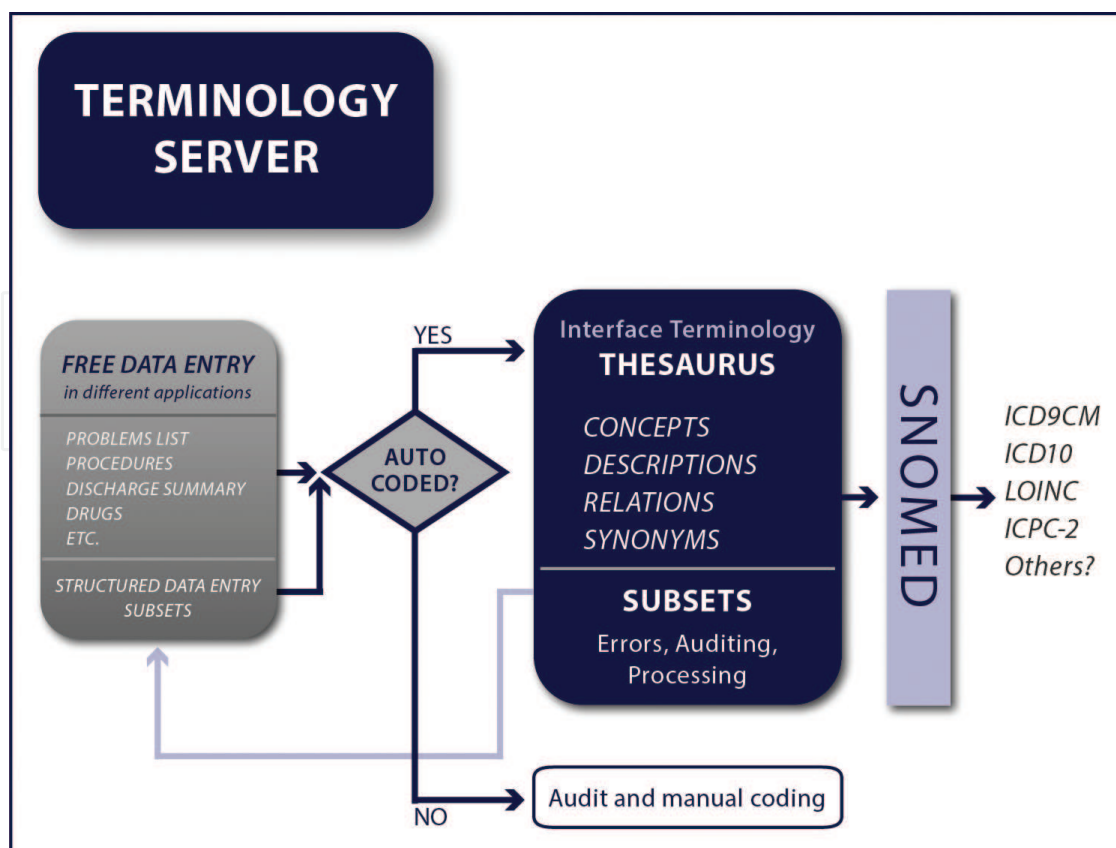


Figure 2. Schema of terminology server.

In 2010, remote Terminology Services (RTS) provided by HIBA through a transnational and interinstitutional implementation [57].

In 2011, the Startup process take places with the aim to extract the greatest amount of clinical information possible from the existing system (mostly in free text), and add this information into the new clinical data repository by coding it. To this purpose, extracted data were processed by the RTS and coded it when it was possible. This data included allergies, reason(s) for the consultation, habits, risk factors, symptoms and diagnosis entered by physicians in a free text form, and only coded diagnoses when they felt it particularly necessary. With the batch processing of these data, the RTS recognized and auto coded 11,118,760 (78.74%) texts (included valid and not valid text), and did not recognized 3,001,991 (21.26%) of the original data [57].

In 2012, we started creation of natural language processing tools and extension of terminological services to the domain of drugs, practices and procedures.

In 2014, the Department of Health Informatics of HIBA, during an effort to achieve international standards of patient health care, in the context of an accreditation process by the JCI, the hospital implemented a software tool for synchronous disambiguation in the EHR, developed in-house. Studies have shown that while the use of abbreviations helps to save time and space during documentation, its use can bring some disadvantages such as unambiguous meanings that often can confuse other healthcare providers with the consequence of causing errors in patient health care. In this sense, the JCI requires that the use of abbreviations must be controlled and documented. To this end since November 2014, an Abbreviations Regulation Committee was established in our hospital with the aim of being in charge of the management and classification of abbreviations used in historical health records. As result of this implementation, 800 abbreviations were classified as doubtful or ambiguous with a total of 400 replacement variations.

The Synchronous Self-Expanding Abbreviation System (SSAS) that detects abbreviations in a free text field. This system was user-centered design and typical abbreviations and their meanings were collected from different areas of the hospital in its construction. The abbreviations can be “unequivocal” (one meaning), “ambiguous” (more than one meaning) and “undefined” (undefined terms). SSAS detected about 4000 abbreviations (1000 univocal, 5000 Ambiguous and 2500 not defined), decreasing almost 40% in the use of abbreviations post implementation [58]. The interface vocabulary takes context parameters with terminology control such as user preferences, specialty or knowledge domain to make a decision that offers a single SNOMED CT concept. The concept-id retrieved is then used to automatically replace the abbreviation with the preferred term. The use of an interface vocabulary offers flexibility to use abbreviations with the added benefit that comes with a reference ontology [59].

6.4. The actual HIBA's terminology web service description (Table 5)

We provide terminology services to several healthcare organizations in the countries of Argentina, Chile, and Uruguay. These include:

- a thesaurus tailored to the local needs and jargon of the professionals who interact with the EHR,
- SNOMED CT as reference standard for interoperability and to implement CDSS,

- cross maps to ICD-9, ICD-10, LOINC, ICPC-2, ATC,
- creation of different types of refsets according to the needs of the organization,
- a drug composition service modeled after the UK's dm + d model

The interface terminology is based in the use of SNOMED CT which is used as the reference terminology. In this sense SNOMED CT serves as a uniform backend representation allowing our interface terminology to adapt the local needs of the institutes we serve. SNOMED CT is the most comprehensive clinical terminology, provides a semantic network with formal structured meanings, has an extendable model, it is widely adopted as an international standard, and it was designed with EHR implementations in mind.

6.5. Our institutional entry terminology, how does it work?

The institutional entry terminology is composed of concepts and descriptions. We use SNOMED definition of these terms, where concepts represent distinct clinical meanings and descriptions are a phrase used to name a concept. Our institutional entry terminology can be divided in several subsets; examples of these are:

- Problems list terminology
- Procedures terminology
- Findings in chest radiography
- Administration routes for drugs
- State of consciousness description
- Physical examination subset
- Liver failure diagnosis

Service	Description
Inteligente prompting	Perform a preliminary search entering the first three characters.
Term recognition	Search for the text entered in the interface vocabulary and offer the alternative to improve the medical record.
Creation of a new term	Enter new term in the interface vocabulary and it is entered into the audit circuit.
List classification	Return back available classification.
Assign classifier	Valid term plus classification return back the corresponding code.
Assign DRG	From a discharge summary encoded with ICD9-CM and other metadata, returns back DRG code.
List domains	Return the domain available (problems, procedures, medications, etc.).
List domain elements	Returns back terms contained in a domain.

Table 5. Functionalities included into terminological web services provided by HIBA.

Some subsets are very large, including thousands of concepts (i.e. the problems list subset). Others are short lists (i.e. the liver failure subset). Each subset was designed in order to be used as the entry terminology in a specific scenario. Concepts are defined only once, regardless of its inclusion in more than one subset; therefore, accessing liver cirrhosis from the problems list or from the Liver failure subset brings the user to the same concept.

The process of adding concepts to the entry terminology and organizing them in subsets is manual. This is done by trained coders that were previously working with the same information in secondary coding using classifications [10]. Construction of the problems list subset was one of our biggest challenges. We decided to base our work on the historic database of our EHR with more than 2 million free text inputs since 1998. All problems list entries and discharge notes were processed to extract all different textual descriptions. We considered that these texts, entered by our own professionals in a completely unconstrained way, would be representative of the local natural language, including abbreviations and jargon. A manual depuration process, assisted by string normalization functions, led to the creation of the Problems List subset with 24,800 different concepts, with 110,000 descriptions in total.

Other subsets were created using arbitrary lists of concepts selected by the clinical terminology team with user input. New concepts or descriptions were accepted from user interfaces and stored for manual evaluation. The data model for the entry terminology was the standard SNOMED CT data model for concepts, descriptions and subsets [60]. New concepts and descriptions were added to the standard SNOMED CT distribution following official SNOMED rules for creating institutional extensions.

Since 2000, physicians at the HIBA have used an inpatient EHR for creating the discharge summary using free text. The discharge summary is a structured abstract of the hospitalization episode where data are registered for caring and management purposes. We developed and implemented a modification of the discharge summary data entry user interface that allows the selection of already coded terms from a local terminology. To achieve this we had to introduce a more restrictive user interface that requires users to select terms from an existing list. The new system should have functions that can facilitate migration from the previous unconstrained text entry model. Information contained in discharge summary is structured in several domains. This structure has the purpose of collecting all the necessary information to group episodes using DRG. In each of these fields, the physician entered free text descriptions. The previous version of the discharge summary software tried to automatically code the entered text using the terminology server. If the term did not match an existing entry in the local terminology, it was addressed to the terminology team for secondary manual codification. The terminology team reviewed all the discharge summaries, assigned ICD-9CM codes and manually grouped them into a DRG [61]. The availability of online consultation about the terminology and input terms created acceptance among users, and led us to maximize the benefits of free and structured texts [9].

6.6. Reference terminology: functions and system description

As regards as reference terminology functions, our TS allows the entry terminology should be represented in the reference terminology (SNOMED CT Spanish Language Version); new concepts can be created for institutional terms that cannot be represented with a standard SNOMED CT code. The system also provides tools to take advantage of the knowledge stored

in SNOMED CT relationships, like obtaining more refined or more general terms, and means of updating to new versions of SNOMED CT without losing information. We used SNOMED CT Spanish Language Version as the reference terminology, but it is important to note that all different language versions of SNOMED CT share the same concepts and relationships. During the translation process only, new descriptions are added. Both entry and reference terminologies were stored following the SNOMED data model, and using SNOMED tools to represent the concepts of the entry terminology. SNOMED CT defines concepts by its relationships with others, so we created new relationships as part of our SNOMED CT extension. SNOMED CT has around 300,000 concepts, but in a clinical setting, health professionals usually use very detailed expressions, adding modifiers to general concepts, like mild ankle sprain. To prevent the exponential growth of the nomenclature, SNOMED CT avoids including such level of combination with modifiers, providing the general concepts (ankle sprain), the possible modifiers (mild) and the rules to correctly relate them (using the has severity relationship).

Any new concept can be represented using this post-coordination technique, creating more detailed subtypes of existing SNOMED CT concepts. Around 33% of the concepts included in the Problems List subset could be directly mapped with existing SNOMED CT concepts; the other 77% needed the addition of one or more modifiers (post-coordination) in order to fully represent the meaning of the entry terminology concept. This rate of post-coordination was dictated by a very permissive policy allowing the use of any term requested by the users, often very specific or personalized. The total of 24,800 concepts was represented with 45,000 new relationships. In each subset, professionals usually try to enter terms that are not valid for later use. We would like the doctor to record the proper diagnosis or reason for encounter instead. In order to reject these terms and for the invalid terms administration, we tag them and add an information text so the professional understands the coding guidelines of the institution. This module provides the tools for tagging these terms and editing the information.

6.7. Aggregate terminology: functions and system description

Between the aggregate terminology functions, our TS provides output to several standard classifications: ICD-9CM (diagnosis and procedures); ICD-10 (diagnosis); ICPC-2 (diagnosis) (International Classification of Primary Care); ATC (drugs) (Anatomical Therapeutic Chemical Classification); Local billing nomenclatures; Aggregate data according to SNOMED CT hierarchies. All these functions run on a centralized software and data structure. The Terminology Server provides these functions to all existing applications in the Health Information System in the form of Web Services. A terminology maintenance software application should also be developed to administrate the institutional terminology, its relationship with SNOMED CT and the mappings.

About Aggregate terminology, the official SNOMED CT cross maps model was implemented, a multi-classification interface was created as part of the Terminology Maintenance Software to visualize, test and modify mappings from SNOMED to different classifications. An SQL algorithm was designed (Oracle SQL) to aggregate concepts according to knowledge stored in SNOMED CT relationships, like all kinds of diabetes, including diabetes complications and excluding maternal and neonatal diseases. These queries are maintained from a module in the Terminology Maintenance Software.

To code the terms in the EHR by a specific classification, the coding application requests, to select the appropriate classification. The system displays a list of classifications available and the operator must select one of them. The system then assigns the code for each term. Using this mechanism, it is possible to select the classifier ICPC-2 for the epidemiological analysis from a problem list of the outpatient EHR, ICD-9 and ICD-10 for a discharge summary in the inpatient EHR. This mapping is possible because we used the official cross-match offer by our reference terminology (SNOMED) or creates our own mapping by the specific terminology team. From a discharge summary coded in ICD-9, it may apply an assigned DRG Service to obtain the corresponding code.

6.8. Terminology maintenance software

The Terminology Maintenance Software includes the following modules:

- **Entry Terminology Administration:** allows the creation of new concepts, description assignment and modeling of each concept with SNOMED CT.
- **Subset Administration:** creation of new subsets, addition and removal of concepts from the subsets, defining hierarchies for tree interfaces.
- **Pending Concepts or descriptions:** all proposed new concepts or descriptions are stored in a list, waiting to be evaluated and modeled, ordered by the number of proposals.
- **Cross Maps Administration:** existing cross maps can be visualized, edited and tested using this module.
- **Data Extraction Rules Administration:** a software interface to visualize and update SNOMED based data extraction queries.

6.9. Status report

Four trained modelers are maintaining the interface terminology, modeling pending concepts or descriptions, running routine quality control checks and maintaining subsets.

We created an ad-hoc automatic process to recode all historic data in our clinical repository, using string matching algorithms; more than 2,200,000 entries were processed.

Around 85% of the original texts received a concept code of the new entry terminology, 10% of them were recognized as invalid entries: therefore, 75% were finally mapped to SNOMED CT. The coding services are used online by our ambulatory and inpatient medical record, receiving around 55,000 requests each month. The task of creating an institutional entry terminology demands a lot of work, but provides an excellent service to the users, and also isolates the terminology system from SNOMED CT changes in newer versions. Local concepts will always be valid, and in the worst case a correction of modeling against SNOMED CT would be required. We found that SNOMED CT cross maps data to ICD-9 is still not adequate for clinical use in our setting, requiring additional manual work on the maps. This may be caused by a different use of the classification in Argentina and the United States.

Our clinical data extraction process, using rules based in SNOMED CT knowledge data, is very effective; however, these rules should be revised for each new SNOMED CT version, as changes in hierarchies and models may affect its effectiveness.

Further reduction of manual classification coding will require adjustments of mapping specifications and user interface changes, aimed to reduce the number of new concepts proposals and enforcing the selection of existing terms. Due to acceptability issues, we have always tried to minimize user interface constraints, thus implementation of these changes will be a slow process.

By means of a much more detailed implementation, the milestone of our new terminology system is the centralization of knowledge representation. The health information system represents uniformly the clinical data entered at any level of care in the institution.

6.10. Terminology service: experience in other settings

One of the most integrated health network of Chile, Megasalud, was using for a decade an EHR named SiapWin. In 2007, they decided to develop their own HIS allowing longitudinal care of patients treated in the network with the mentoring of the medical informatics expertise of HIBA. On behalf of this project, HIBA decided to modify the functionality of their terminology server to provide terminology services to other institutions.

In the layer of access to information, Web Services developed with JAVA, JDK 1.6 was used. The Web Services (WS) were deployed in a SUN's Glassfish application server, and the data was stored in an Oracle 11 g database. The WS were published in the Internet for the remote access of the applications of other institutions.

First implementation in a Chilean provider in 2008: they had clinical data stored and processed in the historical system: about 14 million of unique text phrases. With the terminology services, more than 11 million (78.74%) of texts were automatically codified. In 8 month about 600,000 pieces of new text were entered. About 89.64% of these new texts were successfully recognized by terminology services Nowadays, we are able to recognize above 90% in all regional implementations [57].

The clinical data stored in the legacy system of Megasalud were 14,120,751 single text phrases enabled to process by the RTS. With the batch processing of these data, the RTS recognized and auto coded 11,118,760 (78.74%) texts (included valid and not valid text), and did not recognize 3,001,991 (21.26%) of the original data. In the period between March 1 and October 1, 2009, the physicians at Megasalud entered 592,249 pieces of text in the problem-oriented EHR, 530,897 (89.64%) of them were successfully recognized in the interface terminology of Megasalud by the utilization of RTS in real time. The remainder 61,352 (10.36%) went under the audit process and manual modeling [57].

We consider great value to provide services to other institutions by our RTS. Creating and maintaining a sharable Spanish interface vocabulary database between different countries is a big task as medical Spanish is a rich vocabulary and there are different ways of naming the same clinical entities (polysemy), and different acronyms and synonyms between countries.

Published WS allow the most of the progress achieved by HIBA in the management of terminological domain. There are several services that can be used to process the text entered by a physician in their distance applications [57].

Some examples of others institutions currently consuming RTS are:

Argentina: Healthcare providers and in progress with the federal government

Chile: Healthcare providers and FONASA (National Agreement)

Uruguay: Healthcare providers and AGESIC (National Agreement)

Colombia: Healthcare providers

Actually, we are translating the thesaurus to the Portuguese, for Brazilian institutions.

Conflict of interest

The authors declare that they have no conflict of interest.

Author details

Daniel Luna*, Carlos Otero, María L. Gambarte and Julia Frangella

*Address all correspondence to: daniel.luna@hospitalitaliano.org.ar

Health Informatics Department, Hospital Italiano de Buenos Aires, Buenos Aires City, Argentina

References

- [1] Bates DW, Evans RS, Murff H, Stetson PD, Pizziferri L, Hripcsak G. Detecting adverse events using information technology. *Journal of the American Medical Informatics Association* [Internet]. 2003;**10**(2):115-128 Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=150365&tool=pmcentrez&rendertype=abstract>
- [2] Bates DW, Boyle DL, Teich JM. Impact of computerized physician order entry on physician time. *Proceedings of the Annual Symposium on Computer Application in Medical Care*. 1994:996. Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2247898&tool=pmcentrez&rendertype=abstract>
- [3] Rector AL. Clinical terminology: Why is it so hard? *Methods of Information in Medicine* [Internet]. Dec 1999;**38**(4-5):239-252. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/10805008>
- [4] Shapiro JS, Bakken S, Hyun S, Melton GB, Schlegel C, Johnson SB. Document ontology: Supporting narrative documents in electronic health records. *AMIA Annual Symposium Proceedings* [Internet]. 2005;**2005**:684-688. Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1560738&tool=pmcentrez&rendertype=abstract>
- [5] "Codify". Merriam-Webster.com. Merriam-Webster, n.d. Web. 19 Apr. 2018
- [6] Evans DA, Cimino JJ, Hersh WR, Huff SM, Bell DS, Group C. Toward a Medical-concept Representation Language. The Canon Group. *Journal of the American Medical Informatics Association*. 1994;**1**(3):207-217

- [7] Cimino J. Desiderata for controlled medical vocabularies in the twenty-first century. *Methods of Information in Medicine*. 2012;**37**(1998):394-403
- [8] Cimino JJ. Terminology tools: State of the art and practical lessons. *Methods of Information in Medicine*. 2001;**40**(4):298-306
- [9] Navas H, Osornio AL, Baum A, Gomez A, Luna D, de Quiros FGB. Creation and evaluation of a terminology server for the interactive coding of discharge summaries. *Studies in Health Technology and Informatics* [Internet]. 2007;**129**(Pt 1):650-654. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/17911797>
- [10] Luna D, De Quirós FGB, Garfi L, Soriano E, Flaherty MO. Reliability of secondary central coding of medical problems in primary care by non medical coders, using the International Classification of Primary Care (ICPC). *Objetives: Materials and Methods: Results*. 2001;**10**(Pt 2):94-97
- [11] Hohnloser JH, Kadlec P, Puerner F. Experiments in coding clinical information: An analysis of clinicians using a computerized coding tool. *Computers and Biomedical Research* [Internet]. Oct 1995;**28**(5):393-401. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/8612401>
- [12] Wong ET, Pryor TA, Huff SM, Haug PJ, Warner HR. Interfacing a stand-alone diagnostic expert system with a hospital information system. *Computers and Biomedical Research* [Internet]. Apr 1994;**27**(2):116-129. Available from: <http://linkinghub.elsevier.com/retrieve/pii/S0010480984710123>
- [13] Osenbloom STRR, Iller RA a M, Ohnson KEBJ, Lkin PELE, Rown STHB, Rosenbloom ST, et al. Interface terminologies: Facilitating direct entry of clinical data into electronic health record systems. *Journal of the American Medical Informatics Association* [Internet]. 2006;**13**(3):277-288. Available from: <http://jamia.oxfordjournals.org/content/13/3/277.abstract> \n<http://www.ncbi.nlm.nih.gov/pubmed/16501181>
- [14] Humphreys BL, Lindberg DA, Schoolman HM, Barnett GO. The unified medical language system: An informatics research collaboration. *Journal of the American Medical Informatics Association* [Internet]. 1998;**5**(1):1-11 Available from: <http://www.pubmed-central.nih.gov/articlerender.fcgi?artid=61271&tool=pmcentrez&rendertype=abstract>
- [15] Humphreys BL, McCray AT, Cheh ML. Evaluating the coverage of controlled health data terminologies: Report on the results of the NLM/AHCPR large scale vocabulary test. *Journal of the American Medical Informatics Association*. 1997;**4**(6):484-500
- [16] National Committee on Vital and Health Statistics (NCVHS). Washington, DC: Report on Uniform Data Standards for Patient Medical Record Information; 2000. p. 65. Available from: <https://www.ncvhs.hhs.gov/wp-content/uploads/2014/08/hipaa000706.pdf>
- [17] Malterud K, Hollnagel H. The magic influence of classification systems in clinical practice. *Scandinavian Journal of Primary Health Care*. 1997;**15**(1):5-6
- [18] Cimino JJ. High-quality, standard, controlled healthcare terminologies come of age. *Methods of Information in Medicine*. 2011;**50**(2):101-104

- [19] AHIMA Project Offers Insights into SNOMED, ICD-9-CM Mapping Process [Internet]. Available from: <http://library.ahima.org/doc?oid=60018#.Wmozc7ziaM> and <http://library.ahima.org/doc?oid=60018#.Wmozc7ziaM8>
- [20] Pan American Health Organization. eHealth in Latin America and the Caribbean: Interoperability Standards Review. Washington, DC: PAHO; 2016. Available from: http://iris.paho.org/xmlui/bitstream/handle/123456789/28189/9789275118818_eng.pdf?sequence=1&isAllowed=y
- [21] Giannangelo K. Healthcare Code Sets, Clinical Terminologies, and Classification Systems. 2nd ed. Chicago: American Health Information Management Association (AHIMA); 2007
- [22] Davoudi S, Dooling JA, Glondys B, Jones TD, Kadlec L, Overgaard SM, Ruben K, Wendicke A. Data quality management model (2015 update). Journal of American Health Information Management Association [Internet]. 2015;**86**(10):62-65. Available from: <http://library.ahima.org/PB/DataQualityModel#.Wmo5FrziaM9>
- [23] Bishop J, Bronnert J, Cook J, Macksoud L, Mastronardi C, Morsch M, et al. Automated coding workflow and CAC practice guidance. Journal of American Health Information Management Association. 2010;**81**(7):51-56
- [24] Chute CG. Clinical classification and terminology. Journal of the American Medical Informatics Association. 2000;**7**(3):298-303
- [25] Elkin PL, Brown SH, Carter J, Bauer BA, Wahner-Roedler D, Bergstrom L, et al. Guideline and quality indicators for development, purchase and use of controlled health vocabularies. International Journal of Medical Informatics [Internet]. Dec 18, 2002;**68**(1-3):175-186. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/12467801>
- [26] Cornet R, Chute C. Health concept and knowledge management: Twenty-five years of evolution. IMIA Yearbook [Internet]. 2016;**25**(Suppl. 1):S32-S41 Available from: <http://www.schattauer.de/index.php?id=1214&doi=10.15265/IYS-2016-s037>
- [27] Alakrawi ZM. Clinical terminology and clinical classification systems: A critique using AHIMA's data quality management model. Perspectives in Health Information Management [Internet]. 2016. Retrieved April 2018. Available from: <http://perspectives.ahima.org/clinical-terminology-and-clinical-classification-systems-a-critique>
- [28] Vladeck BC. Diagnostic-related groups. Journal of the American Medical Association [Internet]. Jun 25, 1982;**247**(24):3314-3315. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/6806489>
- [29] Available from: <http://www.who.int/classifications/icd/factsheet/en/>
- [30] DRG [Internet]. Available from: [https://www.healthlawyers.org/hlresources/HealthLawWiki/Diagnosis-related-group-\(DRG\).aspx](https://www.healthlawyers.org/hlresources/HealthLawWiki/Diagnosis-related-group-(DRG).aspx)
- [31] LOINC [Internet]. Available from: <https://loinc.org/faq/basics/>
- [32] NANDA [Internet]. Available from: <http://www.nanda.org/nanda-international-glossary-of-terms.html>

- [33] ISO/TS 17117:2002(E): Health Informatics—Controlled health terminology—Structure and high-level indicators: Technical Committee ISO/TC 215, Health Informatics; 2002. p. 18. Available from: <https://www.iso.org/obp/ui/#iso:std:iso:ts:17117:en>
- [34] International Standards Organization (ISO). TC 215 – Health Informatics. 2007. (Tc 215). Available from: http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_tc_browse.htm?commid=54960
- [35] Chute CG, Cohn SP, Campbell JR. A framework for comprehensive health terminology systems in the United States: Development guidelines, criteria for selection, and public policy implications. *Journal of the American Medical Informatics Association*. 1998;**5**(6):503-510
- [36] Perspectives in Health Information Management. Fall 2014 introduction. *Perspectives in Health Information Management* [Internet]. 2014;**11**:1a. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/25593567>
- [37] SNOMED CT [Internet]. Available from: <https://www.snomed.org/snomed-ct/what-is-snomed-ct/history-of-snomed-ct>
- [38] Imel M, Campbell J. Mapping from a clinical terminology to a classification. In: AHIMA's 75th Anniversary National Convention [Internet]. 2003. Available from: <http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Mapping+from+a+Clinical+Terminology+to+a+Classification#0>
- [39] Chavis S. Two systems, one direction. *For the Record* [Internet]. 2013;**25**:10. Available from: <http://www.fortherecordmag.com/archives/1013p10.shtml>
- [40] Centers for Disease Control and Prevention. International Classification of Diseases. 10th revision. Clinical Modification (ICD-10-CM). 2016. Internet [Internet]. Available from: <http://www.cdc.gov/nchs/icd/icd10cm.htm>
- [41] Rassinoux AM, Miller RA, Baud RH, Scherrer JR. Compositional and enumerative designs for medical language representation. *Proceedings of the AMIA Annual Fall Symposium* [Internet]. 1997;**1997**:620-624. Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2233357&tool=pmcentrez&rendertype=abstract>
- [42] Rosenbloom ST, Miller RA, Johnson KB, Elkin PL, Brown SH. Interface terminologies: Facilitating direct entry of clinical data into electronic health record systems. *Journal of the American Medical Informatics Association* [Internet]. May 1, 2006;**13**(3):277-288 Available from: <https://academic.oup.com/jamia/article-lookup/doi/10.1197/jamia.M1957>
- [43] Chute CG, Cohn SP, Campbell KE, Oliver DE, Campbell JR. The content coverage of clinical classifications. For the Computer-Based Patient Record Institute's Work Group on Codes & Structures. *Journal of the American Medical Informatics Association*. 1996;**3**(3):224-233
- [44] Duarte J, Castro S, Santos M, Abelha A, Machado J. Improving quality of electronic health records with SNOMED. *Procedia Technology* [Internet]. 2014;**16**:1342-1350 Available from: <http://linkinghub.elsevier.com/retrieve/pii/S2212017314003788>

- [45] Gøeg KR, Chen R, Højen AR, Elberg P. Content analysis of physical examination templates in electronic health records using SNOMED CT. *International Journal of Medical Informatics* [Internet]. 2014 Oct;**83**(10):736-749 Available from: <http://linkinghub.elsevier.com/retrieve/pii/S1386505614001075>
- [46] Lee D, de Keizer N, Lau F, Cornet R. Literature review of SNOMED CT use. *Journal of the American Medical Informatics Association* [Internet]. Feb 1, 2014;**21**(e1):e11-e19. Available from: <https://academic.oup.com/jamia/article-lookup/doi/10.1136/amiajnl-2013-001636>
- [47] Gambarte ML, Osornio AL, Martinez M, Reynoso G, Luna D, de Quiros FGB. A practical approach to advanced terminology services in health information systems. *Studies in Health Technology and Informatics* [Internet]. 2007;**129**(Pt 1):621-625. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/17911791>
- [48] Spackman KA. Normal forms for description logic expressions of clinical concepts in SNOMED RT. *Proceedings of the AMIA Symposium* [Internet]. 2001;**2001**:627-631. Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2243264&tool=pmcentrez&rendertype=abstract>
- [49] Hammond WE, Stead WW, Straube MJ, Jelovsek FR. Functional characteristics of a computerized medical record. *Methods of Information in Medicine* [Internet]. Jul 1980;**19**(3):157-162. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/7412563>
- [50] Chute CG, Elkin PL, Sherertz DD, Tuttle MS. Desiderata for a clinical terminology server. *Proceedings of the AMIA Symposium* [Internet]. 1999;**2**(1):42-46 Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2232621&tool=pmcentrez&rendertype=abstract>
- [51] Kahn CE, Wang K, Bell DS. Structured entry of radiology reports using world wide web technology. *Radiographics*. 1996;**16**(3):683-691
- [52] The Computer-Based Patient Record [Internet]. Washington, D.C.: National Academies Press; 1997. Available from: <http://www.nap.edu/catalog/5306>
- [53] Masarie FE, Miller RA, Bouhaddou O, Giuse NB, Warner HR. An interlingua for electronic interchange of medical information: Using frames to map between clinical vocabularies. *Computers and Biomedical Research* [Internet]. Aug 1991;**24**(4):379-400 Available from: <http://linkinghub.elsevier.com/retrieve/pii/001048099190035U>
- [54] Lopez Osornio A, Gambarte ML, Otero C, Gomez A, Martinez M, Soriano E, et al. Desarrollo de un servidor de terminología clínico. 8mo Simp Informática en Salud – 34 JAIIO. 2005;**34**(April):29-43
- [55] Luna D, Plazzotta F, Otero C, González Bernaldo de Quirós F, Baum A, Benítez S. Incorporación de tecnologías de la información y de las comunicaciones en el Hospital Italiano de Buenos Aires. Spanish; 2012-01. Serie: Documentos de Proyectos No.459, CEPAL; https://www.youtube.com/watch?v=5Pb_3q6TQ58&feature=youtu.be
- [56] Osornio AL, Luna D, Gambarte ML, Gomez A, Reynoso G, de Quirós FGB. Creation of a local interface terminology to SNOMED CT. *Studies in Health Technology and Informatics* [Internet]. 2007;**129**(Pt 1):765-769. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/17911820>

- [57] Luna D, Lopez G, Otero C, Mauro A, Casanelli CT, de Quirós FGB. Implementation of interinstitutional and transnational remote terminology services. AMIA Annual Symposium Proceedings [Internet]. 2010;**2010**:482-486. Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3041368&tool=pmcentrez&rendertype=abstract>
- [58] Rodriguez JF, Lede DR, Pérez D, Benítez S, Luna D. Abbreviations System: Preliminary Results of a Satisfaction Study. Cbis. 2016;**2016**:1-6
- [59] Williams M d PA, Perez D. Building an Snomed CT content extension with abbreviation support. Systematized Nomenclature of Medicine. SNOMED CT EXPO 2015;**2015**
- [60] International Health Terminology Standards Development Organization (IHTSDO). SNOMED CT® Clinical Terms User Guide. January 2010;**2010**:99
- [61] Lopez Osornio A, Luna D, Bernaldo de Quiros FG. Creación de un sistema para la codificación automática de una lista de problemas. 5to Simp Informática en Salud – 31 JAIIO [Internet]. 2002;**2002**:31. Available from: <http://www.sis.org.ar/sis2002/paperssis/SIS30.pdf>

