A Guiding Framework for Ontology Reuse in the Biomedical Domain

Tejal Shah1,2 Fethi Rabhi1 Pradeep Ray1 Kerry Taylor2
CSE, UNSW CSE, UNSW ApuHC, UNSW CCI, CSIRO
tejals@cse.unsw.edu.au f.rabhi@unsw.edu.au p.ray@unsw.edu.au kerry.taylor@csiro.au

Abstract

Though recent research has established the interdependencies between several medical and oral health conditions, e-Health systems for medical and oral domains have been designed and implemented to operate independently. Such disparate systems coupled with different data capture and storage formats have led to the formation of medical-dental silos. In this paper, we advocate development of a knowledge base of formal ontology and rules that can help in reducing such silos. For this purpose, we argue that a cross-domain ontology of medical and oral health concepts should be developed by reusing SNOMED-CT. In order to prevent the replication of some of the inherent problems of SNOMED-CT in the resultant ontology, we have further developed a novel guiding framework for steering through the development process. We anticipate that it will function as a generic framework to guide the development of any biomedical ontology that needs to reuse an ontology such as SNOMED-CT.

1. Introduction

Electronic Health (e-Health) technology has brought about significant transformation by replacing the traditional paper-based clinical practice to Information and Communication Technologies (ICT)-based systems for automatic management (storage, processing, and archiving) of health information. This has considerably improved the work efficiency of the healthcare practitioners and has led to the highly developed e-Health care service sector worldwide. E-Health is a new term where ICT systems are deployed for better management and coordination of information. Marconi [1] quotes: ‘e-health is an application of the Internet and other related technology in the healthcare sector to improve the access, efficiency, effectiveness and quality of clinical and business processes utilized by health care organizations, practitioners, patients, and consumers in an effort to improve the health status of patients’. Accordingly, there are medical and dental ICT systems available for storing, managing and accessing patient records. These systems however, do not possess semantic interoperability for seamless sharing of patient information leading to a separation of the patient care management process in the medical and dental domains, which is an unfavorable outcome. In this paper, we propose an approach that is aimed at enabling patient information to be shared seamlessly across the medical and dental ICT systems.

1.1. Motivation

The call for collaborative patient management by medical and dental practitioners is gaining momentum due to increasing research evidence establishing the interdependencies between medical conditions with oral health conditions. In fact, there are more than 120 medical conditions that first appear in the oral cavity [2] thereby putting dental practitioners at the best position for early identification of underlying debilitating and chronic medical conditions. For example, Gardner syndrome is a fatal systemic condition in many instances; however, if diagnosed early, the prognosis can improve. Oral conditions such as osteomas of jaw bones and impacted teeth are early indicators of the syndrome [3]. Since oral health is a specialty within the dental domain, the sharing of patient information across medical and dental systems is essential. However, the disparate data collection and storage formats of the medical and dental domains coupled with the absence of machine-interpretable ontologies and knowledge bases are major impediments to sharing information and inferencing over it to generate usable knowledge. In fact, healthcare in general, remains as one area where information exchange has been severely limited. As stated in [4], “aside from a few implementations in closed care delivery systems...there are no systems that demonstrate the free interchange of patient information among a multitude of healthcare providers...at multiple locations”. Therefore, one way to address the chasm between the medical and oral health domains is with a cross-domain ontology of medical and oral health terms, which is extended with domain knowledge rules to form a scientific knowledge base that can be reused in expert systems to aid in the decision making process of both medical and dental
practitioners. However, no such ontology covering the medical and oral concepts exists.

There are several formal, comprehensive ontologies available in the medical domain. The Systematized Nomenclature of Medicine-Clinical Terms (SNOMED-CT) [5] is a comprehensive ontology for representing clinical medicine conditions. We therefore argue that a cross-domain ontology across the medical and dental domains should be developed by reusing SNOMED-CT. However, reusing an existing ontology has the disadvantage of duplicating the errors of the parent ontology in the resulting ontology. But unlike the availability of formal methodologies to guide the development of an ontology from scratch, there are very few guiding principles available for reusing an ontology and addressing the problems that are likely to arise during the process. This despite the fact that reuse is a widely preferred approach, where possible, as against building an entirely new and independent ontology. Therefore, motivated by the need to develop a cross-domain ontology via reuse, we propose a guiding framework for the reuse and development process. We anticipate that this framework would serve as an important starting point to formalize ontology reuse approaches in the biomedical domain.

1.2. Contributions

The main contributions of this paper are: (i) analyzing the structure and modeling principles of SNOMED-CT for errors and problems with respect to its use in developing a new ontology; (ii) a guiding framework to steer the development of an ontology based on reuse of an existing ontology such as SNOMED-CT; and (iii) an ontology-based approach to connect the medical and dental domains and enable sharing of and reasoning over patient information. The proposed steps within the framework help to highlight the problems and address them in the early stages of the ontology development process thereby saving considerable time and effort.

The rest of the paper is organized as follows. In section 2, we describe ontologies in general and SNOMED-CT in particular and discuss SNOMED-CT’s problems. In section 3, we propose a guiding framework for the ontology development process, which is followed by section 4 on related work. We conclude the paper in section 5.

2. A general introduction to ontologies

Studer et al. [6] provide a complete yet succinct definition for an ontology by combining several well-known ontology definitions from literature: “a formal, explicit specification of a shared conceptualization”. The core structure of such an ontology is in the form of a triplet, concept1 \rightarrow relationship \rightarrow concept2 where concepts are representation of some real-world phenomena. Further, the concepts and relationships are formally defined to avoid ambiguity and to retain meaning when the information is shared across different applications. This retention of meaning is of paramount importance in healthcare where the language used by the healthcare staff varies not only between domains but also between different geographical regions. Moreover, since formal ontologies are based in description logics, there are reasoners available to classify the ontology as well as generate new knowledge by inferencing over the information available. Similarly, it is these relationships and definitions that can help in the mapping of two or more terminologies or ontologies thereby helping in the process of semantic interoperability and data integration even between and among systems that do not use the same ontology. In brief, the uses of ontologies in general and healthcare in particular, can be summed up as follows [7]: (i) data analysis and knowledge discovery, (ii) education and training, (iii) evidence based practice, (iv) information exchange and data integration, and (v) clinical decision support. There are several examples in literature wherein ontologies have been employed for information sharing and in designing interoperable health systems [8-14].

2.1. SNOMED-CT

SNOMED-CT (hereafter referred as SCT in this paper) is the most widely used healthcare terminology in the world and has the most comprehensive coverage of concepts for representing clinical information [15, 16]. In fact, it contains more than 297,000 active concepts, more than 777,000 English language descriptions and more than 890,000 logical relationships in its January 2013 release [17]. Due to its comprehensive coverage, global usage and the fact that there is no other similarly comprehensive ontology for clinical medicine, SCT is used as an ontology rather than a terminology for all practical purposes. Moreover, like other formal ontologies, it is based in Description Logic (DL), specifically EL+, which corresponds to the to the EL profile of the Web Ontology Language (OWL 2) [18] and hence it can reasoned over. However, EL+ is not a very expressive DL but is the most suitable for large ontologies and is computable in polynomial time [19].

Reusing ontologies is the recommended approach for building a new ontology by most ontologists [20]. This prevents random proliferation of ontologies and at
the same time, saves considerable time and effort. Moreover, reuse enables interconnections, mapping and easier alignments across different ontologies thereby encouraging semantic interoperability. Although SCT is widely used, comprehensive in terms of content coverage and based in DL, it is still mired with problems, which makes reasoning over it for intelligent applications such as decision support, error-prone [21]. Therefore, naively reusing SCT will propagate the problems further and in most instances, would multiply them as well. Hence, having a good understanding of the issues is essential so that they can be addressed early on in the development stages [20]. Accordingly, we next discuss some of the issues that are relevant for building a cross-domain ontology by reusing SCT concepts, relationships and structure as appropriate. This is by no means an exhaustive list of SCT problems but it is beyond the scope of this paper to discuss all the problems.

2.1.1. Size. The mere size of SCT with all the concepts and relationships make it impractical to use for many applications since not all applications require to use all the available content of SCT. In fact, we faced problems while classifying the OWL [22] version of SCT in Protégé 4.1 [23]. Our system specifications included a processor of 2.66 GHz, RAM 4 GB, and OS 64 bit. It took less than a minute for SCT to load into protégé; however classifying was a different matter altogether. With a heap size of 1624288000 bit (194 MB) allocated to protégé, the reasoners Pellet [24] and Fact++ [25] both ran out of memory while Hermit 1.3.6 [26] kept running for greater than 2.5 hours and subsequently crashed. However, SnoRocket [27], which is a specifically developed reasoner to classify the OWL version of SCT, classified it in 1 minute. SnoRocket although is a fast classifier, unlike Pellet, it does not support individuals, instance data and rules. Hence, it does not satisfy our requirements for the present purpose. Subsequently, we increased the heap size to 2 GB after which Pellet could classify SCT in 17 minutes; however, the overall performance of protégé considerably slowed down after this. Others too have reported about this issue [28]. Thus, using SCT in its original form on local machines is quite problematic and largely impractical for most purposes.

2.1.2. Problems in the hierarchy. SCT is poly-hierarchical in that some concepts have multiple inheritance or in other words, multiple parent classes. This often leads to user ambiguity in selection of concepts for representing information. Besides, there is an over-use of the is-a relation [29] to classify concepts as subclasses without distinguishing them based on their more accurate relationships of kind-Of, part-Of, instance-Of and so on. This problem of subsumption is discussed in detail in [30].

2.1.3. Modeling errors. Although the underlying DL of SCT allows for modeling of individuals, SCT does not have any individuals [31]. In fact, several of its is-a or subclass-Of relations should actually be modeled as individuals as per the OWL guideline [32], which suggests that that a concept should be modeled as a subclass when it describes a subset of a class otherwise it should be modeled as an individual. This form of modeling effectively makes SCT only a reference terminology and makes it difficult to reason or query over instance data represented in SCT.

2.1.4. Content coverage. As mentioned previously, SCT is a very comprehensive terminology for referencing general medicine terms and to some extent specialist terms; however, it is not specific to any particular specialty domain and hence we hypothesized that there would be missing concepts and relationships in SCT with respect to both medical and oral health terms. To test this hypothesis and identify the missing concepts, we evaluated SCT using a sample set of terms and relationships. The process and results of this evaluation are discussed further down in this paper.

3. Proposed guiding framework

As mentioned previously, knowledge and understanding of issues of the ontology of choice for reuse is essential to prevent the repetition and propagation of the problems in the resulting ontology. Accordingly, we discussed some of the core issues with SCT, as relevant to our task and based on this understanding, we have developed a framework such that it breaks down the process into detailed steps so that the issues at each step can be addressed before proceeding. Figure 1 shows the framework with the core steps involved while reusing SCT to develop our cross-domain ontology. Next, we discuss the framework using examples from our on-going cross-domain ontology development work.

In figure 1, the processes that have a common output are grouped together. Accordingly, the main steps are: reuse of domain knowledge, term to concept matching, relationship to object property matching and building the knowledge base consisting of the cross-domain ontology and knowledge rules. Apart from these, implementation and evaluation is a vital step in any ontology development process; however, because it is not directly related to the reuse of SCT, it is beyond the scope of this paper to discuss the same.
3.1. Domain knowledge reuse

As mentioned previously, the backbone of an ontology consists of concepts and relationships between these concepts. Formally, the word ‘concept’ refers to ‘class’, ‘subclass’ or ‘individual’ in an OWL ontology and is a representation of some real-world phenomenon or ‘term’ as in the context of our work. Therefore, in order to build as comprehensive an ontology as possible for a particular domain, the relevant medical and oral terms must be extracted from the existing scientific literature such as from peer reviewed publications, government health education websites, and medical and dental textbooks and represented as concepts in the resulting ontology. However, manual mining is cumbersome and can lead to missed terms. Therefore, we used the UMLS interactive MetaMap [33] and SemRep [34], which help in identifying the terms as well as the relationships that exist, which might otherwise be missed. Figure 2 is a screenshot of the analysis of an example statement taken from [35]. We have omitted the details for brevity.

The first line in the upper box is the text that is being analyzed in MetaMap. The circled words show the identified entities or concepts as well as the is-a relation between endometrial carcinoma and malignant neoplasm although the latter term was not mentioned in the original text. Thus, it is easier and quicker to detect hidden relationships in this manner instead of by manually referring to SCT. A further detailed interpretation of the result is obtained in SemRep and is shown in the bottom box. Based on the search criteria, it will identify all the ontologies, terminologies, classifications and thesauri that contain the searched term(s). Thus, when a term is not matched to a concept in SCT, the other ontologies or terminologies that have a representation for the term can be easily identified. Moreover, the UMLS Semantic Type [36] which the concept references is also identified. This would guide us to correctly reference the UMLS Semantic Network [36], which is the upper level ontology referenced by SCT, for the new concepts that will be added to the final ontology. An upper ontology consists of general concepts not specific to a particular domain. Referencing an upper ontology helps in consistent ontology development and to achieve semantic interoperability between different ontologies [36].

In addition to identifying the terms, the literature is used to develop domain knowledge rules in natural language so that the domain experts can provide their feedback on the same. These rules are formalized later.
on using the concepts and relationships from the ontologies. For example, a natural language IF-THEN rule will look like this: “IF a patient is diagnosed with oral papillomatosis THEN the patient should be checked for Cowden syndrome”. This rule is based on the domain knowledge that oral papillomatosis is one of the main symptoms of Cowden syndrome [35].

3.2. Term to Concept matching

As mentioned previously, it is important to evaluate SCT for problem areas and to assess its coverage. However, from what we know about the issues with the poly-hierarchical structure, this step is not trivial. SCT contains many synonyms and similar sounding words across more than one hierarchy. Therefore, simply checking for the availability of term in the form of yes or no will not uncover the problems of incorrect and inconsistent modeling of concepts. Hence, the terms collected as part of the domain knowledge reuse step should be checked for: (i) complete match – when the term’s definition and context of use matches with a concept within SCT, the term is considered to have a complete match; (ii) structural match – when a term matches a concept in SCT but the definition and context of use do not match, it is considered as a structural match; (iii) semantic match – when a term does not have an equivalent concept in SCT but a synonymous concept is available that matches the context of use, the term is considered to have a semantic match; (iv) partial match – a term is said to have a partial match when it matches to a more general or a more specific concept; (v) no match – in case of none of the above, the term is considered not matched.

Accordingly, we assessed 400 terms (medical and oral) identified from part of 3.1 and the results are shown in figure 3. Three domain experts guided us for this purpose and provided feedback at each of the steps we completed. The result that 98 of the matched 360 terms did not have a complete match further reinforces our proposed approach of checking each term not only for a matching word but also for context, relationships and definitions. Next, based on this assessment, we analyzed whether to reuse a concept in its original form and position within SCT, or make changes to it, or to add a new concept (by reusing from elsewhere if available). Further, if the concept existed in more than one hierarchy, we identified the hierarchy most relevant to our context of use and meaning; or consolidated the relevant concepts into a single hierarchy to prevent redundancies in usage and modeling. We also modeled the concepts as classes, subclasses or individuals based on the requirements of the resulting cross-domain ontology.

3.3. Relationship to Object Property matching

The relationships between terms in the real world are referred to as object properties in OWL ontologies and we follow the same convention in this paper. These properties enable the DL reasoners to classify the ontology, which includes ensuring consistency, concept satisfiability, and instance checks; and even identify hidden or non-asserted associations once the ontology is classified. The object properties also give meaning and context to the ontology. This further
Table 1. Some interdependent medical and oral health conditions

<table>
<thead>
<tr>
<th>Medical/systemic Conditions</th>
<th>Relation</th>
<th>Oral Conditions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gastroesophageal reflux disease</td>
<td>{symptoms →}</td>
<td>Enamel erosion, dental sensitivity, foul taste</td>
</tr>
<tr>
<td>Immunosuppression, alcoholism</td>
<td>{causative →}</td>
<td>Squamous cell carcinoma (Oral cancer)</td>
</tr>
<tr>
<td>Diabetes mellitus, pneumonia, stroke, rheumatoid arthritis</td>
<td>{← associated →}</td>
<td>Periodontal disease</td>
</tr>
<tr>
<td>Osteoporosis</td>
<td>{symptoms →}</td>
<td>Reduced jaw bone density, tooth loss</td>
</tr>
<tr>
<td>Ventilator associated pneumonia</td>
<td>{← risk factor}</td>
<td>Poor oral hygiene</td>
</tr>
<tr>
<td>Administration of some narcotic analgesics</td>
<td>{risk factor →}</td>
<td>Xerostomia (dry mouth)</td>
</tr>
<tr>
<td>Infective endocarditis</td>
<td>{← risk factor}</td>
<td>Dental extraction</td>
</tr>
<tr>
<td>Human Immunodeficiency Virus (HIV)</td>
<td>{early indicator →}</td>
<td>Oral hairy leukoplakia</td>
</tr>
<tr>
<td>Asphyxia, sepsis</td>
<td>{← complications}</td>
<td>Ludwig’s Angina</td>
</tr>
<tr>
<td>Gardner’s syndrome</td>
<td>{early indicators →}</td>
<td>Osteoma of jaw bone, impacted teeth</td>
</tr>
<tr>
<td>Scurvy</td>
<td>{symptoms →}</td>
<td>Loss of jaw bone, teeth loss, gingival haemorrhage</td>
</tr>
<tr>
<td>Nutritional deficiency especially of vitamins A, C, D</td>
<td>{symptoms →}</td>
<td>Enamel hypoplasia of the pitted variety</td>
</tr>
</tbody>
</table>

enables the ontology to be consistently reused and be semantically interoperable with other ontologies. Therefore, it is vital to identify from literature the accurate relationships between terms. For example, table 1 gives the relationships between few interdependent medical-oral health conditions as obtained from literature. The importance of the relationships can be clearly seen since the terms by themselves essentially mean nothing. Moreover, accurate identification of these relationships is necessary to correctly evaluate the ontology to be reused for the availability of matching properties. Once the relationships are identified from literature, we suggest dividing the matching process into four categories as follows: (i) complete match – when the relationship extracted from literature and via the domain knowledge rules exist in the same context as an object property within SCT; (ii) super-property match – when a more general property is available within SCT. For example, the ‘symptom’ relationship from table 1 should have an object property symptom-Of or similar in SCT but that is not the case. However, SCT has the property associated-With, which indicates that two concepts do have some relation with each other, but does not specify ‘how’ they are associated. The symptom-Of property is a more specific property between two concepts in that it specifies ‘how’ they are associated. Therefore, associated-With is considered as a super-property match for symptom-Of and the latter automatically becomes the sub-property; (iii) sub-property match – when a more specific property is available for a relationship within SCT, it is considered as a sub-property match; (iv) no match – none of the above.

Our early results of matching 24 relationships to object properties are shown in figure 4. It is of interest to note that 5 properties did not match at all or in other words, the relations defined by these properties are not modeled in SCT. One of the not matched relations was the relation between classes or subclasses to individuals, which is translated to the object property instance-Of. This missing property shows that SCT does not support instances. Hence, integrating SCT in its original format with patient data and reasoning over

Figure 4. Relationship to Object property match
both for use in expert systems is not feasible, a requirement for our purpose of inferencing over shared medical and dental records of a patient.

3.4. Knowledge base

This step involves building the ontology itself and formalizing the domain knowledge rules by using the results obtained from the above steps. An important consideration at this point is the modeling of individuals, a concern in the original model of SCT as discussed before. Since SCT does not model individuals and support assertional knowledge, it is difficult and cumbersome to integrate patient data with SCT, and reason and query over it thereby making it unsuitable for use in expert systems. Hence, it is prudent to correctly model as individuals those terms that are incorrectly modeled as subclasses in SCT using the guideline from [32] so that a concept is modeled as a subclass when it describes a subset of a class otherwise it is modeled as an individual. For example, according to the guideline and as required for the purpose of our ontology, we model Cowden syndrome as an individual as against SCT that models it as a subclass. Further, modeling of the ontology in this way will allow for integration with rules that add to the restricted expressivity of the underlying language, OWL in this case. Rules will increase the richness and inferencing capabilities of the ontology so that it can be used in expert systems, unlike SCT.

The practical application of the knowledge base is next explained with the help of a running example.

3.4.1. Example. Figure 5 is a screenshot taken from protégé of two SCT hierarchies for the conditions, oral papillomatosis and Cowden syndrome and a few object properties. Only the most relevant portions are shown in the figure with the circles representing classes or subclasses and diamonds representing individuals. The dotted sections indicate where this example screenshot differs from the original SCT model with respect to modeling of individuals and addition of new properties. Based on this model, the IF-THEN rule mentioned in section 3.1 can be formalized using the Semantic Web Rule Language (SWRL) [37]. SWRL is semantically compatible with OWL and very expressive and hence our language of choice. The IF-THEN rule is completely expressed using two SWRL rules as follows:

\[
\text{Clinical\_Finding(Oral\_Papillomatosis)} \land \\
\text{Clinical\_Finding(Cowden\_Syndrome)} \rightarrow \\
\text{symptomOf(Oral\_Papillomatosis, Cowden\_Syndrome)}
\]

... [rule 1]

\[
\text{Patient(?p) \land hasClinicalFinding(?p, ?x) \land} \\
\text{Clinical\_Finding(?x) \land symptomOf(?x, ?y) \land} \\
\text{Clinical\_Finding(?y) \rightarrow suspect(?p, ?y)}
\]

... [rule 2]

The \( \land \) indicates conjunction and the rule is of the form antecedent \( \rightarrow \) consequent, that is, if the antecedent is true, then the consequent holds. Accordingly, rule 1 states that oral papillomatosis is a symptom of Cowden syndrome and similarly, rules can be written for the other symptoms of Cowden syndrome. When this rule is triggered, it in turn will trigger rule 2, which states that when a patient is diagnosed with a clinical condition (x) and which in turn is a symptom of another condition (y) then the patient should be suspected with having condition (y).

Figure 5. Example screenshot of SNOMED-CT concepts and object properties in Protégé
As is evident, rule 2 is a general rule for conditions that are related by the ’symptom-Of’ property. Therefore, once any related conditions are specified as in rule 1, and the rule is triggered, it will automatically trigger rule 2 to check for patient records with one of those conditions and if such a record is identified, the inferred knowledge will be integrated with the corresponding patient record. With the above rules running, it is also possible to query the ontology and patient information with a complex query such as mentioned below to identify the symptoms of Cowden syndrome as well to obtain a list of patients who may have been diagnosed with the symptoms.

Query: What are the symptoms of Cowden syndrome and which patients have been diagnosed with those symptoms?

The following information is used to answer the query:

Ontological information: figure 5 and rules 1, 2
Patient information: data from patient records integrated with the ontology via the instance-Of relation

A portion of the formal SPARQL [38] query is:

```
PREFIX cdo: <http://example-crosstownontology.org/cdo/>
SELECT ?condition ?name
WHERE {?
condition :symptomOf cdo:Cowden_Syndrome ;
?name :hasClinicalFinding ?condition ;
rdf:type cdo:Patient . }
```

Figure 6 shows the results for the above query based on simulated patient cases. For simplicity, the results are shown for only one symptom of Cowden syndrome that is, oral papillomatosis. This example shows how ontology and rules work together to represent reusable and shareable knowledge as well as allow reasoning and querying over it to discover new knowledge.

An important consideration throughout the development process is to keep the steps, 3.1 through 3.4, iterative based on the feedback of the domain experts. Further, regular validation of the ontology and rules by DL reasoners such as Pellet will ensure consistency in modeling and correctness of subsumption in the resulting ontology. As part of our future work, we aim to evaluate the above framework by implementing the ontology and rules in a prototype system and obtaining feedback from the domain experts. Specifically, we will compare the new ontology and SCT with respect to coverage for the medical and dental domains and when combined with rules, its ability to represent inter-dependent medical and oral health conditions that SCT alone cannot.

4. Related work

Well-established ontology engineering methodologies such as OntoClean [39], Methontology [40] and On-To-Knowledge [41] do not address ontology reuses issues. On the other hand, there is an increasing body of work in the areas of ontology mapping and alignment, ontology integration, merging, and modularization that is targeted at the ontology reuse process [42, 43]. However, since they are more generic in their approach, they do not address the inherent problems that are also unique to the biomedical domain such as large-scale ontologies that are poly-hierarchical and many of which such as SCT in this case, did not follow the ontology development principles. For example, the problem of multiple inheritance and incorrect modeling requires expert assessment before a certain subset of an ontology is reused. Automatic extraction of fragments or subsets, merging or integration and so on will only duplicate the errors in the resulting ontology. Within the biomedical domain, compounding of such errors is especially dangerous considering the ontologies and knowledge base will be realized in an expert system, or an information management system or a translational research system all of which ultimately deal with human lives. In this work therefore, we have provided a novel framework that strongly advocates identification and understanding of the problems of the ontology to be reused and addressing those along the reuse and development process. Furthermore, the processes outlined can be complemented with the above approaches and the available tools and applications for ontology reuse.
Ontologies have also been increasingly used to address the issues of semantic interoperability in healthcare [8-14, 44-46]. Specifically in [44], the authors use OWL based ontology and reasoning to enable data sharing and interoperability across hospital wards and departments and in [45], the authors use ontology layering by employing SWRL rules to address the issue of semantically heterogeneous e-health information. In [46] on the other hand, the authors translate clinical archetypes, which contain defined clinical concepts, into OWL based archetypes to enable mapping with other ontologies and facilitate semantic interoperability. Moreover, OWL ontologies and SWRL rules have also been used for automated decision support for specific health conditions and disorders such as in [47] for acute cardiac disorders and for prescription recommendations such as for selecting diabetes medication in [48].

Thus, it can be seen that researchers have addressed the issues of ontology reuse, semantic interoperability and decision support in different ways. However, to the best of our knowledge, there have not been any previous efforts to bridge the gap between the medical and dental domains by building a formal knowledge base to enable the sharing and reasoning over patients’ electronic medical and dental information. Our approach therefore is different in that we propose the reuse of an already existing and mature ontology and use the combination of OWL and SWRL to bridge the medical and dental domains that are functioning in isolation for most practical purposes by: (i) developing a formal, cross-domain, reusable knowledge base consisting of scientific evidence-based associations to enable the seamless sharing of patient information while retaining the correct meaning of the shared information, and (ii) increasing the expressivity of the underlying ontology with rules to achieve decision support over the shared patient information for the purposes of clinical diagnosis and recommendations.

5. Conclusion

In this paper, we have discussed the necessity of reducing the gap between the medical and dental domains, specifically the oral health specialty of the latter domain by highlighting some associated conditions that are important considerations for both the medical and dental practitioners. We have then explained how this issue can be addressed by developing a cross-domain ontology and knowledge rules to form a knowledge base that can be reused by expert systems to aid in the diagnostic process. For the purpose of building the cross-domain ontology, we identify SCT as the ontology of choice to be reused for its comprehensiveness and global acceptance. Further, we have briefly discussed some of the problems in SCT with the aim of preventing their duplication in the resultant ontology. Based on these issues, we have proposed a guiding framework that breaks down the whole process into small steps such that the problems can be recognized and addressed. We anticipate that this framework will function as a generic framework to reuse SCT or similar ontology for the development of new biomedical ontologies.

6. References