Open Research Online

The Open University's repository of research publications
and other research outputs

Ontology ranking based on the analysis of concept structures

Conference or Workshop Item

How to cite:

Alani, Harith and Brewster, Christopher (2005). Ontology ranking based on the analysis of concept structures. In: Third International Conference on Knowledge Capture (K-Cap), 2-5 Oct 2005, Banff, Alberta, Canada, ACM, pp. 51–58.

For guidance on citations see FAQs.

© 2005 The Authors

Version: Accepted Manuscript

Link(s) to article on publisher's website:
http://www.k-cap.org/kcap05/

Copyright and Moral Rights for the articles on this site are retained by the individual authors and/or other copyright owners. For more information on Open Research Online's data policy on reuse of materials please consult the policies page.

oro.open.ac.uk

Ontology Ranking based on the Analysis of Concept Structures

Harith Alani
Dept. of Electronics & Computer Science
University of Southampton, UK
h.alani@ecs.soton.ac.uk

Christopher Brewster
Dept. of Computer Science
University of Sheffield, UK
C.Brewster@dcs.shef.ac.uk

ABSTRACT
In view of the need to provide tools to facilitate the re-use of existing knowledge structures such as ontologies, we present in this paper a system, AKTiveRank, for the ranking of ontologies. AKTiveRank uses as input the search terms provided by a knowledge engineer and, using the output of an ontology search engine, ranks the ontologies. We apply a number of classical metrics in an attempt to investigate their appropriateness for ranking ontologies, and compare the results with a questionnaire-based human study. Our results show that AKTiveRank will have great utility although there is potential for improvement.

Categories and Subject Descriptors
H.3.3 [**Information Search and Retrieval**]: Selection process

General Terms
Management, Measurement, Experimentation

Keywords
Ontology ranking, ontology analysis, ranking techniques, semantic similarity

1. INTRODUCTION
Ontologies have been shown to be beneficial for representing domain knowledge, and are quickly becoming the backbone of the Semantic Web. Building ontologies, however, represents a considerable challenge for a number of reasons. It takes a considerable amount of time and effort to construct an ontology, and it necessitates a sophisticated understanding of the subject domain. Thus it is an even greater challenge if the ontology engineer is not familiar with the domain. However, one of the major advantages claimed of ontologies is the potential for the "reuse" of knowledge. So in theory, one should be able to reuse other people's ontologies, and modify, extend, and prune them as required, thereby avoiding the huge effort of starting from scratch.

A number of ontology libraries currently exist, hosting various ontology files. Examples of such libraries include Ontolingua[1], the DAML library[2], the Protègè OWL library[3], etc. However, the ontology search facilities provided by these libraries are at best limited to term search, making it difficult for the user to select the relevant ontologies from others than happened to contain a class with the desired label. As the number of publicly available ontologies increases, this problem is bound to get worse. Thus there is a contradiction in this situation. For a variety of purposes, including the Semantic Web, there is a need for more and more ontologies to be constructed and made available. However, as this occurs, so the re-use of this knowledge becomes an ever greater problem. We will term this the *knowledge re-use conundrum*.

In order to achieve an effective level of knowledge re-use, we need search engines capable of helping us find the ontologies we are looking for. Some ontology search engines have been developed that can provide lists of ontologies that contain specific search terms, such as Swoogle [3] and OntoSearch [20]. Such search engines are a good step forward, but more is required in terms of ontology search if re-use is to become a reality.

Google has surpassed other search engines because of the effectiveness of its page ranking approach, and most likely the same will happen in the near future for ontology search engines. As the number of ontologies that such search engines can find increases, so will the need increase for a proper ranking method to order the returned lists of ontologies in terms of their relevancy to the query. A proper ranking of ontologies could save the user a lot of time and effort. It would reduce the need

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
K-CAP'05, October 2–5, 2005, Banff, Alberta, Canada.

[1] http://www-ksl-svc.stanford.edu:5915/
[2] http://www.daml.org/ontologies/
[3] http://protege.stanford.edu/plugins/owl/owl-library/

to examine in detail each and every ontology returned to find out how well it suits the needs of the knowledge engineer.

Another challenge for ranking ontologies presents itself when searching for multiple terms. For example if searching for “pet” AND “food”, then an ontology that has such classes in good semantic proximity from each other should be favoured over other ontologies where those classes might be much further apart. Various formulae have been presented in the literature to measure similarities of terms within semantic networks, mainly for purposes of query expansion.

This paper presents a new ontology ranking system which applies a number of analytic methods to rate each ontology based on how well it represents the given search terms. Related work concerning ontology searching and ranking is reviewed in the following section. A full description of the architecture and ranking method is given in section 3. A number of experiments are detailed in section 4 and a partial evaluation is described in section 5.

2. RELATED WORK

Ranking has always been at the heart of information retrieval. This became even more apparent given the enormous size of the web and its continuous expansion. Google uses the PageRank [8] method to rank documents based on hyperlink analysis. Swoogle [3] and OntoKhoj [9] rank ontologies also using a PageRank-like method that analyses links and referrals between ontologies in the hope of identifying the most popular ontologies. However, the majority of ontologies available on the Web are poorly connected, and more than half of them are not referred to by any other ontologies at all[3]. Poor connectivity would certainly produce poor PageRank results.

Furthermore, a popular ontology does not necessarily indicates a good representation of all the concepts it covers. Popularity does not necessarily correlate with ‘good’ or appropriate representations of knowledge. For example, supposing an engineer was looking for an ontology about “students,” there could be an ontology about the academic domain that is well connected, and thus popular. If this ontology contains a concept named “Student”, then this ontology will show up high on the list of candidates. However, it could very well be the case that the “Student” class is very weakly represented. That ontology might have become popular due to its coverage of publications and research topics, rather than for it’s coverage of student related concepts.

Similarity measures have often been used in information retrieval systems to provide better ranking of query results. Quillian [10] viewed memory search as a spreading-activation process that tries to find intersections between a number of concepts in a semantic network. This theory became the base for many semantic similarity measures used for various purposes, such as query expansion and network analysis.

Ontologies can be viewed as semantic graphs of concepts and relations, and hence similarity measures can be applied to explore these conceptual graphs. Resnik applied a similarity measures to WordNet to resolve ambiguities [12]. The measure he used is based on the comparison of shared features, which was first proposed in [17]. Another common-feature based similarity is the shortest-path measure, introduced by Rada [11]. He argues that the more relationships objects have in common, the closer they will be in an ontology. Rada used this measure to help rank biomedical documents which were represented in a semantic knowledge-base.

Probability-based measures to explore concept similarities over the Gene ontology was investigated in [6]. Jones and colleagues developed a number of measures to estimate similarity between geographical entities, based on analysing non-common super-classes of concepts in a geographical ontology [5].

Most of the measures above are based on pairwise comparison of concepts (or sets of concepts). However, experiments on measuring similarity between whole ontology structures have also been reported [7][19]. To the best of our knowledge none of such measures have been applied to ranking ontologies, even though some work has been reported on ranking semantic queries using ontologies [15].

Our work mainly focuses on investigating how ontology developers judge ontologies, and what measures are required and suitable for ranking those ontologies.

3. AKTIVERANK

To the best of our knowledge, ontology ranking has only been attempted using link-based analysis (eg [3][9]). This section describes AKTiveRank, a system for ranking ontologies by aggregating a number measures that look into certain structural features of concepts, such as their centrality of the terms in a hierarchy, structural density, and semantic similarity to other concepts of interest.

3.1 System Architecture

Figure 1 shows the current architecture of AKTiveRank. The main component (2) is a Java Servlet that receives an HTTP query from a user or an agent (1). The query contains the terms to search for. Currently it is only possible to search for concepts. In other words, search terms will only be matched with ontology classes, and not with properties or comments.

When a query is received, AKTiveRank queries Swoogle

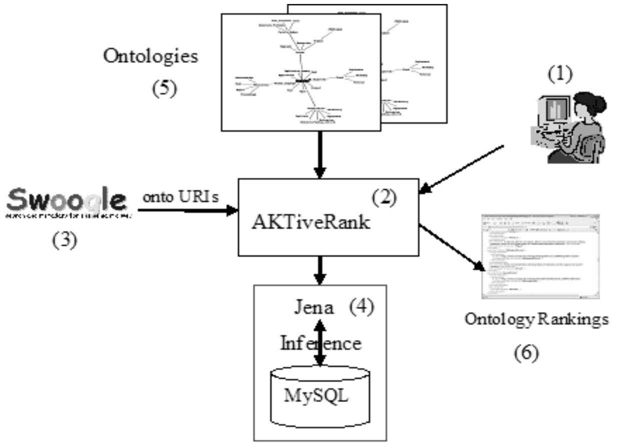

Figure 1: AKTiveRank Architecture.

(3) for the given search terms and scrapes the ontology URIs from the results page returned by Swoogle. Once a list of ontology candidates is gathered from Swoogle, AKTiveRank starts to check whether those ontologies are already stored in the Jena MySQL database backend (4), and if not, load them from the web (5) and add them to the database. The Jena API is used here to read the ontologies and handle the database storage. All the analysis of ontology structures that AKTiveRank performs for the ranking is also undertaken using Jena's API. An inference engine (Racer[4] for OWL Lite and DL, and Jena's own inference engine for OWL FULL) is applied to every ontology loaded from the database before analysed in AKTiveRank.

AKTiveRank then analyses each of the ontology candidates to determine which is most relevant to the given search terms. This analysis will produce a ranking of the retrieved ontologies, and the results are returned to the user as an OWL file containing the ontology URIs and their total ranks.

3.2 The Ranking Approach

AKTiveRank applies four types of assessments (measures) for each ontology to measure the rankings. Each ontology is examined separately. Once those measures are all calculated for an ontology, the resulting values will be merged to produce the total rank for the ontology. The four measures are described in the following.

3.2.1 Class Match Measure

The Class Match Measure (CMM) simply evaluates the coverage of an ontology of the given search terms. AKTiveRank looks for classes in each ontology that have labels matching a search term either exactly (class label identical to search term) or partially (class label "contains" search term).

[4]http://www.sts.tu-harburg.de/ r.f.moeller/racer/

An ontology that contains all search terms will obviously score higher than others, and exact matches are regarded as better than partial matches. For example if searching for "Student" and "University", then an ontology with two classes labelled exactly as the search terms will score more in this measure than another ontology which contains partially matching classes, eg labelled "UniversityBuilding" and "PhDStudent".

Definition 1. Let c[o] be a set of classes in ontology o, and T is the set of search terms.

$$E[o,T] = \sum_{c \in C[o]} \sum_{t \in T} I(c,t)$$

$$I(c,t) = \begin{cases} 1 & : \quad if\ label(c) = t \\ 0 & : \quad if\ label(c) \neq t \end{cases}$$

$$P[o,T] = \sum_{c \in C[o]} \sum_{t \in T} J(c,t)$$

$$J(c,t) = \begin{cases} 1 & : \quad if\ label(c)\ contains\ t \\ 0 & : \quad if\ label(c)\ not\ contain\ t \end{cases}$$

where $E[o,T]$ and $P[o,T]$ are the sets of classes of ontology o that have labels that match any of the search terms t exactly or partially, respectively.

$$CMM[o] = \alpha|E[o,T]| + \beta|P[o,T]|$$

where $CMM[o,\tau]$ is the Class Match Measure for ontology o with respect to search terms τ. α and β are the exact matching and partial matching weight factors respectively. Exact matching is favoured over partial matching if $\alpha > \beta$. In the experiments described in this paper, $\alpha = 0.6$ & $\beta = 0.4$.

3.2.2 Centrality Measure

The Centrality Measure (CEM) is aimed to assess how representative a class is of an ontology. Several approaches have been proposed for the task of defining classes when building an ontology, such as top-down, bottom-up, outside-in [18], or middle-out [14] approaches. Even though all those approaches are valid, but psycholinguistic evidence has shown that middle level concepts tend to be more detailed and prototypical of their categories than classes at higher or lower hierarchical level [13]. Thus we here assume that the more central a class is in the hierarchy, the more likely it is for it to be well analysed and fully represented. The Centrality Measure is meant to estimate just that.

Definition 2. Let $H[c] = max_{p \in P}\{root_c \overset{p}{\rightsquigarrow} botton_c\}$ which is the length of the longest path from the root of the branch that contains class c to its bottom node. $D[c] = max_{p \in P}\{root_c \overset{p}{\rightsquigarrow} c\}$ which is the hierarchical level of class c, or the path length from the root of its branch down to this node, where $c \in E[o, T], P[o, T]$, and $n = |E[o, T]| + |P[o, T]|$.

$$cem[c] = 1 - \left| \frac{D[c] - \frac{H[c]}{2}}{\frac{H[c]}{2}} \right|$$

$$CEM[o] = \frac{1}{n} \sum_{i=1}^{n} cem[c]$$

3.2.3 Density Measure

When searching for a "good" representation of a specific concept, one would expect to find a certain degree of detail in the representation of the knowledge concerning that concept. This may include how well the concept is further specified (the number of subclasses), the number of attributes associated with that concept, number of siblings, etc. All this is taken into account in the Density Measure (DEM). DEM is intended to approximate the representational-density of classes and consequently the level of knowledge detail.

Definition 3. Let $S = \{S_1, S_2, .., S_i, .., S_6\} = \{directRelations[c], indirectRelations[c], instances[c], superclasses[c], subclasses[c], siblings[c]\}$. Note that $directRelations[c]$ includes relations pointing *to* and *from c*. w_i is a weight factor, and $n = |E[o, T]| + |P[o, T]|$ which is the number of matched classes in ontology o.

$$dem[c] = \sum_{i=1}^{6} w_i |S_i|$$

$$DEM[o] = \frac{1}{n} \sum_{i=1}^{n} dem[c]$$

3.2.4 Semantic Similarity Measure

The Semantic Similarity Measure (SSM) calculates how close the classes that matches the search terms are in an ontology. The motivation for this is that ontologies which position concepts further away from each other are less likely to represent the knowledge in a coherent and compact manner. The SSM formula used here is based on the shortest path measure defined in [11]. SSM is measured from the minimum number of links that connects a pair of concepts.

Definition 4. Let $c_i, c_j \in \{classes[o]\}$, and $c_i \overset{p}{\rightsquigarrow} c_j$ is a path $p \in P$ of paths between classes c_i and c_j, where $c_i, c_j \in E[o, T], P[o, T]$

$$ssm(c_i, c_j) = \begin{cases} \frac{1}{length(min_{p \in P}\{c_i \overset{p}{\rightsquigarrow} c_j\})} & : \ if \ i \neq j \\ 0 & : \ if \ i = j \end{cases}$$

$$SSM[o] = \frac{1}{n} \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} ssm[c_i, c_j]$$

3.2.5 Total Score

The total score of an ontology can be calculated once the four measures are applied to all ontologies. Total score is then calculated by aggregating all the measures' values, taking into account their *weights*, which are used to determine the importance of each measure in the ranking.

Definition 5. Let $M = \{M[1], .., M[i], M[4]\} = \{CMM, CEM, DEM, SSM\}$, w_i is a weight factor, and O is the set of ontologies to rank.

$$Score[o \in O] = \sum_{i=1}^{4} w_i \frac{M[i]}{\max_{1 \leq j \leq |O|} M[j]}$$

Measure values are normalised to be in the range (0-1) by dividing by the maximum measure value for all ontologies ($M[j]$).

4. EXPERIMENTS

Lets assume that we need to find an ontology that represents the concepts of "University" and "Student". The list of ontologies returned by Swoogle as a results of the query "university student type:owl" is shown in table 1. "type:owl" simply restricts Swoogle to search for OWL ontologies. Some of the ontologies returned by Swoogle were duplicates (i.e. the same ontology is available under two slightly different URLs). As expected, the same rank was produced by AKTiveRank for all duplicate ontologies, and they were removed from the table to save space. It is worth mentioning that Swoogle returned those duplicated ontologies in very different orders. For example the *exact* same "koala" ontology was returned under 3 URLs in the 2nd[5], 9th[6], and 18th[7] position!

[5] http://protege.stanford.edu/plugins/owl/owl-library/koala.owl
[6] http://www.ling.helsinki.fi/kit/2004k/ctl310semw/Protege/koala.owl
[7] http://www.iro.umontreal.ca/ touzanim/Ontology/koala.owl

Table 1: Swoogle's result for "student university type:owl"

Pos.	Ontology URL
a	http://www.csd.abdn.ac.uk/∼cmckenzi/playpen/rdf/akt_ontology_LITE.owl
b	http://protege.stanford.edu/plugins/owl/owl-library/koala.owl
c	http://protege.stanford.edu/plugins/owl/owl-library/ka.owl
d	http://reliant.teknowledge.com/DAML/Mid-level-ontology.owl
e	http://www.csee.umbc.edu/∼shashi1/Ontologies/Student.owl
f	http://www.mindswap.org/2004/SSSW04/aktive-portal-ontology-latest.owl
g	http://www.mondeca.com/owl/moses/univ2.owl
h	http://www.mondeca.com/owl/moses/univ.owl
i	http://www.lri.jur.uva.nl/∼rinke/aargh.owl
j	http://www.srdc.metu.edu.tr/∼yildiray/HW3.OWL
k	http://www.mondeca.com/owl/moses/ita.owl
l	http://triplestore.aktors.org/data/portal.owl
m	http://annotation.semanticweb.org/ontologies/iswc.owl
n	http://www.csd.abdn.ac.uk/ cmckenzi/playpen/rdf/abdn_ontology_LITE.owl
o	http://ontoware.org/frs/download.php/18/semiport.owl

When AKTiveRank was applied to the resulting list shown in table 1, it produced the values given in Figure 2. Note that Jena failed to parse j and n ontologies, and hence they were dropped from the experiment.

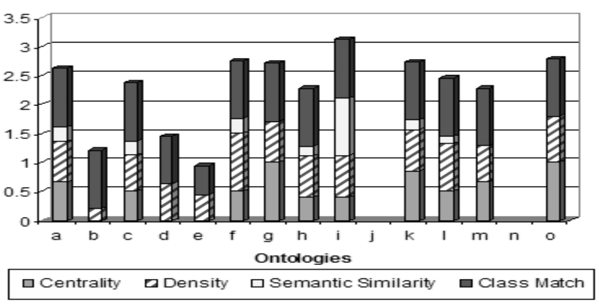

Figure 2: Experiment - student + university.

From the results in figure 2, it can be seen that the values of the Semantic Similarity measure (SSM section 3.2.4) are perhaps the most varied (more distinctive). Ontology i scored the highest value for SSM. This is because the two classes that correspond to *University* and *Student* are directly linked with the property *enrolsAt*. Note that 6 of our ontologies received a SSM of 0.0. This indicates that AKTiveRank did not manage to find any path connecting the two given search terms. Semantic paths that cross via the imaginary *owl:Thing* class are ignored.

The values calculated for Class Match measure (CMM 3.2.1) were more consistent that the other three measures. This is because of the rigidity of this measure, and because most of the ontologies did have a class labelled exactly with the given search terms, apart from ontologies d which had 1 exact match ("University") and 1 partial match ("CollegeStudentPosition"), and ontology e which only had the "Student" class. However, this measure is expected to produce more varied results when searching for a larger set of terms.

As for the values for the Density measure (DEM, section 3.2.3, ontology f scored the highest, followed by l and o. The values of this measure do reflect our expectations where the ontologies with relatively *denser* representations of the classes in question scored higher than others with sparser representation. For example ontology e consists of only 3 classes and it is the smallest is our set of ontologies.

The Centrality measure (CEM, section 3.2.2), which indicates how central a class is in its branch, produces values with a good dispersion of values. From figure 2 we can see that the variety of CEM values resembles that of the Density measure. This enforces the belief that the more central concepts in an ontology tend to have greater detail [13].

4.1 Experiment 1

The weights used to calculate the final scores in this case are 0.25 for all the four measures of centrality, density, similarity, and class match, giving them all equivalent importance. Figure 3 shows the results total score values for our ontologies.

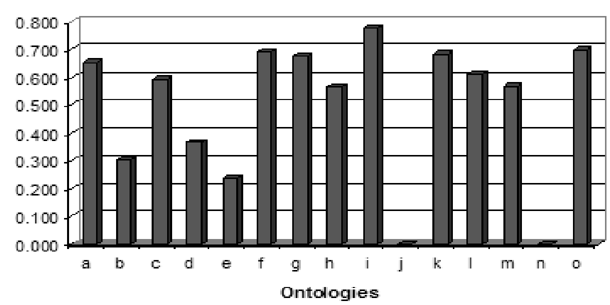

Figure 3: Total scores based on a weight of 0.25 for all four measures

From the results in figure 3, it can be seen that the according to AKTiveRank, ontology i ranked number 1 (with a score of 0.78), followed by f (0.69), and k (0.684).

4.2 Experiment 2

In this case, the ranking weights for the centrality, density, similarity, and class match measures were set to 0.1, 0.2, 0.4, 0.3 respectively, putting more emphasis on the semantic closeness of the selected concepts, as well as their exact match in the ontologies. The new results are shown in figure 4.

4.3 Experiment 3

When using the weights of 0.3, 0.4, 0.2, 0.1 for CEM, DEM, SSM, and CMM respectively, the results were as

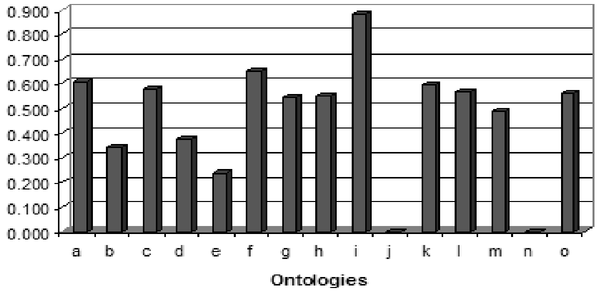

Figure 4: Total scores based on a weight of CEM 0.1, DEM 0.2, SSM 0.4, CMM 0.3

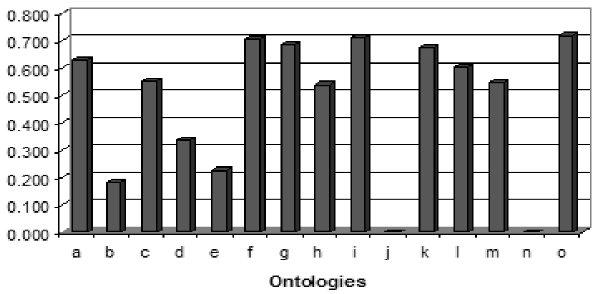

Figure 5: Total scores based on a weight of CEM 0.3, DEM 0.4, SSM 0.2, CMM 0.1

shown in figure 5. The emphasis is now given to the centrality of concepts in their hierarchies, as well as to how *densely* they are represented.

5. EVALUATION

In order to evaluate the utility of the output of AKTiveRank, we needed some independent ranking of the ontologies returned by the search engine. To do this, we designed a paper based questionnaire[8] and asked a small sample computer scientists with a firm grasp of ontologies to answer the questions.

The subjects were presented with a set of instructions, a set of screen shots of the relevant ontologies and a set of seven questions. The questions were designed to be as simple as possible, but also to translate into the measurable parameters used above. In each case, the subjects were asked to rank the top five ontologies from the set presented. They were also given the opportunity to give comments and feedback. The total population sample was only four participants so we cannot make claims of statistical accuracy or significance. The rankings were as follows:

When comparing ranks of experiments 1,2,3 with the ranks generated from our evaluation using the Pearson Correlation Coefficient, we get the results in table 4.

[8]The questions used in the questionnaire are included in Appendix I below.

Table 2: Ranking by human subjects

Onto	Rank	Onto	Rank	Onto	Rank
i	0.666	f	0.419	a	0.419
m	0.368	g	0.348	h	0.348
l	0.344	o	0.344	d	0.337
e	0.334	k	0.331	c	0.322
b	0.251				

Table 3: Ranking Comparison

Onto	Swoogle	Exp1	Exp2	Exp3	Human Rank
a	1	6	3	6	2.5
b	2	12	12	14	13
c	3	8	5	8	12
d	4	11	11	12	9
e	5	13	13	13	10
f	6	3	2	3	2.5
g	7	5	9	4	5.5
h	8	10	8	10	5.5
i	11	1	1	2	1
k	11	4	4	5	11
l	11	7	6	7	7.5
m	11	9	10	9	4
o	11	2	7	1	7.5

The values in table 4 shows that the ranks produced by AKTiveRank are neither completely different, nor an exact match to the ranks produced by our subject (a value of 0 indicates no relation, and 1 is an exact linear relationship between the two datasets, -1 is an exact match but in a reverse order). However, Swoogle's ranks seems to be the least similar to the ranks given by our experts.

A further significant indication of the success of AKTiveRank lies in the fact that both the human ranking and the automated ranks in Experiment 1 and 2 place ontologies *i* and *f* at the top. Of the six top ranked ontologies the two approached overlap by two thirds. We see this as indicative of the validity of our approach. However, the use of human subjects highlighted a number of factors which we discuss in the next section.

6. DISCUSSION

The attempt at obtaining a human ranking of the ontologies highlighted the difficulty of any attempt at evaluation of knowledge resources. The difficulties of evaluating ontologies have already been discussed elsewhere [2, 16]. In this particular case, part of the difficulty arose because the subjects were presented only with a screen capture of the class hierarchy and they expressed the need for information about properties, relations and constraints. It was also noted that it was important to distinguish between the underlying ontology and what may be derived by applying a reasoner. AKTiveRank does apply reasoners to the OWL ontologies when analysing them, but this was not reflected in the screen dumps given in our questionnaire.

There are also contradictory objectives in ontology management. One might assume that ontologies which were more compact were better on the basis they contained

Table 4: Pearson Correlation Comparison of ranks against expert rankings

Rank	Pearson value
Swoogle rank	-0.272
AKTiveRank exp1	0.538
AKTiveRank exp2	0.532
AKTiveRank exp3	0.562

less extraneous information. However, our subjects indicated that it may be useful to have a more complex ontology as this provides the possibility for future expansion with less effort. Our use of the Semantic Similarity Measure implied a preference for an ontology where the relevant concepts are explicitly related in the class hierarchy. However, it may be more effective to represent these relations to the subject in a *graph* rather than in a mainly hierarchical layout.

At least two of the subjects said they would have preferred to view the RDF/OWL format directly which indicates that significantly more information needs to be presented to the subjects of future questionnaires. Furthermore, it indicates that if an ontology ranking system is to facilitate the task of the knowledge engineer, then it needs to be designed in a way so as to provide more information of the right sort. Mere ranking, while it will no doubt be useful, may not be the whole story from a practical perspective.

There appears to be a need to disentangle a number of different parameters which can affect the usability of a knowledge representation. The knowledge engineers perception of the following parameters all may vary depending on the particular application context:

- A compact vs. a spread out ontology.
- An ontology where the concept is not a leaf node as this indicates a more complex representation of the knowledge domain.
- An ontology which is more or less dense (number of siblings, properties, etc.). We have assumed a denser ontology represents a more complex and sophisticated representation of the domain.
- An ontology where the terms requested are closer or further apart. We have assumed that closer is better.
- An ontology where terms occur as class/concept labels or as properties or as relations. We have so far only considered class labels.

In each case, specific needs may over-ride our assumptions and indicate that an ontology engineer's perceptive is extremely complex.

7. CONCLUSIONS

This paper investigated using a number of ontology graph-analysis measures to ranking ontologies. This research has raised a number of questions which need further investigation. The parameters used in the AKTiveRank process need to be reconsidered in the light of the needs of human knowledge engineers. In order to do this, we plan a more extensive human ranking study which will include a larger population of subjects and will try to elicit a greater understanding of the process of ontology evaluation and selection. On the basis of this, we believe a future version of AKTiveRank will be able to rank more effectively and provide a greater range of information about an ontology so as to facilitate its evaluation.

AKTiveRank was unable to handle some ontologies due to the failure of Jena to parse them, most likely due to to syntactical errors. The experiments described in this paper took, on average, 2 minutes to analyse each ontology! This is mainly due to the slow process of retrieving inferred models in Jena, the very little caching and lack of indexing in AKTiveRank, and the inadequacy of existing RDF query languages in dealing with graph queries. We are currently working on linking AKTiveRank to JUNG[9] to speed up path queries on ontologies.

8. ACKNOWLEDGMENTS

This work is supported under the Advanced Knowledge Technologies (AKT) Interdisciplinary Research Collaboration (IRC), which is sponsored by the UK Engineering and Physical Sciences Research Council under grant number GR/N15764/01. The AKT IRC comprises the Universities of Aberdeen, Edinburgh, Sheffield, Southampton and the Open University. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing official policies or endorsements, either express or implied, of the EPSRC or any other member of the AKT IRC. Special thanks to Dr Srinandan Dasmahapatra for his valuable help, to Dr. Daniela Petreli for help with the questionnaire, and to the four scientists who participated in our evaluation.

9. REFERENCES

[1] R. Angles and C. Gutierrez. Querying rdf data from a graph database perspective. In *Proc. of 2nd European Semantic Web conference, ESWC*, pages 346–360, Crete, 2005.

[2] C. Brewster, H. Alani, S. Dasmahapatra, and Y. Wilks. Data driven ontology evaluation. In *Int. Conf. on Language Resources and Evaluation*, Lisbon, Portugal, 2004.

[3] L. Ding, T. Finin, A. Joshi, R. Pan, R. S. Cost, Y. Peng, P. Reddivari, V. C. Doshi, and J. Sachs.

[9] http://jung.sourceforge.net/

Swoogle: A semantic web search and metadata engine. In *Proc. 13th ACM Conf. on Information and Knowledge Management*, Nov. 2004.

[4] C. Halascheck, B. Aleman-Meza, B. I. Arpinar, and A. P. Sheth. Discovering and ranking semantic associations over a large rdf metabase. In *Proc. 30th Very Large DataBases Conf (VLDB'04)*, Toronto, Canada, 2004.

[5] C. B. Jones, H. Alani, and D. Tudhope. *Geographical Terminology Servers: Closing the Semantic Divide*, chapter 11, pages 201–218. Perspectives on Geographic Information Science. M. Duckham, M.F.Goodchild and M.F.Worboys (edt.), Taylor and Francis, 2002.

[6] P. Lord, R. Stevens, A. Brass, and C. Goble. Semantic similarity measures as tools for exploring the gene ontology. In *Proc. Pacific Symposium on Biocomputing*, 2003.

[7] A. Maedche and S. Staab. Measuring similarity between ontologies. In *Proceedings of the European Conference on Knowledge Acquisition and Management (EKAW)*, pages 251–263, Madrid, Spain, 2002. Springer.

[8] L. Page, S. Brin, R. Motwani, and T. Winograd. The pagerank citation ranking: Bringing order to the web. Technical report, Stanford Uni., 1999.

[9] C. Patel, K. Supekar, Y. Lee, and E. Park. Ontokhoj: A semantic web portal for ontology searching, ranking, and classification. In *Proc. 5th ACM Int. Workshop on Web Information and Data Management*, pages 58–61, New Orleans, Louisiana, USA, 2003.

[10] M. Quillian. *Semantic Memory.* In M. Minsky (ed.), Semantic Information Processing, MIT Press, Cambridge, MA., 1968.

[11] R. Rada, H. Mili, E. Bicknell, and M. Blettner. Development and application of a metric on semantic nets. *IEEE Transactions on Systems Management and Cybernetics*, 19(1):17–30, 1989.

[12] P. Resnik. Semantic similarity in a taxonomy: An information-based measure and its application to problems of ambiguity in natural language. *Journal of Artificial Intelligence Research*, 11:95–130, 1999.

[13] E. Rosch. *Principles of Categorization.* in E. Rosch and B. B. Lloyd editors. Cognition and Categorization, Lawrence Erlbaum, Hillsdale, New Jersey, 1978.

[14] D. Skuce. Conventions for reaching agreement on shared ontologies. In *Proc. 9th Banff Knowledge Acquisition for Knowledge-Based Systems Workshop*, Banff Conference Centre, Banff, Alberta, Canada, 1995.

[15] N. Stovanovic, R. Studer, and L. Stovanovic. An approach for the ranking of query results in the semantic web. In *Proc. 2nd Int. Semantic Web Conf. (ISWC03)*, pages 500–516, Sanibel Island, FL, USA, 2003.

[16] Y. Sure, G.-P. Gomez-Perez, W. Daelemans, M.-L. Reinberger, N. Guarino, and N. F. Noy. Why evaluate ontology technologies? because it works! *IEEE Intelligent Systems*, 19(4):74–81, 2004.

[17] A. Tversky. Features of similarity. *Psychological Review*, 84(4), 1977.

[18] M. Uschold and M. Gruninger. Ontologies and semantics for seamless connectivity. *SIGMOD Record*, 33(4), 2004.

[19] P. C. Weinstein and W. P. Birmingham. Comparing concepts in differentiated ontologies. In *Proc. 12th Workshop on Knowledge Acquisition, Modeling and Management (KAW' 99)*, Banff, Alberta, Canada, 1999.

[20] Y. Zhang, W. Vasconcelos, and D. Sleeman. Ontosearch: An ontology search engine. In *Proc. 24th SGAI Int. Conf. on Innovative Techniques and Applications of Artifial Intelligence*, Cambridge, UK, 2004.

Appendix I

The questions used in the questionnaire are included here. The subjects were also provided with introductory material explaining the task, screen shots of the ontologies, as well as with space to make comments after each question. The task was described as consisting of identifying an ontology to be used in a Semantic Web Service to provide information about student numbers in universities. In each case they had to rank the ontologies in the order which best fitted the statement.

1. The ontology has all the terms needed for the application.
2. The ontology is sufficiently compact for the intended application.
3. The ontology has appropriate child concepts for the terms.
4. The ontology has appropriate sibling concepts for the terms.
5. The ontology has the least irrelevant material for the intended application.
6. The ontology relates the terms in a coherent/reasonable manner by having an appropriate ancestor concept.
7. The ontology avoids unnecessary complexity by having the terms sufficiently close to the root.