Experiments to create ontology-based disease models for diabetic retinopathy from different biomedical resources

Arguello Casteleiro, M, Martínez-Costa, C, Des-Diz, J, Fernandez-Prieto, MJ, Wroe, C, Maseda-Fernandez, D, Demetriou, G, Nenadic, G, Keane, J, Schulz, S and Stevens, R

Experiments to create ontology-based disease models for diabetic retinopathy from different biomedical resources

M. Arguello Casteleiro[1], C. Martínez-Costa[2], J. Des-Diz[3], M.J. Fernandez-Prieto[4], C. Wroe[5], D. Maseda-Fernandez[6], G. Demetriou[1], G. Nenadic[1], J. Keane[1], Stefan Schulz[2], and R. Stevens[1]

[1] School of Computer Science, University of Manchester, UK
[2] Institute of Medical Informatics Statistics and Documentation, Medical University of Graz
[3] Hospital do Salnés, Villagarcía de Arousa, Spain
[4] Salford Languages, University of Salford, UK
[5]The BMJ, UK
[6] Midcheshire Hospital Foundation Trust, NHS England, UK

Abstract. According to the World Health Organisation diabetic retinopathy (DR) is a high priority eye disease. This paper investigates a method for creating disease models for DR using the ontologies BioTopLite2 and SNOMED CT and different biomedical resources: 1) consultation notes from anonymised electronic health records; 2) the clinical practice guideline for DR by the American Academy of Ophthalmology; 3) the BMJ Best Practice for DR; and 4) neural language models from Deep Learning (CBOW and Skip-gram) using a 14M PubMed dataset. As SNOMED CT does not contain disease models, the novelty of this study is twofold: a) evaluation of the utility of CBOW and Skip-gram for obtaining DR disease models from the biomedical literature; and b) the proposed method for building ontology-based disease models exploiting SNOMED CT reference sets. In our method, we first propose a representation of SNOMED CT reference sets for DR in OWL by extracting upper modules from SNOMED CT. Secondly, we use content ontology design patterns with BioTopLite2 and SNOMED CT that act as templates to semantically represent clinical content in OWL. We report on the effectiveness of the method.

Keywords. Diabetic Retinopathy, SNOMED CT, ontologies, Deep Learning

1 Introduction

According to the World Health Organization (WHO), diabetic retinopathy (DR) is a high priority eye disease. It can lead to permanent eye damage and its incidence worldwide is expected to increase along with the incidence of diabetes [1]. Examples of evidence-based resources that provide guidance for the diagnosis and management of DR are the Clinical Practice Guidelines (CPG) developed by the American Academy of Ophthalmology (AAO) [2] or the BMJ's Best Practice that covers DR [3].

Brassil et al. [4] acknowledge that "*clinicians often encounter questions in their work setting related to a challenging diagnosis, treatment decisions, or unexpected*

complications, revealing a perceived knowledge gap". This gap provides an opportunity for the use of Clinical Decision Support Systems (CDSS), e.g. evidence-based decision support. There is currently, however, a lack of integration of CDSS into clinical work and the Electronic Health Records (EHRs) [5].

Formal ontologies provide standardized, logic-based descriptions for classes of domain entities in a computationally amenable form. This allows the meaning of entities referred to by data items and characterized by language-specific domain terms to be both precisely known and manipulated, improving consistency. This constitutes a basic requirement for semantic interoperability, i.e. the meaning-preserving communication of information across boundaries of systems, institutions and jurisdictions. Hammond et al. [6] puts forward the notion of semantic interoperability as a "*universal ontology that covers all aspects of health, health care, clinical research, management, and evaluation"*. Hence, semantic interoperability can aid the integration of CDSS and the EHRs as well as the secondary use of clinical data within EHRs.

This study addresses the topic of semantic interoperability by combining the biomedical top-level ontology BioTopLite2 [7] with the ontology underlying the clinical terminology SNOMED CT [8] to formally represent: a) clinical content from EHRs (e.g. anonymised consultation notes for DR), and b) evidence-based clinical content that can be the foundation for building CDSS (e.g. [2] or [3]).

Disease models are formal or semi-formal descriptions that help understand how a disease develops and which treatment approaches can be considered. SNOMED CT is the leading clinical health terminology for use in EHRs and it can be formally represented in the Web Ontology Language (OWL) [9]. However, SNOMED CT does not contain disease models *per se*. In order to develop ontology-based disease models for DR based on the SNOMED CT January 2017 release, two questions need to be addressed: 1) how to acquire a set of SNOMED CT concepts relevant for DR when this SNOMED CT release contained 325143 OWL Classes?; and 2) how to formally represent a disease model for DR?

We propose a method to obtain ontology-based disease models consisting of a two-step process. To validate the proposal, this paper investigates to what extent we can create ontology-based disease models for DR based on BioTopLite2 and SNOMED CT using four different biomedical resources: 1) a small number of anonymised consultation notes; 2) the CPG for DR by the AAO [2]; 3) the BMJ Best Practice for DR [3]; and 4) the neural language models CBOW and Skip-gram of Mikolov et al. [10] produced from Deep Learning using a large corpus of scientific literature (i.e. PubMed [11]).

2 SNOMED CT simple Refsets in OWL with the OWL API

The first step of our method identifies suitable biomedical resources, such as EHRs or evidence bases from where to acquire a set of terms relevant for DR. This step makes use of three design features of SNOMED CT [8]:

1. *SNOMED CT components (i.e. concepts, descriptions and relationships)*; in brief, SNOMED CT descriptions are clinical terms in a given language, di-

vided into (unique) FSNs (Fully Specified Names), preferred terms and synonyms. Each description refers to a SNOMED CT Concept, with Concepts being related by one or more binary relationships to other concepts. Running a Perl script, an OWL ontology can be created for a SNOMED CT release.

2. *SNOMED CT expressions*: either single (pre-coordinated) SNOMED CT concepts or compositional (post-coordinated) expressions constrained by the SNOMED CT Compositional Grammar.
3. *The extensibility mechanism (Refsets)*: A reference set (Refset for short) is a subset of SNOMED CT components. The most basic Refset is the *Simple Refset*, which can fully enumerate a subset of SNOMED CT components.

In this first step, a set of relevant terms for DR from a biomedical resource is mapped to SNOMED CT obtaining a set of pre-coordinated and post-coordinated expressions (i.e. one or more focus concepts). Next, a set of SNOMED CT concept identifiers is created; this is the basis to define a signature of the SNOMED CT ontology that allows the extraction of an ontological module.

We propose to represent a SNOMED CT Simple Refset in OWL as a signature of the SNOMED CT ontology, i.e. a set of OWL Classes. The main benefit of our proposal is avoiding exhaustive enumeration (e.g. all subtypes/descendants of a SNOMED CT concept) and relying on the OWL API [12] to create upper modules (i.e. ModuleType BOT) that guarantee the inclusion of all axioms relevant to the meaning of the OWL Classes included in the signature.

A locality-based module (upper module) contains at least all the (entailed) superclasses of an OWL class included in the signature [13]. Hence, top-level concepts of relevant hierarchies will appear in the upper module extracted [13]. This study uses the FaCT++ reasoner implementation from http://code.google.com/p/factplusplus/.

To validate this step, a set of relevant terms from each of the four biomedical resources is acquired (e.g. terms in the highlights of the BMJ Best Practice for DR). Next, a SNOMED CT signature (one per biomedical resource) is created taking mostly the OWL Classes for the focus concepts of the relevant terms identified. Finally, SNOMED CT Simple Refsets for DR in OWL are created using each signature and applying the OWL API ModuleType BOT; this step is considered successful if smaller subsets of SNOMED CT components are obtained.

3 Content Ontology Design Patterns (Content ODPs)

The second step of our method relies on Ontology Design Patterns (ODPs) that can "*encapsulate in a single named representation the semantics that require several statements in low level ontology languages*" [14]. There are different types of ODPs [15]. This study focuses on Content ODPs, and more specifically, on domain-related ontology patterns [15] to formally represent disease models. This study adopts the ontology framework proposed within the EU SemanticHealthNet project [16] to create Content ODPs. An advantage of the Content ODPs proposed by SemanticHealthNet is that they are independent of any particular EHR specification, such as the openEHR [17] or the HL7 Clinical Document Architecture Release 2 (CDA R2) [18].

The incorporation of concepts from clinical coding systems such as SNOMED CT into the clinical model (e.g. HL7 CDA R2) is known as terminology binding [6] and is far from trivial. However, terminology binding becomes easier when the information represented by clinical models is expressed by Content ODPs. Both evidence-based resources (e.g. CPGs) and EHRs contain clinical statements. Hence, evidence-based decisions and clinical information in EHRs share clinical notions with typically complex semantics that can be formally represented as domain-specific ODPs.

The Content ODPs of this study are based on BioTopLite2 [7] and SNOMED CT [8] in OWL. BioTopLite2 is a biomedical top-level ontology, which introduces basic categories and relations that seek to improve semantic interoperability by clarifying and disambiguating the meaning of domain ontology content and, in our case, the clinical data that they describe. SNOMED CT has close-to-user views of expressions (abbreviated here as common patterns) that represent single diagnostic statements [8].

Table 1. Exemplifying Content ODPs in OWL for SNOMED CT common patterns. Note that the "*clinical life phase*" is a re-interpretation of the SNOMED CT "*clinical finding*"

HL7 CDA R2 section	**SNOMED CT common patterns**	**Content ODPs in OWL Manchester Syntax using BioTopLite2 and a clinical model**
Physical Examination OR Assessment	Clinical finding present	*cm:DiagnosticStatement* and (**btl2:represents** only *cm:ClinicalLifePhase*)
	Clinical finding absent	*cm:DiagnosticStatement* and (not (**btl2:represents** only *cm:ClinicalLifePhase*))
Past Medical History	History of	*cm:DiagnosticStatement* and (**btl2:isPartOf** some *cm:PastHistoryInformation*) and (**btl2:represents** only *cm:ClinicalLifePhase*)
Plan	Procedure not done	*btl2:Plan* and (**btl2:hasRealization** only *cm:ClinicalProcess*)

Table 1 contains examples of Content ODPs in the Manchester OWL Syntax [19] for SNOMED CT common patterns from [20]. As can be seen in Table 1, the Content ODPs can refer to the same clinical model (e.g. HL7 CDA R2) with more than one clinical coding system. All the Content ODPs in OWL within Table 1 refer to diagnostic statements represented as <Observation> HL7 CDA entries within a HL7 CDA document section, where the first column in the table indicates the section for HL7 CDA R2. The same Content ODP can appear in more than one section of a HL7 CDA R2 consultation note (e.g. Physical Examination section or Assessment section) while referring to the same common pattern. In other words, the same Content ODP can be applied to represent diagnostic statements that refer to different processes of care to which an evidence-based resources (e.g. CPGs) will refer.

In Table 1 for the Content ODPs, the prefix *btl2* indicates OWL constructs (Classes in italics and Object Properties in bold) from BioTopLite2 and the prefix *cm* indicates OWL constructs from the clinical model (e.g. HL7 CDA R2 or openEHR). The OWL Class *cm:DiagnosticStatement* is a sub-class of *btl2:InformationObject*.

For the eleven SNOMED CT common patterns listed in [20], it is easy to identify the SNOMED CT hierarchy from which the focus concept will belong. For example, the common pattern from Table 1 *Procedure not done* will have focus concepts from the *Procedure* hierarchy.

To validate the suitability of this step, we obtain the diagnostic statements that underpin the disease models created for DR from four biomedical resources. This step is considered successful if Content ODPs are obtained for each biomedical resource.

4 Results from experiments with different biomedical resources

Each of the following subsections contain: a) examples of the terms from a biomedical resource and its correspondence to SNOMED CT expressions; b) basic statistics of the ontological signature and the locality-based module created; and c) examples of Content ODPs in the Manchester OWL Syntax. It should be noted that a) and b) are obtained in the first step of our method while c) are gained in the second step.

4.1 The Clinical Practice Guideline for DR by the AAO

We start by obtaining terms that appear in the CPG for DR by the AAO [2] within: two tables that gather data collected from clinical trials and epidemiological studies of DR; and one table that summarise the evidence-based management recommendations for patients with diabetes. A total of 18 relevant terms were acquired from [2] corresponding to DR severity level, DR findings, and DR management recommendations.

Table 2. Examples of terms from the CPG for DR by the AAO [2]

Terms from the CPG for DR	Focus concept of SNOMED CT expression	HL7 CDA R2 section
No apparent retinopathy	390834004 (Clinical finding)	Assessment
Mild NPDR	312903003 (Clinical finding)	Assessment
PDR	59276001 (Clinical finding)	Assessment
Microaneurisms	34037000 (Clinical finding)	Physical examination
Intraretinal hemorrhages	28998008 (Clinical finding)	Physical examination
Venous beading	247118003 (Clinical finding)	Physical examination
Panretinal Photocoagulation Laser	312713003 (Procedure)	Plan

Following the first step of the method, the 18 terms from [2] are represented as 15 pre-coordinated and 3 post-coordinated SNOMED CT expressions. The signature has 18 focus concepts as OWL Classes, where 15 are SNOMED CT concepts from the *Clinical finding* hierarchy and 3 from the *Procedure* hierarchy. The locality-based module created has 9.0K OWL Classes and 42.5K axioms.

Following the second step of the method, the SNOMED CT expressions from [2] are transformed into Content ODPs. Table 2 illustrates some terms acquired from [2]

where a row with a white background indicates a pre-coordinated expression, while a row with a grey background indicates a post-coordinated expression. The second column shows the SNOMED CT concept identifier for the focus concept of a SNOMED CT expression and, in brackets, the SNOMED CT hierarchy to which the focus concept belongs. Hence, the second column provides the key information to relate a term from [2] to one of the eleven SNOMED CT common patterns listed in [20]. The last column indicates the HL7 CDA R2 sections where the diagnostic statement may appear.

Combining the information from Table 1 and 2, we can easily create Content ODPs by instantiating the formal concept definitions in OWL from Table 1. Hence, the Content ODP for *Microaneurisms* (Table 2) when absent:

Individual: cs:AbsenceOfMicroaneurismas

Types: *cm:ClinicalFindingAbsent* and (not (**btl2:represents** only *sct:34037000*))

where the prefix *sct* indicates OWL constructs from SNOMED CT in OWL; and the prefix *cs* indicates OWL constructs for diagnostic statements in OWL.

4.2 The BMJ Best Practice DR

We start by acquiring terms that appear in the BMJ Best Practice that covers DR [3]. A total of 29 relevant terms were obtained from [3] corresponding to DR history and examination, DR diagnostic investigations, and DR severity level and management.

Following the first step of the method, the 29 terms from [3] are represented as 24 pre-coordinated and 5 post-coordinated SNOMED CT expressions. The signature has 29 focus concepts as OWL Classes, where 21 are SNOMED CT concepts from the *Clinical finding* hierarchy and 8 from the *Procedure* hierarchy. The locality-based module created has 16.1K OWL Classes and 75.1K axioms.

Following the second step of the method, the SNOMED CT expressions from [3] are transformed into Content ODPs. Table 3 has the same format of Table 2. As shown in Table 3, the BMJ Best Practice for DR [3] indicates if a term for DR history and examination is common or uncommon. We represent formally a common diagnostic statement in OWL as:

Class: *cm:CommonDiagnosticStatement*

EquivalentTo:

cm:DiagnosticStatement and (**btl2:hasPart** some *cm:CommonInformation*)

where *cm:CommonInformation* is a sub-class of *btl2:InformationObject*. We can likewise create the OWL Class *cm:UncommonDiagnosticStatement*.

In Table 3, "*macular thickening*" does not have a SNOMED CT pre-coordinated expression (row with a grey background) and can be represented as the post-coordinated expression 312999006|Disorder of macula of retina|: 42752001|Due to|=89977008|Increased thickness|. In this study, SNOMED CT postcoordinated expressions in OWL have the prefix *sctpost*. The Content ODP for "*macular thickening*" in Table 3 when present (instantiation of the Content ODP from Table 1) is:

Individual: cs:PresenceOfMacularThickening

Types: *cm:ClinicalFindingPresent* and (**btl2:hasPart** some *cm:CommonInformation*) and (**btl2:represents** only *sctpost:312999006_42752001_89977008*)

Table 3. Examples of terms from the BMJ Best Practice for DR [3]

Terms from the BMJ Best Practice DR	Focus concept of SNOMED CT expression	HL7 CDA R2 section
diabetes (common)	73211009 (Clinical finding)	Past Medical History
severe NPDR	312905005 (Clinical finding)	Assessment
microaneurysms (common)	34037000 (Clinical finding)	Physical examination
lipid exudates (common)	247131008 (Clinical finding)	Physical examination
macular thickening (common)	312999006 (Clinical finding)	Physical examination
venous beading (uncommon)	247118003 (Clinical finding)	Physical examination
Vitrectomy	75732000 (Procedure)	Plan

4.3 Anonymised HL7 CDA R2 Consultation Notes for DR

We start by retrieving SNOMED CT coded entries (i.e. terms) that appear in 5 anonymised HL7 CDA R2 consultation notes for DR. A total of 80 <Observation> HL7 CDA entries were obtained from 3 HL7 CDA document sections (i.e. Past Medical history; Physical examination; and Assessment and Plan) of the 5 consultation notes.

Following the first step of the method, the 80 <Observation> HL7 CDA entries refer to 28 unique SNOMED CT focus concepts (no repetitions) that can be represented as 27 pre-coordinated and 1 post-coordinated SNOMED CT expressions. The signature has 28 focus concepts as OWL Classes, where 22 are SNOMED CT concepts from the *Clinical finding* hierarchy and 6 from the *Procedure* hierarchy. The locality-based module created has 1.6K OWL Classes and 7.8K axioms.

Table 4. Examples of terms (coded entries) from HL7 CDA R2 consultation notes for DR

Terms from HL7 CDA R2 consultation notes for DR	Focus concept of SNOMED CT expression	HL7 CDA R2 section
Moderate NPDR	312904009 (Clinical finding)	Assessment
Incipient cataract	52421005 (Clinical finding)	Physical examination OR Assessment
Retinal microaneurysm	34037000 (Clinical finding)	Physical examination
Retinal haemorrhage	28998008 (Clinical finding)	Physical examination
Diabetic retinal venous beading	399866003 (Clinical finding)	Physical examination
Posterior segment fluorescein angiography	252822006 (Procedure)	Plan

Following the second step of the method, the SNOMED CT expressions from the HL7 CDA R2 consultation notes are transformed into Content ODPs. Table 4 has the format of Table 2 and illustrates some of the coded entries retrieved. "*Incipient cataract*" appears in Table 4 and it is found in two consultation notes. Cataract does not appear in the main summary tables of the CPG for DR by AAO [2] or in the highlights of the BMJ Best Practice for DR [3]. However, [2] acknowledges cataract as a side effect/complication of vitrectomy as well as intravitreal injections (i.e. DR treat-

ment) and [3] acknowledges cataract as a complication of patients with diabetes. Hence, EHRs can provide clinical findings that are pertinent for DR while not being within the main summaries or highlights of evidence-based resources.

Combining the information from Table 1 and 4, we can create Content ODPs. The coded entry (i.e. the term) "*incipient cataract*" that appears in Table 4 when present has the following Content ODP in the Manchester OWL Syntax:

Individual: cs:PresenceOfIncipientCataract

Types: *cm:ClinicalFindingPresent* and (**btl2:represents** only *sct:52421005*)

4.4 CBOW and Skip-gram word embeddings from 14M PubMed dataset

PubMed contains important scientific discoveries and findings that can aid healthcare professionals to provide better care. However, the large volume and rapid growth of PubMed makes it difficult to acquire a list of relevant terms for DR directly from PubMed. In this study, we use CBOW and Skip-gram from Deep Learning to get a list of relevant terms for DR using a large-scale dataset of PubMed publications.

The neural language models CBOW and Skip-gram make it feasible to obtain word embeddings (i.e. distributed word representations) from corpora of billions of words. Using similarity measures (i.e. cosine value) we can build up a list with the n top-ranked candidate terms for a target term (e.g. diabetic retinopathy). The experimental set-up (e.g. hyperparameter configuration) to create word embeddings from CBOW and Skip-gram using a 14M PubMed dataset is the same as the one described in [21].

In this study we use three target terms that appear in Table 5. According to the BMJ Best Practice DR, both *microaneurysms* and *lipid exudates* are common findings for DR. For each target term, we limit the list to the fifty candidate terms with the highest cosine value (i.e. the top fifty ranked). A total of 300 candidate terms was obtained. After removing duplicates, the list contains 124 unique candidate terms. Not all the candidate terms are relevant for DR (i.e. false positives). Using the BMJ Best Practice for DR, a medical consultant determined that 113 terms were true positives (tp) and 11 were false positives (fp), which gives 91% of overall precision. The 113 candidate terms considered relevant (tp) relate to the DR history and examination, DR diagnostic investigations, and DR severity level and management.

Table 5. Precision calculated as tp/(tp+fp) per neural language model and target term

	Target terms		
Model	**diabetic_retinopathy**	**diabetic_retinopathy microaneurysms**	**diabetic_retinopathy exudates**
CBOW	86%	86%	88%
Skip-gram	100%	100%	98%

Table 5 shows the precision for each model and target term. Skip-gram outperforms CBOW. The target term in the last column improves CBOW precision.

Following the first step of the method, the 113 tp terms are represented as 35 pre-coordinated (7 of these can act as focus refinement and are not focus concepts) and 5

post-coordinated SNOMED CT expressions. The signature has 34 concepts as OWL Classes: 28 from the *Clinical finding* hierarchy; 4 from the *Procedure* hierarchy; 1 from the *Observable entity* hierarchy; and 1 descendant of the *Morphologically altered structure* OWL Class. Hence, considering the clinicians' views, *neovascularization (morphologic abnormality)* was added to the signature despite of not being a focus concept. The upper module has 9.6K OWL Classes and 45.6K axioms.

Following the second step of the method, the SNOMED CT expressions for the tp terms are transformed into Content ODPs. Some tp terms correspond to SNOMED CT expressions, which have the focus concepts in Table 3 – with the exception of the focus concept for "*venous beading*" (not in the candidate terms list). The Content ODP for planned "*vitrectomy*":

Individual: cs:PlanOfVitrectomy
Types: *btl2:Plan* and (**btl2:hasRealization** only *sct:75732000*)

5 Discussion and Conclusion

The SNOMED CT Refset mechanism is a method for filtering and arranging SNOMED CT concepts for specific domains or use cases [22]. SNOMED CT concepts belonging to a Refset are portable and can be reused among organisations with similar needs. This study has investigated the creation of Simple Refsets for DR in OWL by creating upper modules from ontological signatures using four biomedical resources. The biggest upper module created with the OWL API ModuleType BOT from the SNOMED CT ontology is the one for the BMJ Best Practice DR, which has 5% of the OWL Classes in the SNOMED CT ontology. The smallest upper module is the one from the anonymised HL7 CDA R2 consultation notes, which has 0.5% of the OWL Classes in the SNOMED CT ontology. There are two potential reasons: 1) the very limited number of consultation notes; and 2) the coded entries refer to more specific SNOMED CT concepts. For example "*Diabetic retinal venous beading*" (Table 4) is more specific than "*venous beading*" (Table 2 and 3). Overall, module extraction seems to be an effective means to obtain small subsets of SNOMED CT components.

Although the Content ODPs illustrated in this paper only refer to the OWL Classes included in the ontological signatures created for the SNOMED CT ontology, there is no impediment to roll out the approach presented to the OWL Classes from the upper modules obtained (e.g. descendants of an OWL Class included in the signature). What is essential for the second step of our method to work is complying with the underlying mapping from Table 1 that implies considering the SNOMED CT common patterns from [20] based on the SNOMED CT expression at hand.

Arguably, the best resources to obtain disease models for DR are the CPG for DR by the AAO [2] and the BMJ Best Practice for DR [3]. However, both the CPG for DR by AAO and the BMJ Best Practice for DR need periodic updates from the ever-growing scientific biomedical literature. In 2017, PubMed contains references to more than 27M scientific publications [11], with an average of two papers added per minute in 2016 [23]. This paper has investigated to what extent CBOW and Skip-gram from Deep Learning can derive key clinical content that appears in evidence-based re-

sources like the BMJ Best Practice for DR. The higher precision for Skip-gram (98% to 100%) with the three target terms for DR indicates the potential of the neural language models from Deep Learning to aid periodic updates of evidence-based resources like the BMJ Best Practice.

References

1. WHO priority eye diseases. http://www.who.int/blindness/causes/priority/en/index5.html
2. CPG DR. https://www.aao.org/preferred-practice-pattern/diabetic-retinopathy-ppp-updated-2016
3. BMJ Best Practice DR. http://bestpractice.bmj.com/best-practice/monograph/532.html
4. Brassil, E., Gunn, B., Shenoy, A.M., Blanchard, R.: Unanswered clinical questions: a survey of specialists and primary care providers. JMLA, vol. 105(1), p. 4 (2017).
5. Porat, T., Delaney, B., Kostopoulou, O.: The impact of a diagnostic decision support system on the consultation: perceptions of GPs and patients. BMC Medical Informatics and Decision Making, vol. 17(1), p. 79 (2017).
6. Hammond, W.E., Jaffe, C., Cimino, J.J., Huff, S.M.: Standards in biomedical informatics. In: Biomedical informatics, pp. 211-253. Springer, London (2014).
7. Schulz S, Boeker M, Martinez-Costa C.: The BioTop Family of Upper Level Ontological Resources for Biomedicine. Stud Health Technol Inform, vol. 235, pp. 441-445 (2017).
8. SNOMED CT Starter Guide. https://confluence.ihtsdotools.org/display/DOCSTART/
9. OWL 2. https://www.w3.org/TR/owl2-overview/
10. Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., Dean, J.: Distributed representations of words and phrases and their compositionality. In: NIPS, pp. 3111-3119 (2013).
11. PubMed. https://www.ncbi.nlm.nih.gov/pubmed/
12. Horridge, M., Bechhofer, S.: The owl api: A java api for owl ontologies. Semantic Web, vol. 2(1), pp.11-21 (2011).
13. Grau, B.C., Horrocks, I., Kazakov, Y. Sattler, U.: Just the right amount: extracting modules from ontologies. In: WWW, pp. 717-726 (2007).
14. Egaña, M., Rector, A.L., Stevens, R., Antezana, E.: Applying Ontology Design Patterns in Bio-ontologies. In: EKAW, vol. 5268, pp. 7-16 (2008).
15. Falbo, R.A., Guizzardi, G., Gangemi, A. Presutti, V.: Ontology patterns: clarifying concepts and terminology. In: WOP, pp. 14-26 (2013).
16. Semantic Interoperability for Health Network (SHN). http://www.semantichealthnet.eu/
17. openEHR. http://www.openehr.org/
18. HL7 CDA R2. http://www.hl7.org/implement/standards/product_brief.cfm?product_id=7
19. Horridge, M., Drummond, N., Goodwin, J., Rector, A.L., Stevens, R. Wang, H.: The Manchester OWL syntax. In: OWLed, vol. 216 (2006).
20. Bhattacharyya, S.B.: SNOMED CT Expressions. In Introduction to SNOMED CT, pp. 95-129. Springer, Singapore (2016).
21. Arguello Casteleiro, M., Demetriou, G., Read, W.J., Prieto, M.J.F., Maseda-Fernandez, D., Nenadic, G., Klein, J., Keane, J.A., Stevens, R.: Deep Learning meets Semantic Web: A feasibility study with the Cardiovascular Disease Ontology and PubMed citations. In: ODLS, pp. 1-6 (2016).
22. Lee, D.H., Lau, F.Y., Quan, H.: A method for encoding clinical datasets with SNOMED CT. BMC Medical Informatics and Decision Making, vol. 10 (2010).
23. PubMed statistics. https://www.nlm.nih.gov/bsd/index_stats_comp.html