Utilising Semantic Technologies for Intelligent Indexing and Retrieval of Digital Images

Taha Osman[1], Dhavalkumar Thakker[2] and Gerald Schaefer[3]
[1] *School of Computing & Informatics, Nottingham Trent University, Nottingham, U.K.*
taha.osman@ntu.ac.uk
[2] *School of Computing, University of Leeds, Leeds, U.K.*
d.thakker@leeds.ac.uk
[3] *Department of Computer Science, Loughborough University, Loughborough, U.K.*
gerald.schaefer@ieee.org

Abstract

The proliferation of digital media has led to a huge interest in classifying and indexing media objects for generic search and usage. In particular, we are witnessing colossal growth in digital image repositories that are difficult to navigate using free-text search mechanisms, which often return inaccurate matches as they in principle rely on statistical analysis of query keyword recurrence in the image annotation or surrounding text. In this paper we present a semantically-enabled image annotation and retrieval engine that is designed to satisfy the requirements of the commercial image collections market in terms of both accuracy and efficiency of the retrieval process. Our search engine relies on methodically structured ontologies for image annotation, thus allowing for more intelligent reasoning about the image content and subsequently obtaining a more accurate set of results and a richer set of alternatives matchmaking the original query. We also show how our well-analysed and designed domain ontology contributes to the implicit expansion of user queries as well as the exploitation of lexical databases for explicit semantic-based query expansion.

Keywords: semantic matching, query expansion, image retrieval, ontology engineering, keyword search, knowledge management

1. Introduction

Affordable access to digital technology and advances in Internet communications have contributed to the unprecedented growth of digital media repositories (audio, images, and video). Retrieving relevant media from these ever-increasing repositories is an impossible task for the user without the aid of search tools. Whether we are considering public media repositories such as Google™ images and YouTube™ (Yee et al. 2003) or commercial photo-libraries such as PA Images™ [1], some kind of search engine is required to matchmake the user-query and the available media. This research effort focuses on image/photo retrieval techniques.

Most caption-based public image retrieval engines rely on analysing the text accompanying the image to matchmake it with the user query. Various optimisations were developed including the use of weighting systems where for instance higher regard can be given to the proximity of the keyword to the image location, or advanced text analysis techniques that use term weighting method, which relies on the proximity between the anchor to an image and each word in an HTML file (Fuji and Ishikawa 2005). Similar relevance-analysis and query expansion techniques (Jeon et al. 2003) are used in annotation-enriched image collections, where usually a labour-intensive annotation process is utilised to describe the images with or without the aid of some domain-specific schema (Hare et al. 2006).

Despite the optimisation efforts, these search techniques remain hampered by the fact that they rely on free-text search that, while cost-effective to perform, can return irrelevant results as it primarily relies on the recurrence of exact words in the image caption or the text surrounding the image in an html page. The inaccuracy of the results increases with the complexity of the query. For instance, while performing this research we used the Yahoo™ search engine to look for images of the football player Zico. The search engine returned some good pictures of the player, mixed with photos of cute dogs (as apparently Zico is also

[1] http://www.pressassociation.com/Images/

a popular name for pet dogs). However, when we added the action of scoring to the search text, this seemed to completely confuse the Yahoo search engine and only one picture of Zico was returned, and in that one he was standing still.

Any significant contribution to the accuracy of matchmaking results can only be achieved only if the search engine can "comprehend" the meaning of the data that describes the stored images, for instance, if the search engine can understand that scoring is an act associated with sport activities performed by humans. Semantic annotation techniques have gained wide popularity in associating plain data with "structured" concepts that software programs can reason about (Wang et al. 2006). This effort presents a comprehensive semantic-based solution to image annotation and retrieval as well as deploying query expansion techniques for improving the recall rate. We claim that shrewd analysis of the application domain characteristics, coupled with a subsequently well-designed ontology can significantly contribute to the user query expansion process via direct term replacement or by modifying the domain taxonomy we build for the query. We also present our research into using lexical databases to analyse free-entry queries in our effort to make them compatible with the requirements of our semantic search engine.

The paper begins with an overview of the Semantic web technologies. In Section 3, we review the case study that was the motivation for this work. Sections 4, 5, 6, and 7 detail the implementation roadmap of our semantic-based retrieval system, i.e. ontology engineering, annotation, retrieval, and query expansion. We present our conclusions and plans for further work in Section 8.

2. Overview of the semantic web

2.1. Ontologies (domain conceptualisation)

The fundamental premise of the semantic web is to extend the Web's current human-oriented interface to a format that is comprehensible to software programmes. Naturally this requires a standardised and rich knowledge representation scheme or *Ontology*.

One of the most comprehensive definitions of ontologies is that expressed in (Hare et al. 2006): “Ontology is a *shared conceptualisation of a domain* and typically consists of comprehensive set of concept classes, relationships between them, and instance information showing how the classes are populated in the application domain.” This comprehensive representation of knowledge from a particular domain allows reasoning software to make sense of domain-related entities (images, documents, services, etc.) and aid in the process of their retrieval and use.

2.2. Caption-based semantic annotation

Applied to image retrieval, the semantic annotation of images allows retrieval engines to make more intelligent decisions about the relevance of the image to a particular user query, especially for complex queries. For instance to retrieve images of a ‘pleased’ Wayne Rooney (football star), it is natural to type the keywords ‘Wayne Rooney Happy’ into the Google™ Image Search engine. However, at the time of the experiment, in the first 100 returned images, Wayne Rooney looked happy in only 70% of them, and the search also returned completely irrelevant images such as that of the Ghanaian striker Asamoah Gyan.

The use of Semantic technologies can significantly improve the computer’s understanding of the image objects and their interactions by providing a machine-understandable conceptualisation of the various domains that the image represents. This conceptualisation integrates concepts and inter-entity relations from different domains, such as Sport, People and Emotions relation to the query above (Shadbolt et al. 2006), thus allowing the search engine to infer that Wayne Rooney is a *person* and thus likely to express *emotions* that can be *positive* such as *happiness* and that he is also an *English footballer* playing for Manchester United FC.

2.3. Content-based semantic annotation

The success of caption-based semantic image retrieval largely depends on the quality of the semantic caption (annotation) itself. However, the caption is not always available largely because image annotation is a labour intensive process. In such situations, image recognition techniques are applied, which is better known as content-based retrieval. However, the best content-based techniques deliver only partial success as image recognition is an extremely complex problem (Wu and Yap 2006) (Julia and Schiele 2007), especially in the absence of accompanying text that can aid inferring in the relationship between the recognized objects in the image (Lam and Singh 2006). Highly successful image retrieval techniques such as face recognition rely on a high-quality predefined training set, which might not be available or might not be feasible to compile in a fast response-time setting. Moreover, from a query composition point of view, it is much easier to use a textual interface rather than a visual interface (by providing a sample image or a sketch) (Ying et al. 2007).

3. Case study for semantic-based image retrieval

An opportunity to experiment with our research findings in semantic-based search technology was gratefully provided by PA Images™ . PA Images is a Nottingham-based company which is part of the Press Association Photo Group Company. As well as owning a huge image database in excess of 4 million annotated images that date back to the early 1900s, the company processes a colossal amount of images each day from varying events ranging from sport to politics and entertainment. The company also receives annotated images from a number of partners that rely on a different photo-indexing schema.

More significantly, initial investigation has proven that the accuracy of the results sets that match the user queries do not measure up to the rich repository of photos in the company's library.

The goal of the case study is two-fold. Initially, we intend to investigate the use of semantic technology to build a classification and indexing system that critically unifies the annotation infrastructure for all the sources of incoming stream of photos. Subsequently, we will conduct a feasibility study aiming to improve the end-user experience of their images search engine. At the moment PA Images search engine relies on Free-Text search to return a set of images matching the user requests. Therefore the returned results naturally can go off-tangent if the search keywords do not exactly recur in the photo annotations. A significant improvement can result from semantically enabling the photo search engine. Semantic-based image search will ultimately enable the search engine software to understand the "concept" or "meaning" of the user request and hence return more accurate results (images) and a richer set of alternatives.

It is important here to comment about the dynamics of the retrieval process for this case study as it represents an important and wide-spread application domain where there is a commercial opportunity for exploiting semantic technologies:

1. The images in the repository have not been extracted from the web. Consequently, the extensive research into using the surrounding text and information in the HTML document in improving the quality of the annotation such as in (Wang et al. 2006) (Lam and Singh 2006) is irrelevant.
2. A significant sector of this market relies on fast relay of images to customers. Consequently, this confines advanced but time-consuming image analysis techniques (Wang et al. 2006) to *off-line* aid with the annotation of caption-poor images.
3. The usually colossal amount of legacy images annotated to particular (non-semantic) schema necessitates the integration of these heterogeneous schemas into any new, semantically-enabled and more comprehensive ontologies.

4. Ontology development

4.1. Domain Analysis

Domain analysis is the first and most critical phase in ontology development as it involves interfacing to the domain experts to elicit the knowledge prior to ontology compilation. Our domain analysis started from an advanced point as we had access to the photo agency's current classification system. Hence, we adopted a top-down approach to ontology construction that starts by integrating the existing classification with published evidence of more inclusive public taxonomies[2]. At the upper level, two ontological trees were identified; the first captures knowledge about the event (objects and their relationships) in the image, and the second is a simple upper class that characterises the image attributes (frame, size, creation date, etc.), which is extensible in view of future utilisation of content-based recognition techniques.

A bottom-up approach was used to populate the lower tiers of the ontology class structure by examining the free-text and non-semantic caption accompanying a sample set of sport images. Domain terms were acquired from approximately 65,000 image captions. The terms were purged of redundancies and verified against publicly available related taxonomies such as the media classification taxonomy detailed in (Maedche et al. 2003). An added benefit of this approach is that it allows existing annotations to be seamlessly parsed and integrated into the semantic annotation.

Wherever advantageous, we integrated external ontologies into our knowledge representation. However, more often than not, it is not feasible to consume comprehensive external ontologies that conceptualise a particular domain or a sub-domain. For instance, the unsupervised semantic tagging research in (Maedche et al. 2003) consumes a fairly comprehensive ontology for the soccer domain, and although the majority of concepts fulfil our domain analysis requirements, there is also a significant amount of concepts, primarily aimed at content-based retrieval, that are redundant to our requirements. Hence, taking into consideration that the size of the ontology is critical to the search engine performance, particularly in a fast-response setup, we opted to construct our ontology from scratch. This allows us to fine-tune the concepts and relations to our requirements and more accurately perform the ontology engineering tasks necessary for creating a robust ontology such as normalisation and coverage optimisation. Nevertheless, we did also integrate more generic external ontologies that do not compromise our ontology robustness.

---

[2] http://www.aktors.org/ontology/support#

Bearing in mind the responsiveness requirements of on-line retrieval applications, we applied caching methods to localise the access to reduce its time overhead. Figure 1 represents a subset of our ontology.

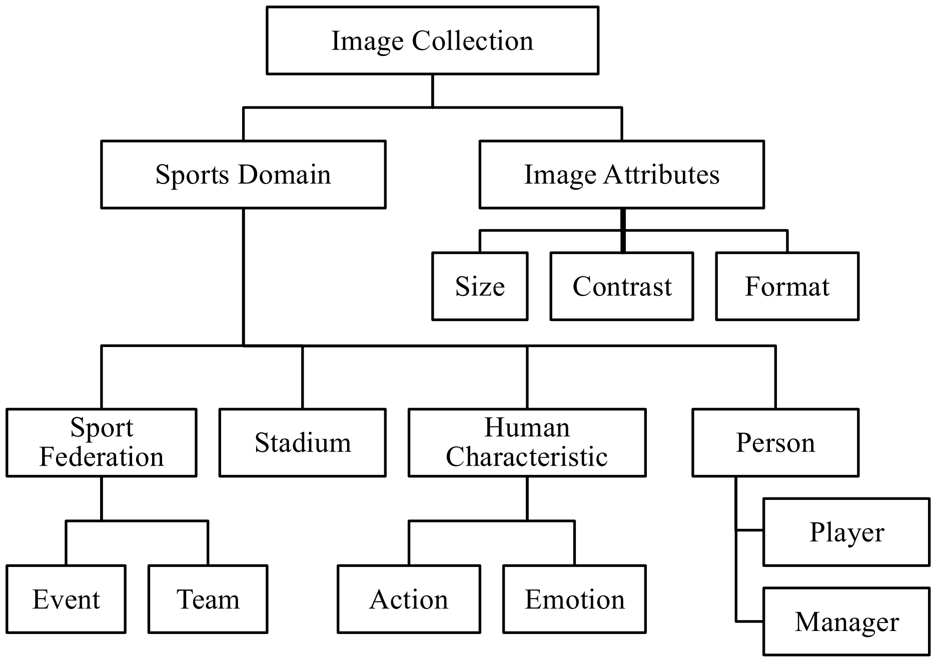

Figure 1 Subset of the ontology tree

4.2. Categorising Data-type and Object properties

All semantic models use two types of properties to build relationships between individuals (classes), Data-type properties, and Object properties. When assigning properties to a class, all its sub-classes inherit their parent class properties.

Deciding on the appropriate type of property to use is not a trivial task. Whereas object properties link individuals of different classes together, data-type properties can only point to immediate values (e.g. text strings), which are meaningless to a reasoning software, except for performing a string-based search. For instance, allocating data-type properties to the person class in order to give each new instance a last name is a correct use of data-type properties because they cannot be reused by another individual. On the other hand, object properties are required to assign someone a nationality since a country is an autonomous object that can have properties such as currency, capital city, language, etc. Hence, a country needs thus to be an instance that can be reused from an already existing ontology.

4.3. Consistency Checking

Unlike database structures, ontologies represent knowledge not data, hence any structural problems will have detrimental effect on their corresponding reasoning agents especially that ontologies are open and distributed by nature, which might cause wide-spread propagation of any inconsistencies (Rector 2003). For instance, in traditional structuring of methodologies, usually the *part-of* relationship is adopted to express relationships between interdependent concepts. So, for players that are *part-of* a team performing in a particular event, Figure 2 below illustrates the commonly taken approach:

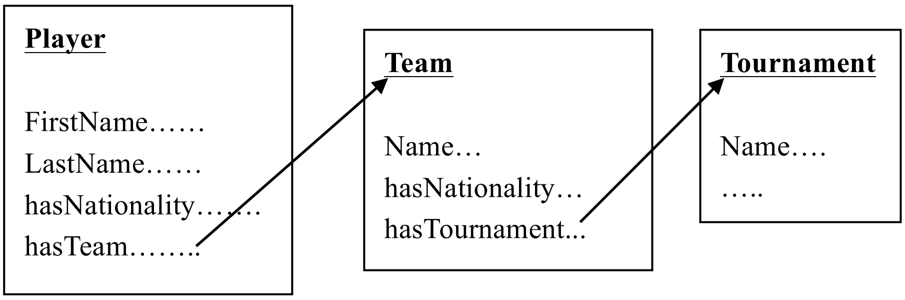

Figure 2 Traditional part-of relationships

However logical the above description appears at first sight, further analysis reveals inconsistency problems. When a player plays for two different teams at the same time (e.g. his club and his national team) or changes clubs every year, it is almost impossible to determine which team the player plays for. Hence, the order of definition (relationship direction) should always be the reversal sequence of the *part-of* relationship as redesigned in figure 3 below:

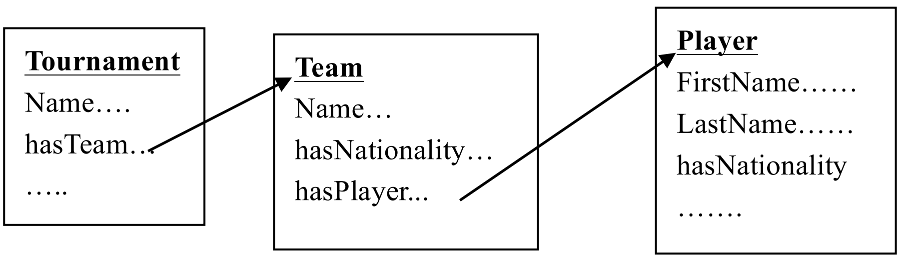

Figure 3 Re-organisation of the player classification

4.4. Coverage

Although consistent, the structural solution in Figure 3 is incomplete as the players' membership is temporal. The same problem occurs with tournaments as from one year to another, teams taking part in a tournament change. This problem can be solved by adding a start and end date for the tournament (see Figure 4 below), rather than by engineering more complex object property solutions. Hence, as far as the semantic reasoner is concerned, the "FIFA World Cup 2002" is a different instance from "FIFA World Cup 2006". The same reasoning can be applied to the class team, as players can change team every season. These considerations, although basic for a human reasoning, need to be explicitly defined in the ontology.

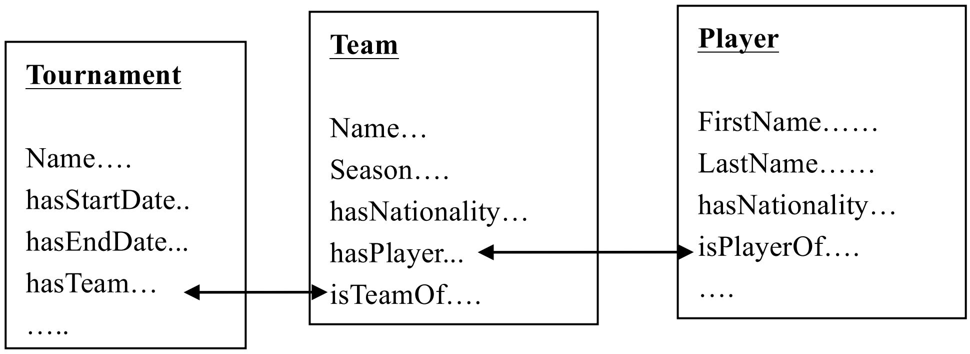

Figure 4 Resolving Coverage problems in ontology

4.5. Normalisation: reducing the redundancy

The objective of normalisation is to reduce redundancy. In ontology design, redundancy is often caused by temporal characteristic that can generate redundant information and negatively affect the performance of the reasoning process.

Direct adoption of the ontology description in Figure 4 above will result in creating new team each season, which is rather inefficient as the team should be a non-temporal class regardless of the varying player's membership or tournament participation every season. Hence, Arsenal or Glasgow Rangers Football clubs need to remain abstract entities. Our approach was to introduce an intermediary temporal *membership* concept that servers as an indispensable link between teams and players, as well as between teams and tournaments as illustrated in Figure 5.

The temporal instances from the Membership class link instances from two perpetual classes as follows:

- memberEntity links to a person (Player, Manager, Supporter, Photographer, etc.)
- isMemberOf refers to the organisation (Club, Press Association, Company, etc.)
- fromPeriod and toPeriod depict membership temporal properties

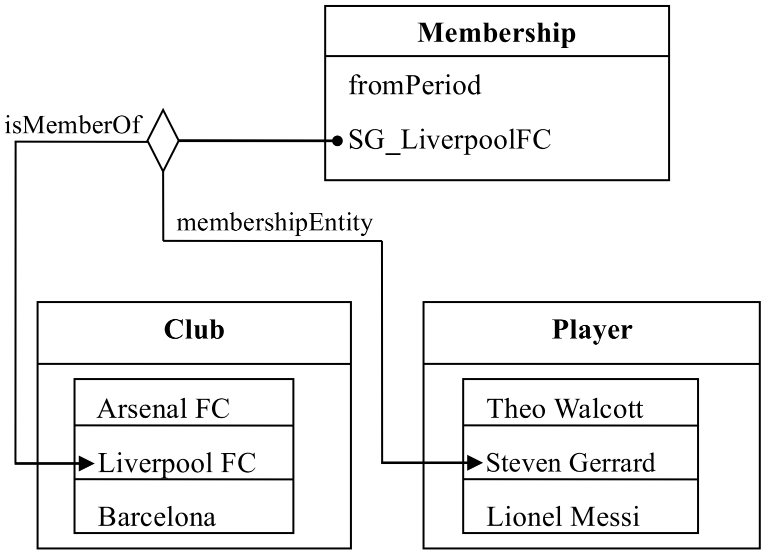

Figure 5 Membership class in the final ontology

5. Image Annotation

The Protégé® ontology editor that was utilised to construct the sport domain ontology. Protégé uses frame-based knowledge representation (Noy and Tu 2003) and adopts OWL as the ontology language. The Web Ontology Language (OWL) (W3C 2004) has become the de-facto standard for expressing ontologies, it adds extensive vocabulary to describe properties and classes and express relations between them (such as disjointness), cardinality (for example, "exactly one"), equality, richer typing of properties, and characteristics of properties (such as symmetry). The Jena (Carroll et al. 2004) Java API was used to build the annotation portal to the constructed ontology.

The central component of the annotation are the images stored (as OWL descriptions) in image library as illustrated in

Figure 6. Each image comprises an object, whose main features are stored within an independent object library. Similarly are the object characteristics, event location, etc. distinct from the image library. This highly modular annotation model facilitates the reuse of semantic information and reducing redundancy.

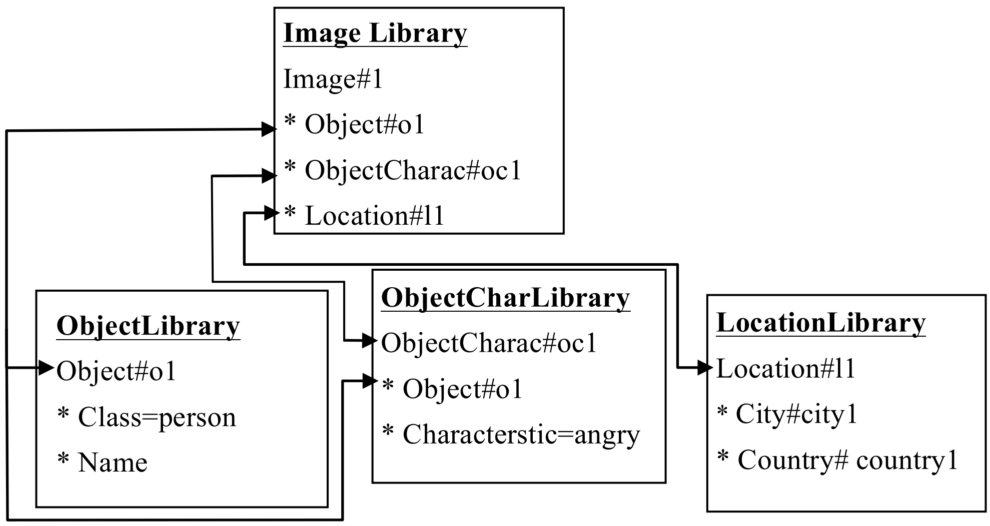

Figure 6 Architecture of the annotation

Taking into account the dynamic motion nature of the sport domain, our research concluded that a variation of the sentence structure suggested in (Carroll et al. 2004) is best suited to design our annotation template. We opted for an "*Actor – Action/Emotion – Object"* structure that will allow the natural annotation of motion or emotion-type relationships without the need to involve NLP techniques (Park and Li 2007). For instance, "Lampard – Smiles – null", or "Mikel – Tackles – Bale". An added benefit of the structure is that it simplifies the task of the reasoner in matching *actor* and *action* annotations with entities that have similar characteristics.

6. Image Retrieval

The image retrieval user interface is illustrated in Figure 7. The search query can include sentence-based relational terms (Actor- Action/Emotion -Object) and/or key domain terms (such as tournament and team). In case multiple terms were selected for the query, the user needs to specify which term represents the *main* search preference (criterion).

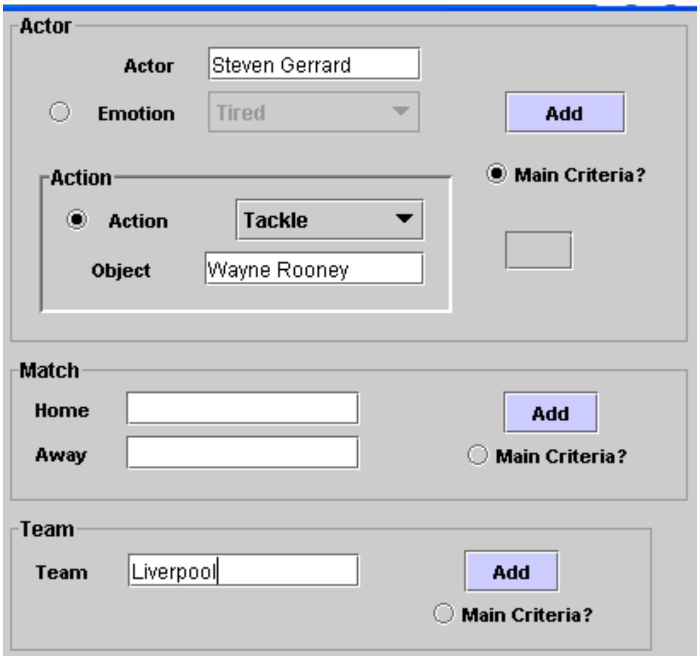

Figure 7 Snapshot of the retrieval interface

For instance, in Figure 7 the relational term (Gerrard Tackles Rooney) is the *primary* search term and team Liverpool is the *secondary* search term. The preference setting is used to improve the ranking of retrieved images.

Figure 8 gives a high level view of the annotation and retrieval mechanism. The semantic description generator allows the annotator to transparently annotate new images and also transforms the user query into OWL format. The semantic reasoning engine applies our matchmaking algorithm at two phases: The first retrieves images with annotations matching all concepts in the query; in the second phase further matchmaking is performed to improve the ranking of the retrieved images in response to user preferences.

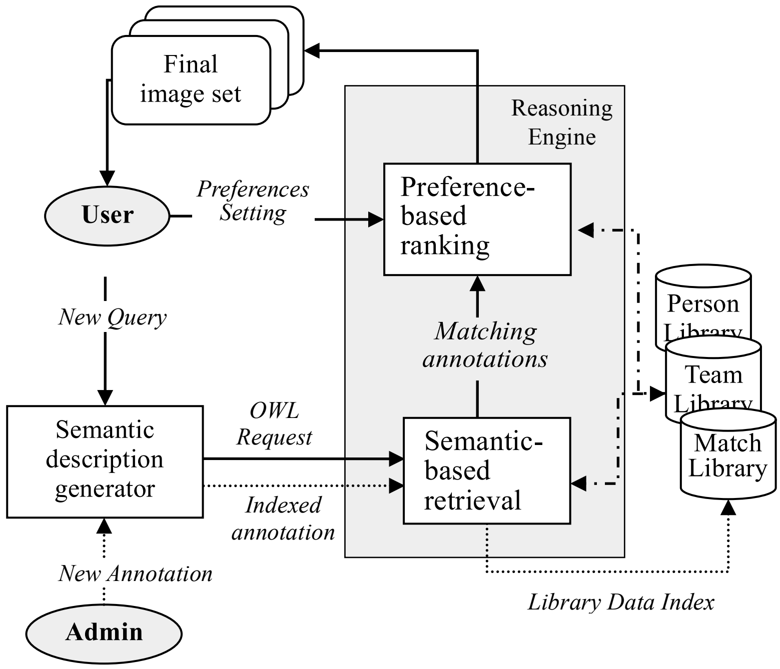

Figure 8 Schematic diagram of the Semantic Web Image Retrieval software

Our reasoning engine uses a nearest neighbour strategy for matchmaking (Osman et al. 2006) (Ching-man Au 2008) to serve both the semantic retrieval and the ranking phases. Our algorithm continues traversing back to the upper class of the ontology and matching instances until there are no super classes in the class hierarchy, i.e. the leaf node for the tree is reached, giving degree of match equal to 0. The degree of match (*DoM*) is calculated according to Equation 1:

$$ADoM = \sum_{i=1}^{n} W_i \frac{MN_i}{GN_i}, \forall W_i \in [0,1] \qquad \textbf{Equation 1}$$

Where the *MN* is the total number of matching nodes in the selected traversal path, and *GN* the total number of nodes in the selected traversal path for a particular matching criterion (*i*). Each criterion is scaled with the importance factor *W* according to the user preferences.

The example below illustrates the operation of the algorithm for a single criterion only where the query is: Object–hasCharacteristic-happy, and image1 and image2 are annotated with Object-hasCharacteristic-happy and Object-hasCharacteristic-smile respectively, the *DoM* for image1 is 1 as the instances match to the level of the leaf node (see Figure 9). However, for image2 instances match to the level of Positive Emotion- Mild class and is one layer lower than the leaf node giving *DoM* = 0.5.

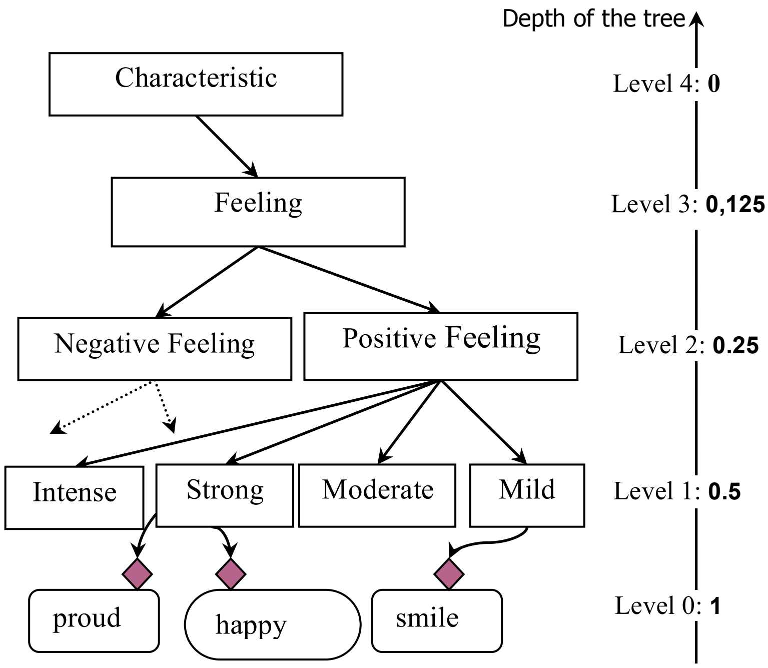

Figure 9 Traversing the Ontology Tree

Although not demonstrated in the above example, the importance factor ***W*** can be used to scale the criticality of the *emotion* concept against other similarity factors such as the player's *name* or *team* for instance.

7. Semantic Web based Query Expansion to achieve better precision and recall

Lately query expansion (QE) techniques have gained a lot of attention in attempting to improve the recall of document and media queries. Query expansion is traditionally considered as a process of supplementing a query with additional terms as the assumption is that the initial query as provided by the user may be an inadequate representation of the user's information needs (Ching-man Au 2008).

QE methods fit naturally into our image retrieval technology as we rely on computing the aggregate degree of match (ADoM) for the semantic relations describing a particular image to determine its match to the original query. Hence, we can easily determine the quality of the retuned results in terms of accuracy and volume and decide whether to apply QE techniques to replace or improve the query concepts to improve the quality of the recall. This is particularly feasible for semantic-based knowledge bases as they provide language expressiveness for specifying the similarity of the concepts (Implicit and Explicit) at different granularity.

Query expansion techniques can be broadly classified into two categories: the first category uses statistical and probabilistic methods (Ching-man Au 2008) to extract frequently occurring terms from successfully recalled documents and image annotations. These terms are then used to expand the keyword set of similar future queries. The Main shortcoming of the statistics-based QE techniques is that they are as good as the statistics they rely on and have similar disadvantages as free-text based search engines in that they lack structure and are difficult to generalise or to reuse for other domains. The second category (Haubold et al. 2006) utilises lexical databases to expand user queries. A lexical database similar to WordNet (Haubold et al. 2006) is employed, in which language nouns, verbs, adjectives and adverbs are organised into synonym sets that can potentially replace or expand the original query concepts. However, when independently deployed, lexical database lack the semantic conceptualisation necessary to interrelate concepts in complex queries and render them comprehensible to search engines.

Semantic relations-based QE technique expands the query with related concepts rather than simple terms. The reminder of this section discusses the design of both the implicit and explicit elements of our semantic-based QE algorithm.

7.1. Implicit query expansion

Taking into account the domain knowledge hardwired into the ontologies, our implicit query expansion mechanism can be considered as a by-product of a well-researched and designed domain ontology.

The “Actor-Action/Emotion-Object” semantic format allows to naturally employing the ontology to find related terms via simple equivalence relations as that of equating the action of *smiling* to the emotion of *happiness*. Taking into account the limited vocabulary of the *sport* domain, in consultation with the domain experts, we decided against the automatic expansion of directly related terms from a lexical public database such as WordNet. Our initial experiments have shown that while that expansion improved image recall, the accuracy of returned results suffered significantly particularly for complex queries where partial replacement of terms might invalidate the semantics of the query.

Using our ontology structure we are also able to expand queries implicitly by analysing more complex relations as in inferring that *Liverpool* is a possible replacement for *Chelsea* as both are *Teams* playing *Football* sport in the *Premier League* in *England*. Moreover, we are able to scale the relatedness of each term in the query tree according to the importance weighting set by the user/domain manager as explained in the pervious section. The implicit query expansion algorithm is implemented in three consecutive steps as detailed in the following subsections.

7.1.1 Computing the relatedness of the expansion terms

Step1: If query has concept $\boldsymbol{C^p}$ *as the primary search concept and* $\boldsymbol{C^s}$ *as the secondary search concept provided by the searcher then we define query expansion on* $\boldsymbol{C^p}$ *as follows:*

Let’s say $\mathbf{C^{p'}}$ is the alternative concept, $\boldsymbol{\delta}$ is the distance between $\mathbf{C^p}$ and $\mathbf{C^{p'}}$ concepts and $\boldsymbol{\Psi}$ is the expected distance between these two concepts implying them related, the expansion function is:

$$\sum_{i=1}^{n}(C^{p} \xrightarrow{\delta_i,\Psi_i} C_i^{'p}) \;, \delta_i \leq \Psi_i \qquad \textbf{Equation 2}$$

The equation implies that concepts $C_i^{p'}$ are related to C^p if they are at acceptable distance from C^p, for instance, if the acceptable distance within a semantic graph is $\Psi = 2$, then for an alternative (neighbouring) concept to be accepted for query expansion, it cannot be more than 2 nodes away ($\delta <= \Psi$) within the semantic graph.

7.1.2 Incorporating semantic properties in implicit query expansion

A major concern in QE techniques is the formalization of relatedness between two concepts in order to select an optimal set of alternatives. For the benefit of the discussion, we feel it is necessary to revisit the following components of Semantic web formalism and their representation in the OWL ontology language:

Taxonomy Relationships (TR): Taxonomy is the concepts classification system facilitated by Semantic Web. *Class* and *Individual* are the two main elements of this structure where a class is simply a name and collection of properties that describe a set of individuals. Examples of relationships between concepts at the taxonomy level are c*lass*, *subclass*, *superclass*, *equivalent class*, i*ndividual*, *sameAs*, *oneOf*, *disjointWith*, *differentFrom*, *AllDifferent*.

Rules based relationships (RR): Semantic Web Rule Language (SWRL) defines rule based semantics using subset of OWL with the sublanguages of Rule Mark-up Language. SWRL extends OWL with horn-like First Order Logic rules to extend the language expressivity of OWL.

We use this relationship formalism to identify explicit and implicit relatedness of concepts. To evaluate implicit relationships we use *subsumption* and *classification* to perform semantic tree traversal and compare the concepts with respect to the semantic network tree as detailed in our image retrieval algorithm earlier. Contrarily, explicit relationship between two concepts always has a Degree of Match (DoM) of 0 or 1 as they explicitly equate or distinct two individuals. For example the *owl:sameAs* equates two individuals to unify two distinct ontology elements while *owl:differentFrom* has exact opposite effect where it makes individuals mutually distinct.

If the taxonomy and rule based implicit and explicit relationship results in **n** number of equivalent concepts represented by $\{C_1, C_2, C_3, \ldots\ldots C_n\}$ or $C^{p'}$, then to calculate *DoM* for these likely replacement concepts we employ another semantic web relationship formalism, which we will refer as property based relationship.

Property Relationships (PR): Properties can be used to state relationships between individuals or from individuals to data values. These relationships are achieved through the data or object type properties. (i.e. *hasTeam*, *hasTournament*, *isMemberOf*)

Step2: Assuming Query preference concept C^p has properties R_i which has value instances I_i^R and the annotation matching the alternative concept $C^{p'}$ has properties R'_i and the value instances $I_i^{R'}$, then we can compare I_i^R and $I_i^{R'}$ semantically using the aggregate degree of match (ADOM) expressed in Equation 1.

7.1.3 Using implicit query expansion to find alternative terms

In this section we illustrate how our QE algorithm works by discussing the following case. If a user is searching for pictures with *England Team* possibly in the *2006 FIFA World Cup tournament*, the system treats *England Team* as user's primary search criterion and *2006 FIFA World Cup Tournament* as secondary search criterion in the query.

Without expanding the query, the retrieval algorithm returns zero results if there are no images annotated with *Team England* (see Table 1). The following section explains the process of expanding query under these circumstances using our algorithm.

England Team ***(C^p)***	
(C^p has properties R_i)	*I_i^R (properties value)*
Has Nationality	*Country (England)*
Has Sport	*Sport (Football)*
IsWinnerOf	*Tournament(Fifawc66)*
hasNationalTeamTournament	*Fifawc66, 70, ...*

Table 1 Preference Concept

In our sports domain ontology implicit subsumption relationship is applied to find relevant primary concepts. For instance, to find alternative terms for *Team England*, the reasoner first retrieves siblings of the

National Team such as *Team Brazil*, *Team Spain*, and then less adjacent siblings of the *Team* instances such as *Team Chelsea* and *Team Barcelona*.

In the following step we compare the relationship as defined in step 2 as illustrated in the Table 2 below:

	Query	Team Brazil		Team Chelsea	
hasNationality	England	Brazil	0	England	1
hasSport	Football	Football	1	Football	1
isWinnerOf	Fifawc 06	Fifawc70	0.5	Prem. 06	0
hasNational TeamTournament	Fifawc 66, 70, …	Fifawc 66,70, …	1	Prem. 93, 94, …	0
DoM		Brazil	2.5	Chelsea	2

Table 2 Comparing relationship

Step3: If the ranked images in stage 2 are $\{X_1, X_2, X_3\ldots\}$, C^s *is the secondary search term in the query provided by the searcher, these ranked images have* C^s *present in their annotation* C^s_X *, then repeat step 2 where* $C^p = C^s$ *and* $C^{p'} = C^s_x$

In our image database this results in images retrieved for the first stage associated with the relevant concepts and they are: Image 1 (Image with Team Brazil in 2006 FIFA world cup), Image 2 (Chelsea – Premiership 2007).

	Query	Image 1		Image 2	
hasTournament	Fifawc 06	Fifawc 06	1	Prem. 07	0
DoM		Team Brazil	2.5	Chelsea	2

Table 3 Analysing secondary terms in the query

7.2. Explicit query expansion

Explicit query expansion involves direct replacement of terms in the user query with terms that were identified as identical by the knowledge domain administrator or the end user. These replacement terms are not part of the ontology infrastructure, but are kept in a separate synonym dictionary that contains one-to-many (USE_FOR) relations between the ontology term and the possible synonyms. For instance, the domain administrator might use the ontology term “Manchester United” to replace the popular term “Man UTD”.

Similarly, users are allowed to cache (USE_FOR) terms on the client-side for exclusive expansion of their queries. The domain administrator has access to the most popular cached nicknames/synonyms and can choose to enter them into the main synonym dictionary.

We considered adding synonyms to the ontology using OWL's owl:sameAs property, but decided against it primarily because of the performance penalties in processing RDF data as opposed to simple text strings. Our initial experiments have shown that the search time increases by a factor of 2.5 if the synonyms are deployed in the ontology. We also think that from a pure semantic engineering point of view, nicknames such as "Man UTD" should not exist as an RDF individual.

Finally, we started considering using NLP techniques to attempt to translate free-entry queries that are not constructed using our domain-tailored retrieval interface (see Figure 4) into our "Actor-Action/Emotion-Object" semantic format to allow for semantic reasoning (Frkovic et al. 2008) (Bhogal et al. 2007).

At the time of writing this paper we succeeded in utilising WordNet lexical database primarily in identifying verbs that might be candidate for the "Action/Emotion" central part of our annotation format. Subsequently, the left part to the verb is further analysed as an "actor" candidate, and the right as an "object" candidate, applying our spelling checker and synonym replacement where appropriate. For instance the free-entry: "Man Utd's Wayne Rooney tackles the French player Zizou" is analysed as detailed in Table 4,5 below.

Man Utd's Wayne Rooney	tackles	the French player Zizou.
Subject part	Verb	Object part

Table 4 Lexical analysis of the query (I)

Man Utd	s	Wayne Rooney	Tackles	the	French	player	Zizou
Manchester United	#	Wayne Rooney	Tackle	###	adj		Zinedine Zidane

Table 5 Lexical analysis of the query (II)

Hence, the sentence analyser infers the request below which can be now fired at our semantic image retrieval engine as illustrated in Table 6.

Actor	Wayne Rooney
Action	Tackle
Object	Zinedine Zidane
Team	Manchester United

Table 6 Inferred request by the sentence analyser

Again because performance is critical for the application area on hand, we opted for on-demand population of our synonym dictionary after verifying the synonyms with WordNet, as opposed to the wholesale approach of uploading WordNet synonyms of our domain terms.

8. Conclusions

In this paper we presented a comprehensive solution for image retrieval applications that takes full advantage of advances in semantic web technologies to coherently implement the annotation, retrieval and query expansion components of the integrative framework. We claim that our solution is particularly attractive to commercial image providers where emphasis is on the efficiency of the retrieval process as much as on improving the accuracy and volume of returned results. For instance, we were reluctant to employ expensive content-based recognition techniques at the retrieval stage and deployed public ontology caching to reduce the reasoning overhead, while designed an efficient query expansion algorithm to improve the quality of the image recall.

The first stage of the development was producing ontologies that conceptualise the objects and their relations in the selected domain. We methodically verified the consistency of our ontology, optimised its coverage, and performed normalisation methods to rid of concept redundancies. Our annotation approach was based on a variation of the “sentence” structure to obtain the semantic-relational capacity for conceptualising the dynamic motion nature of the targeted sport domain. This careful analysis of the domain features allowed us to hardwire application domain knowledge into the ontology and hence implicitly perform query expansion either by simple replacement of equivalent terms or by traversing the ontology tree to modify more complex queries.

The retrieval algorithm is based on the nearest-neighbour matchmaking technique for traversing the ontology tree and can accommodate complex, relationship-driven user queries. The algorithm also provides for user-defined weightings to improve the ranking of the returned images and was extended to embrace query expansion technology in a bid to improve the quality of the recall.

Our efforts in implicit query expansion were greatly aided by our well-structured domain ontology that can be seamlessly deployed to find related terms via simple equivalence relations without compromising the semantics of the overall query. We also presented our initial research into using lexical databases to analyse free-entry queries in our effort to make them compatible with the entry requirements of our semantic search engine.

Although we recognise that image analysis techniques might have a large time overhead for the on-line retrieval process, we intend to research utilizing advances for in semantically-enabled content recognition technology to aid in semi-automating the annotation process of legacy caption-poor images.

References

[1] Yee K et al. (2003) Facated meta data for image search and browsing. Proceedings of the SIGCHI conference on Human factors in computing systems. pp. 401-408.

[2] Fuji A, Ishikawa T (2005) Toward the Automatic Compilation of Multimedia Encyclopaedias: Association Images with Term Descriptions on the Web. In Proceedings of the 2005 International Conference on Web Intelligence – WIC05. Compiègne, France. September 19-22, 2005. 536-542.

[3] Jeon J et al. (2003) Automatic image annotation and retrieval using cross-media relevance models. Proceedings of the 26th annual international ACM SIGIR conference on Research and development in information retrieval. Toronto, Canada.

[4] Hare J et al. (2006) Mind the gap: another look at the problem of the semantic gap in image retrieval. Multimedia Content Analysis, Management, and Retrieval 2006, Vol. 6073, No. 1. (2006).

[5] Wang H, Liu, Chia L (2006) Does ontology help in image retrieval?: a comparison between keyword, text ontology and multi-modality ontology approaches. Proceedings of the 14th annual ACM international conference on Multimedia. pp. 109-112

[6] Shadbolt N, Berners-Lee T, Hall W (2006) The Semantic Web Revisited. IEEE Intelligent Systems, v.21 n.3, pp. 96-101.

[7] Wu K, Yap K (2006) Fuzzy SVM for content-based image retrieval: a pseudo-label support vector machine framework. IEEE Computational Intelligence Magazine, Volume 1, Issue 2, May 2006, pp 10 – 16

[8] Julia V, Schiele B (2007) Semantic modeling of natural scenes for content-based image retrieval. International Journal of Computer Vision 72.2 (2007) pp. 133-157.

[9] Lam T, Singh R (2006) Semantically Relevant Image Retrieval by Combining Image and Linguistic Analysis. Proc. International Symposium on Visual Computing (ISVC), Lecture Notes in Computer Science Vol. 4292, pp. 1686 - 1695, Springer Verlag, 2006

[10] Ying L et al. (2007) A survey of content-based image retrieval with high-level semantics. Pattern Recognition 40.1. pp. 262-282.

[11] Maedche A, Motik B, Stojanovic L (2003) Managing multiple and distributed ontologies on the Semantic Web. The International Journal on Very Large Data Bases – Springer, Volume 12, Number 4 / November, 2003. pp 286-302.

[12] Buitelaar P, Ramaka S (2005) Unsupervised Ontology-based Semantic Tagging for Knowledge Markup. In: Proc. of the Workshop on Learning in Web Search at the International Conference on Machine Learning, Bonn, Germany, August 2005.

[13] Rector A (2003) Modularisation of domain ontologies implemented in description logics and related formalisms including OWL. Proceedings of the 2nd international conference on Knowledge capture. pp. 121 - 128

[14] Noy N, Tu S (2003) The evolution of Protégé-2000: An environment for knowledge-based systems development. International Journal of Human-Computer Studies 58(1), pp. 89–123.

[15] W3C Corporation [online] (2004) OWL Web Ontology Language Overview. http://www.w3.org/TR/owl-features

[16] Carroll J et al. (2004) Jena: implementing the semantic web recommendations. Proceedings of the 13th international World Wide Web conference. New York, USA. 74-83.

[17] Hollink L et al. (2003) Semantic Annotation of Image Collections. In Workshop on Knowledge Markup and Semantic Annotation, KCAP'03.

[18] Park Y, Li Y (2007) Semantic analysis for topical segmentation of videos. 2007 International Conference on Semantic Computing, 17-19 Sept. 2007 , Irvine, CA, USA , 2007, pp. 161-8.

[19] Osman T, Thakker D, Al-Dabass D (2006). Semantic-Driven Matchmaking of Web Services Using Case-Based Reasoning, IEEE International Conference on Web Services (ICWS'06), September 2006. pp. 29-36.

[20] Ching-man Au Y, Gibbins N., Shadbolt N. (2008) A k-nearest-neighbour method for classifying web search results with data in folksonomies. Web Intelligence and Intelligent Agent Technology, 2008. WI-IAT'08. IEEE/WIC/ACM International Conference on. Vol. 1. IEEE, 2008.

[21] Gaihua F et al. (2005) Ontology-based Spatial Query Expansion in Information Retrieval. Lecture Notes in Computer Science, Vol. 3761/2005, pp. 1466-1482.

[22] Lee C, Geunbae G (2005) Probabilistic information retrieval model for a dependency structured indexing system", International Journal on Information Processing and Management, Volume 41 , Issue 2 - March 2005, pp. 161 - 175.

[23] Haubold A, Natsev A, Naphade M (2006) Semantic multimedia retrieval using lexical query expansion and model-based reranking. IEEE International Conference on Multimedia and Expo, July 2006 , Toronto, Canada, pp 1761 - 1764.

[24] Fellbau C (1998) WordNet- An Electronic Lexical Database, The MIT press, Cambridge, MA, USA, May 1998.

[25] Frkovic F, Podobnik V, Trzec K, Jezic, G (2008) Agent-based user personalization using context-aware semantic reasoning. 12th International Conference on Knowledge-Based Intelligent Information and Engineering Systems - KES 2008, pp. 166-73.

[26] Bhogal J, Macfarlane A, Smith P (2007) A review of ontology based query expansion, Information Processing and Management: an International Journal, v.43 n.4, pp. 866-886, July, 2007.