INTEROPERABLE KNOWLEDGE REPRESENTATION IN CLINICAL DECISION SUPPORT SYSTEMS FOR REHABILITATION

L. CECCARONI AND L. SUBIRATS

Abstract. International standards and standard proposals of nomenclatures, ontologies and information models exist in medicine, but the scope of most of them is overly general to cope with the specificities that characterize rehabilitation. Here we carry out an ontology-based exploration of the concepts and relationships in the rehabilitation domain, integrating clinical practice, the clinical investigator record ontology and international standards. The aim of the analysis is to identify potential logical problems with the use of existing models and international standards in representing and reasoning with real clinical data, and to understand whether and how these data might be defined more formally than in the current practice. Our analysis of the relationships among rehabilitation concepts revealed issues related to confusion among classes and their properties, incorrect classifications, overlaps and loss of information. It also suggested properties that should be included in a formal model suitable to be used by decision support systems.

Keywords: knowledge representation, rehabilitation, functional diversity, clinical decision support systems, personalized medicine, evidence-based medicine.

AMS Subject Classification: 68T30

1. Introduction

Intelligent systems, ontologies and interoperable knowledge representations are used, in medicine in general and in physical medicine and rehabilitation (from now on, rehabilitation) in particular, for the purpose of increasing the interoperability among representations of the knowledge about patients’ status. Interoperable representations are especially valuable for certain artificial intelligence systems whose reasoning is based on experience or analogy, because reasoning result improves if it is based on knowledge accumulated through a wide and diverse range of cases spanning as many healthcare organizations as possible. An example are clinical decision support systems (CDSSs) using case-based reasoning (CBR). Nomenclatures, ontologies and information models can be introduced to facilitate automatic knowledge integration when representing patient’s clinical data in these systems.

Information models exist to organize the information about patients and to capture the communication processes of rehabilitation professionals, such as: the virtual medical record (vMR) standard proposal and the ISO/EN 13606 standard. Nomenclatures and ontologies exist to identify and classify the characteristics of a patient, and other related concepts, through standard descriptors, such as: the systematized nomenclature of medicine - clinical terms (SNOMED CT), the international classification of functioning, disability and health (ICF), and the international classification of diseases (ICD). Developing and mapping these information models, nomenclatures (including upper nomenclatures) and ontologies to the medical health records (MHRs) used in healthcare organizations, is the first step towards knowledge integration. Difficulties arise in the development of these models and systems, especially in the standardization of the clinical-knowledge acquisition process (mapping scales and questionnaires into standard
terminologies) using current methodologies (and ICF and SNOMED CT as standards). In the paper, as summarized in Fig. 1, the advantages of mapping electronic health records (EHRs), or patient health records (PHRs), to terminologies and classifications are considered to share and reuse knowledge in the domains of population health, clinical environment, administration and report presentation. In the following sections, the results obtained consisting of advances in the knowledge representation related to computer-based decision support in rehabilitation are presented, using a real-world rehabilitation scenario as case study.

**Figure 1.** Role of classifications and terminologies in healthcare services

The rest of the paper is organized as follows. Section 2 describes a scenario of physical rehabilitation. Sections 3 and 4 explain respectively for what purposes CDSSs and ontologies are developed and used in medicine and rehabilitation. Section 5 describes how problems found in the extraction of standard-based indicators are solved. Section 6 compares potential solutions for EHRs interoperability, and Section 7 describes the architecture of a system which implements the ontology and solves EHRs interoperability. Finally, Section 8 mentions remaining issues about reasoning in the rehabilitation domain, and conclusions are drawn in Section 9.

## 2. Rehabilitation scenario

A scenario of physical rehabilitation is considered here as case study, in which the main goal is to improve the efficacy of rehabilitation of the upper extremity with respect to movement. To achieve this, four sub-goals have been defined:

1. Monitoring the movement of the upper extremity. In the monitoring, several technologies are used, such as T-shirt–mounted sensors and orthotic devices.
2. Performing specific activities of daily living (ADLs) with assistance as needed. In the assistance of movement, the investigator plans the sequence of activities, while investigator agents monitor the movement of the upper extremity and assist patients only when needed. These agents can be therapists or orthotic devices.
3. Performing specific ADLs with continuous feedback. The feedback system gives real-time information about the execution of the activity through virtual reality and interactive video.
4. Programming personalized rehabilitation activities. A knowledge-based system allows personalized planning of sequences of activities on the basis of result indicators, whose values are obtained from activities and other patient’s care information.
The rehabilitation-scenario concepts are depicted in Fig. 2 according to a consensus of experts.

The scenario is composed of five main steps (see Fig. 3): (1) diagnosis, (2) planning, (3) preparation of ADLs, (4) execution of ADLs, and (5) activity reporting.

In this paper, patient is defined adopting Beale and Heard’s top-level ontology [1]: a patient system (or patient) is the object of care (typically one person, but could be more), essentially as a biological or social system (depending on the perspective of the investigator); an investigator system (or investigator) is the investigating, healthcare-providing entity, the totality of rehabilitation professionals and other agents who perform actions related to the care of the patient (including the patient in the role of self-carer or self-medicator, as well as any family members).

3. Decision Support System

To automatically reason about real-world scenarios like this, and to assist professionals in decision making within a standard framework and through so-called clinical decision-support systems (CDSSs), a formal, standard knowledge representation is needed. To achieve it is completely novel and challenging because very few tests, processes and methodologies were designed to be consisted with current standard models [2]. Furthermore, CDSSs are tools which, if used skillfully and respectfully by experienced rehabilitation and health professionals, can help to cover the broad and complex areas of assessment and proposal, allowing professionals to focus not exclusively on test data, but also on individual patients and their environment.

3.1. Categories of CDSSs. In fact, two main categories of CDSSs can be identified: those oriented to assessment and those oriented to proposal. The objectives of the ones oriented to assessment are: evaluation of a patient’s past, current and future status (this includes prognosis: the likely outcome of an illness); quantification of risk; classification of patients according to
their functional diversity. The objectives of the ones oriented to proposal are: risk prevention; definition of therapeutic goals.

Here we deal with the design and development of a CDSS oriented to assessment and, more specifically, to prognosis, a clinical-system environment which requires reasoning under uncertainty. We use for the decision-making process a knowledge-based system (KBS) with case-based reasoning (CBR), with standard indicators used to represent a case (e.g., Emotional functions - b152). Fig. 4 shows a screenshot of the prognosis application’s graphical interface which includes, as sections, the cause of the functional limitation, the rehabilitation treatment, and the four branches of ICF (body functions, activities and participation, environmental factors and body structures). The outcome of the prognosis CDSS is evaluated by investigators.

4. Medical knowledge modeling

Healthcare organizations use several tools to capture information. These tools make use of specific terms, which are sometimes ambiguous: descriptor, grade, index, indicator, parameter, questionnaire, scale, score and test. The terminology used in this paper is defined as follows and is part of an ontology, which we defined (and encoded in OWL 2 [3]) based on standard nomenclatures and ontologies:

- **Index**: a combination of indicators, questionnaires and possibly other indexes. The function representing this combination gives as summarizing result a score.
- **Indicator**: a (subjective or objective) parameter or descriptor used to measure or compare activities and participation, body functions, body structures, environment factors, processes, and results.
- **Questionnaire** (or test): a set of questions answered using a scale.
**Scale**: a mapping between some ordered (qualitative or quantitative) values (or grades) and their description. These values are used to answer questionnaires.

The main metrics of the ontology (which can be accessed at [http://bioportal.bioontology.org/ontologies/47011](http://bioportal.bioontology.org/ontologies/47011)) are as follows: 137 classes; maximum depth of 9; maximum number of siblings of 16; 20 classes with a single subclass, 16 object properties. There are not multiple primitive representations of the same concepts as it is specified in Cimino [4]'s guidelines.

4.1. **Classes**. Most **top-level concepts** come from SNOMED CT [5], e.g.: *Assessment scales* (273249006, subclass of *Staging and scales*), *Physical object* (260787004), *Process* (415178003, subclass of *Observable entity*). Some of them come from ICF, e.g.: *Environmental factors* (e), *Activities and participation* (d). ICF codes have a letter b, d, e or s possibly followed by a number while SNOMED CT codes are composed by a 9-digit number. In selecting and reusing concepts, we always maintained the hierarchical organization and consistency of the original standard ontologies. In any case, top-level categories are overly general to characterize the ontology, which can be better framed, conceptually, through other categories, more related to rehabilitation scenarios, such as: *Diagnosis* (439401001), *Functional independence measure* (FIM) (273469003), *Orthotic devices* (224898003), *Rehabilitation therapy* (52052004).
In Fig. 5, 6 and 7 we show how the ontology is related to and integrate the state of the art. Concepts are encoded as SNOMED CT classes if not otherwise specified. Concepts in bold-frame boxes are encoded as ICF classes (e.g., Body structures). Concepts in grey boxes are reused from Beale and Heard [1]'s clinical investigator record ontology (e.g., Instruction). Underlined relations are reused from the ICD (specifically its 11th revision, ICD-11) (e.g., has localization). Concepts in white boxes and non-underlined relations represent an extension by
the authors of existing approaches based on clinical practice (e.g., has indicator). Relations in the figure, if not otherwise specified, are is a. At the top-level of the ontology, there are the following concepts, as shown in Fig. 5.

- **Historical** (following Sowa [6]'s top-level categories), which subsumes:
  - Observation: information created by an act of observation, measurement, questioning, or testing of the patient or related substance;
  - Action: a record of intervention actions that have occurred, due to instructions or otherwise;
  - Environment: information on the context of the patient;
- **Interpretation of findings**: inferences of the investigator using the personal and published knowledge bases about what the observations mean, and what to do about them; includes all diagnoses, assessments, plans, goals;
- **Instruction**: instructions, based on interpretations of findings, sufficiently detailed so as to be directly executable by investigator agents (people or machines), in order to accomplish a desired intervention.

The Interpretation of findings category corresponds to Sowa [6]'s Description category, to the notion of hypothesis in general science and to Rector [7]'s Meta-observations. The Instruction category corresponds to Sowa [6]'s Script category. Assessment (see Fig. 5) relates to past, current or projected states of affairs. Proposal relates to desired ones. A Diagnosis is the attachment of a label to a group of observed signs and symptoms, which designates it (in the understanding of the investigator) as being a particular phenomenon. A Differential diagnosis allows for multiple possibilities, due to the lack of sufficient information or understanding to attach one label. A Goal, such as monitoring the upper limb while performing the Activity of daily living (ADL) (129025006) Dressing (129003000) (d540), is a statement about what the desired state of the patient system should be, while a Plan is a statement about how to get there.

### 4.2. Object properties.

Object properties represent relationships between two classes or instances. Apart from properties is a, is composed of and modifies, object properties of the proposed ontology are described in Table 1. Some of them are based on ICD-11 [7].

<table>
<thead>
<tr>
<th>Object property</th>
<th>Domain</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>has contraindication</td>
<td>ADL</td>
<td>Substance</td>
</tr>
<tr>
<td>has disease (ICD-11)</td>
<td>Person, ADL</td>
<td>Disease, Participation, Body functions, Body structures</td>
</tr>
<tr>
<td>has goal</td>
<td>Person</td>
<td>Disease</td>
</tr>
<tr>
<td>has indicator</td>
<td>Diagnosis, ADL</td>
<td>Observation parameter</td>
</tr>
<tr>
<td>has interpretation</td>
<td>Diagnosis</td>
<td>Indicator</td>
</tr>
<tr>
<td>has location (ICD-11)</td>
<td>Observation</td>
<td>Body structures</td>
</tr>
<tr>
<td>has occupation</td>
<td>Person</td>
<td>Occupation</td>
</tr>
<tr>
<td>has technology</td>
<td>ADL</td>
<td>Physical object</td>
</tr>
<tr>
<td>is assisted by, is assisted as needed by</td>
<td>ADL</td>
<td>Person, Physical object</td>
</tr>
<tr>
<td>is executed by</td>
<td>ADL</td>
<td>Person</td>
</tr>
<tr>
<td>is manifestation of (ICD-11)</td>
<td>Disease</td>
<td>Observation parameter</td>
</tr>
<tr>
<td>is programmed by</td>
<td>ADL</td>
<td>Person</td>
</tr>
<tr>
<td>is used</td>
<td>ADL</td>
<td>Observation parameter</td>
</tr>
</tbody>
</table>
Fig. 5, 6 and 7 show the most used relationships among concepts, and the context in which they are used in a typical scenario. Continuous lines represent the relation \textit{has subclass}. Discontinuous lines represent: the object properties defined in Table 1, the relation \textit{composed of} or the relation \textit{modifies}.

4.3. \textbf{Indicators.} Indicators are the main classes used for representing the status of a patient and potentially number in the thousands. The two main classes used in rehabilitation are \textit{process indicators} and \textit{result indicators}. A process indicator is used to assess whether a task is being performed correctly. A result indicator is used to assess the performance in carrying out an activity or whether the objectives of the activity have been achieved. To facilitate human practice, only a selection of indicators, grouped into \textit{core sets}, are used. Core sets can be formed according to \textit{functionality}, \textit{pathology} or \textit{rehabilitation process}. Core sets are useful because human investigators can process only a fraction of the categories found in relevant ontologies such as ICF and SNOMED CT [8]. Core sets already exist for several pathologies, such as \textit{multiple sclerosis}, \textit{spinal cord injury} (SCI) or \textit{traumatic brain injury}, though finding the core categories for rehabilitation processes and moving from a pathology-based approach to one based on functionality and rehabilitation is needed. A methodology to define core sets of standard-based, rehabilitation-based indicators from existing questionnaires and non-standard indicators (observation parameters) would include at least the following steps:

1. search for questionnaires and non-standard indicators about rehabilitation;
2. prioritization and selection of questionnaires and non-standard indicators based on literature-relevance, and coverage of rehabilitation processes and indicator types;
3. extraction of standard-based indicators from selected questionnaires and non-standard indicators;
4. aggregation of indicators according to a taxonomy of rehabilitation processes;
5. selection of a core set of these indicators.

Steps 1 and 2 have been already carried out by the authors and other researchers in previous research; in this paper we contribute to step 3, while the fourth and fifth steps belong to the medicine domain and to future work.

5. \textbf{Extraction of standard-based indicators}

Methodologies exist to extract indicators from clinical questionnaires and encode them using international standards. If it is necessary to combine indicators and their values, additional methodologies might be needed to carry out this combination. These methodologies for combination do not currently exist for rehabilitation or their scope is limited. To encode indicators into international standards, the World Health Organization’s ICF is considered first because it provides a comprehensive specification of health-related human functioning in the domains of (i) body functions and structures (e.g., \textit{sensory}, \textit{neuromusculoskeletal} and \textit{movement-related functions}), (ii) activities and participation, ranging from basic (e.g., dressing and eating) to complex (e.g., working and living independently), and (iii) environmental factors that provide a context for understanding functioning, functional diversity and health. If no suitable category is found in the ICF to define a concept, SNOMED CT, which includes top-level categories, is considered. The methodology to encode indicators into ICF can be found in Cieza et al. [9][10], while the methodology to encode them into SNOMED CT is described by Subirats and Cecconari [11]:

1. Using any search-capable SNOMED CT interface (e.g., [12]), search for the concept that you want to encode.
2. If there is no exact match, search for a synonym.
3. If there are no synonyms, use a combination of hypernyms and hyponyms to find concepts that are modeled in SNOMED CT.
(4) Check if the type of the concepts found in SNOMED CT properly models the concept to be encoded. There are 19 types of concepts in SNOMED CT, such as clinical finding, physical object, social context, physical force, substance or procedure. In the previous examples, the type should be *procedure* in all cases.

See Table 2 for an example.

Limits exist in these methodologies if resulting indicators need to be used by formal reasoning systems, because in several cases branches of standard classes are heterogeneous in the relationships used and some of the leaves of a sub-tree represent properties or parts of the parent concept rather than subclasses. In particular, an analysis of the relationships within ICF revealed problems related to confusion between classes and their properties, incorrect classifications and overemphasis on subsumption [13].

<table>
<thead>
<tr>
<th>Questionnaire (FIM) item</th>
<th>SNOMED CT</th>
<th>ICF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dressing upper body</td>
<td>Ability to dress (165235000)</td>
<td>Dressing (d540)</td>
</tr>
<tr>
<td></td>
<td>Puting on clothes (d5400)</td>
<td>Dressing (d540)</td>
</tr>
<tr>
<td>Dressing lower body</td>
<td>Ability to dress (165235000)</td>
<td>Dressing (d540)</td>
</tr>
<tr>
<td></td>
<td>Ability to put on footwear (284978003)</td>
<td>Putting on footwear (d5402)</td>
</tr>
<tr>
<td>Toileting</td>
<td>Toileting (129004006)</td>
<td>Toileting (d530)</td>
</tr>
<tr>
<td>Transfers: bed/chair/ wheelchair</td>
<td>Chair/bed transfer ability (165236004)</td>
<td>Transferring oneself (d420)</td>
</tr>
<tr>
<td>Transfers: toilet</td>
<td>Ability to transfer between wheelchair and toilet (302274006)</td>
<td>Transferring oneself (d420)</td>
</tr>
<tr>
<td>Locomotion: stairs</td>
<td>Climbing stairs (129016000)</td>
<td>Climbing (d4551)</td>
</tr>
</tbody>
</table>

### 5.1. Difficulties of mapping clinical questionnaires into standard terminologies and ontologies.

Several problems exist with the standardization of observation parameters (questionnaires’ indicator names and values) into SNOMED CT and ICF. One main difficulty is that there is a mismatch between the way SNOMED CT categorizes its concepts and the questionnaire items. SNOMED CT makes a distinction between *Clinical finding* and *Observable entity*. *Findings* are observations that are meaningful by themselves whereas *observables* need to have values to complete their meaning. A questionnaire item should be mapped only to SNOMED CT observables (see Table 2). Combinations of a questionnaire item and its values map to findings. If we take, for example, the *Transfers: toilet* item, the complete mapping to SNOMED CT and ICF is shown in Table 3. (Here we see that the extremes of the scale can be mapped to findings, as an alternative to a mapping using only observables.) The mapping to ICF doesn’t suffer from the problem of category mismatch, as ICF categories are neutral with respect to functional diversity.

Sections 5.1.1 and 5.1.2 deal with specific issues found in the standardization of indicators to ICF and SNOMED CT, respectively. Observation parameters of questionnaires FIM, *spinal cord independence measure* (SCIM), *patient competency rating scale* (PCRS), *extended Glasgow outcome scale* (GOSE), *hospital anxiety and depression scale* (HAD), from the *Historical* class of 350 patients who suffer from neurological diseases, are considered.

### 5.1.1. Problems found in the standardization of observation parameters to ICF.

Overlaps and loss of information are found when implementing the methodology of ICF-encoding of Cieza et al. [9][10].
Table 3. Encoding of indicator values into SNOMED CT and ICF

<table>
<thead>
<tr>
<th>Questionnaire item’s value</th>
<th>SNOMED CT</th>
<th>ICF</th>
</tr>
</thead>
<tbody>
<tr>
<td>&quot;Transfers: toilet&quot; 7</td>
<td>Able to transfer between wheelchair and toilet</td>
<td>&quot;Transferring oneself&quot; 0</td>
</tr>
<tr>
<td>(302275007)</td>
<td></td>
<td>&quot;Transferring oneself&quot; 1/1/2/2/3</td>
</tr>
<tr>
<td>6/5/4/3/2</td>
<td>Ability to transfer between wheelchair and toilet with modified independence / supervision / minimal assistance / moderate assistance / maximal assistance</td>
<td></td>
</tr>
<tr>
<td>&quot;Transfers: toilet&quot; 1</td>
<td>Unable to transfer between wheelchair and toilet (302276008)</td>
<td>&quot;Transferring oneself&quot; 4</td>
</tr>
</tbody>
</table>

When the content of an item is not explicitly named in the corresponding ICF category, but at the same time is included, then the item is linked to this ICF category and the additional information not explicitly named by the ICF is documented. The problem here is that when, e.g., functions of body structures are linked only to the activity or body function, body structures cannot be mapped and distinguished. For example, the ICF standardizations of Dressing the upper body and Dressing the lower body of FIM are the same (Dressing, d540) or they are not based on the original body structures (Putting on clothes, d5400, and Putting on footwear, d5402). A solution would be to add relations representing the additional information not explicitly named by the ICF, e.g., the relation has localization to link to Body structures.

The response options of an item are linked if they refer to additional constructs. Depending on the answer-option chosen, the item is standardized to some indicators or others. For example, the standardization of Dressing the upper body of SCIM is: Support and relationships (e3) if the answer is Requires total assistance or Requires partial assistance; Dressing (d540) if it is Independent; Assistive products and technology for personal use in daily living (e1151) if it is Independent but requires adaptive devices. Items, independently of their nature, are standardized to e3 if the answer is Requires total assistance. This causes an aggregation of diverse content into one indicator, with potential loss of semantics. A solution would be to keep the information found in the question and answer, and add a relation representing the degree of assistance, as shown for SNOMED CT in Table 3.

Items are linked into high level categories. If the content of an item is standardized into high-level ICF categories, it can cause a loss of information as higher categories are a generalization of the concept and also do not usually appear in core sets (see section 4.3). For example, the option Requires total assistance of nearly all items of SCIM is standardized to e3. In the ICF core sets of SCI, there is a lower-level category Health professionals (e355) (which is not a completely satisfactory standardization), but the category e3 is not there. A solution would be to form new core sets according to rehabilitation processes (an initiative which is already ongoing) and taking into account this generality issue both in core-set definitions and in mappings of questionnaires.

If the information provided by the meaningful concept is not sufficient for making a decision about the most precise ICF category it should be linked to, the meaningful concept is assigned not definable (nd), personal factor (pf), not covered by ICF (nc) or health condition (hc). For example, in PCRS, the item Problem in accepting criticism is standardized to nd. A solution would be to use SNOMED CT to complement the ICF, e.g., in this case, with the class Tends to be sensitive to criticism (286846009).

5.1.2. Problems found in the standardization of observation parameters to SNOMED CT. In the mapping of MHR observation parameters to SNOMED CT, the main difficulty is the standardization of the values of the indicators. Specifically, problems appear when different types of items are mapped to the same standard concept and their values need to be combined. Combining
values, and changing data types, has the disadvantage that information is (potentially) lost. For example, the observable entity *Climbing stairs* (129016000) is the standardization of SCIM and FIM concepts. The range of values of FIM’s *Locomotion: stairs* is 1 to 7, while the range of SCIM’s *Stair management* is 0 to 3. Another example is the concept *Anxiety* (48694002), which is measured by GOSE as a Boolean, while in HAD it is presented as a subscale with a range of values between 0 and 21.

6. Interoperability and information models

With respect to interoperability, to exchange EHRs in a standard way, we have taken into account the *Smart Open Services for European Patients* (epSOS), the *virtual medical record* (vMR) and the *ISO/EN 13606* initiatives. The *patient summary* is part of epSOS [14] and includes evaluated persons’ relevant health information. Relevant information is understood as the minimal set of personal health data (160 descriptors) that are of interest to health professionals to assist citizens, and the ignorance of which could pose a risk to the health of the evaluated person. The epSOS’ main objective is to assist practitioners in unscheduled care and its general policy is to adopt international standards, such as HL7 v2. vMR, a simplified view of HL7 v3, is still in the process of improvement and evolution by the vMR Project Team [15], and there are several CDSSs that use it [16]. vMR is designed to reduce development costs and time responses in CDSSs. As a consequence, although it is not widely adopted in hospitals and there are not many tools available to facilitate its implementation (unlike the information model ISO/EN 13606 [17]), we considered it the most appropriate information model for CDSSs today. Table 4 shows some examples of the mapping between CDSS concepts and vMR objects, where the data types used are:

- **concept descriptor** (CD), a reference to a concept defined in an external code system, terminology, or ontology;
- **entity name** (EN), a name for a person, organization, place or thing;
- **timestamp** (TS), a quantity specifying a point on the axis of natural time; a point in time is most often represented as a calendar expression; and
- **interval timestamp** (IVLTS), a set of consecutive values of an ordered base datatype.

These data types are a simplified version of ISO 21090 data types, which is an implementable specification based on the abstract HL7 v3 data types specification.

<table>
<thead>
<tr>
<th>CDSS object (SNOMED CT code)</th>
<th>Reference information model (RIM)</th>
<th>Attribute</th>
<th>Data type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Name (3714840003)</td>
<td>Person</td>
<td>Name</td>
<td>EN [0..*]</td>
</tr>
<tr>
<td>Date of birth (1840990003)</td>
<td>Evaluated person</td>
<td>Birth time</td>
<td>TS [0..1]</td>
</tr>
<tr>
<td>Problem (or disorder, 64572001)</td>
<td>Problem base</td>
<td>Problem code</td>
<td>CD</td>
</tr>
<tr>
<td>Event (272379006)</td>
<td>Adverse event base</td>
<td>Adverse event code</td>
<td>CD</td>
</tr>
<tr>
<td>Date of diagnosis (432213005)</td>
<td>Problem base</td>
<td>Diagnostic event time</td>
<td>IVLTS</td>
</tr>
<tr>
<td>Body functions, activities and participation, environmental factors and body structures</td>
<td>Observation base</td>
<td>Observation focus</td>
<td>CD</td>
</tr>
</tbody>
</table>
7. ARCHITECTURE OF THE KNOWLEDGE-BASED SYSTEM

The architecture of a healthcare institution’s knowledge-representation system which uses the described ontology and interchanges knowledge with one or more health institutions using vMR is illustrated in Fig. 8. This architecture includes the following modules:

- **MHR**: Relevant information derived from each clinical event systematically and sequentially recorded. A *medical health record* is a set of both written and graphic documents, referring to episodes of illness of a patient, and the activity that is generated because of these episodes, stored in electronic form. These clinical data records, mainly from questionnaires, are stored in MHRs.
- **Knowledge storage system**: MHR’s observation parameters standardized into ICF and SNOMED CT indicators.
- **Interface engine**: Engines which encode/decode into/from HL7 v3 messages among healthcare institutions. HL7 v3 is *object-oriented* (OO), uses the *Unified Modeling Language* (UML) and is based on a data model called the *Reference Information Model* (RIM).
- **Aggregator**: Health-institutions knowledge aggregated before storage.
- **Investigator system**: System which analyzes the knowledge and interprets findings about patients.

**Figure 8.** Architecture of a system which implements the ontology and vMR

8. ISSUES IN REASONING

For the decision-making process, we use a *knowledge-based system* (KBS) with *case-based reasoning* (CBR). The CBR reasoning cycle includes the typical retrieve, reuse, revise and retain phases. In the first phase (retrieve), which is implemented using jColibri, the *most similar* patients to the evaluated patient are retrieved and ranked, using a *k-nearest neighbor* (k-NN) similarity measure. Applying reasoning such as CBR is always possible within a healthcare institution because of data homogeneity. Efforts are ongoing (of which this research is part) to adopt common, standard-based representations, therefore allowing the exploitation of data among all compliant institutions. However, even with the multiplication of potential data-sets to be used as case libraries, several issues remain with reasoning in the rehabilitation domain and we mention here some of them:

- New standardized indicators need to be made compatible with existing, historical data series.
The use and weight of demographic variables (such as disease, years from the injury, cause or age) needs to be considered.

Even if standard indicators are used, different patients may be evaluated using different indicators, making their comparison difficult to be defined.

Indicators may not be independent of each other and this dependence in classes and values needs to be taken into account.

The most relevant categories for representing patients with respect to their functional diversity (instead of diseases) need to be determined, using core sets or otherwise.

Similarity calculation among patients is used in the retrieve phase of CBR to provide decision support. About weight assignment for similarity calculation, after several experiments and consultation with medical investigators, we can suggest that the more specific a category, the more weight it should have; and categories which do not appear in core sets should have weight equal to (or close to) 0. The weight given to ICF categories which appear in core sets depends on their level: a relative weight of 1 should be given to 2nd level categories, a weight of 2 to 3rd level categories, and a weight of 3 to 4th level categories. Finally, in approaches using indicators standardized into SNOMED CT, as there are no SNOMED CT core sets, ICF core sets could be used to assign weights.

9. Conclusion

The framework of the research presented in this paper is the progress in clinical decision support using formal reasoning paradigms. To this aim, some previous steps need to be taken, such as: having clear, standard, interoperable knowledge organized in ontologies, nomenclatures (such as ICF and SNOMED CT) and information models (such as vMR); and having a mapping system between clinical questionnaires and indicators expressed in standard terminology (both in terms of classes and in terms of values). Interoperable representations are valuable for a number of artificial intelligence systems whose reasoning is based on experience or analogy, because reasoning result improves if it is based on knowledge accumulated through a wide and diverse range of cases spanning as many healthcare organizations as possible. We analyzed the mapping from observed parameters in the clinical practice to ICF and to SNOMED CT observables and findings; we detected problematic issues and we provided suggestions towards solutions that can improve the current situation, showing the potential of introducing artificial intelligence techniques in medical assessment and proposal.

10. Acknowledgment

The research described in this paper arises from a Spanish research project called Rehabilita (Disruptive technologies for the rehabilitation of the future), which is funded by the Centre for Industrial Technological Development (CDTI), under the CENIT program, in the framework of the Spanish government’s INGENIO 2010 initiative. This work is also partly supported by the Catalonia Competitiveness Agency (ACCIÓ). The opinions expressed in this paper are those of the authors and are not necessarily those of Rehabilita project’s partners, CDTI or ACCIÓ.

References


Luigi Ceccaroni - obtained an M.Sc. degree in Information-Technology Languages and Systems by the Universitat Politécnica de Catalunya - BarcelonaTech (UPC)(Spain); and completed his PhD in Artificial Intelligence (first class honors) at UPC. His main research interests combine: ontologies; recommendation systems; semantic tools; Web services; the semantic Web; personalization; and application of artificial intelligence to healthcare and environmental sciences. During his career he has worked as a member of research staff at the Network Agent Research (NAR) group at Fujitsu Laboratories of America; as director of research at TMT Factory (Barcelona, Spain) and as a senior member of research staff of the Software Department (LSI) at UPC. He has also been adjunct professor of Artificial Intelligence at UPC. He is currently working as a senior member of research staff at Barcelona Digital Technology Centre.
Laia Subirats - holds a B.Sc. degree in Telecommunications Engineering (first class honors) from Pompeu Fabra University (Spain) and a M.Sc. in Telematics Engineering by Universitat Politècnica de Catalunya - BarcelonaTech (UPC) (Spain) and Carlos III University of Madrid (Spain). During her degree she has been working in some projects in disciplines such as ontologies, software engineering, social networks, reputation and e-learning; both in national and international centers (Telefónica I+D and European Organization of Nuclear Research in Geneva, Switzerland) and initiatives such as Google Summer of Code. She is currently working in Barcelona Digital Technology Centre in her Ph.D. in Artificial Intelligence about intelligent systems for decision support in rehabilitation.