


# A Methodology for Ontology Reuse: The Case of the Abdominal Ultrasound Ontology

Nur Zareen Zulkarnain, Centre for Advanced Computing Technology, Fakulti Teknologi Maklumat Dan Komunikasi, Universiti Teknikal Malaysia Melaka, Melaka, Malaysia

Farid Meziane, University of Salford, Manchester, United Kingdom

 <https://orcid.org/0000-0001-9811-6914>

## ABSTRACT

There is an abundance of existing biomedical ontologies such as the National Cancer Institute Thesaurus and the Systematized Nomenclature of Medicine-Clinical Terms. Implementing these ontologies in a particular system however, may cause unnecessary high usage of memory and slows down the systems' performance. On the other hand, building a new ontology from scratch will require additional time and efforts. Therefore, this research explores the ontology reuse approach in order to develop an Abdominal Ultrasound Ontology by extracting concepts from existing biomedical ontologies. This article presents the reader with a step by step method in reusing ontologies together with suggestions of the off-the-shelf tools that can be used to ease the process. The results show that ontology reuse is beneficial especially in the biomedical field as it allows for developers from the non-technical background to build and use domain specific ontology with ease. It also allows for developers with technical background to develop ontologies with minimal involvements from domain experts.

## KEYWORDS

Biomedical Ontology, Bioportal, Ontology Reuse, Ontology

## INTRODUCTION

In previous work, Zulkarnain et al. (2015a) proposed the development of an architecture to support ultrasound report generation and standardisation. The proposed architecture and report were validated by radiologists and specialists at the 2015 UK Radiological Congress held in the City of Liverpool (Zulkarnain et al., 2015b) and the 2016 British Medical Ultrasound Annual Scientific Meeting and Exhibition held in the City of York (Zulkarnain et al., 2016a). In the medical and radiography fields, ultrasound reports are the main media used to communicate the results of an ultrasound examination from a sonographer or radiologist to a referring clinician. It was reported that images alone are of limited value since the outcomes of any ultrasound investigation are based on the findings during the scan (Boland, 2007). Indeed, many features and quantitative data are collected during the ultrasound examination such as tissue characterisation and various measurements and it is this information that is communicated via the reports.

The use of information technology (IT) in the medical field such as electronic patients' records and some decision support systems has allowed for a better understanding of some pathologies, health management and patient care. However, the integration of IT in the radiology field is limited particularly in the reporting phase. During the development of the standard report and its validation, it was highlighted by the radiologists and clinicians that the main issue is the variations in the reporting

DOI: 10.4018/IJIT.2019100101

styles. These variations were noticed in the structure of the reports as well as in the terminologies used. These variations may impact on the way a report is interpreted and in turn affect the decision-making process and the way a patient is managed (Zulkarnain et al., 2015a). Radiologists and clinicians believe that the solution to this problem resides in using structured reporting with the support of an ontology as its knowledge base (Kahn, et al., 2009). The use of an ontology will allow the standardisation of the terminology used during reporting, allowing a better exploitation of these reports by computerised tools for knowledge discovery, classification and predictions.

The work reported in this paper, is the development of an ontology that will complement and support the computerised standard report developed by Zulkarnain et al., (2015a). Furthermore, this paper is an extension and consolidation of the works reported by Zulkarnain et al. (2015a, 2015b, 2016a, 2016b).

There are different approaches to develop ontologies. We can use existing ontologies, develop one from scratch or adapt and reuse existing ontologies. Each approach has its advantages and disadvantages. For example, in one hand reusing an existing ontology will require no resources for its development but may be too large for a specific application and its integration to the rest of the system could be problematic. On the other hand, developing a new one will require a lot of efforts during its development but may fit better to the requirements of the new application. In this research, we are specifically interested in abdominal ultrasound reporting, and the existing ontologies evaluated were found to be large. For example, using the National Cancer Institute Thesaurus (NCIT) would require a large storage, as it contains as many as 118,941 classes, and more time to process. In this research we adopted the ontology reuse approach for developing the Abdominal Ultrasound Ontology (AUO) to be used in the ultrasound reporting system.

This paper first reviews the methods that have been used in past ontology reuse works and then assesses the possibility of reusing one of the three established biomedical ontologies namely the Foundational Model of Anatomy (FMA), the Radiology Lexicon (RadLex) or the Systematized Nomenclature of Medicine-Clinical Terms (SNOMED CT). We then propose a methodology to reuse biomedical ontologies together with the existing tools that can be used to facilitate the reuse process. The methodology aids ontology developers in: (i) selecting suitable ontologies for reuse based on their corpus; (ii) selecting the concepts to reuse from the selected ontologies; (iii) evaluating the developed ontology with minimal help from the domain experts. It is anticipated that the development of the AUO will serve two main purposes in the standardisation of the ultrasound reporting system: (i) it will be used to standardize the development of ultrasound reports and enforce the use of a standard terminology and (ii) to analyse ultrasound reports written in Natural Language (English free-text) with the aim of automatically transforming them into a structured format.

The remaining of the paper is organised as follows. Section 2 reviews some of the previous works on ontology reuse and discusses their suitability to be used in the current research. In section 2, we review some of the existing biomedical ontologies and how well they cover the terms and lexicon used in ultrasound reporting. The proposed reuse methodology will be described in detail in section 3. We discuss the results of using the proposed methodology in developing the AUO in section 4 and conclude the paper in section 5.

## RELATED WORK

Ontology reuse is the process where parts of existing ontologies are used in the development of new ones (Bontas, 2005). The reused parts are very often manipulated to meet the requirements of a new application (Katsumi and Gruninger, 2016). Such an approach reduces the development cost of an ontology as the need for domain experts is minimal (Alani et al., 2006) and increases interoperability (Simperl, 2009) as the new and old ontologies will share features and concepts. However, existing tools are not providing the necessary support for the ontology reuse process (Maedche et al, 2003; Simperl, 2009). Most ontology reuse methodologies proposed (Alani, 2006; Caldarola et al., 2015;

Capellades, 1999; Fernández-López et al., 2013; Russ et al., 1999; Shah et al., 2013; Simperl, 2009, Uschold et al., 1998) follow the following four steps: (i) Ontology selection for reuse, (ii) Concept selection, (iii) Concept customization and (iv) Ontology integration.

The purpose of the first step, ontology reuse, is to select the ontology to be reused. The selection is based on a set of criteria according to the requirements of the new ontology and includes the language used in the development of the ontology, its reasoning capabilities and how well it covers the terminology used in the specific application domain. We note that in this phase, if required by the needs of the new ontology, several ontologies can be selected for reuse. The selection step is followed by the concepts selection phase that are then translated and customised to suit the needs of the new ontology in the concept customization step. The final step involves the integration of the newly developed ontology into the new system or application.

In the ontology reuse system proposed by Alani (2006), all the new terms needed for the new ontology are first listed to determine which existing ontology should be selected for reuse. The system then searches for relevant ontologies online using the terms that have been listed and from the obtained result, the ontologies will then be ranked and only the first few will be analysed and selected for reuse. Once the ontology for reuse has been selected, the system will determine whether to reuse the whole ontology or only a segment of it. This step could be considered as concept selection where relevant concepts are being selected. In his work, Alani (2006) has decided to reuse several ontologies for one concept. Thus, each group of related concepts needs to be customized to ensure that the concepts have a standardised format so that they can be merged as one concept. Since the concept is reused from several ontologies, each concept contains different properties which resulted in additional knowledge representation. Finally, the ontology is automatically evaluated.

Another example of ontology reuse is in the development of the Oral-Systemic Health Cross-Domain Ontology (OSHCO) by Shah et al. (2013). In developing OSCHO, the first step taken by Shah et al. was to determine the scope of the ontology by recognizing the domain of coverage, the intended use and what questions should OSCHO be able to answer. Once the scope has been determined, Shah et al. used a tool in Bioportal<sup>1</sup> and submitted several domain related terms to test the domain coverage of several ontologies that have the potential to be reused. This has resulted in SNOMED CT being the best candidate. Compared to Alani, Shah et al. reused only one ontology, SNOMED CT where they selected the concepts needed then added other relevant concepts not included in SNOMED CT. Since they reused only one ontology, the new ontology developed, OSCHO, closely follows the model of SNOMED CT where the new concepts added are customized to follow the same model.

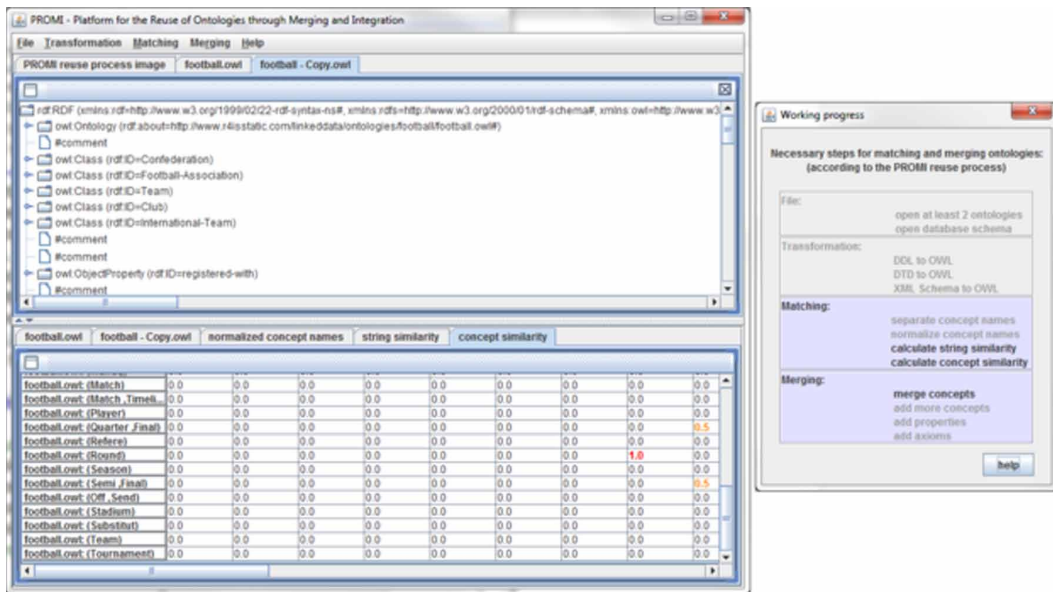
Russ et al. (1999) in their work aim to develop an aircraft ontology that can be used by several applications to ensure knowledge sharing between them. During the development of their work, they realized that there exist two ontologies that are related to the aircraft domain.

Furthermore, some concepts exist in one ontology but not in the other. Thus, Russ et al. decided to merge both ontologies to create a more complete one. The first step taken by Russ et al. is selecting the ontologies to reuse which are the time ontology (which is publicly available) as well as the two existing aircraft ontologies. The time ontology is written in Ontolingua while the two aircraft ontologies were written in Loom. After both aircraft ontologies were merged, the time ontology is translated to Loom so that it could be integrated into the aircraft ontology. The same approach was used by Caldarola et al. (2015) in developing an ontology in the food domain where metadata were manually translated to better understand the concepts.

Polionto (Ortiz, 2011) is another example of an ontology developed using ontology reuse. Ortiz in his work developed Polionto by reusing two ontologies in the political domain where the first ontology is in Portuguese and the other is in English. He first selects relevant concepts from both ontologies to reuse and compares them to their corpus. Those selected concepts are then translated and integrated to create a multilingual political ontology called Polionto.

Comprehending the need for a tool that will assist the process of ontology reuse, Bontas and her colleague developed a tool named PROMI that is able to perform the steps required in an ontology

Figure 1. The user interface of PROMI

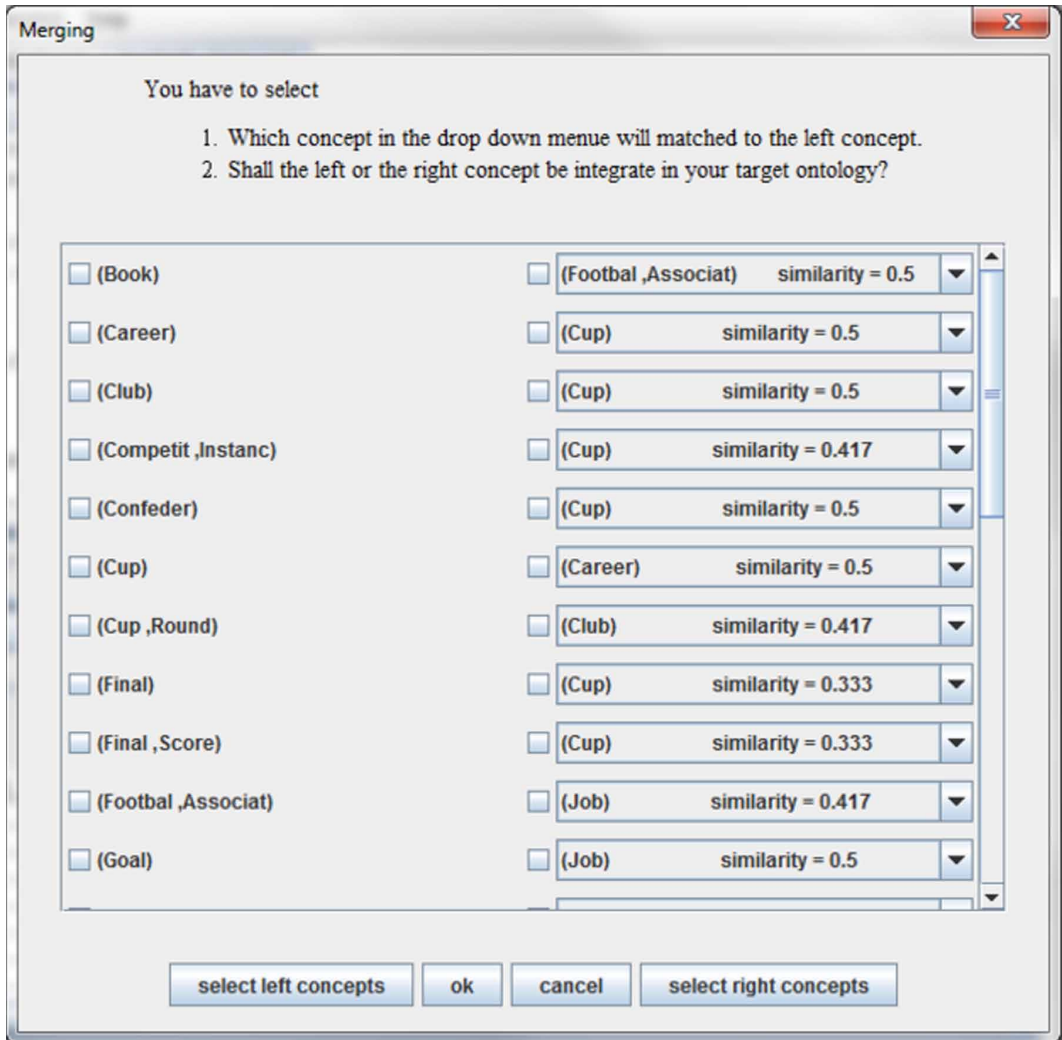


reuse process (Bontas, 2007) as seen in Figure 1. PROMI begins by prompting the users to upload at least two ontologies to be reused. Then, the language of the ontologies will be examined in order to decide whether the transformation to OWL language should be performed. Next, the matching of the concepts was conducted where PROMI first separates and normalises the concept names before calculating the string and concept similarity. Finally, the concepts will be merged, and users are allowed to include more concepts, properties and axioms as necessary.

The development of PROMI is an immense step forward in assisting the process of ontology reuse. However, there are several limitations in using PROMI as a tool to assist ontology reuse. First, PROMI does not provide the facility to assess the suitability of an ontology to be reused for a specific domain. Instead, it begins by prompting the users to upload ontologies which they have selected beforehand. It also does not recommend any ontologies for reuse based on the domain which means that the users will need to use their own expertise in selecting suitable ontologies. Second, the concept matching between the two ontologies depends heavily on the similarity measures chosen by the users. For example, the similarity measure for the word “tournament” and “competition” was 0 using the Euclidean Distance measure and 0.143 using the Hamming Distance measure. Therefore, the users need to have some knowledge about the differences between these measures for them to be able to select the most appropriate one. Finally, in order to merge concepts in PROMI, users will need to select from a list of concepts and its equivalent candidate concepts which have a certain degree of similarity. This is shown in Figure 2. PROMI could be a beneficial tool in assisting the process of ontology reuse. However, there are still several features that can be included and improved so that it can ease the ontology process even more and allows the usage on large ontologies to be more efficient.

From the examples mentioned above, it can be seen that the ontology reuse methodologies used in previous works is roughly similar to the four steps mentioned above. In this paper, these steps were also used as a guideline in developing an ontology reuse methodology for the biomedical domain. The methodology presented in this paper will allow for the development of a new ontology by reusing multiple existing ontologies and suggest tools that would help in each step of the methodology. The difference between our approach and the existing ones is the incorporation of off-the-shelf tools to aid in the required tasks. The main motivation in proposing this methodology is to allow novice

Figure 2. The process of merging concepts in PROMI



developers especially from the non-technical background to create and use ontology in their field by alleviating the notoriously painstaking task of developing a domain specific ontology from scratch. Our proposed methodology would also allow for ontology to be developed with only minimal involvement of domain experts.

### Review of Existing Biomedical Ontologies

Many ontologies are developed in the biomedical field as it has a large number of terminologies and concepts. In building a medical ultrasound reporting system, the first step would be to investigate if there is an existing ontology that can be used, or if there are ontologies that can be reused instead of building a new one from scratch. As a first step, we need to review related suitable existing ontologies. After careful research, three ontologies have been initially selected namely: the Foundational Model of Anatomy (FMA), the Systematized Nomenclature of Medicine - Clinical Terms (SNOMED CT) and the Radiology Lexicon (RadLex).

This initial selection was purely based on their coverage, language and popularity in the biomedical community. In selecting one ontology to be adopted we first look at the domain of each ontology. FMA covers the concepts and relationship in the structural organization of the human body from the macrocellular to microscopic levels (Rosse and Mejino, 2003) while SNOMED CT covers more and includes clinical findings, chemical substance scales and other miscellaneous health information (Benson, 2010). RadLex on the other hand incorporates many complex radiology related domains from basic science to imaging technology (Rubin, 2008). This shows that both FMA and SNOMED CT cover a wider area compared to RadLex. However, RadLex would be more domain specific in relation to developing an ontology that covers the abdominal ultrasound process.

To select the best ontology to be reused, we adopted the metrics provided in BioPortal as shown in Table 1 to compare the three ontologies identified. From here, we can see that all three ontologies are too large to be implemented for one specific system as this could result in slow computing time and taking up a lot of memory space. Thus, we have decided to reuse only the relevant concepts from all three ontologies and merge them into a new one with the aim of achieving a wider coverage and less resources consumption.

### The Proposed Methodology

The most important part in reusing ontologies is to select the most suitable ones. From the initial review, FMA, SNOMED CT and RadLex have been identified as potential ontologies to be reused. However, these ontologies will need to be evaluated by comparing them with our corpus of 100 medical ultrasound reports to ensure that these ontologies are the right ones for reuse and provide the right coverage.

In this section, the different steps of the developed ontology reuse methodology will be described in detail. The terms of the corpus constructed from the ultrasound reports will be used for selecting suitable ontologies. These terms are then used to select the relevant concepts from these ontologies and merge them into a single one before being evaluated by domain experts. Figure 3 summarises the methodology steps.

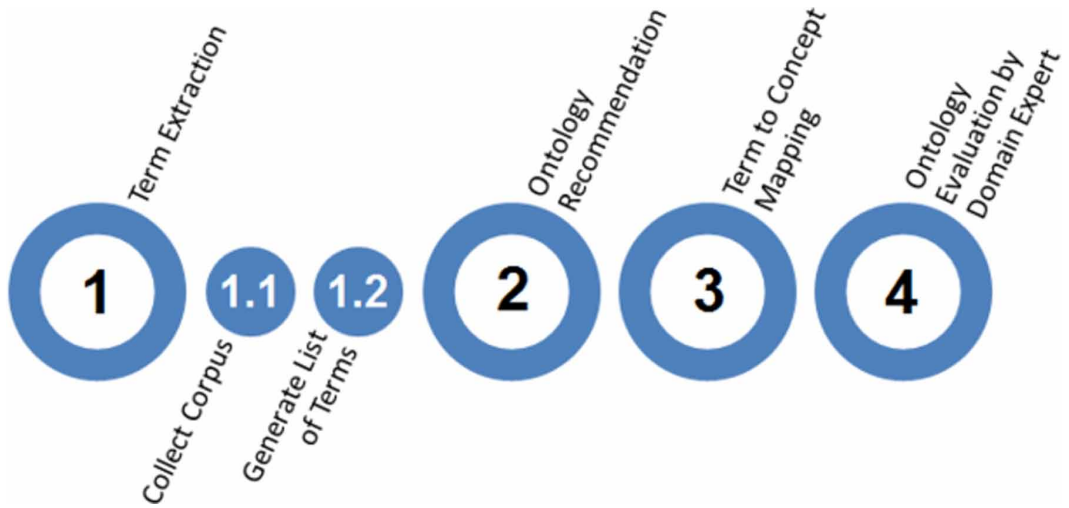
### Term Extraction

The scope and domain of the Abdominal Ultrasound Ontology (AUO) is to model the taxonomy, pathology, equipment and other terms related to abdominal ultrasound reporting. 100 sample abdominal ultrasound reports have been collected and used to construct our corpus. These reports have an average word count of 70.82, average number of words per sentence of 11.15 and average token size of 6.45 characters. These sample reports were obtained from the Directorate of Radiology of the University of Salford. Once the corpus is constructed, the next step is to extract all the relevant terms to generate a list of biomedical and technical terms to be included in the ontology. Two biomedical extraction applications were used during the extraction process namely: TerMine<sup>2</sup> and BioTex<sup>3</sup> to define a subset of the most suitable terms for standard reporting systems. 49 out of the 100 sample reports were submitted to both applications and the results obtained are summarised in Table 2.

Table 1. Comparison between FMA, SNOMED CT and RadLex

	FMA	SNOMED CT	RadLex
Number of classes	104,258	324,129	46,140
Number of individuals	0	0	46,140
Number of properties	172	152	96
Format	OWL / OBO	OWL	OWL

Figure 3. Ontology reuse methodology



The BioTex application extracted more terms (761 terms) compared to only 241 terms for TerMine. The superiority of BioTex is due to its ability to extract both multi-words and single-words as TerMine only extracts multi-words. Hence, BioTex was chosen as the term extraction application to be used. For example, if the sentence “Normal liver with no focal lesions seen” was submitted to both applications, TerMine will only extract one multi-word term which is “focal lesion” while BioTex will extract not only the multi-word term but also “liver” which is a single word term. If single-word terms such as “liver”, “kidney” and “spleen” are not extracted, the ontology developed would be incomplete. Terms which are extracted from BioTex were also validated using the Unified Medical Language System (UMLS) which is a set of documents containing health and biomedical vocabularies and standards. Using BioTex we managed to extract 1119 terms from the 100 sample ultrasound reports.

### Ontology Recommendation

Once the list of terms is obtained, the next step is to select the ontology. Three criteria were set for the selection of the ontology and these can be summarised as follows:

Table 2. Comparison of biomedical term extraction using TerMine and BioTex

	TerMine	BioTex
Language	English	English
License	Open	Open
POS Tagger	GENIA Tagger / Tree Tagger	Tree Tagger
Terms Found	241 (GENIA Tagger)	761
	232 (Tree Tagger)	
Extraction Type	Multi-word extraction	Multi-word extraction and
		Single-word extraction

- (1) **Ontology coverage** - To which extent does the ontology covers the terms extracted from the corpus?
- (2) **Ontology acceptance** - Is the ontology being accepted in the biomedical field and how often is it used?
- (3) **Ontology language** - Is the ontology written in OWL, OBO or other ontology format? How widely is the format being used in the ontology community?

Ontology coverage has the highest weightage when determining whether an ontology is suitable for reuse or not. An ontology that contains most of the concepts needed will preserve the model of the ontology. This model will then be followed by concepts taken from other ontologies. The next important criteria is the acceptance of the ontology within the biomedical community as the level of acceptance indirectly shows the quality of the ontology (McDaniel et al., 2016). The higher the acceptance score, the more the ontology is being used thus promoting interoperability between different systems. Finally, the format used to develop the ontology is also an important criterion to avoid translating the ontology to another format.

Using the three defined criteria, FMA, SNOMED CT and RadLex were initially selected as the most suitable candidates for reuse. As a further step, we have used the ontology recommender system provided by BioPortal which is an open ontology library that contains ontologies with domains that range from anatomy, phenotype and chemistry to experimental conditions (Noy et al., 2009). There are several frameworks that can be used to select suitable ontologies for reuse such as the one proposed by Trokanas and Cecelja (2016) which uses similarity measures to calculate the compatibility of an ontology for reuse (candidate ontology) and the ontology they would like to expand (primary ontology). However, we decided to use the recommender provided by BioPortal as it is specialised for the biomedical domain and all three ontologies to be investigated are available on BioPortal.

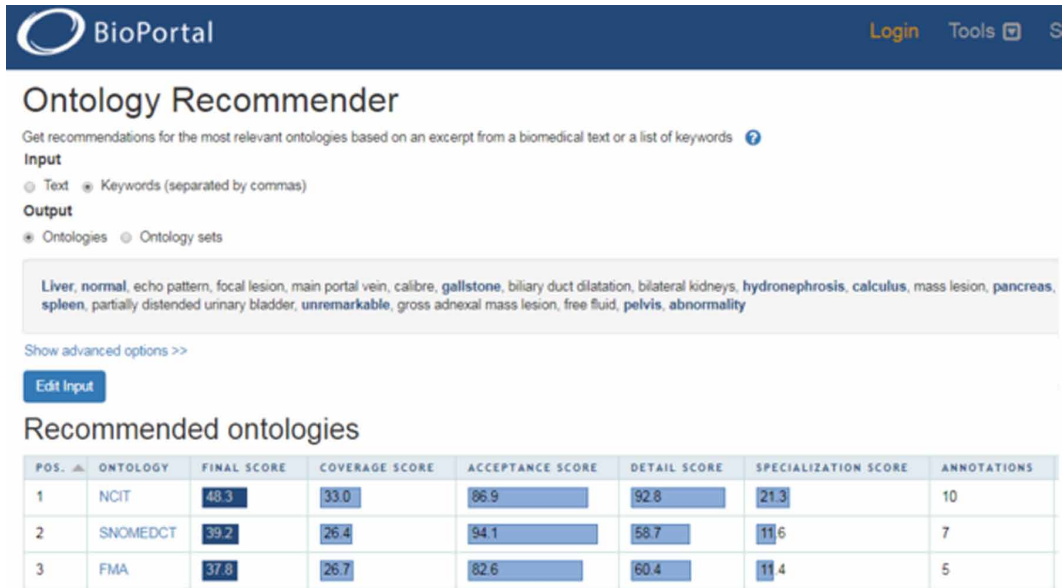
Once a user provides a list of terms, the ontology recommender available on BioPortal uses these terms to suggest the most suitable ontologies for reuse. The recommender uses four criteria namely: (i) Coverage, (ii) Acceptance, (iii) Detail and (iv) Specialization. Based on BioPortal analysis, these are the criteria that are the most relevant for recommending ontologies for reuse (Jonquet et al., 2010). The terms can be submitted as a paragraph or a list of terms to the recommender and as a result, it will return a list of 25 recommended ontologies ranked from the highest to the lowest scores (see Figure 4). The ranking of all the ontologies available in the BioPortal repository that meets all the criteria is computed by giving scores to four metrics namely coverage, acceptance, knowledge detail and specialization. The final score is calculated based on the following formula:

$$\text{FinalScore} = (\text{CoverageScore} * 0.55) + (\text{AcceptanceScore} * 0.15) + (\text{KnowledgeDetailScore} * 0.15) + (\text{SpecializationScore} * 0.15)$$

The coverage score is given based on the number of terms in the input that are covered by the ontology. It is given the highest weightage which is 0.55 compared to the other three metrics which are all given a weightage of 0.15 because ontology coverage is seen as an important factor in determining the suitability of reusing an ontology for a certain corpus. The acceptance score indicates how well-known and trusted the ontology is in the biomedical field. It is given based on the total of website visits to the ontology as well as whether or not the ontology exists in UMLS. Knowledge detail score on the other hand indicates the level of details in the ontology; i.e. does the ontology have definitions, synonyms or other details. This is also an important metric in determining the quality of an ontology. This is because an ontology that only has a hierarchy of terms and contains no other details such as definitions would merely be seen more as a taxonomy rather than an ontology. Thus, it would have less purpose as a knowledge base in a system or application. Lastly, the specialization score is given based on how well the ontology covers the domain of the input. This is different compared to the coverage score because the size of the ontology is also given attention since the specialization



Figure 4. BioPortal's ontology recommender



score ranks domain specific ontologies higher compared to general ontologies. In this research, the specialization score is not as important as the coverage score since the aim is to achieve the highest coverage even though we would need to reuse more than one ontology. An example of this scoring system in action is shown in Figure 4 where 21 terms extracted from a sample ultrasound report were submitted.

There is however a limitation when using BioPortal as it allows submission of only 500 words. To process all the 1119 terms that were extracted from our reports, we have to customize the existing recommender system to come up with another system by manipulating the data from BioPortal's ontology recommender API (BioPortal, n.d.). We first developed the recommender system that would submit all 1119 terms to BioPortal's API and give a recommendation of 25 ontologies ranked according to its final score just like how it would be in BioPortal's recommender. However, it seems that the 1119 terms were too big for the recommender's server to handle and causes the recommender to crash midway without giving any result.

To overcome this, we decided to develop the recommender system that will allow all 1119 terms to be submitted but will process the terms one by one instead. Each term will be submitted to the API to get a list of ontology recommendations and their scores. The ontology with the highest score will be selected as the ontology recommended for the individual term. The frequency of all ontologies recommended will then be counted. The ontology with the highest frequency will be reused first, followed by the second and the following ontology until we achieve the optimum coverage. Figure 5 summarizes the used process.

Figure 6 shows an excerpt of the result from processing 1119 terms using the recommender system we have developed. The recommender system allows us to submit quite a huge number of terms to be processed and the result will then be stored for analysis. This is seen as a better alternative because it is impossible to straightaway submit all 1119 terms and get a result since the API server would not be able to handle such a huge number. Figure 6(a) shows an excerpt of all the list of terms submitted and the ontology recommended for each term based on the final score they got. This score is calculated by BioPortal using the same formula that they used for their recommender.

Figure 5. Algorithm for getting ontology recommendations for term list

```

Input: A list of terms, T
Output: 1. Ontology recommendation, O[n]
           and score, S[n] for each term, T[n]
           2. Ranking of ontology recommendation, O[n]
           with its frequency, F
1 foreach T[n] do
2   submit URL =
     "http://data.bioontology.org/recommender?apikey=
     &input=" & T[n];
3   if Result != Empty then
4     return O[n] = 0;
5   else
6     return O[n] = acronym [0] and S[n] =
       evaluationScore[0];
7   end
8 end
9 foreach O[n] do
10  count F of each O[n];
11  sort O[n] from highest F to lowest;
12 end
    
```

Figure 6. (a) Ontology recommendation for each term (b) Ranking of ontology recommended

(a)	1	Keywords	Ontology Name	Score	(b)	1	Ontology Name	Frequency
	2	scan	NCIT	0.929		2	NCIT	476
	3	ultrasound	NCIT	0.859		3	SNOMEDCT	207
	4	biliary dilatation	MEDDRA	0.836		4	RADLEX	48
	5	biliary duct	LOINC	0.809		5	LOINC	44
	6	gallbladder wall	NCIT	0.838		6	MESH	42
	7	renal mass	SNOMEDCT	0.894		7	RCD	25
	8	abnormality	NCIT	0.853		8	MEDDRA	16
	9	echotexture	RADLEX	0.779		9	OMIM	12
	10	flow	NCIT	0.798		10	NIFSTD	10
	11	bile duct	NCIT	0.929		11	CRISP	9
	12	biliary obstruction	NCIT	0.909		12	SWEET	9
	13	fatty infiltration	SNOMEDCT	0.855		13	CCO	7
	14	focal lesions	SNOMEDCT	0.602		14	DCM	7
	15	gall stones	MESH	0.96		15	SOPHARM	6
	16	hepatic veins	MESH	0.909		16	WHO-ART	6
	17	liver texture	NCIT	0.889		17	HUPSON	5
	18	normal appearances	SNOMEDCT	0.816		18	ONTOAD	5
	19	abdomen and pelvis	SNOMEDCT	0.898		19	EFO	5
	20	biliary duct dilatation	LOINC	0.674		20	NMR	4

After all terms have been submitted to the recommender, the frequency of the ontology recommended for each term will be counted and sorted from highest to lowest. The recommender system has ranked NCIT as the ontology with the highest frequency – which have been recommended for 476 terms; followed by SNOMED CT (207) and RadLex (48) as shown in Figure 6(b). This has proven that our initial selection of reusing FMA, SNOMED CT and RadLex was not very accurate since FMA does not even appear in the top 20 of the ontologies recommended for this corpus. The reason for this could be because FMA is very large and broad so it loses a lot of points in the specialization score. Its broad domain also means that it covers only the general terms available in the corpus causing it to also lose points in the coverage score hence resulting in it not being recommended for reuse.



should be extracted together with its subclasses or children instead of parents and ancestors. They argue that immediate parents and ancestors are unimportant and extracting them would increase the risk of creating an ontology that is equal to the ontology being reused. However, we believe that parents and ancestors are important in connecting concepts so that they would not be floating. If we were to take “spleen” and its subclass as one module, “kidney” and its subclass as another module, as well as “millimeter” and its subclass as another module, it would be hard to group these modules under the same category. Furthermore, the ontology being developed is very specific to the abdominal ultrasound domain. Thus, reusing parents and ancestors of a concept reduces the risk of the ontology being as large as the original one. Once all terms have been searched, this process will then be repeated for the remaining recommended ontologies which are SNOMED CT and RadLex.

A walkthrough example: Consider a list of terms that contains three words which are “gallbladder”, “duct dilation” and “gallstone”. We first take the first word “gallbladder” and query if the concept exist in NCIT. This returns an exact match where there exists a concept in NCIT with the preferred label “gallbladder.” Thus, this concept together with all its knowledge detail will be reused. We then look to see if the concept has any parents or ancestors. “Gallbladder” has a parent “organ” and an ancestor “anatomic structure, system, or substance” which both will be reused. Since “gallbladder” has an exact match in NCIT, it is removed from the term list.

We then query the second word “duct dilation” in NCIT which returns a partial match which consist of the word “duct” and another word “dilation”. Both concepts will be reused together with their knowledge details as well as their parents and ancestors. However, different to “gallbladder”, “duct dilation” would not be removed from the term list since it is only a partial match. The final word in the list, “gallstone” is then queried which gives a synonym match to the concept “gallbladder stone” in NCIT.

“Gallbladder stone” together with all its knowledge detail including synonyms will be reused. All its parents and ancestors will also be reused, and the term will be removed from the term list. Once all terms have been queried in NCIT, we then see if there are any terms left in the term list. This give us the term “duct dilation” which is queried to see whether there is any match with the second ontology recommended which is SNOMED CT. This returns a synonym match with the concept “dilation of duct”. Thus, the concept “dilation of duct” is reused and the term to concept mapping process is finally complete.

Katsumi and Gruninger (2016) in their work discussed ontology modelling in ontology reuse where they stressed that for an ontology to be considered as being reused from another existing ontology, it should have at least a small fragment of the existing ontology. In determining which model an ontology should follow, Katsumi and Gruninger states that it depends on the strength of the existing ontology. If the existing ontology is weaker than the new ontology that is being developed, then the new ontology may only map some part of the existing ontology’s model. However, if the existing ontology is stronger than the new ontology being developed, then the new ontology may follow the model of the existing ontology.

In the case of several ontology being reused, Katsumi and Gruninger suggested that some part of the new ontology should follow some part of the existing ontology being reused. They however, fail to define what are the criteria that classifies an ontology as stronger or weaker compared to the other ontology. In this research, three ontologies were reused namely NCIT, SNOMED CT and RadLex. Since NCIT has the highest frequency by far (478), compared to the other two (207 and 48 respectively), we assumed that NCIT is stronger than SNOMED CT and RadLex. This has prompted us in following the modelling of NCIT in the development of the Abdominal Ultrasound Ontology (AUO).

When merging concepts reused from SNOMED CT and RadLex into the ontologies reused from NCIT, we would first find suitable parents for the concept. If no such parent exists, the parent and ancestors of the concept will then be reused according to the modelling of NCIT. If no match is found in any of these three ontologies, a new concept will then be created with the help of domain experts.

Figure 8. Term to Concept mapping guide

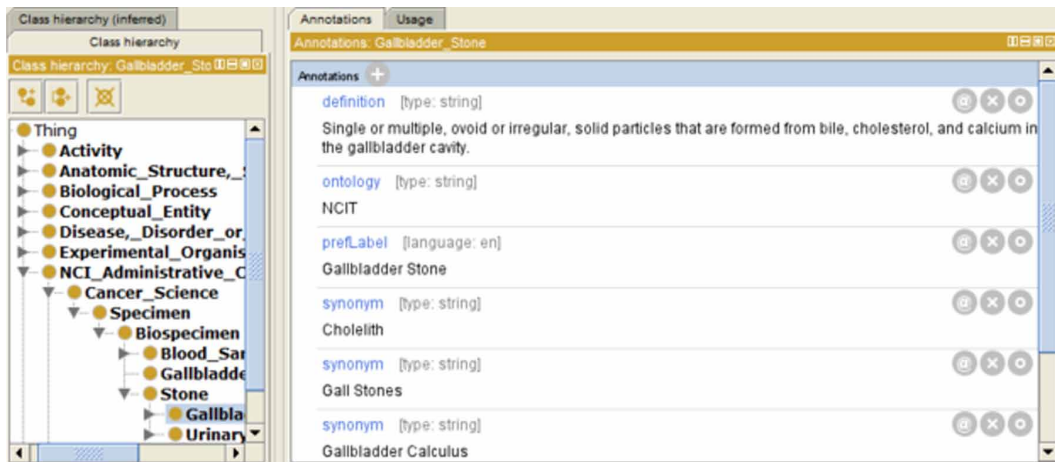
```
Input: 1. A list of terms, T
       2. A list of recommended ontologies, Ontology
1 foreach T[n] do
2   foreach Ontology[j] do
3     if T[n] matched in Ontology[j] then
4       if match is PrefLabel or synonym
5         then
6           | Reuse the concept
7         else
8           | Reuse the partial match concept
9         end
10        if concept have parent or ancestor
11         then
12          | Reuse the parent or ancestor
13        end
14      else
15        | Leave term in T
16      end
17    end
18  end
19 end
```

The new concepts created will be carefully integrated into AUO and will contain the same level of knowledge details as other concepts in AUO. Expertise from the domain experts are needed at this level to give the correct definition and add relevant synonyms to these new concepts. All concepts in AUO were annotated with their original ontology so that it is easier to make any reference in the future if necessary. Figure 9 shows a snapshot of the Abdominal Ultrasound Ontology as displayed in Protégé.

### Ontology Evaluation

The developed AUO was validated by domain experts mainly to check that the relationships between concepts and their definitions are correct. During the evaluation phase, the domain experts together with the ontology developers, went through the whole ontology. Some inconsistencies were highlighted but overall all the experts were happy with the information generated and stored in the ontology. With regards to vocabulary coverage, the experts believe that 92.7% ontology coverage is good enough to cover all the important concepts needed for an abdominal ultrasound reporting system. For the remaining 7.3% of the terms that had no match in the ontology, some of them were caused by human error such as spelling mistakes made by the reporters. There are also several terms that the domain expert believes could be omitted as these words should not be in an ultrasound report for good practice. Examples of such words are “comet tail”, “NAD”, and “hepato petal”. These words might be understood by the radiologist but might make no sense to others (Edwards et al., 2014). The main objective of using this ontology reuse methodology is to achieve as much coverage as possible and reduce the need for domain experts in developing the ontology. If an ontology were to be built from scratch, domain experts will be needed from the very first step in designing the ontology. However, with ontology reuse, domain experts were only needed at the end of the process to verify that everything is correct and to assist in adding few new concepts into the ontology.

Figure 9. Snapshot of the Abdominal Ultrasound Ontology as displayed in Protégé



## RESULT AND DISCUSSION

The evaluation and testing of the methodology were divided into two phases. Phase 1 tries to prove that ontology reuse provides a wider coverage compared to using only one existing ontology. The testing was performed using only 49 sample ultrasound report which was also used as the training data in developing AUO. Phase 2 looks at how well AUO perform when being evaluated with existing and new sample ultrasound reports. It also looks at how difficult it is to include additional concepts when there are new requirements.

### Phase 1: 49 Sample Ultrasound Reports

The coverage of AUO was first tested on a set of 49 sample ultrasound reports with a total of 761 terms in comparison with two existing ontologies which are NCIT and SNOMED CT. A term to concept matching was performed using the 761 terms extracted from the sample ultrasound report corpus and the result are shown in Figure 10. The developed AUO gave the highest number of concept match compared to reusing only one ontology. Between NCIT and SNOMED CT, NCIT has the higher concept match with a total of 151 PrefLabel matches, 79 synonyms matches and 438 partial matches. SNOMED CT on the other hand has only 98 PrefLabel matches, 104 synonyms matches and 431 partial matches. The reason SNOMED CT has lower PrefLabel matches compared to synonyms is because of its naming convention. For example, the preferred label for “kidney” is “kidney structure” and “entire gallbladder” for “gallbladder”. When writing reports, radiologist often use simpler words like “kidney” and “gallbladder” instead of “kidney structure” and “entire gallbladder” thus, when term to concept matching was performed, SNOMED CT returned more synonym matches then PrefLabel.

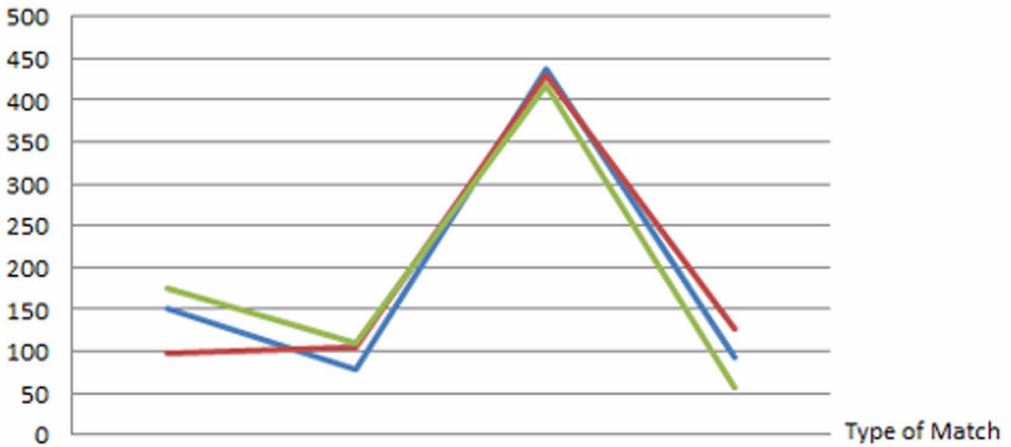
Compared to NCIT and SNOMED CT, AUO returns the highest total matches with 176 PrefLabel matches, 111 synonym matches and 418 partial matches. The reason AUO returns the most number of matches is because the ontology reuse methodology selects the best matches from different ontologies and merge them into the AUO. Its exhaustive mapping in several ontologies based on the ontology rank has ensured that almost all terms in the corpus were covered by AUO. Whenever possible, a PrefLabel match will be inserted in the ontology. If not, a synonym match will be added then only partial matches are included to ensure the ontology has a wide coverage of the corpus.

As expected, and confirmed by the results, it is better to reuse several ontologies then using a single one as this offers a better coverage. Figure 11 shows the percentage of total match and no match in all three ontologies. If ontology reuse was done by mapping the 761 terms against NCIT, there



Figure 10. Breakdown of total match according to type against NCIT, SNOMED CT and Abdominal Ultrasound Ontology (AUO)

Number of Matches

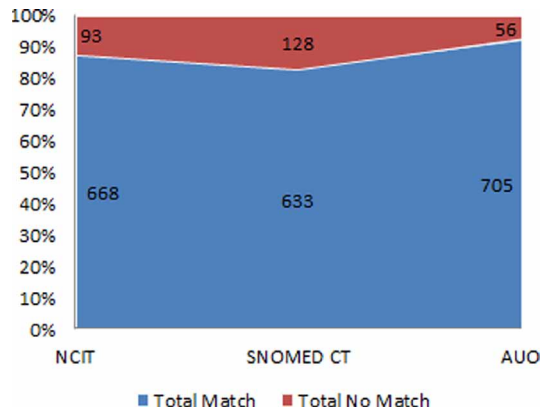


PrefLabel Match	Synonym Match	Partial Match	No Match	
151	79	438	93	— NCIT
98	104	431	128	— SNOMED CT
176	111	418	56	— AUO

will only be an 87.8% of coverage. If the mapping were done against SNOMED CT, the percentage of coverage would be only 83.2% which is lower than NCIT. However, the percentage of coverage increases to 92.6% when several ontologies were reused; which in this case are NCIT, SNOMED CT, and RadLex. The percentage of no match is also very small (7.4%) which means that the AUO covers almost all the terms in the corpus. The reason there is still 7.4% of no match is because there are several terms in the corpus that the domain experts believe are poor usage of terms to describe findings in an ultrasound report. The domain expert believes that this is bad practice and the medical ultrasound experts are now slowly cutting down the usage of such words thus making it irrelevant to be in the AUO. Another reason for the 7.4% of no match is spelling errors made by ultrasound reporters. This is not a concern for now but for future work, we could consider using the ontology to also correct and understand these errors.

NCIT has a total of 118,941 classes while SNOMED CT has 324,129 classes. However, there are only 668 and 633 matches respectively for each NCIT and SNOMED CT regarding abdominal ultrasound terminology. On the other hand, AUO has only 509 classes which is less than 0.5% of either NCIT or SNOMED CT but still managed to have 705 matches which is more than the matches NCIT and SNOMED CT each gets. This is because of the specialization of the ontology. Since the ontology has an intended purpose in an application, it is much better and more efficient to build a domain specific ontology through reuse. It definitely would not be efficient to store a large ontology such as NCIT and SNOMED CT and use only less than 0.6% of it. This is because it would take a lot of storage space and it will also slow down the application since the application will need to go through the whole ontology to find a match. Thus, the better way to develop an ontology-based application is to build a new domain specific ontology through ontology reuse methodology.

Figure 11. Percentage of total match and no match in NCIT, SNOMED CT and AUO



## Phase 2: 100 Sample Ultrasound Reports

Phase 1 of the testing and evaluation process of the ontology reuse methodology has proven that ontology reuse provides a wider coverage compared to using only one existing ontology. Phase 2 looks at how well AUO performs in relation to the term to concept mapping when being evaluated using 100 sample reports; where 49 of the sample reports were used for training purpose and the other 51 were a new sample ultrasound reports which were not used in developing AUO. Phase 2 also looks at how much work is needed to update the ontology to include additional concepts based on the requirements of the new corpus.

Terms extraction which was performed on the 100 sample reports returned 1119 terms which is 358 more than the terms extracted in phase 1. The number of matches for the 761 terms has already been found out in Phase 1. In order to find out the number of matches the rest of the terms got, we checked whether the remaining 358 terms exist in AUO. Figure 12 shows the percentage of each type of matches as well as no match for all 358 terms when compared with AUO.

Overall, there is a quite high percentage of total match which is 72.1% where 237 out of the 358 terms were found as partial match (66.2%), 15 as synonym match (4.2%) and 6 as prefLabel match (1.7%). This shows that the AUO is adequate enough to cover most terms in the biomedical domain since the total of no match is quite small which is 27.9%. In abdominal ultrasound reporting, the content of the reports usually contains words which are more or less similar in semantic even though different terms were used especially in the case of normal ultrasound reports. Only in abnormal ultrasound reports we could find a higher variety of words being used. With the existence of AUO as a knowledge base, the usage of different words in writing ultrasound reports would not have mattered because the semantic of the words are more important than the words themselves and because of the similarity of the contents, a small number of the corpus used as training data still gives quite a good number of total match percentage.

After comparing the 358 terms with AUO, we found that there are 100 terms without a match in AUO. A good methodology should be able to allow for it to adapt to any changes when required. Thus, in order to reduce the percentage of no match, the ontology reuse process was done again on all 358 terms to include new concepts into AUO. Firstly, we need to check whether the addition of 51 new sample reports changes the ontology recommended for reuse. Fortunately, NCIT still comes up as the best ontology to reuse followed by SNOMED CT and RadLex. Next, all 358 terms were then compared with NCIT so that term to concept mapping could be performed. Once the process was completed, it was then repeated with SNOMED CT and RadLex. After the addition of the new concepts and synonyms such as “adnexal masses”, “incidental finding” and “morphology” to AUO, the total number of terms without a match was reduced from 100 to 26, which is a 74% reduction.



Figure 12. Percentage of total prefLabel, synonym, partial and no match for 358 new terms when compared to AUO

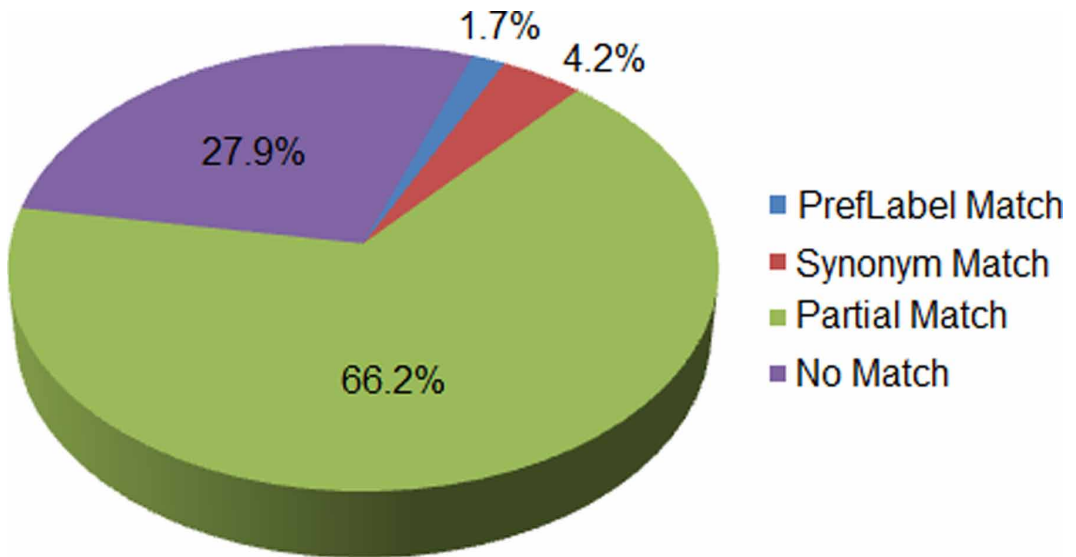


Figure 13 shows the percentage of total match of all three types of matches as well as the percentage of no match for the 1119 terms after new concepts have been added to AUO.

The total match percentage after the addition of new concepts has increased from 72.1% to 92.7% where 692 out of the 1119 terms were found as partial match (61.8%), 139 as synonym match (12.4%) and 206 as prefLabel match (18.4%). Before the addition of the new concepts, the percentage of partial match was higher at 66.2% compared to the current percentage of 61.8%. This number was reduced after the term to concept mapping was done because some of the partial matches were found as either a synonym or prefLabel match in NCIT, SNOMED CT or RadLex. For example, the word “right ovary” was first labelled as partial match since AUO does not have a concept called “right ovary” but it has two concepts which are “right” and “ovary”. However, after performing term to concept mapping again, “right ovary” was found as a prefLabel match in NCIT thus the concept was added to AUO. This causes the percentage of synonym and prefLabel match to also increase from 4.2% and 1.7% to 12.4% and 18.4% respectively. Figure 14 shows the comparison of the percentage of prefLabel, synonym, partial and no match for phase 1, phase 2 (before the extension) and phase 2 (after the extension).

Phase 2 of the testing and evaluation of AUO shows that the number of sample ultrasound reports used as testing data was adequate in ensuring that AUO would be able to cover most abdominal ultrasound sample reports. This also proves that the domain expert’s assumption that the 92.7% total matches of the first version of AUO is indeed enough to cover most abdominal ultrasound sample reports. It also shows that the ontology reuse methodology proposed in this research is adaptable since it is fairly easy for new concepts to be added according to new requirements from additional sample reports.

## CONCLUSION

Ontology reuse can be beneficial in developing domain specific ontologies for application system whereby it reduces development time and redundancy. The lack of proper methodology and tools in reusing ontology has hindered this effort. Thus, this paper proposed a methodology to reuse ontology together with supporting tools that would make the ontology reuse process much easier.

Figure 13. Percentage of total prefLabel, synonym, partial and no match for all 1119 terms when compared to AUO

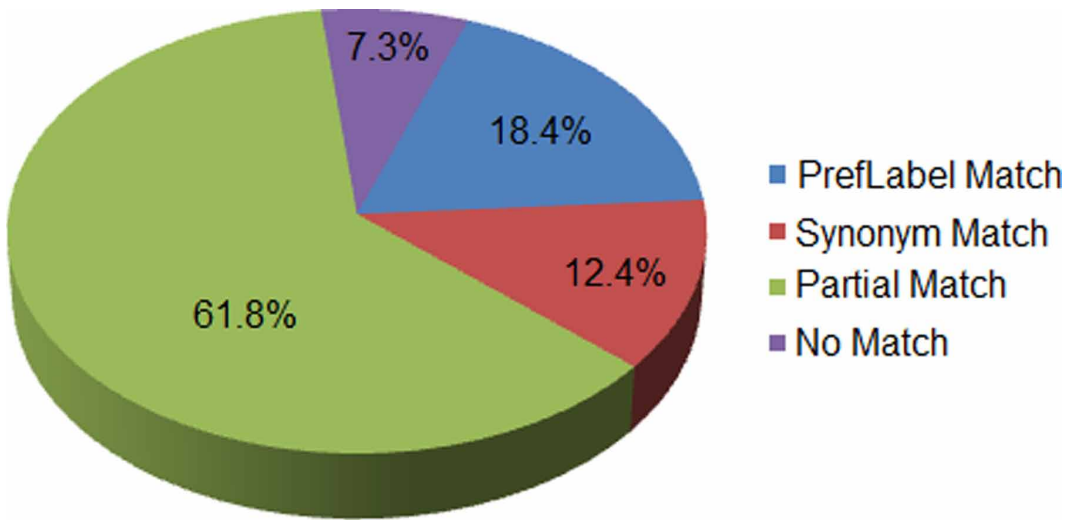
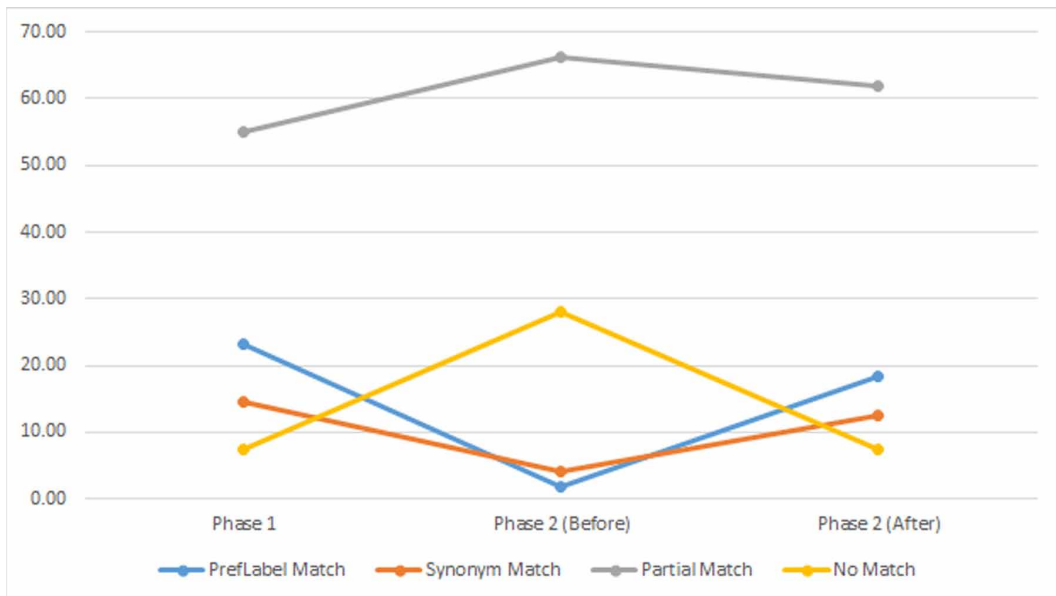


Figure 14. Comparison of the percentage of prefLabel, synonym, partial and no match for phase 1, phase 2 (before the extension) and phase 2 (after the extension)



The development of AUO using this methodology has proven that ontology reuse is beneficial in developing a small domain specific ontology which has wide coverage of the terminology used in the application system compared to using a large general domain ontology. It also proves that the methodology is adaptable as it allows for changes to be made easily when there are new requirements from the corpus. It is hoped that the proposed ontology reuse methodology would encourage more usage of ontology in medical system without the development of similar domain ontologies that would cause redundancy.

## REFERENCES

- Alani, H. (2006). Position paper: Ontology construction from online ontologies. In *Proceedings of the 15th International Conference on World Wide Web, WWW '06*, Edinburgh, Scotland (pp. 491–495). New York, NY: ACM. doi:10.1145/1135777.1135849
- Alani, H., Brewster, C., & Shadbolt, N. (2006). Ranking Ontologies with AKTiveRank, pages 1–15. Springer Berlin Heidelberg, Berlin, Heidelberg. doi:10.1007/11926078\_1
- Benson, T. (2010). *Principles of Health Interoperability HL7 and SNOMED*. London: Springer.
- BioPortal (n.d.). API Documentation. <http://data.bioontology.org/documentation>
- Boland, G. (2007). Enhancing the radiology product: The value of voice-recognition technology. *Clinical Radiology*, 62(11), 1127. doi:10.1016/j.crad.2007.05.014 PMID:17920876
- Bontas, E. P. (2007). *A contextual approach to ontology reuse*. PhD thesis, Freie Universität Berlin.
- Bontas, E. P., Mochol, M., & Tolksdorf, R. (2005). Case studies on ontology reuse. In *Proceedings of the 5th International Conference on Knowledge Management, IKNOW05*. Academic Press.
- Caldarola, E. G., Picariello, A., & Rinaldi, A. M. (2015). An approach to ontology integration for ontology reuse in knowledge based digital ecosystems. In *Proceedings of the 7th International Conference on Management of Computational and Collective intelligence in Digital EcoSystems, MEDES '15*, Caraguatubá, Brazil (pp. 1-8). New York, NY: ACM. doi:10.1145/2857218.2857219
- Capellades, M. A. (1999). Assessment of reusability of ontologies: a practical example. In *Proceedings of AAAI1999 Workshop on Ontology Management* (pp. 74–79). AAAI Press.
- Doran, P., Tamma, V., & Iannone, L. (2007). Ontology module extraction for ontology reuse: An ontology engineering perspective. In *Proceedings of the Sixteenth ACM Conference on Conference on Information and Knowledge Management, CIKM '07*, Lisbon, Portugal (pp. 61-70). New York, NY: ACM. doi:10.1145/1321440.1321451
- Edwards, H., Smith, J., & Weston, M. (2014). What makes a good ultrasound report? *Ultrasound*, 22(1), 57–60. doi:10.1177/1742271X13515216 PMID:27433194
- Fernández-López, M., Gómez-Pérez, A., & Suárez-Figueroa, M. C. (2013). Methodological guidelines for reusing general ontologies. *Data & Knowledge Engineering*, 86, 242–275. doi:10.1016/j.datak.2013.03.006
- Jonquet, C., Musen, M. A., & Shah, N. H. (2010). Building a biomedical ontology recommender web service. *Journal of Biomedical Semantics*, 1(S1), S1. doi:10.1186/2041-1480-1-S1-S1 PMID:20626921
- Kahn, C. E. Jr, Langlotz, C. P., Burnside, E. S., Carrino, J. A., Channin, D. S., Hovsepian, D. M., & Rubin, D. L. (2009). Toward best practices in radiology reporting. *Radiology*, 252(3), 852–856. doi:10.1148/radiol.2523081992 PMID:19717755
- Katsumi, M., & Gruninger, M. (2016). What is ontology reuse? In Ferrario, R. and Kuhn, W., (Eds.), *Formal Ontology in Information Systems: Proceedings of the 9th International Conference, FOIS 2016* (pp. 9-22). IOS Press.
- Lossio-Ventura, J. A., Jonquet, C., Roche, M., & Teisseire, M. (2014). Biotex: A system for biomedical terminology extraction, ranking, and validation. In *Proceedings of the 2014 International Conference on Posters & Demonstrations Track ISWC-PD'14*, Riva del Garda, Italy (pp. 157–160). CEUR-WS.org.
- Maedche, A., Motik, B., Stojanovic, L., Studer, R., & Volz, R. (2003). An infrastructure for searching, reusing and evolving distributed ontologies. In *Proceedings of the 12th International Conference on World Wide Web, WWW '03*, Budapest, Hungary (pp. 439–448). New York, NY: ACM. doi:10.1145/775152.775215
- Martínez-Romero, M., Jonquet, C., Oñáiztegui, A., Connor, M. J., Graybeal, J., Pazos, A., & Musen, M. A. (2017). Nbo ontology recommender 2.0: An enhanced approach for biomedical ontology recommendation. *Journal of Biomedical Semantics*, 8(1), 21. doi:10.1186/s13326-017-0128-y PMID:28592275

- McDaniel, M., Storey, V. C., & Sugumaran, V. (2016). The role of community acceptance in assessing ontology quality. In E. Métais, F. Meziane, M. Saraee et al. (Eds.), *Natural Language Processing and Information Systems: 21st International Conference on Applications of Natural Language to Information Systems, NLDB 2016*, Salford, UK, June 22-24 (pp. 24-36). Springer International Publishing. doi:10.1007/978-3-319-41754-7\_3
- Mejino, J. L. Jr, Rubin, D. L., & Brinkley, J. F. (2008). FMARadLex: An application ontology of radiological anatomy derived from the foundational model of anatomy reference ontology. *AMIA ... Annual Symposium Proceedings - AMIA Symposium*. AMIA. PMID:18999035
- Noy, N. F., Shah, N. H., Whetzel, P. L., Dai, B., Dorf, M., Griffith, N., & Musen, M. A. et al. (2009). Bioportal: Ontologies and integrated data resources at the click of a mouse. *Nucleic Acids Research*, 37, W170–W173. doi:10.1093/nar/gkp440 PMID:19483092
- Ortiz, A. (2011). Polionto: Ontology reuse with automatic text extraction from political documents. In *Proceedings of the 6th Doctoral Symposium in Informatics Engineering* (pp. 309–320). Academic Press.
- Rosse, C., & Mejino, J. L. V. Jr. (2003). A reference ontology for biomedical informatics: The Foundational Model of Anatomy. *Journal of Biomedical Informatics*, 36(6), 478–500. doi:10.1016/j.jbi.2003.11.007 PMID:14759820
- Rubin, D. L. (2008). Creating and curating a terminology for radiology: Ontology modeling and analysis. *Journal of Digital Imaging*, 21(4), 355–362. doi:10.1007/s10278-007-9073-0 PMID:17874267
- Russ, T., Valente, A., MacGregor, R., & Swartout, W. (1999). Practical experiences in trading off ontology usability and reusability. In *Proceedings of the 12th Workshop on Knowledge Acquisition, Modeling and Management, KAW'99* (pp. 16–21). Academic Press.
- Shah, T., Rabhi, F., & Ray, P. (2013). OSHCO: A crossdomain ontology for semantic interoperability across medical and oral health domains. In *Proceedings of the IEEE 15th International Conference on e-Health Networking, Applications Services, HEALTHCOM'13, Lisbon, Portugal*, October 9-12 (pp. 460–464). IEEE. doi:10.1109/HealthCom.2013.6720720
- Shah, T., Rabhi, F., Ray, P., & Taylor, K. (2014). A guiding framework for ontology reuse in the biomedical domain. In *Proceedings of the 47th Hawaii International Conference on System Sciences, HICSS 2014* (pp. 2878–2887). IEEE. doi:10.1109/HICSS.2014.360
- Simperl, E. (2009). Reusing ontologies on the semantic web: A feasibility study. *Data & Knowledge Engineering*, 68(10), 905–925. doi:10.1016/j.datak.2009.02.002
- Trokanas, N., & Cecelja, F. (2016). Ontology evaluation for reuse in the domain of process systems engineering. *Computers & Chemical Engineering*, 85, 177–187. doi:10.1016/j.compchemeng.2015.12.003
- Uschold, M., Clark, P., Healy, M., Williamson, K., & Woods, S. (1998). An experiment in ontology reuse. In *Proceedings of the 11th Knowledge Acquisition Workshop, KAW'98*. Academic Press.
- Zulkarnain, N. Z., Crofts, G., & Meziane, F. (2015a). An architecture to support ultrasound report generation and standardisation. In *Proceedings of the 8th International Conference on Health Informatics, HEALTHINF 2015*, Lisbon, Portugal (pp. 508–513). SCITEPRESS. doi:10.5220/0005252505080513
- Zulkarnain, N. Z., Crofts, G., & Meziane, F. (2015b). A proposed model for the standardisation of ultrasound report. *Poster presented at the UK Radiological Congress, ACC Liverpool, UK*. Academic Press.
- Zulkarnain, N. Z., Crofts, G., & Meziane, F. (2016a). Automatic classification of abdominal ultrasound findings as normal, abnormal or inconclusive. *Poster presented at the British Medical Ultrasound Annual Scientific Meeting & Exhibition, York, UK*. Academic Press.
- Zulkarnain, N. Z., Meziane, F., & Crofts, G. (2016b). A methodology for biomedical ontology reuse. In E. Métais, F. Meziane, M. Saraee et al. (Eds.), *Natural Language Processing and Information Systems: 21st International Conference on Applications of Natural Language to Information Systems, NLDB 2016*, Salford, UK, June 22-24 (pp. 3-14). Springer International Publishing.

## ENDNOTES

1. <http://bioportal.bioontology.org>
2. <http://www.nactem.ac.uk/software/termin>
3. <http://tubo.lirmm.fr/biotex/>

*Nur Zareen Zulkarnain received her BTech. (Hons) degree from Universiti Teknologi PETRONAS, Malaysia in 2011, an MSc. from The University of Manchester, UK in 2012 and a PhD. in Computer Science from the University of Salford, UK in 2017. Currently, she is a Senior Lecturer in the Department of Intelligent Computing & Analytics in Universiti Teknikal Malaysia Melaka where her teaching and research mainly involves artificial intelligence in the area of informatics and natural language processing.*

*Farid Meziane holds a chair in Data and Knowledge Engineering and a PhD in Computer Science from the University of Salford and is the head of the Informatics Research Centre. He is the co-Chair of the International Conference on Application of Natural Language to Information Systems (NLDB) and has served on the programme committees of over 20 conferences. He is on the editorial board of 5 international journals. His research interests are in the broad area of data and knowledge engineering. This includes data mining, information extraction and retrieval, Big Data, and the semantic web.*