Towards the semantic enrichment of Computer Interpretable Guidelines: a method for the identification of relevant ontological terms

Manuel Quesada-Martínez, PhD[1,2], Mar Marcos, PhD[3], Francisco Abad-Navarro, MSc[1], Begoña Martínez-Salvador, PhD[3], Jesualdo Tomás Fernández-Breis, PhD[1]
[1]Departamento de Informática y Sistemas, Universidad de Murcia, IMIB-Arrixaca, Spain;
[2]Center of Operations Research (CIO), Miguel Hernández University of Elche, Spain;
[3]Dept. of Computer Engineering and Science, Universitat Jaume I, Spain

Abstract

Clinical Practice Guidelines (CPGs) contain recommendations intended to optimize patient care, produced based on a systematic review of evidence. In turn, Computer-Interpretable Guidelines (CIGs) are formalized versions of CPGs for use as decision-support systems. We consider the enrichment of the CIG by means of an OWL ontology that describes the clinical domain of the CIG, which could be exploited e.g. for the interoperability with the Electronic Health Record (EHR). As a first step, in this paper we describe a method to support the development of such an ontology starting from a CIG. The method uses an alignment algorithm for the automated identification of ontological terms relevant to the clinical domain of the CIG, as well as a web platform to manually review the alignments and select the appropriate ones. Finally, we present the results of the application of the method to a small corpus of CIGs.

Introduction

According to the latest definition, Clinical Practice Guidelines (CPGs) are "statements that include recommendations intended to optimize patient care that are informed by a systematic review of evidence and an assessment of the benefits and harms of alternative care options"[1]. CPGs have the potential to facilitate the translation of clinical research results into practice, and to improve the quality and outcomes of healthcare. Computer-Interpretable Guidelines (CIGs) can be defined as formalized versions of CPG contents for use as decision-support systems. The emergence of CIGs was motivated by the interest in making CPG recommendations available to clinicians in an easier and immediate way, compared to CPGs in text form. The benefits of the use of CIGs in clinical settings are documented in the literature, and include improved CPG compliance and increased efficiency[2].

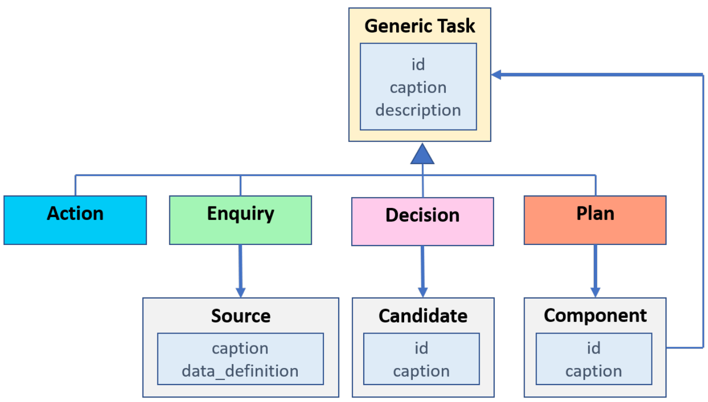
CPGs are rich in knowledge of very varied type. In line with this, CIG representation languages provide a wide range of modeling constructs. Peleg *et al.* identify several dimensions falling into two main categories: structuring in plans of decisions and actions, and linking to patient data and medical concepts[3]. The latter includes the definition of a domain model comprising classification hierarchies, concept definitions, and relationships between concepts. Such a domain model is of particular relevance when it comes to deploy the CIG within the clinical information system. The reason is that CPGs (and thus CIGs) often refer to data at a rather high level of abstraction, which implies that they are not directly available and thus should be derived from the data in the Electronic Health Record (EHR) e.g. using hierarchies or definitions of abstract concepts.

In this context, we aim at the semantic enrichment of the CIG using an OWL ontology purposely designed for the clinical domain of the CIG, including ontological terms e.g. for the clinical data that are referred to as well as the diagnostic and therapeutic interventions that are used. This ontology could be used e.g. as starting point for the conceptual and abstraction definitions required for the interoperability with the EHR[4,5]. As a first step, in this paper we describe a method to support the development of such an ontology starting from a CIG. The method uses an alignment algorithm for the automated identification of ontological terms relevant to the clinical domain of the CIG, based on both the natural language descriptions of CIG elements and existing ontologies. Rather than starting from the CPG text, we have chosen to use the CIG to take advantage of the text analysis and formalization efforts made by the CIG modeler. Specifically, the method works with models represented using the PROforma CIG language[6], and uses an adapted version of the OntoEnrich framework[7] algorithms to identify alignments with the SNOMED CT ontology[8]. We also describe a web platform designed to allow the expert (typically, a knowledge engineer proficient in both the PROforma language and the CIG domain) to review the terms identified by the algorithms, directly accessing the official SNOMED CT browser[9], and select the appropriate ones. Finally, we present the results of the application of the method to a small corpus of 9 CIGs.

Materials and Methods

The PROforma CIG language

There is a number of proposals for representing CPGs (and clinical protocols) in a computer-interpretable format[10]. In this work we focus our attention in the PROforma language[6], which is an executable process description language grounded in a task-based formalism. Tasks represent actions or activities to be performed by an external agent (e.g. clinician, nurse) or by the execution engine itself. There are four types of tasks, namely enquiries, decisions, actions and plans (see example in Figure 1). Enquiries request data from the environment, which have to be entered by the user or read from a database. Decisions are activities where a choice has to be made among different candidates. Actions represent activities that have to be performed in the external environment (e.g. perform blood glucose level test). Finally, plans group together several lower-level tasks. Using plans, tasks can be organized hierarchically from an initial root plan. In addition to tasks, data play a major role, e.g. enquiries refer to sources, which are data items whose value has to be supplied.




```
action :: 'PSA_test_action_8' ;
  caption :: 'PSA_test_action' ;
end action.
```

```
enquiry :: 'PSA_after_Treatment_enquiry_3' ;
  caption :: 'PSA_after_Treatment_enquiry' ;
  source :: 'PSA_level_Evaluation_after_Treatment' ;
    caption :: 'PSA_level_Evaluation_after_Treatment';
      description :: '' ;
      data_definition ::
        'PSA_level_Evaluation_after_Prostatectomy_type';
  source :: 'PSA_level_value_after_Treatment' ;
    caption :: 'PSA_level_after_Treatment' ;
      description :: '' ;
      data_definition :: 'realType' ;
end enquiry.
```

```
plan  :: 'VLRisk_Therapy_for_LE_Lower_20y' ;
  caption :: 'VLRisk_Therapy_for_LE_Lower_20y' ;
    precondition ::
       result_of( VLRisk_LifeExpectancy_Decision )
         = LE_Lower_20y;
    component :: 'Active_Surveillance_plan_1' ;
      caption :: 'Active_Surveillance_plan' ;
      task_definition :: 'Active_Surveillance_plan_1' ;
      number_of_cycles ::1;
      ltwh :: 248,120,48,36;
end plan.
```

Figure 1: The PROforma task model with examples.

The PROforma language has been used for building and deploying a range of decision support systems, CPGs, and other clinical applications. Here, our goal is to analyze the content of PROforma CIGs in order to define a method for the enrichment of such CIGs. With a view to this enrichment, we are primarily interested in those aspects where clinical background knowledge plays a fundamental role, such as tests, test results, diagnosis, etc. Consequently our method focuses on the task and data elements of the PROforma language (see details below).

A method for the semantic enrichment of PROforma CIGs

In this work, we propose the use of ontologies[11] for bridging the gap between CIGs and other resources/systems, such as databases of clinical trials and EHRs. Ontologies can be defined as controlled vocabularies that allow describing the meaning of data (its semantics) in a human and machine-readable way. They are used more and more often to aid processing of information in biomedical research and in healthcare systems[12]. Biomedical ontologies, such as SNOMED CT, can be found in repositories like BioPortal[13].

The goal of our method is to obtain a CIG domain model which could be used e.g. to achieve the semantic interoperability between the CIG and the EHR using the concept definitions it contains[5]. Figure 2 shows the pipeline of the proposed method. It takes as input a CIG file codified in the PROforma language and a source biomedical ontology codified in OWL. The method computes the alignments between the CIG content and the ontology. Conceptually, an alignment expresses the relationship between the CIG content and the ontology. For example, an alignment between the action "PSA_test_action" (see Figure 1) and the SNOMED CT term "Prostate specific antigen (substance)" (SNOMED CT id 102687007) shows some relationship between the action represented in the CIG and the substance defined in SNOMED CT. Then the alignments are validated by a human expert, and based on this, a module of the ontology is automatically extracted. This module not only contains the concepts found by the alignments based on the CIG, but also other concepts that are logically related with them in the ontology. For example, since the SNOMED CT term "Prostate specific antigen (substance)" is related with the term "Peptide hydrolase (disposition)" (SNOMED CT id 741072006) through the attribute "has disposition", the CIG-based ontology module would contain such term too. This module is an independent OWL file which could be finally used by the human expert for creating a custom domain model to enrich the CIG.

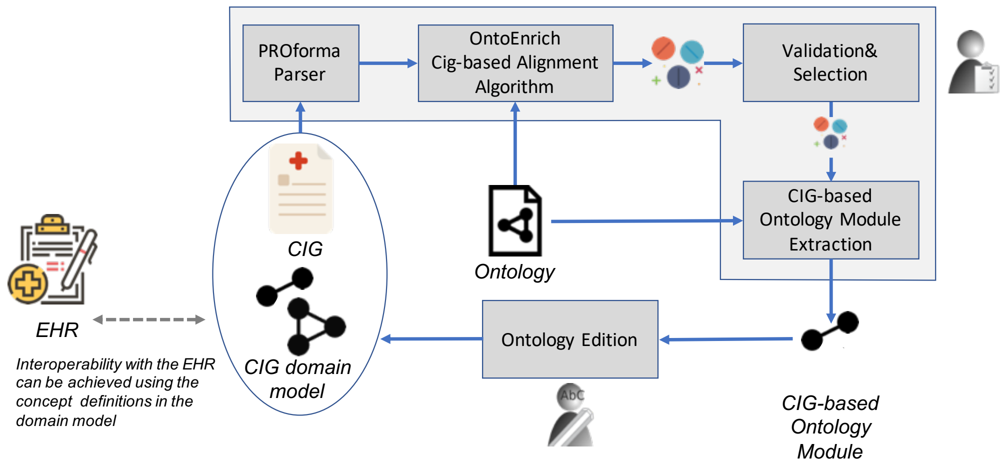


Figure 2: Summary of the different stages of the proposed method.

The proposed method exploits the CIG content by using a textual alignment algorithm. For that, we have analyzed which content of the CIG file is relevant. The analysis of the PROforma format lets us determine which fields should be relevant if we want to capture the CIG content expressed in natural language. Each PROforma task has a unique alphanumeric attribute that works as identifier. Moreover, each task may contain a set of attributes, such as: "caption", "description", "precondition", "data_definition" and so on (see Figure 1). Since our method applies a textual alignment algorithm, we target attributes containing natural language information. For example, given the plan "VL-Risk_Therapy_for_LE_Lower_20y" of Figure 1, the attributes "number_of_cycles" or "ltwh" can be considered meaningless, while other like the identifier or the caption could be useful.

OntoEnrich CIG-based alignment algorithm

From a lexical point of view, a PROforma CIG can be described as follows. A CIG file will be characterized by a set of tasks $\{T_1, ..., T_n\}$. In turn, each task T_i will be lexically described by a set of attributes $\{A_1, ..., A_m\}$. In order to define a method generic enough to be adapted to different CIG modelers, we define a configurable set $\{AT_1, .., AT_k\}$ with the attribute types that will be analyzed. For example, the "id" and "caption" of the task "plan" (Figure 1) contain a natural language description useful for its alignment with external ontologies, consequently the "id' and "caption" attributes will be included in the AT set. This lexical model of the CIG will be used for linking each T_n with a set of alignments $\{M_1, ..., M_s\}$. Finally, each alignment M_k will be characterized by its description in natural language and the associated Uniform Resource Identifier (URI). For example, the action "Biopsy_of_Suspicious_Nodes_action" will be related with SNOMED CT through the alignment "Biopsy (procedure)", whose URI is `http://snomed.info/id/86273004`.

Alignment algorithms have been traditionally used for finding similar terms defined in two or more ontologies[14]. An alignment could be obtained using different approaches like purely lexical, using semantics relations, or as a combination of both. In this work, we propose the use of an extension of the OntoEnrich framework algorithm[7]. OntoEnrich alignment is lexical and includes some stages taken from the Natural Language Processing (NLP) pipelines such as as tokenization and lemmatization. The lexical context of the CIGs is defined by the attributes, which can be in turn of different nature. This forces us to use an alignment algorithm based on tokens, to achieve more flexibility. As a result, even though there are no exact alignments for a complete sentence we are able to find alignments for some parts of it. For example, we found no alignments for the sentence "Biopsy_of_Suspicious_Nodes_action" however we found one for "Biopsy (procedure)".

OntoEnrich algorithms had been previously used for finding relationships between the terms of two ontologies. In this work these algorithms have been adapted to PROforma CIG elements as sketched above. Figure 3 shows a schematic view of the OntoEnrich algorithm, including the NLP pipeline, together with an illustrative example of its application to a CIG for the assessment of colorectal cancer risk (CRcaTriage CIG). In the process, OntoEnrich builds a graph of lemmas that is subsequently exploited for obtaining the mappings (further details can be read in our previous work[7]). OntoEnrich does integrate different NLP methods, and the user can configure which ones to use. For example, the Stanford core NLP tools[15] could be used to process the natural language descriptions. However, note that OntoEnrich does not use external services, such as BioPortal alignment services[13], for comparing the textual descriptions in the attributes.

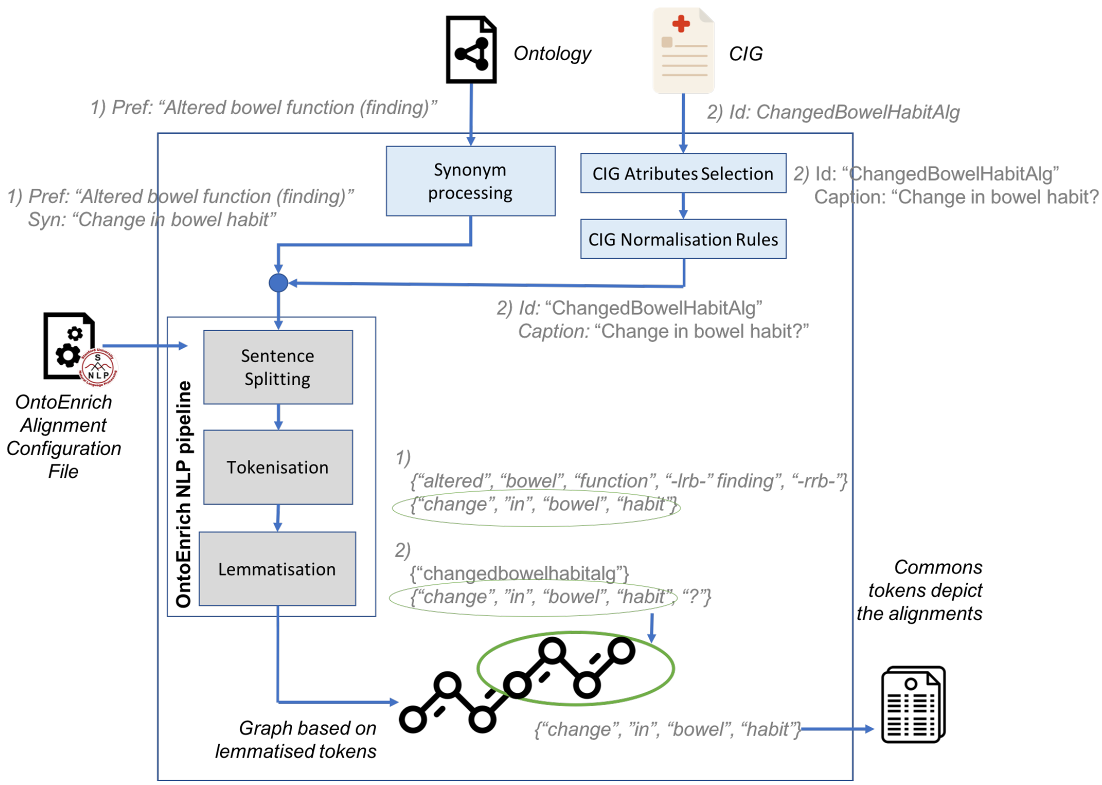


Figure 3: Schematic view of the OntoEnrich alignment algorithm with an illustrative example.

The following adjustments have been made to the OntoEnrich algorithms:

- We parse the CIG file and obtain the set of a attributes $\{A_1, ..., A_m\}$. Based on the output, a customized set of text normalization rules (using regular expressions) is defined and applied for cleaning the CIG attributes. For example, in Figure 1 we observe that the captions of the "enquiry" and "action" attributes contain the type of task as suffix, therefore a rule for removing the words "_enquiry" or "_action" can be added.
- The cleaned attributes and the source OWL ontology, SNOMED CT in our case, are the inputs to the alignment algorithm. The algorithm not only takes into account the SNOMED CT preferred terms but also the synonyms. For instance, the string "Change in bowel habit" appears in one of the synonyms of the SNOMED CT preferred

term “Altered bowel function (finding)”, hence this SNOMED CT term is proposed as an alignment of the PROforma element with id “ChangedBowelHabitAlg”, whose “caption” is “Change in bowel habit?”. It should be pointed out that an alignment occurs when one token in the attributes, or a consecutive combination of them, is found as label of one ontology term or its synonyms. In other words, partial alignments are allowed in the CIG side but not in the SNOMED CT side.

- We define a set of filters that let us select, for each task, the most specific alignments in $M_1, ..., M_s$, that is, we remove those alignments of tokens that are contained in others. For example, the alignment “Scores (qualifier value)” obtained for the “Gleason_score_plan” is removed since the alignment “Gleason score (observable entity)” is more specific.

The advantages of applying a token-based alignment algorithm is that we obtain a large amount of ontological terms that are lexically related with the CIG content. However, this textual alignment lacks information about the context, which could make some alignments not specific enough. For this reason, we offer the expert the possibility to confirm or reject the alignments found at the end of the process.

CIG-based ontology module extraction

Finally, those alignments validated by the expert are the *seed signature* S for the automatic extraction of a locality-based module M[16]. This module includes a subset of the axioms in an ontology θ, and is extracted from θ for a set S of concepts (concept or property names) in the alignments. Informally, everything that can be inferred from the ontology θ about the topic consisting of the concepts in S and M, is already known by its module M. Figure 5 shows an example of the axiomatic description of the alignment “Biopsy”. Conceptually, the module will include not only the SNOMED CT term “Biopsy (procedure)” but also its related axioms. Regardless of this, note that a single module will be extracted for each seed signature, which is derived from the corresponding CIG.

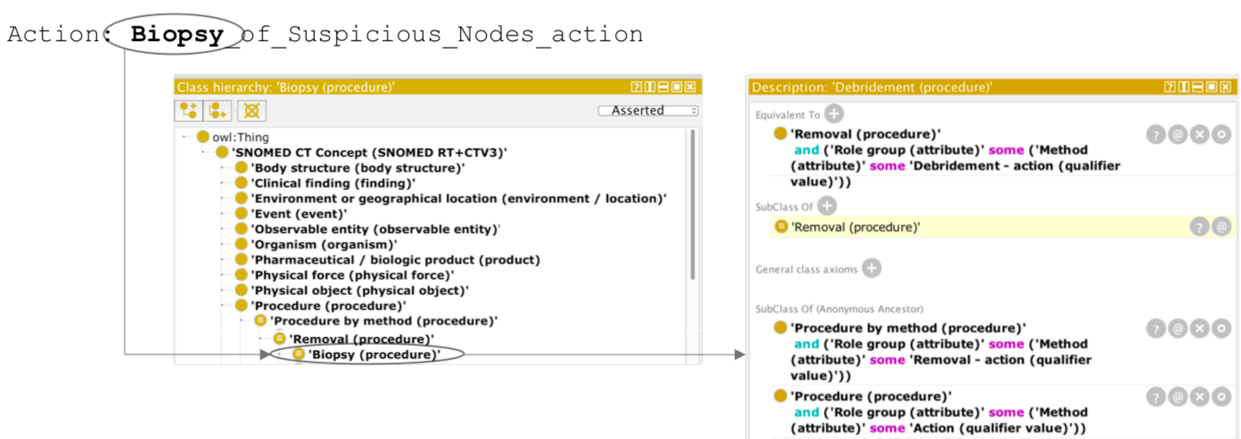


Figure 4: Example of an alignment found in the SNOMED CT ontology for the PROforma task “Biopsy_of_Suspicious_Nodes_action”.

Results and discussion

The method has been implemented in Java and encapsulated as a library. It receives as input a source ontology processable by the OWL API library and a PROforma CIG. For the former we use the SNOMED CT OWL version released on January 2018. The selection of SNOMED CT was done based on an initial experiment consisting in searching for alignments between the content of a sample CIG and a set of BioPortal ontologies. Since SNOMED CT was the ontology that produced the highest number of alignments, we concluded that it was the most promising ontology for our purposes. Another argument in favor of the selection of this ontology was the expertise of the authors.

Those parts of the method that do not require an interaction with the user are executed automatically. Additionally, a web platform has been created to facilitate the analysis of the alignments as well as to select the appropriate ones after confirming their validity. This should be manually done by an expert knowledgeable of the CIG domain. Both the platform and the results described in this section can be accessed in the URL `http://sele.inf.um.es/cig-enrichment`.

We applied the method to a corpus of 9 PROforma models. The majority of these models come from the OpenClinical.net repertoire[10], which is an open access repository of CIG applications (with associated documentation) developed by different authors. The corpus includes 2 other PROforma CIGs previously authored by researchers of this work, which model respectively a pathway for the management of coexistent chronic heart failure and chronic obstructive pulmonary disease[17] and a CPG for the management of prostate cancer. This implies that various modeling aspects (naming conventions, etc.) differ across CIGs. Note also that the corpus covers a variety of specialities, including Emergency Medicine, Cardiology, Mental Health, Oncology, and Respiratory Medicine.

Based on the the manual analysis of this corpus, we configured the attribute set as AT={id, caption, description, range}. Table 1 shows the number of elements for each CIG together with the number of alignments found, grouped by type of task. The number of alignments reveals that the content of our CIG corpus is lexically present in a relevant ontology like SNOMED CT. This is a promising result, since it shows that part of the content expressed in natural language within this set of CIGs could be semantically enriched using terms from this ontology.

Table 1: Results of the alignments found in the corpus of 9 PROforma CIGs. Columns #ele. and #alig. contain, respectively, the number PROforma elements and the number alignments.

	Action		**Enquiry**		**Decision**		**Plan**		**Data**	
	#ele.	**#alig.**	**#ele.**	**#alig.**	**#ele.**	**#alig.**	**#ele.**	**#alig.**	**#ele.**	**#alig.**
Asthma (Assessment and treatment of asthma)	19	67	3	27	13	120	11	20	25	136
CHF-COPDpathway (Management of coexistent CHF and COPD)	66	214	31	82	5	8	26	84	11	34
Cough (Diagnosis and treatment of chronic cough)	7	66	5	27	3	40	10	44	11	59
CRcaTriage (Colorectal cancer risk)	2	0	4	6	1	0	1	3	10	31
Depression (Diagnosis and management of depression)	8	14	3	10	4	20	4	14	22	115
HeadInjury (Work-up and management of acute head injury)	13	6	9	11	4	1	9	15	37	69
PCguideline (Management of prostate cancer)	102	338	59	234	19	14	60	156	20	118
STIK (Soft tissue injuries of the knee)	6	12	6	0	8	0	7	20	14	18
TBScreening (Screening for tuberculosis)	6	18	5	7	2	4	2	6	8	18

Preliminary analysis of the alignments

An in-depth examination of the results shown in Table 1 can be done through our web platform. Figure 5 shows an example of how the CIG elements and their alignments are presented to the expert. Using this functionality, a preliminary analysis of the results was carried out by two knowledge modelers who were familiar with the CIGs and their respective domains. Note that no formal protocol was established for this analysis and hence it cannot be regarded as a validation. For instance, the two modelers could assess very differently the accuracy of alignments that were not specific enough. A series of illustrative examples from this preliminary analysis are described below.

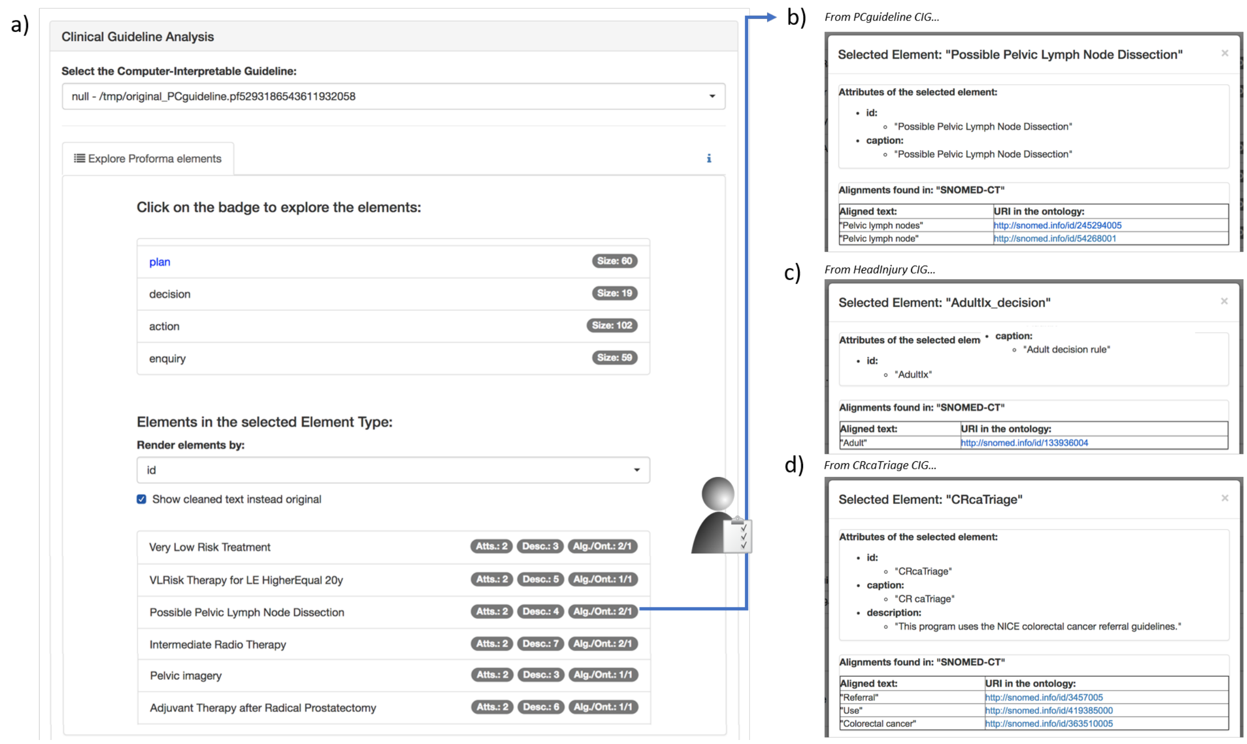


Figure 5: Composition of screenshots showing the alignment results as presented by the web platform: a) Alignment results for the plans of PCguideline CIG; b) Alignments of "Possible Pelvic Lymph Node Dissection"; c) Alignments of "AdultIx"; and d) Alignments of "CRcaTriage".

The user can focus on the inspection of the context defined by the alignments found for a certain PROforma task, which contains lexical content codified in the AT set. The advantages of this presentation can be observed in the examples of Figures 5 b), c) and d). The NLP alignment strategy based on tokens enables the algorithm to detect e.g. singular and plural, such as in the case of "Possible Pelvic Lymph Node Dissection", which is aligned with both "Pelvic lymph nodes" and "Pelvic lymph node" (see Figure 5 b)). Other cases have been completely covered by the alignments. An outstanding example is "fatigue_lack_of_energy_or_tired_all_the_time", from the Depression CIG. This PROforma term has been aligned with "Fatigue (finding)", "Lack of energy (finding)" and "Tired all the time (finding)", all of which are correct.

In a number of cases the experts reported an expected SNOMED CT alignment which was not found by the method. For example, while the SNOMED CT term "Loss of interest (finding)" was proposed as alignment for the element "loss_of_interest_or_pleasure_in_usual_activities", from the Depression CIG, the term "Loss of capacity for enjoyment (finding)", whose synonym is "Loss of pleasure", was not proposed. The reason is that currently we only search for partial alignment of consecutive tokens. This could be solved by considering other non consecutive combinations. The alignments proposed for the "CRcaTriage" plan of CRcaTriage CIG are an example of alignments labelled as partially correct by the experts (see Figure 5 d)). Here the SNOMED CT alignment "Colorectal cancer" was proposed based on the "description" attribute, and it was deemed correct by the experts. However, the SNOMED CT term "Triage (procedure)" suggested by the experts was not detected as an alignment because our current implementation of the NLP pipeline does not identify several tokens in the "CRcaTriage" string. This could be solved by tuning our algorithm to take into account this kind of camel-case style, which would increase our recall.

Another problem is related to the use of acronyms. For example, in the previous example the acronym “CRca” stands for colorectal cancer. Although this acronym is not considered in SNOMED CT, the information codified in the “description” attribute allows us to detect the right alignment with “Colorectal cancer”. In contrast, no alignments were proposed for the action “PSA_Test_action_2” of the PCguideline CIG, because although the acronym “PSA” is included in the SNOMED CT synonym “PSA - Prostate specific antigen”, it does not exactly match the label. This has led us to consider whether partial alignments should be allowed in the ontology side. However, ontology labels in principle describe the terms they represent without ambiguity, therefore fragmenting them would make the alignment lose part of its context.

This preliminary analysis has served us to assess the strengths and weaknesses of our method, as well as to identify some points of improvement that would increase its performance. With respect to the practical utility of the tools, both the automation of the process and the availability of a web platform to analyse the alignments have been highlighted by our knowledge modelers as an important advance with respect to the current situation, which would facilitate the difficult and time-consuming task of developing a CIG-based domain model from scratch.

Conclusions and future work

In this paper we present a method with a view to the semantic enrichment of CIGs represented in the PROforma language. The method is based on an algorithm that automatically computes the alignment between the natural language contents of the CIG and the SNOMED CT ontology, as well as on a web platform through which the expert can review the resulting alignments and select the most appropriate ones to extract a custom ontology module. Note that although the platform currently shows an OWL module assuming that all the alignments are selected, it is ready to automatically extract a module based only on those selected by the expert.

As mentioned before, we chose to use the CIG as source material to take advantage of the CPG analysis and formalization efforts made by the modeler. This is very important in itself, because the CPG usually includes extensive information on many more aspects than those relevant to CIG representation. Thus, the use of the entire CPG text would not only dismiss the modeling efforts but also add unnecessary noise. As a consequence of this decision, our method does not require to deal with negative sentences, because negation rarely occur in CIGs, as opposed to CPGs[18]. In any case, since our alignment algorithm is based on tokens, negative sentences would be partially covered.

The alignment results obtained so far are promising according to our knowledge modelers, although there is room to improve the performance of the algorithm. With respect to the overall method, our modelers agree that it could facilitate and speed up to a great extent the development process of a custom ontology to enrich the CIG. However, it remains to be determined whether the ontology modules obtained in this way are manageable, e.g. the number of terms related to the selected alignments might be excessive.

In the literature, the enrichment/extension of CIGs has been proposed in few papers. The GLARE framework considers an enrichment with exceptions, or unexpected situations that could occur and that require a deviation from the standard execution of the CIG[19]. Likewise, the MobiGuide project proposes an extension with CIG customized contexts, induced by either personal or technological concepts, defining which changes must be made to the CIG under specific circumstances[20]. In contrast, our ideas for the enrichment of the CIG involve supplementary knowledge that can be used e.g. for the interoperability with the EHR or even for educational purposes.

As future work we intend to improve our alignment algorithm so that it can deal with the problems of the current version. On the other hand, we plan to perform a principled and comprehensive validation. As explained before, our analysis of results lacked a defined protocol including aspects such as initial training and assessment guide for evaluators. Additionally, the validation of our method requires considering a larger CIG corpus and involving a greater number of evaluators. Apart from this, we are also interested in performing a comparative analysis of the alignment results obtained using a set of ontologies beyond SNOMED CT, e.g. those publicly available in repositories such as BioPortal. Over the longer term, we will consider to carry out a case study on the coverage of the CIG ontologies obtained with our method for the purpose of CIG-EHR interoperability.

Acknowledgements

This work has been partially funded by the Spanish Ministry of Economy, Industry and Competitiveness, by the European Regional Development Fund (ERDF) Programme and by the Fundación Séneca through grants TIN2014-53749-C2-1-R, TIN2014-53749-C2-2-R and 19371/PI/14.

References

[1] Institute of Medicine. *Clinical Practice Guidelines We Can Trust*. The National Academies Press, Washington, DC, 2011.

[2] A. Latoszek-Berendsen, H. Tange, H.J. van den Herik, and A. Hasman. From clinical practice guidelines to computer-interpretable guidelines a literature overview. *Methods Inf Med*, 49(6):550–570, 2010.

[3] M. Peleg, S. Tu, J. Bury, P. Ciccarese, J. Fox, R. A Greenes, R. Hall, P. D Johnson, N. Jones, A. Kumar, et al. Comparing computer-interpretable guideline models: a case-study approach. *J Am Med Inform Assoc*, 10(1):52–68, 2003.

[4] M. Marcos, J.A. Maldonado, B. Martínez-Salvador, D. Moner, D. Boscá, and M. Robles. An archetype-based solution for the interoperability of computerised guidelines and electronic health records. In *In Proc. of the 13th Conference on Artificial Intelligence in Medicine, AIME 2011.*, pages 276–285, 2011.

[5] M. Marcos, J.A. Maldonado, B. Martínez-Salvador, D. Boscá, and M. Robles. Interoperability of clinical decision-support systems and electronic health records using archetypes: A case study in clinical trial eligibility. *J Biomed Inform*, 46(4):676–689, 2013.

[6] D. R. Sutton and J. Fox. The Syntax and Semantics of the PROforma Guideline Modeling Language. *J Am Med Inform Assoc*, 10(5):433–443, oct 2003.

[7] M. Quesada-Martínez, J. T. Fernández-Breis, R. Stevens, and N. Aussenac-Gilles. Ontoenrich: A platform for the lexical analysis of ontologies. In *Proc. of the International Conference on Knowledge Engineering and Knowledge Management*, pages 172–176. Springer, 2014.

[8] SNOMED International. SNOMED CT. `https://www.snomed.org/snomed-ct`, 2018 (acc. March 8, 2017).

[9] SNOMED International. SNOMED International SNOMED CT Browser. `http://browser.ihtsdotools.org/`, 2018 (acc. March 8, 2017).

[10] OpenClinical CIC. OpenClinical.net. `https://www.openclinical.net/`, 2015 (acc. March 8, 2017).

[11] N. Guarino. *Formal Ontology in Information Systems: Proc. of the 1st International Conference June 6-8, 1998, Trento, Italy*. IOS Press, Amsterdam, The Netherlands, 1st edition, 1998.

[12] B. M. Konopka. Biomedical ontologies - a review. *Biocybern Biomed Eng*, 35(2):75 – 86, 2015.

[13] P. L. Whetzel, N. F. Noy, N. H. Shah, P. R. Alexander, C. Nyulas, T. Tudorache, and M. A. Musen. Bioportal: enhanced functionality via new web services from the national center for biomedical ontology to access and use ontologies in software applications. *Nucleic Acids Res*, 39(suppl_2):W541–W545, 2011.

[14] A. Ghazvinian, N. F. Noy, and M. A. Musen. Creating mappings for ontologies in biomedicine: simple methods work. In *AMIA Annu Symp Proc*, volume 2009, page 198. American Medical Informatics Association, 2009.

[15] Stanford CoreNLP. a suite of core NLP tools. *URL http://nlp. stanford. edu/software/corenlp. shtml (Last accessed: 2013-09-06)*, 2016.

[16] University of Manchester. Modularity: How do locality-based modules work? `http://owl.cs.manchester.ac.uk/research/modularity/`, 2017 (acc. March 8, 2017).

[17] E. Lozano, M. Marcos, B. Martínez-Salvador, A. Alonso, and J.R. Alonso. Experiences in the development of electronic care plans for the management of comorbidities. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 5943 LNAI:113–123, 2010.

[18] S. Gindl, K. Kaiser, and S. Miksch. Syntactical Negation Detection in Clinical Practice Guidelines. In S.K. Andersen, G.O. Klein, S. Schulz, J. Aarts, and M.C. Mazzoleni, editors, *eHealth Beyond the Horizon - Get IT There*, pages 187–192, Studies in Health Technology and Informatics, Volume 136, 2008. IOS Press.

[19] G. Leonardi, A. Bottrighi, G. Galliani, P. Terenziani, A. Messina, and F. D. Corte. Exceptions Handling within GLARE Clinical Guideline Framework. *AMIA Annu Symp Proc*, 2012:512–521, Nov 2012.

[20] A. Goldstein, E. Shalom, D. Klimov, Y. Shahar, A. Fux, M. Peleg, P. Sofer, S. Quaglini, S. Panzarasa, L. Sacchi, and N. Larburu. Deliverable D3.1: Extended CIG format. Technical report, MobiGuide project, 2013 (accessed March 8, 2018). `http://www.mobiguide-project.eu/en/downloads/category/10-public-deliverables?download=68:d3-1-extended-cig-format`.