

Advances to Semantic Interoperability Through CPR Ontology Enrichment Extracting from SOAP Framework Reports

David Mendes and Irene Pimenta Rodrigues

Universidade de Évora - Portugal

Abstract

We present the work done as a contribution to use an enriched ontology as the support for Semantic Interoperability among clinicians and systems in healthcare providing environments. Clinical practice ontologies are the next generation workhorse for automatic reasoning using Semantic Web techniques and tools in the healthcare sub-domain. Ontology instance acquisition from semi-structured data that renders a full picture of the general clinical practice is crucial for solid enrichment of an Ontology that is designed to embrace the generality of information located in EHR systems. These systems communicate syntactically using HL7 standardized messaging but must evolve to semantic interoperability based in a well formed standardized semantic where CPR acts as a Knowledge Representation infrastructure. Automated acquisition is absolutely a must given the enormous amounts of information available in the mentioned sources. Recent efforts directed to solve the overwhelming complexity of HL7 V3 CDA archetype, like the greenCDA template proposal, along with computability gained with OWL DL ontologies reasoning is leading to the possibility of development of foundations for strong Clinical Decision Support tools and Computable Semantic Interoperability representations in the Semantic Web. As an intermediate step to acquisition from standardized messaging we present the ontology population/enrichment taken from the widespread framework for communication that is the SOAP (Subjective, Objective, Assessment, Plan) clinical encounters documenting system.

Keywords: *Computable Semantic Interoperability; CPR Ontology; Automatic Ontology Enrichment; SOAP Framework*

1 Introduction

Several technologies related to the Semantic Web as well as scientific knowledge and standardization efforts have been developing very recently at an astonishing space that lead us to consider that we have come of age of gathering them all together and produce valuable contributions to the sub-domain of clinical practice automated reasoning. Most of the developments introduced have been maturing for years, or even decades, but finally in the last 2 years the convergence of the mentioned fields are rendering availability of usable products that will allow artificial intelligence researchers and noticeably natural language processing researchers [1] to build

upon.

2 Work done

We have been developing efforts to extract clinical information from texts in Portuguese in order to represent them in computable forms able to reason about using Semantic Web tools and techniques [2]. Previously there was work done in proper selection of an adequate form Knowledge Representation (KR) suitable for the task of supporting the possibility of automated reasoning about clinical practice [3]. Meanwhile, we presented our proposal of ontology population from HL7 V2.xml messaging not yet specifying the target ontologies but

mainly reviewing and focusing on the acquisition¹ process [4].

pro-70 [5] standardized by the World Wide Web Consortium (W3C) in 2009 by reasons that we now explain briefly here for self containment of the present article.

3 Experimental results

In the Portalegre district in Portugal the Unidade Local
 30 de Saúde do Norte Alentejano (ULSNA), E.P.E. (2) has
 in its objectives the provision of primary and secondary
 health care to the population. ULSNA is a health-care
 providing regional system that includes 2 hospitals (José
 Maria Grande in Portalegre and Santa Luzia in Elvas)
 35 and the primary care centers in all the district counties.
 A group of clinicians chosen by our trial investigator
 Dr. Carlos Baeta provided us some dozens of clinical
 reports de-identified according to safe-harbor princi-
 40 ples, as reviewed in [9], from the SAM system in use
 both in the Primary Healthcare units and in the Hos-
 pitals. These clinicians are mainly cardiologists from
 the hospitals but also general medicine (primary care)
 physicians that normally use reports like those provided
 to communicate between them. Using the clinical data
 45 that is available for us we intend to take advantage of
 the tooling presented here to reach the objectives pre-
 sented in the Introduction section. The adequacy of
 the method and tools proposed here, however, are to
 be evaluated by the clinicians themselves. We are cur-
 50 rently in the process of assessing the manual processes
 with a selected group of clinicians both from primary
 and secondary care that will after developing their own
 TMs, at their own pace, integrate them in the complete
 ontology acquisition workflow to be able to assess the
 55 reasoning capabilities that can be made available. We
 are developing a process of quality assessment that will
 be presented as soon as available and processed, ex-
 pected during 2012. After this controlled group we
 intend that the generated TMs are included in a corpus
 60 that will ease extremely the adherence of the rest of the
 community and try to expand to regional level which
 shall then be a straightforward step to achieve.

4 CPR in 2011 as adequate KR for Clinical Decision Support and other automated reasoning endeavours

When developing our previous work [3] we faced sev-
 eral options in choosing an adequate Ontology and we
 chose Computer-based Patient Record Ontology (CPR),
 115

4.1 Theoretical Considerations

The medical practice sub-domain we want to represent
 75 is a many faceted science that renders complexities with
 difficult issues to be addressed such as Temporality,
 Location, Granularity, High ambiguity in free text ter-
 minology, Jargon plagued with acronyms and even per-
 sonal nicknames for example. We have to abide to solid
 80 design foundations for proper Ontology alignment and
 interoperability. Well formed ontologies are able to sup-
 port a variety of secondary uses not anticipated when
 the ontology was originally conceived [6] and we may
 have to pick among available ontologies that have is-
 85 sues of overlapping, ambiguity, non-completeness and
 more. Trying to figure out the availability of such an
 ontology suitable for our purposes, we tried to find
 an available standard ontology according to the Ontol-
 ogy Realism principles enunciated in [7] and with the
 90 freedom to be extendable according to anyone's par-
 ticular needs. The ontologies shall be in accordance
 to the OBO Foundry principles. We just try to bring
 together the latest Software Engineering principles to
 the Ontology Engineering field. With the loose cou-
 95 pling availability and configurable service intermixing,
 we picked what we could spot has low-hanging fruit
 to develop our systems rendering them sub-optimal but
 demonstrable of the validity of the concepts and easily
 extendable/tunable with better ontology support, finer
 100 Web Service provisioning but most of all with better
 clinical judgement about the "smart choices" that have
 to be taken to better populate the ontology given the
 scarce source structure of the original clinical episodes
 texts as seen in section 5.

4.2 Classes to populate

We try to follow the CPR archetypes depicted in Fig-
 ure 1 as close as possible when acquiring from free-
 text. However, some of the more complex classes like
cpr:medical_problem are fairly hard to formulate and
 110 we have to admit that the gleaning presented in the
 present work renders a rather simplistic representation.
 The formalization of a medical practice theory however,
 the purpose of the CPR ontology, renders the possibil-
 ity of important automated reasoning from the simple
 115 acquisition that we are proposing.

A much more systematic representation has to be
 used as source of information for the rendered ontolo-
 gies to approach more the reality of clinical practice

¹In the present paper the term "acquisition" is always referring to clinical terms for ontology instance creation

²<http://www.ulsna.min-saude.pt/>

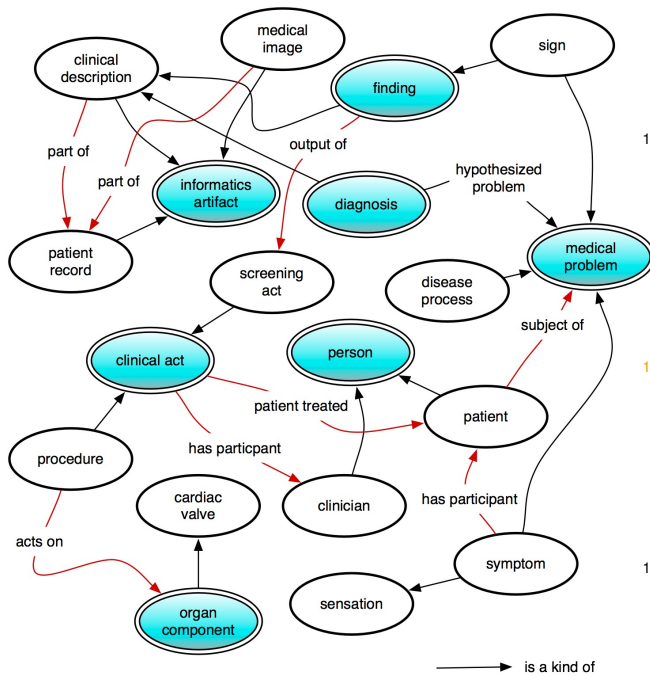


Figure 1: CPR Patient Record Archetypes

for a given clinician, medical device, service, specialty,
 120 healthcare providing institution or system. Some proposals that can contribute to this enhancement of our work are presented ahead in the Future Work section.

5 Automated acquisition from Clinical Episodes Text

125 As reviewed in [4] the state-of-the-art for acquisition from clinical text has enjoyed strong developments in recent years. In the mentioned paper we present a proposal for automated acquisition from HL7 messaging but here we are delving into the more generic possibility
 130 of extracting from free text present in most reporting systems used by clinicians. Going from clinical episodes text that is usually presented in a human friendly format to one adequate for computer processing involves a fair amount of text processing to handle different situa-
 135 tions because: (1) Reports aggregate information from different clinical episodes that are not uniquely identified or not even individually dated, (2) the clinician is only identified by his/her name if any identification is made at all, (3) the information conveyed in free text
 140 is intended only to be understandable by fellow practitioners or even by the clinician him/herself making use of pragmatic jargon normally plagued with acronyms and nicknames abundant in their specific community, (4) text is profoundly intermixed with decorative ele-
 145 ments for better legibility, normally in PDF or HTML files, (5) the clinicians natural language may be other

than English without concepts defined in a foundational thesaurus like SNOMED CT or FMA for instance that don't exist in their particular language and finally, (6)
 150 the time spanning and snapping of the processes depicted in natural language is often difficult to extract and represent formally. We intend to collect our information from SOAP reports like the de-identified sample in Figure 2.

5.1 SOAP PDF Processing

We take advantage of the fact that the report depicts a clinical encounter in a semi-structured way to manipulate into a more tractable source. The Subjective, Objective, Assessment, Plan (SOAP) framework, used
 160 to structure progress notes to facilitate problem specific, clinical decision making by physicians, is a well known, canonical structure in the medical domain. The underlying structure of the SOAP report induces some very important assumptions to be true. We find sections that
 165 can be associated with Subjective, the symptoms section S where we extract directly into a **cpr:symptom record**, medications found here are those administered only during the patient visit; Objective, the objective section O that are sign records **cpr:sign-finding** that
 170 we take as generator for **cpr:clinical_finding** in the Assessment section; Assessment, the analysis section A which are the clinical investigation acts whose outputs can be consequence of any of physiological or patho-
 175 logical processes, and finally, Plan, the plan section P where the **cpr:therapeutic-act** can be extracted with all the timing, posology and prescriptions registered in that particular clinical encounter, medications here are prescribed for discharge [8]. Aggregating the instances collected so far we finally engage in the more complex **cpr:medical_problem** development. We now take advantage of the fact that we have to translate from jargon to English to customize our centralized TMM (Translation Memory Managers) like the **Google translator toolkit**³ or **mymemory translation services**⁴ enhanced with our own Translation Memories and Glos-
 saries. We use the architecture presented ahead in section "The full Software architecture picture" to do all the juggling involved in workflow processing of our
 190 source documents. We start from a PDF document, export to XML, de-identify according to determined legal ruling (like HIPAA Safe-harbor) [9], refine the clinical jargon using automated translations with the aid of CAT (Computer Aided Translation) software with previously
 195 trained Translation Memories in TMX standardized for-
³<https://translate.google.com/toolkit/>
⁴<http://mymemory.translated.net/>

HEALTH CENTER
PONTE DE SOR
MAIN OFFICE

| | | |
|---------------------------|------------|-------------------|
| Paciente 5689_SOAP | | *XXXXXXXX* |
| Birth Date XX-XX-XXXX | (XX Years) | *XXXXXXXX* |
| XXXXXXXXXX | | |
| XXXX XXXXXXXXX | | |

Registo Clínico da Consulta

SPEC. **12/07/2010 18:13** **Dr.(a) Carlos Baeta**

Dr.(a) Carlos Baeta

S_{OAP} Back-external pain episodes:
-Since 2 years.
-Without anginal features.
- With palpitations and facial flushing
Eco - N
Has maintained variable hypertension (140/90)
Keep episodes of palpitations.

O_{AP}

A_{OP} Holter(27/05/10)-RS; 51 a 119: M-75; ESSV infrequent
T3, T4, TSH - N; AVM and Catecholamines - N

P_{OPA} Cordarone - 1 tablet per day
Repeat Holter within 6 months for assessment of the need to Arrhythmology query

| Comercial Name | Qt. |
|--|-----|
| 1 Amiodarona [Cordarone] , 200 mg, Comprimido, Blister - 60 unit(s) Posol.: 1 tablet per day (6 per week) | 1 |

| CARDIOLOGY | |
|------------|---|
| HOLTER | 1 |

Figure 2: SOAP report de-identified sample

mat and then finally convert into raw text to proceed to Semantic Annotation. The pre-processing workflow may be roughly depicted in Figure 3.

5.2 The adequate annotation workflow

A set of sequential steps must be used to go from the pure text to the extracted concept instance. Those steps workflow can be configured declaratively using the software architecture shown in the specific architecture section. The translation steps involved are: (1) Manual translation (that is indispensable for the translator tutoring) with the precise clinicians validation of their jargon adequately translated into English, (2) PDF to raw text, or to structured (XML), converting for adequate documents cleansing. The remote annotation steps are: (3) NER (Named Entity Recognition) of all the patient names, clinician names, addresses, symptoms, signs and prescriptions with all the acronyms, units, time-spanning and time-snapping involved with the usual short forms that a particular doctor usually uses. In our particular situation we maintain a local cache of the previously identified vocabulary to check exact concept matching, (4) WSD (Word Sense Disambiguation) where terms can be disambiguated without technical clinical expertise. Most of them however have to be disambiguated according to the previously acquired concepts in our controlled vocabulary, (5) EAV (Extraction of Attributes and Values) is the final pure, single language, task that has to be performed and in which we need the tooling that this paper refers to filter the concepts from the annotated text to extract concept instances, (6) Semantic annotation using the interface of BioPortal, either manually using the interactive interfaces or automatically with the Web Services available. Given the array of Web Services that can semantically annotate bio-medical concepts in English, we chose to use an evolutionary approach for use of the BioPortal annotator. We first use the annotator Web Interface after manual preprocessing for the TM tutoring and later a fully automatic workflow based in Web Services orchestration.

6 Ontology driven annotation

The annotation step is done using annotation provided by semantically aware REST annotators in BioPortal⁵. This service renders a two step annotation process that builds upon user-selected ontologies to perform the second step of semantic annotation expansion by trying to

correlate the term identified in the first step among the selected biomedical ontologies:

The Semantically Expanded Annotations are then returned by the REST Web Service in TXT, CSV or XML. With this latest option we can proceed to extract into a CPR instance. To define what elements of the returned XML are suitable for CPR instances we further enhance the XML with GRDDL entries specifying the appropriate XSLT transform to apply (2 XML lines). Unfortunately the announced web service provisioning of annotations in OWL format is not yet available from BioPortal but we have the proposed structure and that gives us the ability of developing the transforms for OWL converting as soon as that format becomes available.

7 Restricting to Clinical Practice using only SNOMED CORE

One important and pragmatic restriction that we impose is in the use of SNOMED CORE to restrain the terminology mapping. Instead of using the BioPortal SNOMED CT that would render us the choice among more than 311,000 concepts we intend to explore only a consolidated view of clinical practice terms and we now have the possibility to use only SNOMED CORE that turns our solution much more manageable while not losing relevant terminology. Available since July 2009, the CORE (Clinical Observations Recording and Encoding) Problem List Subset of SNOMED CT was derived based on data-sets submitted by seven large scale health-care institutions. The purpose of the CORE Project is to define a Unified Medical Language System subset that is most useful for documentation and encoding of clinical information at a summary level, such as problem list, discharge diagnosis or reasons of encounters. The most frequently used terms (covering 95% of usage volume) from these institutions are mapped to the corresponding SNOMED CT concepts where such concepts exist. The Subset contains about 5,000 SNOMED CT concepts and this downsizing permits the reasoners to classify much more effectively the resulting annotations.

This important constraint that we are imposing is a practical solution based in what views are currently available but does not take into proper account recent research that is evolving regarding issues in automated extraction like modularization and segmentation as introduced in [10].

⁵<http://biportal.bioontology.org>

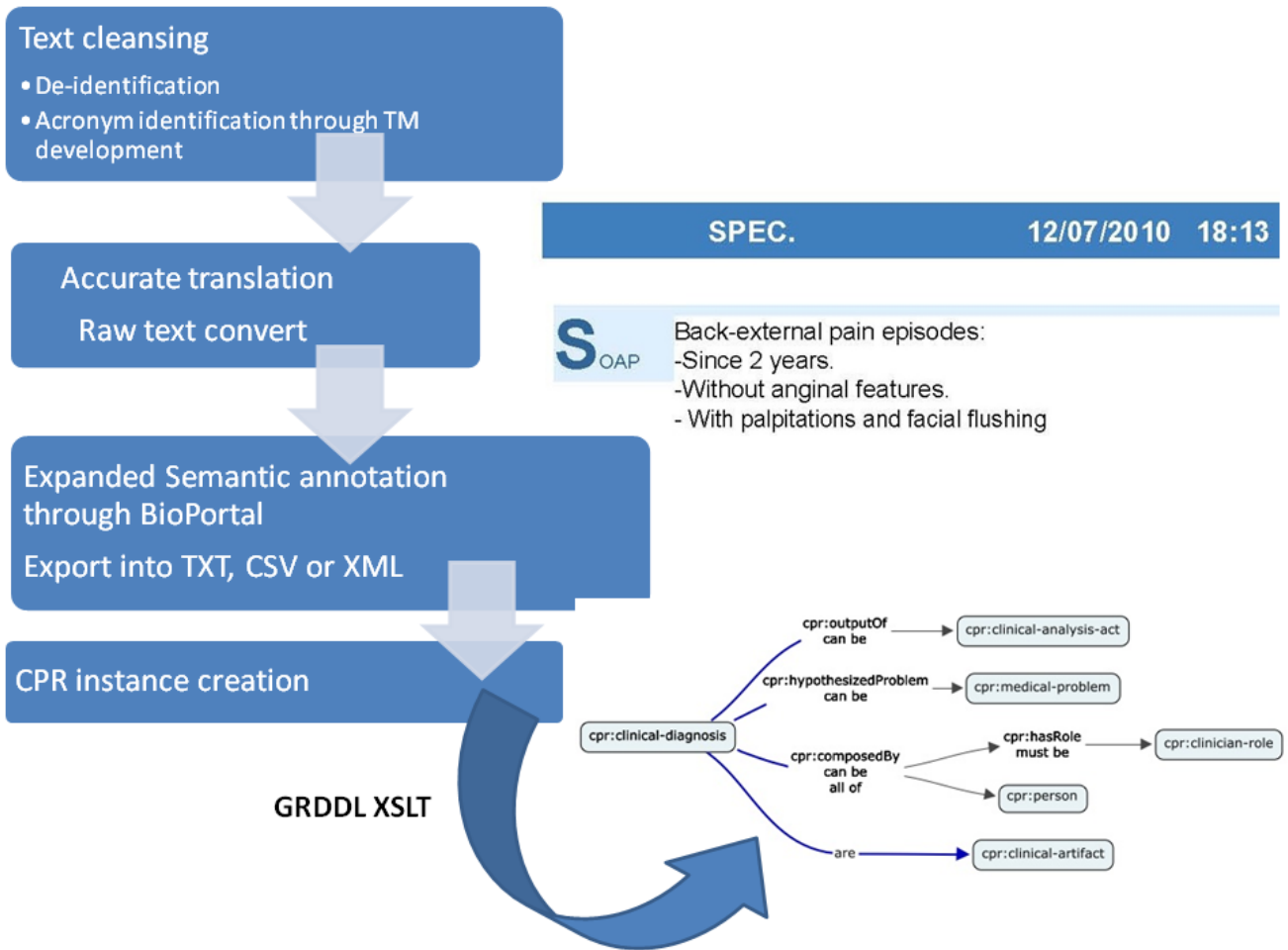


Figure 3: SOAP text processing workflow

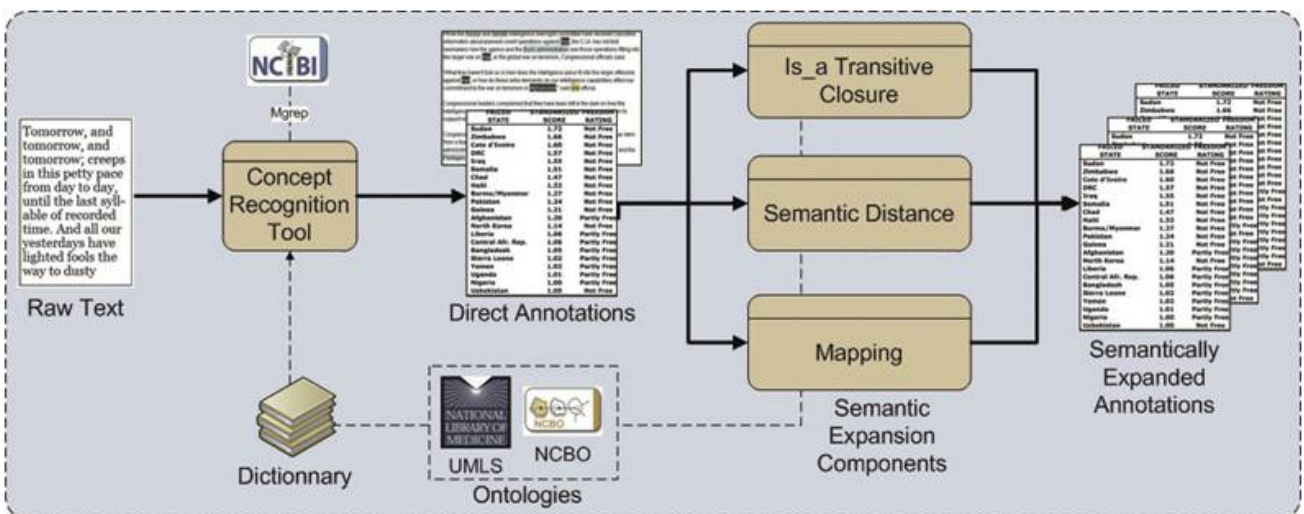


Figure 4: BioPortal Ontology Driven Annotation Workflow

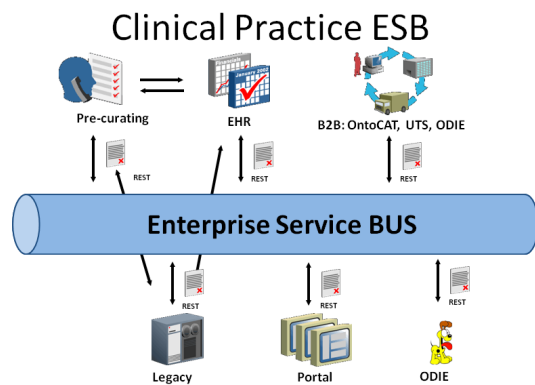


Figure 5: CP-ESB Architecture

7.1 Automatic transform generation

Based in the availability of CPR class archetypes and the XML structured responses from the REST Web Services we developed XSLT mappings that are invoked through the use of GRDDL mechanism. In the case of XSLT we just use plain XML processing but we can even use some Java based framework that has the ability to process Ontologies and GRDDL like Jena⁶ to integrate in the Java Enterprise Edition based system presented in the next section.

7.2 The full software architecture picture

As presented in [11] we build our proposal based in a lightweight messaging bus that we call Clinical Practice – Enterprise Service Bus (CP-ESB). This RESTful based hub is responsible for orchestrating all the communications between Web Service invocations allowing a high degree of customization and plug-and-play ability that renders our proposal very flexible and future proof:

Reviewing the above referred points from SOAP we can now summarize in the following table the complete acquisition points that fit in the timeline for clinical process acquisition proposed by Scheuermann et al in 2009 [12]. Each identified point generates a gleaning entry in `soap_to_cpr.xslt` transform file summarized in Table 1.

As explained in the “Classes to populate” section we can only collect a thin representation when using our SOAP approach but the application of this methodology by the enormous number of clinicians that use this globally widespread framework delivers interesting results like the mere inferencing of the actual relations between

collected instances from S,O,A and the consequences described in P.

8 Path to facilitate CSI

CSI is not dependent of any particular implementation, data structure or archetype in any of the different points that have to share the knowledge. What is important is a shareable meaning foundation that has to be agreed by all. We have previously argued [14] that an important contribution is brought by BRIDG⁷. The Biomedical Research Integrated Domain Group (BRIDG) Model is a collaborative effort engaging stakeholders from the Clinical Data Interchange Standards Consortium (CDISC), the HL7 Regulated Clinical Research Information Management Technical Committee (RCRIM TC), the National Cancer Institute (NCI) and its Cancer Biomedical Informatics Grid (caBIG[®]), and the US Food and Drug Administration (FDA). The BRIDG model is an instance of a Domain Analysis Model (DAM). The goal of the BRIDG Model is to produce a shared view of the dynamic and static semantics for the domain of protocol-driven research and its associated regulatory artefacts. In the mentioned paper we sustain that in Healthcare, every system involved must agree in the meaning of a clinical concept. This is the fundamental concept of Shared Semantics that has to be realized to obtain CSI in the Health sub-domain of knowledge. We present the HL7 RIM inadequacy for ontology mapping and how to circumvent it using the BRIDG DAM. This approach will serve as guideline to alleviate the evident differences between CPR structure and the underlying models structure like the RIM or SOAP. Several options that are taken pursue the DAM.

9 Future work

In 2007 the work done under the W3C - Semantic Web Healthcare and Life Sciences Interest Group auspices⁸ rendered amazing contributions done with the most interesting technological landscape available at the time. Some realities that have since evolved and we could use for the present work are:

- (a) The Problem-Oriented Medical Record Ontology that was started to be developed in 2006 was the target of the fruitful GRDDL efforts that appeared during 2007 in which base we are now working on. This efforts however were rdfs based for the OWL developments were still incipient. Only in

⁶<http://jena.sourceforge.net/index.html>

⁷<http://www.bridgmodel.org/>

⁸<http://www.w3.org/2001/sw/hcls/>

| Instance Type | SOAP Section |
|------------------------|----------------------------------|
| cpr:person | H: Header (if not de-identified) |
| cpr:patient | H |
| cpr:symptom | S |
| cpr:sign-finding | O |
| cpr:clinical-finding | A |
| cpr:clinical_diagnosis | A |
| cpr:therapeutic-act | P |
| cpr:medical_problem | All |

Table 1: XSLT transforms for XML OBA Annotation to CPR instance

365 November 2009 W3C standardized a OWL based ontology called the CPR that we use in our work. **9.2 Using only refined templates of HL7 V3 like greenCDA instead of the full CDA**

(b) For a given HL7 Clinical Document Architecture (CDA) document the GRDDL gleaning presented in 2007 with its associated transform⁹ creates one cpr v0.5 instance with one **cpr:patient.record** with as many **cpr:screening-acts** and associated **cpr:clinical-descriptions** as episodes are referred in the CDA source. This approach, though impressive technology, lacks the granularity that can be rendered now with further developments that CPR has gained in its latest versions.

Taking advantage of the developments done so far we intend to build upon what was then achieved and propose work to be developed like:

380 9.1 Extend our proposal to use GRDDL for gleaning HL7 documents into CPR

The same architecture that we are proposing for gleaning from BioPortal's Open Biomedical Annotator (OBA) output may be, as well, used to extract resources from CDA documents compliant with RIM V3. However it seems obvious that the complexity that the specification conforms could easily "hinder the trees and not show the forest". Many recent proposals are being developed to apply restricted templates to make RIM tractable. The most notorious, that we intend to make reference to here, is the greenCDA¹⁰ project that already has delivered results. In fact, the greenCDA specification has been officially balloted via HL7 and the version 1.0 is already available for download¹¹.

⁹<http://www.w3.org/2001/sw/grddl-wg/td/hl7-rim-to-pomr.xslt>

¹⁰http://wiki.hl7.org/index.php?title=GreenCDA_Project

¹¹http://www.hl7.org/documentcenter/ballots/2010SEP/downloads/CDAR2_IG_GREENMOD4CCD_R1_I1_2010SEP.zip

The idea was that the CDA was capable of representing any aspect of the medical record for any purpose. That sounds like a noble idea but in practice it creates a fixed overhead for even the simplest data exchange. The greenCDA project has developed a pragmatic methodology for creating simplified schemas that can be transformed directly to or from normative CDA. The initiative can be briefly described as a simple to use, XML construct that incorporates structured and unstructured clinical summary information. Simpler than CDA but full featured. There are tools available to convert CDA into Green CDA and also additional tools are being developed to enable easy creation of Green CDA constructs by navigating the RIM, selecting attributes, and selecting associations to consolidate to make the XML flatter. We feel that the development of XSLT for GRDDL processing from greenCDA documents is a very viable process of rendering a quasi complete picture of clinical practice extracting from HL7 messaging. We already have available the XML-Schema and transforms for back and forth CDA conversion and very easily will develop the OBA annotator output of greenCDA documents to CPR GRDDL transforms.

420 9.3 Restrict CPR to be OWL DL well formed or even further

Unfortunately, CPR is not currently OWL DL syntax conformant. For the current reasoners to work effectively with OWL based ontologies they have to be OWL DL to be validatable in linear time. Although our clinical practice ontologies are restricted in their size most of the times, the reasoners check the validity a-priori and tend to refuse to work when some non-conformities are found. One interesting line of research may be the adequate restriction of CPR to be OWL DL conformant without loss of the needed expressiveness for clinical practice representation. We found even arguable that

with adequate restrictions, like elimination of transitivity [13] a Horn SHIQ ontology can be developed⁴⁸⁰
⁴³⁵ from CPR and then we reach the possibility of applying
 Consequence-Driven Reasoners that have been shown
 recently to be very attractive for dealing with huge data-
 sets.

10 Conclusion

⁴⁴⁰ We presented our proposal for using the enriched CPR
 ontology, automating the acquisition with resource to
 SOAP framework reports, as support for Computable⁴⁹⁰
 Semantic Interoperability. Being a framework directed
⁴⁴⁵ to human interpretation it lacks the completeness that
 can be achieved with more formal ways of representing
 clinical encounters directed to computer consumption
 like the various standardized protocols and archetypes
 developed for computer to computer interaction through
 messaging like ISO HL7 27931:2009 or the Archetypes⁴⁹⁵
⁴⁵⁰ in HL7 V3 RIM. We demonstrated the workflow to
 evolve from a raw SOAP report in natural language to a
 CPR instance creation both manually and automatically
 through the use of the available Web Services that “on-
 tology driven” annotate with resource of “feeder ontolo-
⁴⁵⁵ gies” that we criteriously select. Our final architecture
 has the possibility to automatically process the source
 documents into the ontology through the suggested soft-
 ware components that is easily expanded/refined to in-
 corporate new sources of clinical information, enhanced
⁴⁶⁰ ways of annotation and different target ontologies in a
 plug-and-play manner.⁵⁰⁵

References

1. Demner-Fushman D, Chapman WW, McDonald⁵¹⁰
⁴⁶⁵ CJ. What can natural language processing do for
 clinical decision support?. *Journal of Biomedical
 Informatics* 42 (5), Biomedical Natural Language
 Processing. 2009; pp. 760-772.
2. Mendes D, Rodrigues IP. Automatic Ontology Pop-⁵¹⁵
⁴⁷⁰ ulation extracted from SAM Healthcare Texts in
 Portuguese. In: *Jornadas de Informática da Univer-
 sidade de Évora 2011*. ISBN: 978-989-97060-2-6,
 URL <http://host.di.uevora.pt/docs/actas-jjue2011.pdf>
3. Mendes D, Rodrigues IP. Well Formed Clinical⁵²⁰
⁴⁷⁵ Practice Ontology Selection. In: *Jornadas de In-
 formática da Universidade de Évora 2011*. ISBN:
 978-989-97060-2-6, URL: [http://host.di.uevora.pt/
 docs/actas-jjue2011.pdf](http://host.di.uevora.pt/docs/actas-jjue2011.pdf)
4. Mendes D, Rodrigues IP. A Semantic Web prag-
 matic approach to develop Clinical ontologies,
 and thus Semantic Interoperability, based in HL7
 v2.xml messaging. In: *HCist 2011 - Proceed-
 ings of the International Workshop on Health and
 Social Care Information Systems and Technolo-
 gies*. Springer-Verlag - book of the CCIS series
 (Communications in Computer and Information
 Science).
5. W3C, 2009. Computer-based Patient Record
 Ontology URL [http://code.google.com/p/
 cprontology](http://code.google.com/p/cprontology).
6. Smith B, Scheuermann RH. Ontologies for clinical
 and translational research: Introduction. *Journal
 of biomedical informatics* 44. 2011; 3-7. URL
<http://www.ncbi.nlm.nih.gov/pubmed/21241822>.
7. Smith B, Ceusters W. Ontological realism: A
 methodology for coordinated evolution of scien-
 tific ontologies. *Applied ontology* 5 (3-4), 2010;
 pp. 139-188.
8. Mowery D, Wiebe J, Visweswaran S, Harkema H,
 Chapman WW. Building an automated soap classi-
 fier for emergency department reports. *Journal of
 Biomedical Informatics*. 2011.
9. Meystre S, Friedlin M, South FJ, Shen BR, Samore
 S. Automatic de-identification of textual docu-
 ments in the electronic health record: a review
 of recent research. *BMC medical research method-
 ology*. 2010; pp. 10-70.
10. Ogbuji C, Arabandi S, Zhang S, Zhang GQ. Seg-
 menting and Merging Domain-specific Ontology
 Modules for Clinical Informatics. In: Galton A,
 Mizoguchi R (Eds.), *Frontiers in Artificial Intelli-
 gence and Applications: Formal Ontology in Infor-
 mation Systems - Proceedings of the Sixth Interna-
 tional Conference (FOIS 2010)*; pp. 414-428.
11. Mendes D, Rodrigues IP. Clinical Practice On-
 tology Development Best Practices. *Journal of
 Biomedical Informatics*. 2012; To be published.
12. Scheuermann RH, Ceusters W, Smith B. Toward an
 Ontological Treatment of Disease and Diagnosis.
AMIA Summit on Translational Bioinformatics.
 San Francisco, CA. 2009; pp. 116-120.
13. Kazakov Y. Consequence-driven reasoning for
 Horn SHIQ ontologies. *IJCAI*. 2009; pp.
 2040-2045.

- 525 14. Mendes, D. & Rodrigues, I.P. The BRIDG model
as the most authoritative resource in Shared Se-
mantics for Ontologies development in Healthcare
practice. In Proceedings of the International Con-
ference on Knowledge Management and Informa-
530 tion Sharing, pp. 1-9.

Conflicts of Interest

None declared.

Correspondence

David Mendes
535+351 963564161
dmendes@uevora.pt