AUTHOR(S): NKISI-ORJI, I., WIRATUNGA, N., HUI, K.-Y., HEAVEN, R. and MASSIE, S.

TITLE: Taxonomic corpus-based concept summary generation for document annotation.

YEAR: 2017

Publisher citation: NKISI-ORJI, I., WIRATUNGA, N., HUI, K.-Y., HEAVEN, R. and MASSIE, S. 2017. Taxonomic corpus-based concept summary generation for document annotation. In Kampus, J., Tsakonas, G., Manolopoulos, Y., Iliadis, L. and Karydis, I. (eds). Lecture notes in computer science [online], 10450, research and advanced technology for digital libraries: proceedings of the 21st international conference on theory and practice of digital libraries (TPDL 2017), 18-21 September 2017, Thessaloniki, Greece. Cham: Springer [online], pages 49-60. Available from: https://doi.org/10.1007/978-3-319-67008-9_5

OpenAIR citation: NKISI-ORJI, I., WIRATUNGA, N., HUI, K.-Y., HEAVEN, R. and MASSIE, S. 2017. Taxonomic corpus-based concept summary generation for document annotation. In Kampus, J., Tsakonas, G., Manolopoulos, Y., Iliadis, L. and Karydis, I. (eds). Lecture notes in computer science, 10450, research and advanced technology for digital libraries: proceedings of the 21st international conference on theory and practice of digital libraries (TPDL 2017), 18-21 September 2017, Thessaloniki, Greece. Cham: Springer, pages 49-60. Held on OpenAIR [online]. Available from: https://openair.rgu.ac.uk

Taxonomic corpus-based concept summary generation for document annotation

Ikechukwu Nkisi-Orji[1], Nirmalie Wiratunga[1], Kit-Ying Hui[1], Rachel Heaven[2], and Stewart Massie[1]

[1] Robert Gordon University, Aberdeen, UK
{i.o.nkisi-orji✉, n.wiratunga, k.hui, s.massie}@rgu.ac.uk
[2] British Geological Survey, Nottingham, UK
reh@bgs.ac.uk

Abstract. Semantic annotation is an enabling technology which links documents to concepts that unambiguously describe their content. Annotation improves access to document contents for both humans and software agents. However, the annotation process is a challenging task as annotators often have to select from thousands of potentially relevant concepts from controlled vocabularies. The best approaches to assist in this task rely on reusing the annotations of an annotated corpus. In the absence of a pre-annotated corpus, alternative approaches suffer due to insufficient descriptive texts for concepts in most vocabularies. In this paper, we propose an unsupervised method for recommending document annotations based on generating node descriptors from an external corpus. We exploit knowledge of the taxonomic structure of a thesaurus to ensure that effective descriptors (concept summaries) are generated for concepts. Our evaluation on recommending annotations show that the content that we generate effectively represents the concepts. Also, our approach outperforms those which rely on information from a thesaurus alone and is comparable with supervised approaches.

Keywords: Taxonomy, Text Annotation, Information Discovery

1 Introduction

Digital library resources that were not born-digital are increasingly being made available for electronic access through mass digitisation efforts. Unlocking the content of such resources to enhance search and browse remains a challenge as facilities for content linking and navigation are often absent. Semantic annotation plays an important role in this regard by mapping the content of documents to unambiguous concepts from controlled vocabularies (or thesauri). The thesaurus models an organisation of knowledge in a domain and when used to annotate documents, is expected improve organisation, access and dissemination [1]. Accordingly, several digital repositories have controlled vocabularies from which authors and annotators select concepts to annotate or tag digital content. Popular thesauri for knowledge organisation include the Medical Subject Headings

(MeSH) and Library of Congress Subject Headings (LCSH). The selection of concepts for use in annotation largely rely on manual efforts which is tedious, time-consuming, and lacks scalability. Controlled vocabularies can contain several thousand concepts making it difficult to find the right concepts for annotation. Although it may not be possible to fully automate the annotation process as it is quite subjective, the ability to recommend a useful subset of concepts will reduce the burden on annotators.

Semantic annotation can be done at different levels of granularity (e.g. entire documents, sections/chapters, or specific terms) and approaches for recommending annotations differ accordingly. While a high-level understanding of content may be sufficient when annotating an entire document, areas such as named entity recognition, word sense disambiguation, and co-reference resolution are more pertinent to annotating specific terms. This work focuses on the annotation of segments of documents (e.g. chapters and sections) which is especially useful for books and other publications that can cover a range of domain topics. Digital agents can reuse such annotations in bespoke ways such as to meet an information need by dynamically assembling a document using relevant segments of other documents. The strategies for annotating segments of documents can be generalised for annotating entire documents. Accordingly, we treat segments of documents as individual documents. Also, we use the terms thesaurus and controlled vocabulary interchangeably and in either case, assume a taxonomy of domain concepts.

The most effective approaches for recommending annotations rely on reusing the concepts that were assigned to documents in an annotated corpus which have features that are similar to the document being annotated. However, such supervised methods make it difficult to recommend concepts that do not appear in the annotated corpus. Also, an annotated corpus has to be created before use for annotating new documents. Alternative approaches that do not require an annotated corpus often rely on the use of thesaurus-based features (concept terms, synonyms, and descriptions) to recommend annotations. Relying on thesaurus-based features can lead to poor results as controlled vocabularies often lack sufficient textual content that effectively describe concepts. In this work, we use a corpus-based approach for generating descriptive texts for concepts (concept summaries) which are subsequently used for recommending annotations for documents. Our main contribution is the generation of concept summaries from a corpus which are sufficiently descriptive of the concepts in a thesaurus. A key process in our approach is the use of knowledge of semantic relatedness between the concepts of a thesaurus to identify the documents from which concept summaries are extracted. In our evaluation, we use generated concept summaries for recommending annotations and compare its performance to alternative approaches.

The remainder of this paper is organised as follows: section 2 reviews relevant literature; section 3 presents our corpus-based approach for generating concept summaries and recommending document annotations; section 4 is an experimental evaluation which compares our approach to alternative approaches; and section 5 concludes with an outline for future work.

2 Related work

We categorise the popular approaches for recommending annotations in the literature as either supervised or unsupervised methods as shown in Figure 1.

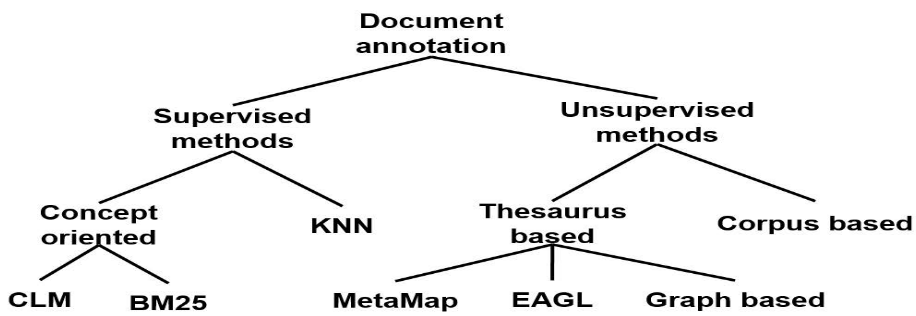

Fig. 1. Document annotation approaches.

The supervised methods reuse the annotations of previously tagged documents that share some similarity to the document being annotated. The intuition is that a new document can inherit some or all the annotations that were assigned to similar previously seen documents. Concept-oriented approaches generate concept summaries by merging all the documents that have been annotated with a concept. Concept summaries are then indexed so that a document to be annotated forms a query for which relevant summaries are retrieved. The corresponding concepts for top ranked summaries are recommended as annotations. Popular approaches for retrieving relevant concept summaries are CLM and BM25 [2]. CLM uses a language model for retrieval while BM25 uses the Okapi BM25 ranking function [3]. K-nearest neighbour (KNN) is a state-of-the-art supervised approach and is used in systems such as the Medical Text Indexer (MTI) [4, 5]. Instead of generating concept summaries, KNN adds each annotated document to an index. When a new document is to be annotated, it uses a document ranking function to retrieve K most similar documents. The annotations of retrieved documents form candidate concepts to be recommended. One variant of KNN ranks candidate concepts by combining the relevance scores of documents for which they form annotations [6, 7]. Another variant passes the features of candidate concepts to a machine classifier which determines which concepts to put forward for annotation [8]. Some features that are used by a classifier include the proportion of retrieved documents that were annotated with the concept and if the concept appears in the title or content of a document. Experimental results show that KNN or hybrids of it are most effective in recommending annotations [2, 8]. However, supervised approaches cannot be used when a corpus of annotated documents does not exist. It is also difficult to effectively recommend concepts that appear sparingly or are absent from the annotated corpus.

Unsupervised methods do not require an annotated corpus when recommending annotations. They rely on thesaurus-based indicators or resources that are generated from external sources. MetaMap parses the document to be annotated to identify exact and partial mappings to concept terms. Identified mappings form candidate concepts for annotation and using several linguistic principles, a ranking of concepts is generated [9]. Some considerations for ranking candidate concepts include the number of times it appears in a document and if it was a partial or complete match. Similar to the supervised concept-oriented approaches, EAGL generates concept summaries so that the concepts whose summaries are most similar to a document are recommended. The concept summaries are generated by merging the textual features of each concept (e.g. synonyms and descriptions). Concept summaries are indexed and a variety of retrieval approaches can be used to retrieve summaries (e.g. vector space model). Experiments show that EAGL performs better than MetaMap [2]. However, both approaches perform poorly when compared with the supervised methods. In MetaMap, the inability to disambiguate terms in documents was cited as one reason for poor performance [9]. EAGL is fast and efficient but controlled vocabularies often lack sufficient textual content to generate effective concept summaries.

In [10], documents are annotated with DBpedia concepts using a graph-based approach. First, the key terms in a document are identified and linked to corresponding DBpedia concepts. Titles of DBpedia entries form concept terms while corresponding textual contents provide textual context for disambiguating terms in documents. The DBpedia graph structure is then analysed to identify central nodes which connect the concepts that were linked to the document. These central concepts are used to annotate the document. Although the results are promising, this approach is suitable if the intent is to annotate with DBpedia or a similar knowledge resource with rich textual content. When using a different thesaurus, candidate concepts may not have equivalent DBpedia entries. This is especially true in specialised domains whose concepts may not have DBpedia entries. An analysis of the geoscience-related concepts used in this work showed that over 50% of the concepts have no corresponding DBpedia entries. Also, DBpedia often conflate concepts (e.g. "Rocks" and "Rock type" point to the same article). It may be desirable to maintain subtle differences in specialised domains. In this work, we adopt an approach that is similar to EAGL but augment thesaurus-based concept summaries with node descriptors from an external corpus. The use of external corpus such as Wikipedia has helped in generating additional useful information to aid the alignment/matching of concepts from different taxonomies [11].

3 Corpus-based concept summaries

We use an external corpus to generate concept summaries for the concepts in a thesaurus. When generating a concept's summary, other concepts in its neighbourhood are used for disambiguation forming a semantic filter which ensures that the summary generated is relevant to a concept. This relies on a taxonomic structure of the thesaurus to measure semantic relatedness between concepts. A high-level overview of the process for generating a concept summary is presented in Figure 2. We summarise the process in the following steps:

1. The concept term (textual label) is issued as a query to retrieve documents from a corpus. We refer to this concept term as *query concept*.
2. The documents retrieved in step 1 are mapped to the thesaurus to identify the concepts expressed in them. We refer to the set of concepts that are identified in a document as *document concepts*.
3. Each document in step 2 is re-ranked based on the semantic overlap between the query concept and document concepts. We use a semantic relatedness algorithm to measure semantic overlap.
4. The query-biased snippets of top ranked documents in step 3 are extracted and merged to form a concept summary.
5. Steps 1 - 4 is repeated for all the concepts in the thesaurus generating a corpus of concept summaries.

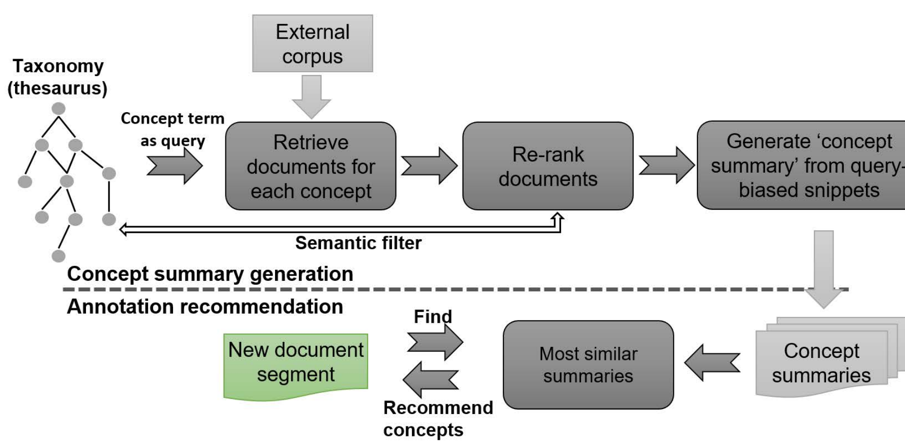

Fig. 2. Overview of concept summary generation which is used for annotation recommendation.

Generated concept summaries are indexed for use in recommending annotations. In order to annotate a document, the most similar concept summaries are retrieved using the BM25 ranking function which is a state-of-the-art term vector based model. Concepts are recommended in the order of ranking of their summaries. The remainder of this section describes the steps above in more detail.

3.1 Discover candidate source documents

Our intent is to generate concept summaries from documents which are most relevant to the concepts of a thesaurus. First, we identify a set of documents that are potentially relevant to each concept. Accordingly, each concept term is issued as query to a corpus that is assumed to contain documents which are relevant to the thesaurus. The documents that are retrieved for the query concept form candidate sources for generating its summary. Concept terms are often very short making it difficult to appropriately represent an information need. Due to reasons such as the presence of polysemous terms (e.g. rock: music or stone?), some of the documents that are retrieved for a concept may not be relevant. Therefore, we introduce a semantic re-ranking step to identify a subset of retrieved documents that we are more relevance to the concept.

3.2 Semantic re-rank of documents

Semantic re-rank measures the degree to which a document's concepts cluster about the query concept. The intuition is that a document's relevance increases as its concepts cluster closer to query concepts on the taxonomy. To identify document concepts, we match concept terms from the thesaurus to a keyword index of documents. Both concept terms and the keyword index are stemmed to maximise match discovery. Considering that there may be polysemous terms in the keyword index and the likely introduction of errors by conflating words through stemming, we impose the requirement that a document should also contain a *semantic context* of the concept before it is deemed present. The semantic context of a concept is the set of all concepts that are directly linked to it in the thesaurus [12]. For example, the semantic context of "rock" in a geological thesaurus may include "igneous rock" and "sedimentary rock". A document that describes the music genre "rock" is unlikely to contain those semantic contexts. The outcome of mapping documents to the thesaurus is a bag-of-concepts representation for each document.

Next, we estimate the semantic closeness of a document's concepts to the query concept by cumulating pairwise semantic relatedness measures. We use the Wu and Palmer algorithm [13] as shown in equation 1 to measure relatedness between concepts which correlates well with human judgments of relevance [14]. The algorithm preserves the specificity cost and specialisation cost properties which are important when comparing the nodes of a taxonomy. Specificity cost property requires that relatedness between neighbouring concepts increase with greater taxonomic depth while specialisation cost property requires that further specialisation implies reduced relatedness [15]. Wu and Palmer requires finding the most specific common subsumer (MSCS) of a concept pair being compared which is the most distant concept from the root node that subsumes them.

$$rel(c_i, c_j) = \frac{2 * n(c_i, c_j)}{n(c_i) + n(c_j) + 2 * n(c_i, c_j)} \quad (1)$$

where c_i and c_j are concepts being compared, $n(c_i)$ is minimum node count from c_i to the MSCS, $n(c_j)$ is minimum node count from c_j to the MSCS, and $n(c_i, c_j)$ is minimum node count from the MSCS to the root node.

In other words, the query concept forms a central node on the taxonomy from which document concepts are measured. Let x denote query concept and C_d denote the concepts of document d. A cumulation of pairwise semantic relatedness measures between x and C_d as shown in equation 2 determine the semantic relevance score of d.

$$semScore(d, x) = \sum_{x, c_i \in C_d} rel(x, c_i) \quad (2)$$

Afterwards, the documents that were retrieved for a query concept are sorted by semantic relevance scores.

3.3 Generate concept summaries

The final step of concept summary generation is the extraction of relevant content from top ranked documents. We generate and extract document *snippets* which are short textual summaries in search result listings for the purpose of determining relevance prior to viewing entire documents. We use a dynamic snippet generation approach that scores the sentences of a document with respect to a query and retrieves the most relevant sentences. This query-biased snippet generation approach has been shown to be effective in extracting useful document summaries and is adopted by several search engines [16]. We use the BM25 ranking function to identify relevant sentences for snippet generation. Snippets of top K documents are then merged to create a concept's summary.

4 Evaluation

We compare our approach for recommending annotations using corpus-based concept summaries (CCS) to alternative approaches in the literature.

4.1 Dataset and experiment setup

The evaluation dataset is from 1,948 document sections in 30 geological documents (mostly geology memoirs) which were manually annotated by domain experts in a project aimed at enhancing access content. These documents are book-like containing multiple sections[3]. Figure 3 is an example of a document section that is annotated with two concepts. The entries "value" and "scheme" refer to concepts and their source thesaurus respectively.

[3] An example of documents used http://pubs.bgs.ac.uk/publications.html?pubID=B01745

```xml
<section>
    <sectioninfo>
        <indexterm scheme="CHRONOSTRAT" value="KU"> </indexterm>
        <indexterm scheme="AMF" value="5283"> </indexterm>
    </sectioninfo>
    <title>MID-CRETACEOUS TO END CRETACEOUS REGIONAL SHELF SUBSIDENCE</title>
    <para>Crustal extension had effectively ceased by mid-Cretaceous times (e.
    established over the Manx region, with structural demarcation between bloc
    Ireland implies deposition of Chalk across the entire Manx region. The max
    </para>
</section>
```

Fig. 3. Example of an annotated document section.

We selected 3 controlled vocabularies that were used to annotate the documents – BGS Geoscience Thesaurus (THESAURUS)[4], BGS Geochronology (CHRONOSTRAT)[5] and BGS Lexicon of Named Rock Units (LEXICON)[6]. Concepts from these vocabularies were used 701 times (276 unique concepts) to annotate 397 document sections making an average of 1.8 concepts per document section. We use these concepts (110 from THESAURUS, 122 from LEXICON, and 44 from CHRONOSTRAT) and subset of document sections for our evaluation. We randomly select 2/3 of the dataset for training in supervised approaches and for parameter tuning, and we report results on remaining 1/3 for all approaches.

Wikipedia was used as the corpus for generating concept summaries and the vector space model (with BM25 ranking) for retrieving articles for query concepts. Specifically, we extracted a subset of Wikipedia (286,766 articles) that were tagged with one or more terms from the "Earth sciences" sub-category hierarchy. Articles from the Earth sciences category align with our evaluation dataset. We generated concept summaries from 5-sentence snippets of top 10 ranked articles as determined using the training dataset.

Alternative approaches for recommending annotations which we compare are:

- SUP_{BM25}: Supervised approach which generates concept summaries using the content of all documents that were annotated with a concept. BM25 ranking function ($k1 = 1.2$, $b = 0.75$) is used to identify summaries that are most similar to an unseen document from indexed concept summaries. The concepts are recommended in the order of the relevance scores of retrieved summaries.
- SUP_{LM}: Language model approach that is similar to SUP_{BM25}. We use a language model based on Dirichlet similarity ($\mu = 3500f$) to retrieve summaries from the concept summary index. Dirichlet similarity uses Dirichlet priors for Bayesian smoothing and is a popular language model retrieval approach.
- SUP_{KNN}: KNN approach that indexes annotated documents and notes the annotating concepts. When a new document is to be annotated, the most

[4] http://www.bgs.ac.uk/discoverymetadata/13603129.html
[5] http://data.bgs.ac.uk/doc/Geochronology.html
[6] http://data.bgs.ac.uk/doc/Lexicon.html

similar documents are retrieved from the index. The concepts that were used to annotate top K ($K = 10$) retrieved documents are ranked by summing the relevance scores of their respective documents. Although some previous works have used a language model for retrieval, BM25 ranking gave the best results which we report.
UNSUP$_{\mathrm{EAGL}}$: Unsupervised approach that generates concept summaries from concept terms, synonyms, definitions, and other textual content in a thesaurus. Concept summary indexing and the process for recommending annotations are similar to SUP$_{\mathrm{BM25}}$.
CCS$_{\mathrm{Lite}}$: A variant of our approach (CCS) which does not re-rank documents that are retrieved for a concept from the corpus. This enables us to evaluate the impact of semantic re-rank in CCS.

All document indexing and ranking functions were implemented on Elasticsearch using its Java API[7]. We use mean average precision (MAP), recall and F1 measures to compare approaches. MAP combines the precision and ranking quality of recommended annotations in a single performance measure making it easier to compare different systems (see equation 3). Precision is the proportion of recommended annotations that are correct. Recall is the proportion of correct annotations that are included in recommended annotations. We measure recall at 5, 7 and 10 top recommended concepts for annotation. F1 measure is the harmonic mean of precision and recall.

$$MAP = \frac{\sum_{d=1}^{|D|} AP(d)@n}{|D|} \tag{3}$$

D is the set of all documents being annotated, $AP(d)@n = \sum_{k=1}^{n} P(k)/min(m,n)$ is the average precision (AP) of recommended concepts for document d, $P(k)$ is the precision at position k of ordered concepts recommended, m is the number of relevant annotations for d, n is the maximum number of recommended concepts being evaluated. We set $n = 10$ for all approaches.

4.2 Results and discussion

The results of different methods for recommending annotation are presented in Table 1. CCS outperformed UNSUP$_{\mathrm{EAGL}}$ on all the evaluation metrics used. Also, results of CCS are not very far from those of supervised approaches. CSS outperformed CCS$_{\mathrm{Lite}}$ highlighting the utility of re-ranking documents before generating concept summaries. Unsurprisingly, the difference between CCS and CCS$_{\mathrm{Lite}}$ is minimal given that we used a subset of Wikipedia that is mostly relevant to the domain. We expect the impact of semantic re-ranking to be more pronounced when using a more diverse corpus where there is greater possibility of encountering polysemous terms.

[7] Elasticsearch Java API http://www.elastic.co/guide/en/elasticsearch/client/java-api/5.2

Table 1. Mean average precision (MAP), recall (R) and F-measure (F1) of the approaches that we compared for recommending document annotations.

	MAP	F1@5	R@5	R@7	R@10
$\mathrm{SUP}_{\mathrm{BM25}}$	0.2967	**0.2489**	**0.4935**	0.5176	**0.5691**
$\mathrm{SUP}_{\mathrm{LM}}$	0.2632	0.2192	0.4508	0.4874	0.5177
$\mathrm{SUP}_{\mathrm{KNN}}$	**0.3093**	0.2412	0.4767	**0.5192**	0.5387
$\mathrm{UNSUP}_{\mathrm{EAGL}}$	0.2221	0.1258	0.3749	0.4007	0.4394
CCS	0.2647	0.2045	0.4419	0.4860	0.5345
$\mathrm{CCS}_{\mathrm{Lite}}$	0.2469	0.2074	0.4409	0.4754	0.5157

The F1 measures are low for all the methods because there are only few concepts that annotate each document. As an illustration, consider a document that is annotated with one concept which is correctly included in the top 5 recommendations. The precision will be 0.2 (1/5) and F1 0.33. This appears to be low even though the correct annotation was recommended. Recall value is more relevant in this case since it shows the proportion of the correct annotations that an approach was able to discover. The choice of concepts for annotating documents is quite subjective and attaining high recall values remain a challenge [8].

The results that are obtained for the other approaches in our evaluation mostly agree with previous comparisons [2]. The performance of $\mathrm{UNSUP}_{\mathrm{EAGL}}$ was weak due to insufficient textual content in the controlled vocabularies. CHRONOSTRAT and THESAURUS describe only concept terms and synonyms (or alternative spellings), while LEXICON includes some descriptive text. As expected, the supervised approaches performed strongly. $\mathrm{SUP}_{\mathrm{BM25}}$ was overall best in retrieving the right annotations as shown in recall values. $\mathrm{SUP}_{\mathrm{KNN}}$ ranked correct concepts slightly better as MAP values indicate.

In Figure 4, we show the performance of different approaches as the proportion of training and test dataset vary. The training dataset simulates the proportion of documents that were annotated prior to recommending annotations for test documents. A random function was used to split the documents and recall (R@5) in the figure show performances on the test dataset. The performance of CCS remained fairly similar and was only outperformed by the supervised approaches when over 50% of the dataset was used for training.

Although the supervised approaches may be better at recommending annotations, unsupervised approaches remain relevant for generating an initial set of annotated corpus. Supervised and unsupervised approaches are usually combined to form hybrid document annotation systems.

5 Conclusion

In this work, we introduced a corpus-based approach for generating descriptive textual content (concept summaries) for the concepts of a thesaurus. We used

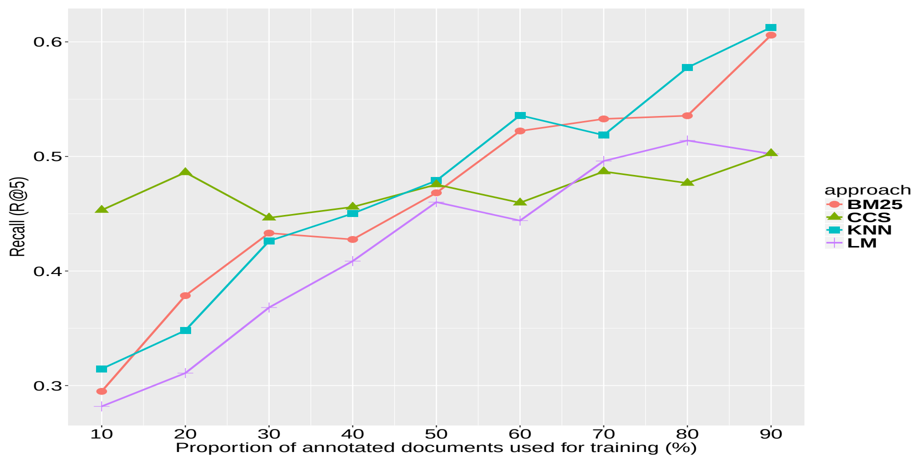

Fig. 4. Performances (R@5) with varying proportion of dataset used for training.

semantic knowledge from the thesaurus to identify the best documents for generating concept summaries. Concept summaries were then used to recommend annotations for documents. Our goal was to overcome the limitations of unsupervised thesaurus-based approaches which suffer from insufficient descriptive texts for effective use in recommending document annotations. Evaluation using a manually annotated corpus showed that this objective was achieved and that our results were somewhat comparable with the supervised approaches.

Future work will explore alternative ways of using concept summaries for recommending annotations. For example, the graph-based approach in [10] can be applied using concept summaries as DBpedia articles. Also, an approach that is based on explicit semantic analysis (ESA) will be explored. Instead of recommending annotations based on the similarity of term vectors, the ESA approach utilises concept vectors. Finally, we have assumed an independence of document sections by treating them as separate documents. In reality, the rest of the document from which a section is extracted can provide useful information for determining the right concepts for annotating the section.

Acknowledgement

This work is partly funded by the British Geological Survey (BGS) through the BGS University Funding Initiative (BUFI). We are grateful for the valuable comments of our reviewers.

References

1. Berlanga, R., Nebot, V., Pérez, M.: Tailored semantic annotation for semantic search. Web Semantics: Science, Services and Agents on the World Wide Web (2014)

2. Trieschnigg, D., Pezik, P., Lee, V., De Jong, F., Kraaij, W., Rebholz-Schuhmann, D.: MeSH Up: Effective MeSH text classification for improved document retrieval. Bioinformatics **25**(11) (2009) 1412–1418
3. Robertson, S.E., Walker, S., Beaulieu, M., Gatford, M., Payne, A.: Okapi at TREC-4. NIST Special Publication SP (1996) 73–96
4. Große-Bölting, G., Nishioka, C., Scherp, A.: A comparison of different strategies for automated semantic document annotation. In: Proceedings of the 8th International Conference on Knowledge Capture, ACM (2015) 8
5. Aronson, A.R., Mork, J.G., Gay, C.W., Humphrey, S.M., Rogers, W.J.: The NLM indexing initiative's Medical Text Indexer. Medinfo **11**(Pt 1) (2004) 268–72
6. Giannopoulos, G., Bikakis, N., Dalamagas, T., Sellis, T.: Gontogle: a tool for semantic annotation and search. The Semantic Web: Research and Applications (2010) 376–380
7. Huang, M., Névéol, A., Lu, Z.: Recommending MeSH terms for annotating biomedical articles. Journal of the American Medical Informatics Association **18**(5) (2011) 660–667
8. Dramé, K., Mougin, F., Diallo, G.: Large scale biomedical texts classification: A kNN and an ESA-based approaches. Journal of Biomedical Semantics **7**(1) (2016) 40
9. Aronson, A.R.: Effective mapping of biomedical text to the UMLS Metathesaurus: The MetaMap program. In: Proceedings of the AMIA Symposium, American Medical Informatics Association (2001) 17
10. Hulpus, I., Hayes, C., Karnstedt, M., Greene, D.: Unsupervised graph-based topic labelling using DBpedia. In: Proceedings of the Sixth ACM International Conference on Web Search and Data Mining, ACM (2013) 465–474
11. Hertling, S., Paulheim, H.: Wikimatch: using wikipedia for ontology matching. In: Proceedings of the 7th International Conference on Ontology Matching-Volume 946, CEUR-WS. org (2012) 37–48
12. Fernández, M., Cantador, I., López, V., Vallet, D., Castells, P., Motta, E.: Semantically enhanced Information Retrieval: An ontology-based approach. Web Semantics: Science, Services and Agents on the World Wide Web **9**(4) (2011) 434–452
13. Wu, Z., Palmer, M.: Verbs semantics and lexical selection. In: Proceedings of the 32nd Annual Meeting on Association for Computational Linguistics, Association for Computational Linguistics (1994) 133–138
14. Hliaoutakis, A., Varelas, G., Voutsakis, E., Petrakis, E.G., Milios, E.: Information retrieval by semantic similarity. International Journal on Semantic Web and Information Systems **2**(3) (2006) 55–73
15. Knappe, R., Bulskov, H., Andreasen, T.: Perspectives on ontology-based querying. International Journal of Intelligent Systems **22**(7) (2007) 739–761
16. Leal Bando, L., Scholer, F., Turpin, A.: Query-biased summary generation assisted by query expansion. Journal of the Association for Information Science and Technology **66**(5) (2015) 961–979