Proceedings of the 2007 ASME International Mechanical Engineering Congress & Exposition

IMECE07

November 11-15, 2007, Seattle, Washington, USA

IMECE2007-42531

ABET ALIGNMENT OF LEARNING RESOURCES IN THE ENGINEERING PATHWAY DIGITAL LIBRARY

Jia-Long Wu, Department of Mechanical Engineering, UC Berkeley

Alice Agogino, Department of Mechanical Engineering, UC Berkeley

ABSTRACT

The *Engineering Pathway* (EP) digital library (www.engineeringpathway.com) strives to provide quality educational resources for learners of all age levels. ABET Engineering Criteria has been the driving force behind modern engineering education reform since its introduction at the turn of the century. In order to help engineering educators and administrators meet the challenges of developing and teaching a learning outcomes-focused curriculum, *EP* is linking existing resources to appropriate ABET criteria. This paper summarizes the research behind using our ABET alignment process where computational linguistics and information retrieval tools are used to augment the ABET alignment process. Experts then review these recommendations and make corrections where needed. By taking this approach, we not only shorten the time to align existing learning resources; but also improve the scalability by aligning new resources as they are being submitted. The technologies can also be applied to the development of thesauri and recommender systems that can be tailored to individual faculty needs.

1 INTRODUCTION

The *Engineering Pathway* (*EP*) digital library (www.engineeringpathway.com) [6] is a portal to high-quality teaching and learning resources in applied science and math, engineering, computer science/information technology and engineering technology, for use by K-12 and university educators and students. Funded by the National Science Foundation (NSF), it is the engineering "wing" of NSF's National Science Digital Library (NSDL). EP combines the learning resources from the NEEDS (National Engineering Education Digital-library System) [12] and the K-12 *TeachEngineering* digital library [21] EP was developed to provide learning resources that would meet the educational needs of the twenty-first century. EP is using as a strategic guide in a report recently issued by the National Academy of Engineering – *Educating the Engineer of 2020: Adapting Engineering Education to the New Century* [11]. Realizing the importance of lifelong learning in today's engineering education, *EP* strives to not only provide quality learning resources for learners of all age levels (with over 10,000 K-12, higher education and professional development resources); but also serves as the information hub for educators and school administrators to find and exchange tips and materials for advancing their engineering courses and curricula. In addition to the regular digital library features of searching and browsing resources, EP has built portal pages for educators to find tips on teaching and designing courses, and for administrators to find resources on curriculum planning and accreditation. A screenshot of EP is shown in Fig. 1.

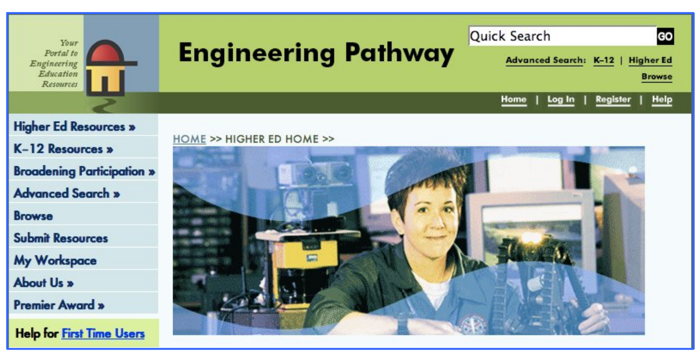
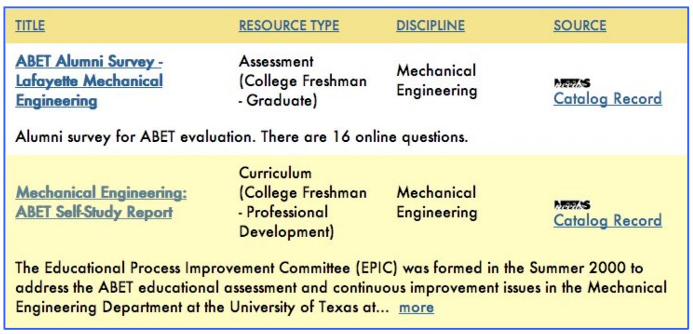


Fig. 1. Higher Education

In order to help engineering educators and administrators meet the challenges of developing and teaching a learning outcomes-focused curriculum, *EP* is building an ABET Accreditation Series collection that will serve as a repository of exemplars to engineering accreditation. *EP* is currently working with ABET and trained ABET reviewers on obtaining self-study reports containing exemplars, disaggregating these reports and creating appropriate metadata records for communicating the resource to engineering educators. Fig. 2 shows an example of some of the resources in this ABET self-study collection. In addition to developing the ABET collection, *EP* was challenged with how to link its large legacy of approximately 10,000 existing resources to appropriate ABET learning outcomes criteria. Table 1 shows a list of resources currently cataloged in *EP* for each ABET accredited discipline.



TITLE	RESOURCE TYPE	DISCIPLINE	SOURCE
ABET Alumni Survey - Lafayette Mechanical Engineering	Assessment (College Freshman - Graduate)	Mechanical Engineering	Catalog Record
Alumni survey for ABET evaluation. There are 16 online questions.			
Mechanical Engineering: ABET Self-Study Report	Curriculum (College Freshman - Professional Development)	Mechanical Engineering	Catalog Record
The Educational Process Improvement Committee (EPIC) was formed in the Summer 2000 to address the ABET educational assessment and continuous improvement issues in the Mechanical Engineering Department at the University of Texas at... more			

Fig. 2. Example ABET Self-Study Resources in the Engineering Pathway Digital library.

Improving the efficiency of document retrieval from digital document repositories is a key research issue in information science. Use of keyphrases to assist the information retrieval processes has been proven to increase the retrieval efficiency of the system. Fagan [7] and Mitra [10] both show that phrase-based indexing helps to increase the precision of the overall retrieval, especially at lower relevance ratings.

Finding a set of key phrases for a document repository that is suitable for both indexing and document retrieving is a challenge for large-scale systems. Whereas small document repositories can rely on human expertise in manually scanning their documents and selecting the most useful phrases, exclusive use of human tagging is not feasible for large document repositories, such as the *EP* digital library. In addition, prior work has shown that human tagging of keyphrases in "keyword" fields often leaves out key contextual phrases that are needed for consolidated repositories [22]. Thus there is a motivation to find computational support for identifying effective key phrases from digital document repositories.

We summarize the research behind our ABET alignment process where computational linguistics and information retrieval tools are used to augment human expertise for ABET alignment of resources in the *Engineering Pathway* digital library.

Table 1. Engineering Education Learning Resources Cataloged in EP Organized by ABET Discipline. The number of resources in each category listed in brackets.

Computing [1,622]

- Computer Science [1,470]
- Information Systems [118]
- Information Technology [233]

Engineering [7,246]

- Aerospace Engineering [200]
- Agriculture Engineering [200]
- Architecture Engineering [146]
- Bio/Bio-Medical Engineering [261]
- Ceramic Engineering [99]
- Chemical Engineering [361]
- Civil Engineering [344]
- Computer Engineering [1,235]
- Construction Engineering [77]
- Electrical Engineering [557]
- Engineering Management [146]
- Engineering Mechanics [97]
- Environmental Engineering [399]
- General Engineering, Engineering Science [2,525]
- Geological Engineering [137]
- Industrial Engineering [227]
- Manufacturing Engineering [179]
- Materials Engineering [304]
- Mechanical Engineering [763]
- Mineral and Mining Engineering [93]
- Naval Architecture and Marine Engineering [75]
- Nuclear Engineering [100]
- Ocean Engineering [95]
- Petroleum Engineering [72]
- Software Engineering [166]
- Surveying and Geomatics Engineering [63]

2 THEORETICAL FRAMEWORK AND LITERATURE REVIEW

The research behind our ABET alignment process builds on prior research in computational linguistics, specifically keyphrase extraction and information retrieval. A brief summary of the theoretical underpinnings follows.

Several key phrase extraction techniques have been proposed and implemented successfully in different contexts. Some exploit the syntactic nature in languages [4] and others approach the problem by studying the statistical significance of the key phrases in the document sets they reside [9,13]. Bookstein and Picard have both found that keyphrases are often clustered together and appear in a similar context with other key phrases more often than not. Bookstein [4] proposes two measures to evaluate the candidate keyphrases. The *condensation clustering* measure reflects the degree of "clumping" of candidate phrases in the context and the *linear clustering measure* recognizes the patterns of textual units within which the candidate phrases reside. Results from Picard's paper [14] suggest that content-bearing key phrases display higher first and second order similarity between one another and are less similar to the noise phrases.

Wu [22] proposed an automated keyphrase extraction algorithm using a non-dominated sorting multi-objective genetic algorithm [5] that uses the "clumping" property of keyphrases proposed by Bookstein [4] to formulate the multiple objectives. The objective was to find the smallest phrase set that has the best precision, as measured by average condensation clustering. Keyphrases were retrieved from a collection of mechanical design conference papers and the results were presented to domain experts for evaluation. The automatically generated keyphrase list was compared to author-assigned keyphrases. Ninety percent of the generated phrases were deemed appropriate for use in a thesaurus for engineering design. Of note is that 80% of the author-assigned keyphrases used in the documents showed up in the generated list, leaving 20% that were considered effective phrases by experts, but not included in the author list. This implies that keyphrase identification can be used to both reduce the workload involved, but can also expand the list beyond that of human experts.

Unfortunately, the corpus of documents for the ABET alignment was not large enough to justify the full procedure proposed by Wu [22]. However, we do use the results to motivate the use of automatic keyphrase generation to augment those generated by human experts. In this paper, we use keyphrases generated by both human experts and computation to compose an initial list of keyphrases. The QTag probabilistic part-of-speech (POS) tagger [17] is used to tag the ABET learning outcome criteria from ABET EC 2000 documents [1]. Heuristic rules are then applied to extract all possible noun phrases from the documents automatically. This combined list of human and computer-generated keyphrases is used in the Apache Lucene [3] full text search algorithm on the *EP* corpus of metadata. The keyphrase extraction algorithm is then applied to this expanded corpus of documents, thus adding to the list of keyphrases generated. This recursive process is called *relevance feedback* in the information retrieval literature [8,23].

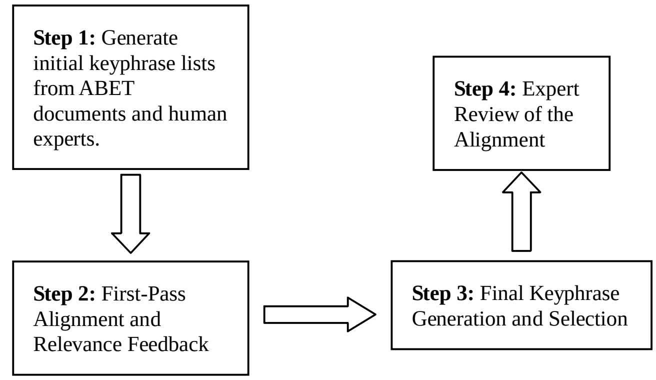


Fig. 3. Process Flow Chart

3 METHODOLOGY AND IMPLEMENTATION

There are four major steps in aligning higher education engineering education resources in EP with ABET learning outcome criteria. Details on each step are described in the following sub-sections and the complete flow chart of the alignment process is show in Figure 3.

3.1 Step 1: Seed Keyphrase List Generation

To generate an initial keyphrase list for ABET alignment, we took ABET EC2000 documents as the corpus and applied a probabilistic part-of-speech tagger called QTag [17] on the corpus. The lexicon we used with QTag was trained using about a million English words. We have found it to be accurate in identifying the correct part-of-speech for the words when used it in previous experiments.

With the corpus properly tagged with part-of speech information, we applied heuristic rules to extract noun phrases as our initial keyphrase list from the corpus. Past research [7] has shown that noun phrases are good representations of the content they reside in and hence, they are generally good content-bearing phrases, or keyphrases as we refer to them in this paper. Parsing single word noun phrases from the tagged corpus is straight-forward, but a procedure is needed to identify multi-word noun phrases. Table 2 shows the rules we used for parsing multi-word noun phrases. For example, "NN" meaning two noun words together will form a new two-word noun phrase. Part-of-speech tags used in QTag is also listed in Table 2 for reference.

Because we focused on ABET documentation for learning outcomes general criteria (a)-(k), the corpus of documents was quite small. Hence, the keyphrases that could be found solely through the computational linguistic approach outlined in section 3.1 was not sufficient for aligning EP resources with ABET criteria.

To augment the generated list, we asked human experts who are familiar with ABET EC 2000 to identify additional keyphrases. These keyphrases were then merged with those generated by computation linguistic algorithms as the "seed" keyphrases for use in the next step.

Table 2. Noun Phrase Extraction Rules and POS Tag List

Noun Phrase Type	**Abbreviation Definitions**
NN NN	NN *noun, common singular (action)*
NN NNS	NNS *noun, common plural (actions)*
OD NN	OD *number, ordinal (fourth)*
OD NNS	JJ *adjective, general (near)*
JJ NN	JJR *adjective, comparative (nearer)*
JJ NNS	JJS *adjective, superlative (nearest)*
JJR NN	RB *adverb, general (chronically, deep)*
JJR NNS	RBR *adverb, comparative (easier, sooner)*
JJS NN	RBS *adverb, superlative (easiest, soonest)*
JJS NNS	
RB JJ	
RBR JJ	
RBS JJ	
RB JJR	
RBR JJR	
RBS JJR	
RB JJS	
RBR JJS	
RBS JJS	

3.2 Step 2: First-pass Alignment and Relevance Feedback

We then used the Lucene full text search engine [3] to match these "seed" keyphrases to all documents in the EP digital library for each of the ABET general criteria (a)-(k). This first-pass attempt at aligning *EP* resources with ABET general criteria yielded too many "false positive" matches – the initial keyphrase list lacked precision. For example the use of keyphrase "laboratory experience" or just "experiments" for criterion (b) "Design and conduct experiments, analyze and interpret data" did identify a number of relevant learning resources, but it also identified many more general resources for government or research laboratories that were not particularly targeted for teaching and learning.

To improve the precision of the results, we applied a relevance feedback technique used in information retrieval to improve the quality of the keyphrase lists. The idea of relevance feedback is to refine or expand the initial query with the initial search results that are relevant. There are several models of relevance feedback and the most common one is often referred to as *explicit feedback*, where the relevancy of the returned documents is obtained directly from expert users. Users may simply decide if the document is related to their information needs, or they could even rank them with scores in some scales. Research has shown that this approach could improve the results precision by 40 to 60 percent [18,19]. However, this requires significant amount of work by the human experts.

The approach we took uses a variant of the relevance feedback algorithm called *pseudo feedback* that does not require human feedback. The idea behind this algorithm is that the top ranked results computed are assumed to be relevant to the initial query and are used to refine or expand the search, thus using the relevance ranking of the search engine to give feedback. Past research has shown that this algorithm is highly effective in some settings, especially with large dataset [2]. It also has the benefit of not requiring any human feedback and it matches well with our goal of minimizing intervention of human experts. The number of top ranked results to use in relevance feedback is a difficult problem in itself. In this paper, we took the top 50 percent of the initial search results if the number of resources returned by the initial query was more than 200. For search results with less than 200 resources, we took the top 100 resources.

3.3 Step 3: Final Keyphrases Generation and Selection

Using the top ranked *EP* resource candidates from Step 2, we then applied the same noun phrase extraction algorithm described in section 3.1 to the corpus comprised of abstracts of these *EP* resources. Human experts familiar with ABET (a)-(k) criteria selected new keyphrases from this pool of generated noun phrases. This final list was larger, yet more precise than the list generated in Steps 1 and 2. For example "laboratory experience" was removed from the list for criterion (b) and replaced with more specific terms, such as "laboratory safety", "laboratory instruction, "laboratory principles", "product testing", "laboratory protocols", "reliability testing", etc.

3.4 Step 4: Expert Review of the Alignment

The final lists of keyphrases were then fed into the Lucene full text search engine to retrieve *EP* resources that served as candidates for alignment to respective ABET (a)-(k) criteria. Two human experts reviewed the list for each criterion and the resulting resources were aligned accordingly in the *Engineernig Pathway* digital library.

4 RESULTS

The results in Table 3 show the number of keyphrases generated in Steps 1 and 3, along with the increase between steps for all criteria except (a) – "apply mathematics, science, and engineering", as this one was judged to be too general for tagging in the *EP* digital library. The editors of the library were concerned that most of its resources could be related to this criterion and thus reserved alignment of (a) to targeted resources that were developed explicitly to show the real world applications of basic mathematics and physics, such as *Project Links* at the Rensselaer Polytechnic Institute [16].

The average increase in keyphrases was 355%, with the largest increase at 1,640% for criterion (h) – "understanding global, economic, environmental, and societal context of

Table 3. Number and Change of Keyphrases in Steps 1 and 3 for ABET Learning Outcomes – General Criteria (a)-(k).

ABET Criterion	Number of Step 1 Keyphrases	Number of Step 3 Keyphrases	Increase in Number of Keyphrases
b	8	35	338%
c	28	195	596%
d	10	24	140%
e	6	12	100%
f	9	18	100%
g	9	10	11%
h	10	174	1,640%
i	2	6	200%
j	4	23	475%
k	8	4	-50%
Total	94	501	
Average			**355%**

Table 4. Number of *EP* Resources Identified in Steps 3 and 4 for ABET Learning Outcomes – General Criteria (a)-(k).

ABET Criterion	*EP* Resources Identified in Step 3	*EP* Resources Accepted after Step 4 Review	Precision Rate
b	284	224	79%
c	929	607	65%
d	132	113	86%
e	87	65	75%
f	209	175	84%
g	125	103	82%
h	388	286	74%
i	55	28	51%
j	169	143	85%
k	90	63	70%
Total	2,468	1,807	
Average			**75%**

engineering. There was one category in which the number of keyphrases at the end of Step 3 was actually reduced – criterion (k) which focuses on use of modern engineering tools in engineering practice. The more general keyphrases, such as "professional education" yielded imprecise results and were pruned from the final list.

Table 4 shows the number of *EP* learning resources identified as candidates for alignment at Step 3 for each ABET criteria (a)-(k). A total of 2,468 were identified as candidates in Step 3, reduced to 1,897 by the human evaluators in Step 4. Although the overall average acceptance rate was 75%. the following criteria had acceptance rates over 80%:

- (d) Work effectively in multi-disciplinary teams
- (f) An understanding of professional and ethical responsibility
- (g) Communicate effectively
- (j) Integrate knowledge of contemporary issues

Criterion (i) "Engage in life-long learning " had the lowest precision with an acceptance rate of 51%. This criterion also had the lowest number of resources identified in *EP* overall as well.

5 EVALUATION

One *Engineering Pathway* collection is an initiative of the National Academy of Engineering's Center for the Advancement of Scholarship on Engineering Education (CASEE) – the *Peer Reviewed Research Offering Validation of Effective and Innovative Teaching (PR2OVE-IT)* collection [15]. *PR2OVE-IT*

Table 5 Evaluation Results with PR2OVE-IT Resources in EP

ABET Criterion	*PR2OVE-IT* Editor-Alignment	Editor-Aligned Resources Identified in Step 3	Recall Rate (%)	All *PR2OVE-IT* Resources Identified in Step 3
b	11	10	91%	57
c	31	25	81%	136
f	2	2	100%	12
j	2	2	100%	8

aims to consolidate and characterize scholarly research publications in engineering education to provide engineering faculties a venue for consulting current and innovative teaching practices. *PR2OVE-IT* currently has cataloged over 500 publications in engineering education research and its editors have characterized them by interventions/ instructional practices and student learning outcomes.

In April 2007, *EP* acquired the catalog records of all *PR2OVE-IT* publications. As the *PR2OVE-IT* publications were categorized properly by the ABET general criteria 3(a)-(k), it provides an ideal test data set for our ABET alignment work. In Table 5 we show the number of *PR2OVE-IT* resources that were found to be candidates for alignment with each of the ABET criteria and the precision rate of those that were already characterized as one of the criteria by human editors. The ABET criteria tested were those provided in the *PR2OVE-IT* collection at the time of the analysis.

The recall rates of pre-characterized *PR2OVE-IT* resources are high across the board, meaning the keyphrases generated through our approach could accurately identify most of the ABET-aligned resources in the collection. The keyphrases also helped us identify additional *PR2OVE-IT* resources that might be investigated further for alignment with ABET criteria. Based on an average precision rate of over 90% in Table 4, we are confident that our approach does generate ABET alignment that otherwise would be lost. Further validation of these additional ABET alignments is planned to be one of our next steps in expanding this research.

6 CONCLUSIONS AND FUTURE RESEARCH

The *Engineer 2020* report issued by NAE highlighted the challenges we are facing today in building quality curricula and courses in engineering education. The *Engineering Pathway* digital library strives to help educators and school administrators meet these challenges by building an extensive collection of digital engineering education resources and community services to find and exchange tips and materials for improving engineering courses and curricula. One of the key elements in *EP* is aligning its resources to ABET Engineering Criteria 2000, which has been a major driving force behind engineering curriculum reform since its introduction.

Aligning the existing 10,000 resources one-by-one in *EP* would take enormous efforts and would not be scalable for the few trained ABET reviewers available to us. In this paper we fused computational linguistic algorithms and human expert knowledge to build a set of keyphrases for ABET criteria to help us identify potential ABET alignments, narrowing the candidate list down to a much smaller number for our reviewers to evaluate. Although the total number of resources was around 10,000 the largest review occurred for criterion (c) Design a system, component, or process, where 929 resources were reviewed to yield 607 alignments. Although a large number, it cut the number of reviews required down by a factor of ten. The number of reviews required for the other criteria was substantially smaller.

We used QTag POS tagger and heuristic rules to find noun phrases from a corpus of ABET EC 2000 documents and combined those with phrases recommended by human experts to form the initial set of ABET criteria keyphrases. The first-pass alignments were made by feeding “seed” keyphrases into the Lucene full-text search engine to find related *EP* resources for each of the ABET criteria. Pseudo-relevance feedback was utilized on the abstracts of the resources in the first-pass alignments to improve the quality of the ABET keyphrases. Human experts then reviewed and finalized the keyphrase lists for all criteria. Final ABET alignment recommendations were made and the results were evaluated by the experts.

We found that the precision rate of the ABET alignment recommendations made by this algorithm are generally high. On average, 75% of the alignments were deemed acceptable with some criteria having 80% or higher acceptance rate. This is an extremely high success rate by information retrieval standards. The results were also encouraging when we compared the recommended alignments with the expert-tagged alignments for selected resources in the *PR2OVE-IT* collection. Our algorithm managed to recall over 90% of those previously tagged for ABET alignment by its editors. We were also able to identify additional resources from this collection that might be re-reviewed for additional alignment. Table 6 shows the statistics of EP learning resources that are aligned to EC2000 program outcomes a through k, including the alignments found through this research. The complete list of aligned resources can be accessed through “Browse Learning Resources” page on *EP* at http://www.engineeringpathway.com/ep/browse/abet_criteria/.

The final keyphrases identified also have value for other features in the *Engineering Pathway* digital library. They could be used to create a thesaurus that could be used for adding keywords or for browsing of resources. They might even be used during the cataloging process to recommend possible ABET-alignment as new catalog records are being created.

The general approach outlined in this paper could be used for indexing and alignment of resources in any large text-based data set where human expertise is required but needs to be minimized, including large enterprise information systems or product design archives [20].

There are a number of research questions left open. The pseudo-relevance feedback used a 50% cut-off in search results from the Lucene search prioritization algorithm. We plan on experimenting with different cutoff points to optimize the impact on the quality of the final keyphrases. Setting the bar higher would likely yield higher precision and lower recall, and vice versa.

The ultimate test of validity for the project will be in the evaluation of engineering faculty who are looking for learning resources to help them in lectures and curriculum development.

We will monitor use patterns of online activity and conduct surveys at ABET-related workshops.

Table 6. Engineering Education Learning Resources Cataloged in EP Aligned to ABET Program Outcomes and Assessment. The number of resources in each category listed in brackets.

(a) Apply mathematics, science, and engineering [162]
(b) Design and conduct experiments, analyze and interpret data [348]
(c) Design a system, component, or process [798]
(d) Work effectively in multi-disciplinary teams [219]
(e) Identify, formulate, and solve engineering problems [126]
(f) Understand professional and ethical responsibility [221]
(g) Communicate effectively [164]
(h) Understand global, economic, environmental, and societal context [418]
(i) Engage in life-long learning [61]
(j) Integrate knowledge of contemporary issues [175]
(k) Use modern engineering tools in engineering practice [117]

ACKNOWLEDGMENTS

This research was partially funded by the National Science Foundation Grant DUE-0532922, under the NSDL (National Science Digital Library) Program. The authors wish to thank Jonathan Hey and John Prados for helping out in the final evaluation of the candidate resources for alignment in Step 4.

REFERENCES

1. ABET: http://www.abet.org
2. J. Allan, "Relevance feedback with too much data," Proc. of the 18th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, Seattle, WA, USA, 1995, pp. 337-343.
3. Apache Lucene: http://lucene.apache.org/index.html
4. Bookstein, S.T. Klein, T. Raita, "Clumping Properties of Content-Bearing Words," Journal of the American Society for Information Science, 49(2), 1998, pp. 102-114.
5. K. Deb, S. Agrawal, A. Pratap and T. Meyarivan, "A Fast Elitist Non-dominated Sorting Genetic Algorithm for Multi-objective Optimization: NSGA-II," Proceedings of the Parallel Problem Solving from Nature VI (PPSN-VI), 2000, pp. 849-858.
6. EngineeringPathway: http://www.engineeringpathway.com
7. J.L. Fagan, "The Effectiveness of a Nonsyntactic Approach to Automatic Phrase Indexing for Document Retrieval," Journal of the American Society for Information Science, 40(2), 1989, pp. 115-132.
8. F.A. Grootjen and Th.P. van der Weide, "Conceptual Relevance Feedback," Proceedings of the 2002 IEEE International Conference on Systems, Man and Cybernetics, (NLPKE 2002), Hammamet, Tunis, 2002.
9. B., Krulwich, & C., Burkey, "The Infofinder Agent-Learning User Interests through Heuristic Phrase Extraction," IEEE Intelligent Systems & Their Applications, 12(5), 1997, pp. 22-27.
10. M. Mitra, C. Buckley, A. Singhal and C. Cardie, "An Analysis of Statistical and Syntactic Phrases," Proceedings of RIAO97, Computer-Assited Information Searching on the Internet, Montreal, Canada, 1997, pp. 200-214.
11. National Academy of Engineering (NAE), Educating the Engineer of 2020: Adapting Engineering Education to the New Century , National Academy Press, 2005. Committee Report).
12. NEEDS – National Engineering Digital library System: http://www.needs.org.
13. Nevill-Manning, I. Witten and G. Paynter, "Browsing in Digital Libraries: a Phrase-Based Approach," Proceedings of 1997 International Conference on Digital Libraries, 1997, pp. 230-236.
14. J. Picard, "Finding content-bearing terms using term similarity," Proceedings of Ninth Conference of the European Chapter of the Association for Computational Linguistics, Bergen, Norway, 1999.
15. PR2OVE-IT: http://www.pr2ove-it.org
16. Project Links: http://links.math.rpi.edu/
17. QTag Algorithm: http://www.english.bham.ac.uk/staff/omason/software/qtag.html
18. I. Ruthven and M. Lalmas, "A Survey on the Use of Relevance Feedback for Information Access Systems,"
19. G. Salton. Automatic Text Processing: The Transformation, Analysis, and Retrieval of Information by Computer. Addison-Wesley, Reading, MA, 1989.

20. S. Song, A. Dong and A. M. Agogino, “Modeling Information Needs in Engineering Databases Using Tacit Knowledge,” Journal of Computing and Information Science in Engineering, 2, No. 3, pp. 199-207, 2002.

21. TeachEngineering: http://www.teachengineering.com

22. J.-L. Wu and A.M. Agogino. Automating Keyphrase Building with Multi-Objective Genetic Algorithms. Proceedings of the Hawaii International Conference on System Science, HICSS, 2003.

23. J. Xu and W. B. Croft, "Query Expansion Using Local and Global Document Analysis,” Proceedings of the 19th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 4-11, 1996.