emerald**insight**

Program
An algorithm for suffix stripping
M.F. Porter,

Article information:

Users who downloaded this article also downloaded:

(2006),"The Porter stemming algorithm: then and now", Program, Vol. 40 Iss 3 pp. 219-223 https://doi.org/10.1108/00330330610681295

(1972),"A STATISTICAL INTERPRETATION OF TERM SPECIFICITY AND ITS APPLICATION IN RETRIEVAL", Journal of Documentation, Vol. 28 Iss 1 pp. 11-21 https://doi.org/10.1108/eb026526

For Authors

If you would like to write for this, or any other Emerald publication, then please use our Emerald for Authors service information about how to choose which publication to write for and submission guidelines are available for all. Please visit www.emeraldinsight.com/authors for more information.

About Emerald www.emeraldinsight.com

Emerald is a global publisher linking research and practice to the benefit of society. The company manages a portfolio of more than 290 journals and over 2,350 books and book series volumes, as well as providing an extensive range of online products and additional customer resources and services.

Emerald is both COUNTER 4 and TRANSFER compliant. The organization is a partner of the Committee on Publication Ethics (COPE) and also works with Portico and the LOCKSS initiative for digital archive preservation.

*Related content and download information correct at time of download.

An algorithm for suffix stripping

M.F. Porter

Computer Laboratory, Corn Exchange Street, Cambridge

ABSTRACT

The automatic removal of suffixes from words in English is of particular interest in the field of information retrieval. An algorithm for suffix stripping is described, which has been implemented as a short, fast program in BCPL. Although simple, it performs slightly better than a much more elaborate system with which it has been compared. It effectively works by treating complex suffixes as compounds made up of simple suffixes, and removing the simple suffixes in a number of steps. In each step the removal of the suffix is made to depend upon the form of the remaining stem, which usually involves a measure of its syllable length.

1. INTRODUCTION

Removing suffixes from words by automatic means is an operation which is especially useful in the field of information retrieval. In a typical IR environment, one has a collection of documents, each described by the words in the document title and possibly the words in the document abstract. Ignoring the issue of precisely where the words originate, we can say that a document is represented by a vector of words, or *terms*. Terms with a common stem will usually have similar meanings, for example:

> CONNECT
> CONNECTED
> CONNECTING
> CONNECTION
> CONNECTIONS

Frequently, the performance of an IR system will be improved if term groups such as this are conflated into a single term. This may be done by removal of the various suffixes -ED, -ING, -ION, -IONS to leave the single stem CONNECT. In addition, the suffix stripping process will reduce the total number of terms in the IR system, and hence reduce the size and complexity of the data in the system, which is always advantageous.

Many strategies for suffix stripping have been reported in the literature.[eg 1-6] The nature of the task will vary considerably depending on whether a stem dictionary is being used, whether a suffix list is being used, and of course on the purpose for which the suffix stripping is being done. Assuming that one is not making use of a stem dictionary, and that the purpose of the task is to improve IR performance, the suffix stripping program will usually be given an explicit list of suffixes, and, with each suffix, the criterion under which it may be removed from a word to leave

130

a valid stem. This is the approach adopted here. The main merits of the present program are that it is small (less than 400 lines of BCPL), fast (it will process a vocabulary of 10,000 different words in about 8.1 seconds on the IBM 370/165 at Cambridge University), and reasonably simple. At any rate, it is simple enough to be described in full as an algorithm in this paper. (The present version in BCPL is freely available from the author. BCPL itself is available on a wide range of different computers, but anyone wishing to use the program should have little difficulty in coding it up in other programming languages.) Given the speed of the program, it would be quite realistic to apply it to every word in a large file of continuous text, although for historical reasons we have found it convenient to apply it only to relatively small vocabulary lists derived from continuous text files.

In any suffix stripping program for IR work, two points must be borne in mind. Firstly, the suffixes are being removed simply to improve IR performance, and not as a linguistic exercise. This means that it would not be at all obvious under what circumstances a suffix should be removed, even if we could exactly determine the suffixes of the words by automatic means.

Perhaps the best criterion for removing suffixes from two words W_1 and W_2 to produce a single stem S, is to say that we do so if there appears to be no difference between the two statements 'a document is about W_1,' and 'a document is about W_2'. So if W_1='CONNECTION' and W_2='CONNECTIONS' it seems very reasonable to conflate them to a single stem. But if W_1='RELATE' and W_2='RELATIVITY' it seems perhaps unreasonable, especially if the document collection is concerned with theoretical physics. (It should perhaps be added that RELATE and RELATIVITY *are* conflated together in the algorithm described here.) Between these two extremes there is a continuum of different cases, and given two terms W_1 and W_2, there will be some variation in opinion as to whether they should be conflated, just as there is with deciding the relevance of some document to a query. The evaluation of the worth of a suffix stripping system is correspondingly difficult.

The second point is that with the approach adopted here, i.e. the use of a suffix list with various rules, the success rate for the suffix stripping will be significantly less than 100%, irrespective of how the process is evaluated. For example, if SAND and SANDER get conflated, so most probably will WAND and WANDER. The error here is that the -ER of WANDER has been treated as a suffix when in fact it is part of the stem. Equally a suffix may completely alter the meaning of a word, in which case its removal is unhelpful. PROBE and PROBATE for example, have quite distinct meanings in modern English. (In fact these would not be conflated in our present algorithm.) There comes a stage in the development of a suffix stripping program where the addition of more rules to increase the performance in one area of the vocabulary causes an equal degradation of performance elsewhere. Unless this phenomenon is noticed in time, it is very easy for the program to become much more complex than is really necessary. It is also easy to give undue emphasis to cases which appear to

be important, but which turn out in practice to be rather rare. For example, cases in which the spelling of the root of the word changes with the addition of a suffix, as in DECEIVE/DECEPTION, RESUME/ RESUMPTION, INDEX/INDICES, occur much more rarely in real vocabularies than one might at first suppose. In view of the error rate that must in any case be expected, it did not seem worthwhile to try and cope with these cases.

It is not obvious that the simplicity of the present program is any demerit. In a test on the well known Cranfield 200 collection [7] it gave an improvement in retrieval performance when compared with a very much more elaborate program which has been in use in IR research at Cambridge since 1971[2, 6]. The test was done as follows: the words of the titles and abstracts in the documents were passed through the earlier suffix stripping system, and the resulting stems were used to index the documents. The words of the queries were reduced to stems in the same way, and the documents were ranked for each query using term coordination matching of query against document. From these rankings, recall and precision values were obtained using the standard recall cutoff method. The entire process was then repeated using the suffix stripping system described in this paper, and the results were as follows:

earlier system		present system	
precision	recall	precision	recall
0	57.24	0	58.60
10	56.85	10	58.13
20	52.85	20	53.92
30	42.61	30	43.51
40	42.20	40	39.39
50	39.06	50	38.85
60	32.86	60	33.18
70	31.64	70	31.19
80	27.15	80	27.52
90	24.59	90	25.85
100	24.59	100	25.85

Clearly the performance is not very different. The important point is that the earlier, more elaborate system certainly performs no better than the present, simple system.

(This test was done by Prof. C.J. van Rijsbergen.)

2. THE ALGORITHM

To present the suffix stripping algorithm in its entirety we will need a few definitions.

A *consonant* in a word is a letter other than A, E, I, O and U, and other than Y preceded by a consonant. (The fact that the term 'consonant' is defined to some extent in terms of itself does not make it ambiguous.) So in TOY the consonants are T and Y, in SYZYGY they are S, Z and G. If a letter is not a consonant it is a *vowel*.

132

A consonant will be denoted by c, a vowel by v. A list ccc ... of length greater than 0 will be denoted by C, and a list vvv ... of length greater than 0 will be denoted by V. Any word, or part of a word, therefore has one of the four forms:

$$CVCV \ldots C$$
$$CVCV \ldots V$$
$$VCVC \ldots C$$
$$VCVC \ldots V$$

These may all be represented by the single form

$$[C] \, VCVC \ldots [V]$$

where the square brackets denote arbitrary presence of their contents. Using $(VC)^m$ to denote VC repeated m times, this may again be written as

$$[C] \, (VC)^m \, [V].$$

m will be called the *measure* of any word or word part when represented in this form. The case $m=0$ covers the null word. Here are some examples:

$m=0$ TR, EE, TREE, Y, BY.
$m=1$ TROUBLE, OATS, TREES, IVY.
$m=2$ TROUBLES, PRIVATE, OATEN, ORRERY.

The *rules* for removing a suffix will be given in the form

$$(\text{condition}) \; S1 \; \rightarrow \; S2$$

This means that if a word ends with the suffix S1, and the stem before S1 satisfies the given condition, S1 is replaced by S2. The condition is usually given in terms of m, e.g.

$$(m>1) \; \text{EMENT} \; \rightarrow$$

Here S1 is 'EMENT' and S2 is null. This would map REPLACEMENT to REPLAC, since REPLAC is word part for which $m=2$.

The 'condition' part may also contain the following:

*S — the stem ends with S (and similarly for the other letters).

v — the stem contains a vowel.

*d — the stem ends with a double consonant (e.g. –TT, –SS).

*o — the stem ends cvc, where the second c is not W, X or Y (e.g. –WIL, –HOP).

And the condition part may also contain expressions with *and, or* and *not*, so that

$$(m>1 \text{ and } (\text{*S or *T}))$$

tests for a stem with $m>1$ ending in S or T, while

$$(\text{*d and not } (\text{*L or *S or *Z}))$$

tests for a stem ending with a double consonant other than L, S or Z. Elaborate conditions like this are required only very rarely.

133

In a set of rules written beneath each other, only one is obeyed, and this will be the one with the longest matching S1 for the given word. For example, with

SSES → SS
IES   → I
SS   → SS
S   →

(here the conditions are all null) CARESSES maps to CARESS since SSES is the longest match for S1. Equally CARESS maps to CARESS (S1='SS') and CARES to CARE (S1='S').

In the rules below, examples of their application, successful or otherwise, are given on the right in lower case. The algorithm now follows:

Step 1a

SSES	→ SS	caresses	→ caress
IES	→ I	ponies	→ poni
		ties	→ ti
SS	→ SS	caress	→ caress
S	→	cats	→ cat

Step 1b

(m>0)	EED → EE		feed	→ feed
			agreed	→ agree
(*v*)	ED →		plastered	→ plaster
			bled	→ bled
(*v*)	ING →		motoring	→ motor
			sing	→ sing

If the second or third of the rules in Step 1b is successful, the following is done:

AT → ATE		conflat(ed)	→ conflate
BL → BLE		troubl(ing)	→ trouble
IZ → IZE		siz(ed)	→ size
(*d and not (*L or *S or *Z))			
→ single letter		hopp(ing)	→ hop
		tann(ed)	→ tan
		fall(ing)	→ fall
		hiss(ing)	→ hiss
		fizz(ed)	→ fizz
(m=1 and *o) → E		fail(ing)	→ fail
		fil(ing)	→ file

The rule to map to a single letter causes the removal of one of the double letter pair. The –E is put back on –AT, –BL and –IZ, so that the suffixes –ATE, –BLE and –IZE can be recognised later. This E may be removed in step 4.

Step 1c

(*v*) Y → I	happy	→ happi
	sky	→ sky

134

Step 1 deals with plurals and past participles. The subsequent steps are much more straightforward.

Step 2

(m>0)	ATIONAL	→ ATE	relational	→ relate
(m>0)	TIONAL	→ TION	conditional	→ condition
			rational	→ rational
(m>0)	ENCI	→ ENCE	valenci	→ valence
(m>0)	ANCI	→ ANCE	hesitanci	→ hesitance
(m>0)	IZER	→ IZE	digitizer	→ digitize
(m>0)	ABLI	→ ABLE	conformabli	→ conformable
(m>0)	ALLI	→ AL	radicalli	→ radical
(m>0)	ENTLI	→ ENT	differentli	→ different
(m>0)	ELI	→ E	vileli	→ vile
(m>0)	OUSLI	→ OUS	analogousli	→ analogous
(m>0)	IZATION	→ IZE	vietnamization	→ vietnamize
(m>0)	ATION	→ ATE	predication	→ predicate
(m>0)	ATOR	→ ATE	operator	→ operate
(m>0)	ALISM	→ AL	feudalism	→ feudal
(m>0)	IVENESS	→ IVE	decisiveness	→ decisive
(m>0)	FULNESS	→ FUL	hopefulness	→ hopeful
(m>0)	OUSNESS	→ OUS	callousness	→ callous
(m>0)	ALITI	→ AL	formaliti	→ formal
(m>0)	IVITI	→ IVE	sensitiviti	→ sensitive
(m>0)	BILITI	→ BLE	sensibiliti	→ sensible

The test for the string S1 can be made fast by doing a program switch on the penultimate letter of the word being tested. This gives a fairly even breakdown of the possible values of the string S1. It will be seen in fact that the S1-strings in step 2 are presented here in the alphabetical order of their penultimate letter. Similar techniques may be applied in the other steps.

Step 3

(m>0)	ICATE	→ IC	triplicate	→ triplic
(m>0)	ATIVE	→	formative	→ form
(m>0)	ALIZE	→ AL	formalize	→ formal
(m>0)	ICITI	→ IC	electriciti	→ electric
(m>0)	ICAL	→ IC	electrical	→ electric
(m>0)	FUL	→	hopeful	→ hope
(m>0)	NESS	→	goodness	→ good

Step 4

(m>1)	AL	→	revival	→ reviv
(m>1)	ANCE	→	allowance	→ allow
(m>1)	ENCE	→	inference	→ infer
(m>1)	ER	→	airliner	→ airlin
(m>1)	IC	→	gyroscopic	→ gyroscop
(m>1)	ABLE	→	adjustable	→ adjust
(m>1)	IBLE	→	defensible	→ defens

135

(m > 1) ANT	→	irritant	→ irrit
(m > 1) EMENT	→	replacement	→ replac
(m > 1) MENT	→	adjustment	→ adjust
(m > 1) ENT	→	dependent	→ depend
(m > 1) and (*S or *T)) ION	→	adoption	→ adopt
(m > 1) OU	→	homologou	→ homolog
(m > 1) ISM	→	communism	→ commun
(m > 1) ATE	→	activate	→ activ
(m > 1) ITI	→	angulariti	→ angular
(m > 1) OUS	→	homologous	→ homolog
(m > 1) IVE	→	effective	→ effect
(m > 1) IZE	→	bowdlerize	→ bowdler

The suffixes are now removed. All that remains is a little tidying up.

Step 5a

(m > 1) E	→	probate	→ probat
		rate	→ rate
(m = 1 and not *o) E →		cease	→ ceas

Step 5b

(m > 1 and *d and *L) →		single letter	control
→		controll	→
		roll	→ roll

The algorithm is careful not to remove a suffix when the stem is too short, the length of the stem being given by its measure, m. There is no linguistic basis for this approach. It was merely observed that m could be used quite effectively to help decide whether or not it was wise to take off a suffix. For example, in the following two lists:

list A	list B
RELATE	DERIVATE
PROBATE	ACTIVATE
CONFLATE	DEMONSTRATE
PIRATE	NECESSITATE
PRELATE	RENOVATE

−ATE is removed from the list B words, but not from the list A words. This means that the pairs DERIVATE/DERIVE, ACTIVATE/ ACTIVE, DEMONSTRATE/DEMONSTRABLE, NECESSITATE/ NECESSITOUS, will conflate together. The fact that no attempt is made to identify prefixes can make the results look rather inconsistent. Thus PRELATE does not lose the −ATE, but ARCHPRELATE becomes ARCHPREL. In practice this does not matter too much, because the presence of the prefix decreases the probability of an erroneous conflation.

136

Complex suffixes are removed bit by bit in the different steps. Thus GENERALIZATIONS is stripped to GENERALIZATION (Step 1), then to GENERALIZE (Step 2), then to GENERAL (Step 3), and then to GENER (Step 4). OSCILLATORS is stripped to OSCILLATOR (Step 1), then to OSCILLATE (Step 2), then to OSCILL (Step 4), and then to OSCIL (Step 5). In a vocabulary of 10,000 words, the reduction in size of the stem was distributed among the steps as follows:

Suffix stripping of a vocabulary of 10,000 words

Number of words reduced in step 1:		3597
"	2:	766
"	3:	327
"	4:	2424
"	5:	1373
Number of words not reduced:		3650

The resulting vocabulary of stems contained 6370 distinct entries. Thus the suffix stripping process reduced the size of the vocabulary by about one third.

ACKNOWLEDGEMENTS

REFERENCES

1. LOVINS, J.B. Development of a Stemming Algorithm. *Mechanical Translation and Computational Linguistics*. 11 (1) March 1968 p. 22–31.
2. ANDREWS, K. The Development of a Fast Conflation Algorithm for English. *Dissertation for the Diploma in Computer Science*, Computer Laboratory, University of Cambridge, 1971.
3. PETRARCA, A.E. and LAY W.M. Use of an automatically generated authority list to eliminate scattering caused by some singular and plural main index terms. *Proceedings of the American Society for Information Science, 6* 1969 p. 277–282.
4. DATTOLA, Robert T. *FIRST: Flexible Information Retrieval System for Text*. Webster N.Y: Xerox Corporation, 12 Dec 1975.
5. COLOMBO, D.S. and NIEHOFF R.T. *Final report on improved access to scientific and technical information through automated vocabulary switching*. NSF Grant No. SIS75–12924 to the National Science Foundation.
6. DAWSON, J.L. Suffix Removal and Word Conflation. *ALLC Bulletin*, Michaelmas 1974 p. 33–46.
7. CLEVERDON, C.W., MILLS J. and KEEN M. *Factors Determining the Performance of Indexing Systems* 2 vols. College of Aeronautics, Cranfield 1966.

137

This article has been cited by:

1. Marcos Mouriño-García, Roberto Pérez-Rodríguez, Luis Anido-Rifón, Manuel J. Fernández-Iglesias, Víctor M. Darriba-Bilbao. 2018. Cross-repository aggregation of educational resources. *Computers & Education* **117**, 31-49. [Crossref]

2. Shirin Noekhah, Naomie Binti Salim, Nor Hawaniah Zakaria. A Comprehensive Study on Opinion Mining Features and Their Applications 78-89. [Crossref]

3. Bekir Parlak, Alper Kursat Uysal. On Feature Weighting and Selection for Medical Document Classification 269-282. [Crossref]

4. Rupam Gupta, Anjali G. Jivani. Analyzing the Stemming Paradigm 333-342. [Crossref]

5. Moses Olaifa, Sunday Ojo, Tranos Zuva. Hybrid Service Discovery Approach for Peer-to-peer Web Service Systems 281-293. [Crossref]

6. Moutaz Wajih Hamadeh, Sherief Abdallah. Discover Trending Topics of Interest to Governments 366-373. [Crossref]

7. Christopher Baechle, Ankur Agarwal, Xingquan Zhu. 2017. Big data driven co-occurring evidence discovery in chronic obstructive pulmonary disease patients. *Journal of Big Data* **4**:1. . [Crossref]

8. Manuel W. Bickel. 2017. A new approach to semantic sustainability assessment: text mining via network analysis revealing transition patterns in German municipal climate action plans. *Energy, Sustainability and Society* **7**:1. . [Crossref]

9. Chintan Amrit, Tim Paauw, Robin Aly, Miha Lavric. 2017. Identifying child abuse through text mining and machine learning. *Expert Systems with Applications* **88**, 402-418. [Crossref]

10. Marija Djokic-Petrovic, Vladimir Cvjetkovic, Jeremy Yang, Marko Zivanovic, David J. Wild. 2017. PIBAS FedSPARQL: a web-based platform for integration and exploration of bioinformatics datasets. *Journal of Biomedical Semantics* **8**:1. . [Crossref]

11. Kamil Topal, Mehmet Koyutürk, Gültekin Özsoyoğlu. 2017. Effects of emotion and topic area on topic shifts in social media discussions. *Social Network Analysis and Mining* **7**:1. . [Crossref]

12. Matthieu Picault, Thomas Renault. 2017. Words are not all created equal: A new measure of ECB communication. *Journal of International Money and Finance* **79**, 136-156. [Crossref]

13. Pavlina Fragkou. 2017. Applying named entity recognition and co-reference resolution for segmenting English texts. *Progress in Artificial Intelligence* **6**:4, 325-346. [Crossref]

14. Constantina Alina Hossu, Ioan Cristian Ioja, Mihai R. Nita, Tibor Hartel, Denisa L. Badiu, Anna M. Hersperger. 2017. Need for a cross-sector approach in protected area management. *Land Use Policy* **69**, 586-597. [Crossref]

15. Ansamma John, P.S. Premjith, M. Wilscy. 2017. Extractive multi-document summarization using population-based multicriteria optimization. *Expert Systems with Applications* **86**, 385-397. [Crossref]

16. Liyin Shen, Hang Yan, Hongqin Fan, Ya Wu, Yu Zhang. 2017. An integrated system of text mining technique and case-based reasoning (TM-CBR) for supporting green building design. *Building and Environment* **124**, 388-401. [Crossref]

17. Chuan Liu, Wenyong Wang, Guanghui Tu, Yu Xiang, Siyang Wang, Fengmao Lv. 2017. A new Centroid-Based Classification model for text categorization. *Knowledge-Based Systems* **136**, 15-26. [Crossref]

18. Cataldo Musto, Pasquale Lops, Marco de Gemmis, Giovanni Semeraro. 2017. Semantics-aware Recommender Systems exploiting Linked Open Data and graph-based features. *Knowledge-Based Systems* **136**, 1-14. [Crossref]

19. Martin Haselmayer, Marcelo Jenny. 2017. Sentiment analysis of political communication: combining a dictionary approach with crowdcoding. *Quality & Quantity* **51**:6, 2623-2646. [Crossref]

20. Massimo Aria, Corrado Cuccurullo. 2017. bibliometrix : An R-tool for comprehensive science mapping analysis. *Journal of Informetrics* **11**:4, 959-975. [Crossref]

21. David C. Anastasiu, George Karypis. 2017. Efficient identification of Tanimoto nearest neighbors. *International Journal of Data Science and Analytics* **4**:3, 153-172. [Crossref]

22. Yung-Jan Cho, Pei-Wen Fu, Chi-Cheng Wu. 2017. Popular Research Topics in Marketing Journals, 1995–2014. *Journal of Interactive Marketing* **40**, 52-72. [Crossref]

23. Akos Rona-Tas, Antoine Cornuéjols, Sandrine Blanchemanche, Antonin Duroy, Christine Martin. 2017. Enlisting Supervised Machine Learning in Mapping Scientific Uncertainty Expressed in Food Risk Analysis. *Sociological Methods & Research* **86**, 004912411772970. [Crossref]

24. Alon Zweig, Gal Chechik. 2017. Group online adaptive learning. *Machine Learning* **106**:9-10, 1747-1770. [Crossref]

25. Pu Li, Bao Xiao, Wenjun Ma, Yuncheng Jiang, Zhifeng Zhang. 2017. A graph-based semantic relatedness assessment method combining wikipedia features. *Engineering Applications of Artificial Intelligence* **65**, 268-281. [Crossref]

26. Todd Bodnar, Matthew L. Dering, Conrad Tucker, Kenneth M. edu. 2017. Using Large-Scale Social Media Networks as a Scalable Sensing System for Modeling Real-Time Energy Utilization Patterns. *IEEE Transactions on Systems, Man, and Cybernetics: Systems* **47**:10, 2627-2640. [Crossref]

27. Francesc Valls, Ernesto Redondo, David Fonseca, Ricardo Torres-Kompen, Sergi Villagrasa, Nuria Martí. 2017. Urban data and urban design: A data mining approach to architecture education. *Telematics and Informatics* . [Crossref]

28. Dustin S. Stoltz, Marshall A. Taylor. 2017. Paying with change: The purposeful enunciation of material culture. *Poetics* **64**, 26-39. [Crossref]

29. Mathias Kraus, Stefan Feuerriegel. 2017. Decision support from financial disclosures with deep neural networks and transfer learning. *Decision Support Systems* . [Crossref]

30. Nikita Basov, Julia Brennecke. Duality Beyond Dyads: Multiplex Patterning of Social Ties and Cultural Meanings 87-112. [Abstract] [Full Text] [PDF] [PDF]

31. Nihar M. Ranjan, Rajesh S. Prasad. 2017. Automatic text classification using BPLion-neural network and semantic word processing. *The Imaging Science Journal* **4**, 1-15. [Crossref]

32. Upendra Kumar, Aishwarya N. Reganti, Tushar Maheshwari, Tanmoy Chakraborty, Björn Gambäck, Amitava Das. 2017. Inducing Personalities and Values from Language Use in Social Network Communities. *Information Systems Frontiers* **38**. . [Crossref]

33. Marcos Antonio Mouriño García, Roberto Pérez Rodríguez, Luis Anido Rifón. 2017. Wikipedia-based cross-language text classification. *Information Sciences* **406-407**, 12-28. [Crossref]

34. Ahmed I. Saleh, Mohammed F. Al Rahmawy, Arwa E. Abulwafa. 2017. A semantic based Web page classification strategy using multi-layered domain ontology. *World Wide Web* **20**:5, 939-993. [Crossref]

35. Dima Badawi, Hakan Altınçay. 2017. Termset weighting by adapting term weighting schemes to utilize cardinality statistics for binary text categorization. *Applied Intelligence* **47**:2, 456-472. [Crossref]

36. Damiano Spina, Johanne R. Trippas, Lawrence Cavedon, Mark Sanderson. 2017. Extracting audio summaries to support effective spoken document search. *Journal of the Association for Information Science and Technology* **68**:9, 2101-2115. [Crossref]

37. Georgia M. Kapitsaki. 2017. Creating and utilizing section-level Web service tags in service replaceability. *Service Oriented Computing and Applications* **11**:3, 285-299. [Crossref]

38. Li Wei Yuan, Lei Su, Yin Zhang, Guang Fang, Peng Shu. 2017. Cloud-based learning system for answer ranking. *Cluster Computing* **20**:3, 2253-2266. [Crossref]

39. Nicola Ferro. 2017. What Does Affect the Correlation Among Evaluation Measures?. *ACM Transactions on Information Systems* **36**:2, 1-40. [Crossref]

40. SANDRA KÜBLER, CAN LIU, ZEESHAN ALI SAYYED. 2017. To use or not to use: Feature selection for sentiment analysis of highly imbalanced data. *Natural Language Engineering* **3**, 1-35. [Crossref]

41. Gianni Costa, Riccardo Ortale. 2017. Machine learning techniques for XML (co-)clustering by structure-constrained phrases. *Information Retrieval Journal* **43**. . [Crossref]

42. Muhammad Afzal, Maqbool Hussain, Robert Brian Haynes, Sungyoung Lee. 2017. Context-aware grading of quality evidences for evidence-based decision-making. *Health Informatics Journal* **328**, 146045821771956. [Crossref]

43. Muslim Chochlov, Michael English, Jim Buckley. 2017. A historical, textual analysis approach to feature location. *Information and Software Technology* **88**, 110-126. [Crossref]

44. Ibrahim Jameel Mujhid, Joanna C. S. Santos, Raghuram Gopalakrishnan, Mehdi Mirakhorli. 2017. A search engine for finding and reusing architecturally significant code. *Journal of Systems and Software* **130**, 81-93. [Crossref]

45. Juyoung An, Namhee Kim, Min-Yen Kan, Muthu Kumar Chandrasekaran, Min Song. 2017. Exploring characteristics of highly cited authors according to citation location and content. *Journal of the Association for Information Science and Technology* **68**:8, 1975-1988. [Crossref]

46. Tien-Duy B. Le, Ferdian Thung, David Lo. 2017. Will this localization tool be effective for this bug? Mitigating the impact of unreliability of information retrieval based bug localization tools. *Empirical Software Engineering* **22**:4, 2237-2279. [Crossref]

47. Christopher Klinkmüller, Ingo Weber. 2017. Analyzing control flow information to improve the effectiveness of process model matching techniques. *Decision Support Systems* **100**, 6-14. [Crossref]

48. Ashley Muddiman, Natalie Jomini Stroud. 2017. News Values, Cognitive Biases, and Partisan Incivility in Comment Sections. *Journal of Communication* **67**:4, 586-609. [Crossref]

49. Armando López-Cuevas, José Ramírez-Márquez, Gildardo Sanchez-Ante, Kash Barker. 2017. A Community Perspective on Resilience Analytics: A Visual Analysis of Community Mood. *Risk Analysis* **37**:8, 1566-1579. [Crossref]

50. Shikai Guo, Rong Chen, Hui Li. 2017. Using Knowledge Transfer and Rough Set to Predict the Severity of Android Test Reports via Text Mining. *Symmetry* **9**:8, 161. [Crossref]

51. Anwar Ali Yahya. 2017. Swarm intelligence-based Approach for Educational Data Classification. *Journal of King Saud University - Computer and Information Sciences* . [Crossref]

52. Jacob Foley, Paul Kwan, Mitchell Welch. 2017. A Web-Based Infrastructure for the Assisted Annotation of Heritage Collections. *Journal on Computing and Cultural Heritage* **10**:3, 1-25. [Crossref]

53. Md Abul Bashar, Yuefeng Li, Yan Shen, Yang Gao, Wei Huang. 2017. Conceptual annotation of text patterns. *Computational Intelligence* **1**. . [Crossref]

54. Daniel J. Hopkins. 2017. The Exaggerated Life of Death Panels? The Limited but Real Influence of Elite Rhetoric in the 2009–2010 Health Care Debate. *Political Behavior* **74**. . [Crossref]

55. Jochen Trommer, Sebastian Bank. 2017. Inflectional learning as local optimization. *Morphology* **70**. . [Crossref]

56. Shirish Singh, Praveen Kumar. 2017. User specific context construction for personalized multimedia retrieval. *Multimedia Tools and Applications* **16**. . [Crossref]

57. Jason West. 2017. Validating curriculum development using text mining. *The Curriculum Journal* **28**:3, 389-402. [Crossref]

58. Obada Alhabashneh, Rahat Iqbal, Faiyaz Doctor, Anne James. 2017. Fuzzy rule based profiling approach for enterprise information seeking and retrieval. *Information Sciences* **394-395**, 18-37. [Crossref]

59. Pushpalatha K, Ananthanarayana V S. 2017. A tree based representation for effective pattern discovery from multimedia documents. *Pattern Recognition Letters* **93**, 143-153. [Crossref]

60. Despoina Chatzakou, Athena Vakali, Konstantinos Kafetsios. 2017. Detecting Variation of Emotions in Online Activities. *Expert Systems with Applications* . [Crossref]

61. Rastislav Krchnavy, Marian Simko. Sentiment analysis of social network posts in Slovak language 20-25. [Crossref]

62. Rao Muhammad Adeel Nawab, Mark Stevenson, Paul Clough. 2017. An IR-Based Approach Utilizing Query Expansion for Plagiarism Detection in MEDLINE. *IEEE/ACM Transactions on Computational Biology and Bioinformatics* **14**:4, 796-804. [Crossref]

63. Swapna Gottipati, Venky Shankararaman, Melvrick Goh. Mining Capstone Project Wikis for Knowledge Discovery 371-380. [Crossref]

64. Jingwei Li, Dan Lin, Anna Cinzia Squicciarini, Jin Li, Chunfu Jia. 2017. Towards Privacy-Preserving Storage and Retrieval in Multiple Clouds. *IEEE Transactions on Cloud Computing* **5**:3, 499-509. [Crossref]

65. Martin Krallinger, Obdulia Rabal, Anália Lourenço, Julen Oyarzabal, Alfonso Valencia. 2017. Information Retrieval and Text Mining Technologies for Chemistry. *Chemical Reviews* **117**:12, 7673-7761. [Crossref]

66. Allen H. Huang, Reuven Lehavy, Amy Y. Zang, Rong Zheng. 2017. Analyst Information Discovery and Interpretation Roles: A Topic Modeling Approach. *Management Science* . [Crossref]

67. Shimi Naurin Ahmad, Michel Laroche. 2017. Analyzing electronic word of mouth: A social commerce construct. *International Journal of Information Management* **37**:3, 202-213. [Crossref]

68. Divya Padmanabhan, Satyanath Bhat, Shirish Shevade, Y. Narahari. 2017. Multi-Label Classification from Multiple Noisy Sources Using Topic Models. *Information* **8**:2, 52. [Crossref]

69. Wolfgang Glänzel, Sarah Heeffer, Bart Thijs. 2017. Lexical analysis of scientific publications for nano-level scientometrics. *Scientometrics* **111**:3, 1897-1906. [Crossref]

70. Davit Khachatryan, Brigitte Muehlmann. 2017. Determinants of successful patent applications to combat financial fraud. *Scientometrics* **111**:3, 1353-1383. [Crossref]

71. Davide Falessi, Massimiliano Di Penta, Gerardo Canfora, Giovanni Cantone. 2017. Estimating the number of remaining links in traceability recovery. *Empirical Software Engineering* **22**:3, 996-1027. [Crossref]

72. Shun Li, Junhao Wen, Fengji Luo, Min Gao, Jun Zeng, Zhao Yang Dong. 2017. A New QoS-Aware Web Service Recommendation System Based on Contextual Feature Recognition at Server-Side. *IEEE Transactions on Network and Service Management* **14**:2, 332-342. [Crossref]

73. Christopher R. Reynolds, Kealan Exley, Matthieu A. Bultelle, Inaki Sainz de Murieta, Richard I. Kitney. 2017. Debugging experiment machinery through time-course event sequence analysis. *Engineering Biology* **1**:1, 51-54. [Crossref]

74. Savvas Papagiannidis, Eric W.K. See-To, Dimitris G. Assimakopoulos, Yang Yang. 2017. Identifying industrial clusters with a novel big-data methodology: Are SIC codes (not) fit for purpose in the Internet age?. *Computers & Operations Research* . [Crossref]

75. Yu Zhao, Shaohua Wang, Ying Zou, Joanna Ng, Tinny Ng. Automatically Learning User Preferences for Personalized Service Composition 776-783. [Crossref]

76. Pradeep K. Venkatesh, Shaohua Wang, Ying Zou, Joanna W. Ng. A Personalized Assistant Framework for Service Recommendation 92-99. [Crossref]

77. Chris Mills, Gabriele Bavota, Sonia Haiduc, Rocco Oliveto, Andrian Marcus, Andrea De Lucia. 2017. Predicting Query Quality for Applications of Text Retrieval to Software Engineering Tasks. *ACM Transactions on Software Engineering and Methodology* **26**:1, 1-45. [Crossref]

78. Abdul Shahid, Muhammad Tanvir Afzal. 2017. Section-wise indexing and retrieval of research articles. *Cluster Computing* **84**. . [Crossref]

79. MazisPanagiotis, Panagiotis Mazis, TsekrekosAndrianos, Andrianos Tsekrekos. 2017. Latent semantic analysis of the FOMC statements. *Review of Accounting and Finance* **16**:2, 179-217. [Abstract] [Full Text] [PDF]

80. Abdulgabbar Saif, Nazlia Omar, Mohd Juzaiddin Ab Aziz, Ummi Zakiah Zainodin, Naomie Salim. 2017. Semantic concept model using Wikipedia semantic features. *Journal of Information Science* **4**, 016555151770623. [Crossref]

81. Christian Sternitzke. 2017. Interlocking Patent Rights and Value Appropriation: Insights From the Razor Industry. *IEEE Transactions on Engineering Management* **64**:2, 249-265. [Crossref]

82. Abhinav Singh, Conrad S. Tucker. 2017. A machine learning approach to product review disambiguation based on function, form and behavior classification. *Decision Support Systems* **97**, 81-91. [Crossref]

83. Partha Mukherjee, Gondy Leroy, David Kauchak, Srinidhi Rajanarayanan, Damian Y. Romero Diaz, Nicole P. Yuan, T. Gail Pritchard, Sonia Colina. 2017. NegAIT: A new parser for medical text simplification using morphological, sentential and double negation. *Journal of Biomedical Informatics* **69**, 55-62. [Crossref]

84. Benjamin Fuller, Mayank Varia, Arkady Yerukhimovich, Emily Shen, Ariel Hamlin, Vijay Gadepally, Richard Shay, John Darby Mitchell, Robert K. Cunningham. SoK: Cryptographically Protected Database Search 172-191. [Crossref]

85. Liang Xu, Wensheng Dou, Chushu Gao, Jie Wang, Jun Wei, Hua Zhong, Tao Huang. SpreadCluster: Recovering Versioned Spreadsheets through Similarity-Based Clustering 158-169. [Crossref]

86. Poyraz Umut Hatipoglu, Anil Atvar, Yusuf Oguzhan Artan, Oguzhan Sereflisan, Ali Demir. Software requirement traceability analysis using text mining methods 1-4. [Crossref]

87. Chunyang Chen, Zhenchang Xing, Ximing Wang. Unsupervised Software-Specific Morphological Forms Inference from Informal Discussions 450-461. [Crossref]

88. He Jiang, Xiaochen Li, Zijiang Yang, Jifeng Xuan. What Causes My Test Alarm? Automatic Cause Analysis for Test Alarms in System and Integration Testing 712-723. [Crossref]

89. Cao Langcai, Li Zhihui, Liu Yuanfang. Research of text clustering based on improved VSM by TF under the framework of Mahout 6597-6600. [Crossref]

90. Fabio Palomba, Pasquale Salza, Adelina Ciurumelea, Sebastiano Panichella, Harald Gall, Filomena Ferrucci, Andrea De Lucia. Recommending and Localizing Change Requests for Mobile Apps Based on User Reviews 106-117. [Crossref]

91. Isidoros Perikos, Ioannis Hatzilygeroudis. Aspect based sentiment analysis in social media with classifier ensembles 273-278. [Crossref]

92. Andrea Di Sorbo, Sebastiano Panichella, Carol V. Alexandru, Corrado A. Visaggio, Gerardo Canfora. SURF: Summarizer of User Reviews Feedback 55-58. [Crossref]

93. Lothar Walter, Alfred Radauer, Martin G. Moehrle. 2017. The beauty of brimstone butterfly: novelty of patents identified by near environment analysis based on text mining. *Scientometrics* **111**:1, 103-115. [Crossref]

94. Louis-Gregory Strolger, Sophia Porter, Jill Lagerstrom, Sarah Weissman, I. Neill Reid, Michael Garcia. 2017. The Proposal Auto-Categorizer and Manager for Time Allocation Review at the Space Telescope Science Institute. *The Astronomical Journal* **153**:4, 181. [Crossref]

95. Le Nguyen Hoai Nam, Ho Bao Quoc. 2017. The Hybrid Filter Feature Selection Methods for Improving High-Dimensional Text Categorization. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems* **25**:02, 235-265. [Crossref]

96. Mourad Sarrouti, Said Ouatik El Alaoui. 2017. A passage retrieval method based on probabilistic information retrieval model and UMLS concepts in biomedical question answering. *Journal of Biomedical Informatics* **68**, 96-103. [Crossref]

97. Yinying Wang, Alex J. Bowers, David J. Fikis. 2017. Automated Text Data Mining Analysis of Five Decades of Educational Leadership Research Literature. *Educational Administration Quarterly* **53**:2, 289-323. [Crossref]

98. Ruggero Sainaghi, Paul Phillips, Emma Zavarrone. 2017. Performance measurement in tourism firms: A content analytical meta-approach. *Tourism Management* **59**, 36-56. [Crossref]

99. Minyoung Kim. 2017. Simultaneous Learning of Sentence Clustering and Class Prediction for Improved Document Classification. *The International Journal of Fuzzy Logic and Intelligent Systems* **17**:1, 35-42. [Crossref]

100. Mahmud Hasan, Mehmet A Orgun, Rolf Schwitter. 2017. A survey on real-time event detection from the Twitter data stream. *Journal of Information Science* **12**, 016555151769856. [Crossref]

101. Wei Wang, Hongwei Wang, Yuan Song. 2017. Ranking product aspects through sentiment analysis of online reviews. *Journal of Experimental & Theoretical Artificial Intelligence* **29**:2, 227-246. [Crossref]

102. Emily M. Cody, Jennie C. Stephens, James P. Bagrow, Peter Sheridan Dodds, Christopher M. Danforth. 2017. Transitions in climate and energy discourse between Hurricanes Katrina and Sandy. *Journal of Environmental Studies and Sciences* **7**:1, 87-101. [Crossref]

103. Vicky Chen, John Paisley, Xinghua Lu. 2017. Revealing common disease mechanisms shared by tumors of different tissues of origin through semantic representation of genomic alterations and topic modeling. *BMC Genomics* **18**:S2. . [Crossref]

104. Xin Xia, David Lo, Ying Ding, Jafar M. Al-Kofahi, Tien N. Nguyen, Xinyu Wang. 2017. Improving Automated Bug Triaging with Specialized Topic Model. *IEEE Transactions on Software Engineering* **43**:3, 272-297. [Crossref]

105. Marcel Dunaiski, Gillian J. Greene, Bernd Fischer. 2017. Exploratory search of academic publication and citation data using interactive tag cloud visualizations. *Scientometrics* **110**:3, 1539-1571. [Crossref]

106. V. B. Singh, Sanjay Misra, Meera Sharma. 2017. Bug Severity Assessment in Cross Project Context and Identifying Training Candidates. *Journal of Information & Knowledge Management* **16**:01, 1750005. [Crossref]

107. Qiwei He, Bernard P. Veldkamp, Cees A. W. Glas, Theo de Vries. 2017. Automated Assessment of Patients' Self-Narratives for Posttraumatic Stress Disorder Screening Using Natural Language Processing and Text Mining. *Assessment* **24**:2, 157-172. [Crossref]

108. Chandrakant D. Patel, Jayesh M. Patel. GUJSTER: A rule based stemmer using dictionary approach 496-499. [Crossref]

109. Alberto Alexander Gayle, Motomu Shimaoka. 2017. Evaluating the lexico-grammatical differences in the writing of native and non-native speakers of English in peer-reviewed medical journals in the field of pediatric oncology: Creation of the genuine index scoring system. *PLOS ONE* **12**:2, e0172338. [Crossref]

110. Jyh-Jian Sheu, Ko-Tsung Chu, Nien-Feng Li, Cheng-Chi Lee. 2017. An efficient incremental learning mechanism for tracking concept drift in spam filtering. *PLOS ONE* **12**:2, e0171518. [Crossref]

111. Laura Moreno, Gabriele Bavota, Massimiliano Di Penta, Rocco Oliveto, Andrian Marcus, Gerardo Canfora. 2017. ARENA: An Approach for the Automated Generation of Release Notes. *IEEE Transactions on Software Engineering* **43**:2, 106-127. [Crossref]

112. Sunghoon Lim, Conrad S. Tucker, Soundar Kumara. 2017. An unsupervised machine learning model for discovering latent infectious diseases using social media data. *Journal of Biomedical Informatics* **66**, 82-94. [Crossref]

113. Lixin Chen. 2017. Do patent citations indicate knowledge linkage? The evidence from text similarities between patents and their citations. *Journal of Informetrics* **11**:1, 63-79. [Crossref]

114. Mukul Gupta, Pradeep Kumar, Bharat Bhasker. 2017. Clustering of Micro-Messages Using Similarity Upper Approximation. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems* **25**:01, 53-79. [Crossref]

115. Zhe Liu, Bernard J. Jansen. 2017. ASK: A taxonomy of accuracy, social, and knowledge information seeking posts in social question and answering. *Journal of the Association for Information Science and Technology* **68**:2, 333-347. [Crossref]

116. Ankhtuya Ochirbat, Timothy K. Shih, Chalothon Chootong, Worapot Sommool, W.K.T.M. Gunarathne, Hai-Hui Wang, Zhao-Heng Ma. 2017. Hybrid occupation recommendation for adolescents on interest, profile, and behavior. *Telematics and Informatics* . [Crossref]

117. Mahmood Yousefi-Azar, Len Hamey. 2017. Text summarization using unsupervised deep learning. *Expert Systems with Applications* **68**, 93-105. [Crossref]

118. Gianni Costa, Riccardo Ortale. 2017. XML Clustering by Structure-Constrained Phrases: A Fully-Automatic Approach Using Contextualized N-Grams. *International Journal on Artificial Intelligence Tools* **26**:01, 1760002. [Crossref]

119. Kisik Song, Karp Soo Kim, Sungjoo Lee. 2017. Discovering new technology opportunities based on patents: Text-mining and F-term analysis. *Technovation* **60-61**, 1-14. [Crossref]

120. Adelina Ciurumelea, Andreas Schaufelbuhl, Sebastiano Panichella, Harald C. Gall. Analyzing reviews and code of mobile apps for better release planning 91-102. [Crossref]

121. Fernando Loizides, Sam Jones. 2017. A methodology for digitally exploring electronic publication content. *Information Services & Use* **36**:3-4, 211-214. [Crossref]

122. Thomas Lansdall-Welfare, Saatviga Sudhahar, James Thompson, Justin Lewis, Nello Cristianini. 2017. Content analysis of 150 years of British periodicals. *Proceedings of the National Academy of Sciences* **114**:4, E457-E465. [Crossref]

123. Ali Jahanian, Shaiyan Keshvari, S. V. N. Vishwanathan, Jan P. Allebach. 2017. Colors -- Messengers of Concepts. *ACM Transactions on Computer-Human Interaction* **24**:1, 1-39. [Crossref]

124. Lizhou Feng, Youwei Wang, Wanli Zuo. 2017. Novel feature selection method based on random walk and artificial bee colony. *Journal of Intelligent & Fuzzy Systems* **32**:1, 115-126. [Crossref]

125. Masoud Makrehchi, Mohamed S. Kamel. 2017. Extracting domain-specific stopwords for text classifiers. *Intelligent Data Analysis* **21**:1, 39-62. [Crossref]

126. Arazo M Mustafa, Tarik A Rashid. 2017. Kurdish stemmer pre-processing steps for improving information retrieval. *Journal of Information Science* 016555151668361. [Crossref]

127. G. Shankaranarayanan, Roger Blake. 2017. From Content to Context. *Journal of Data and Information Quality* **8**:2, 1-28. [Crossref]

128. Sherin Moussa. An Approach for Opinion-Demographic-Topology Based Microblog Friend Recommendation 815-824. [Crossref]

129. Alon Ben-Ari, Howard Chansky, Irene Rozet. 2017. Preoperative Opioid Use Is Associated with Early Revision After Total Knee Arthroplasty. *The Journal of Bone and Joint Surgery* **99**:1, 1-9. [Crossref]

130. Henning Femmer, Daniel Méndez Fernández, Stefan Wagner, Sebastian Eder. 2017. Rapid quality assurance with Requirements Smells. *Journal of Systems and Software* **123**, 190-213. [Crossref]

131. Longinos Marin, Pedro J. Martín, Alicia Rubio. 2017. Doing Good and Different! The Mediation Effect of Innovation and Investment on the Influence of CSR on Competitiveness. *Corporate Social Responsibility and Environmental Management* . [Crossref]

132. Fei Xie, Xindong Wu, Xingquan Zhu. 2017. Efficient sequential pattern mining with wildcards for keyphrase extraction. *Knowledge-Based Systems* **115**, 27-39. [Crossref]

133. Olivier Toubia, Oded Netzer. 2017. Idea Generation, Creativity, and Prototypicality. *Marketing Science* **36**:1, 1-20. [Crossref]

134. Carlo De Medio, Fabio Gasparetti, Carla Limongelli, Matteo Lombardi, Alessandro Marani, Filippo Sciarrone, Marco Temperini. Towards a Characterization of Educational Material: An Analysis of Coursera Resources 547-557. [Crossref]

135. S. Sarkar, B. Tak. 2017. Entity search techniques for expediting entitlement resolution in technology support services. *IBM Journal of Research and Development* **61**:1, 8:86-8:97. [Crossref]

136. Archana Chougule, Vijay Kumar Jha, Debajyoti Mukhopadhyay. Adaptive Ontology Construction Method for Crop Pest Management 665-674. [Crossref]

137. Miroslav Smatana, Peter Butka. Hierarchical Topic Modeling Based on the Combination of Formal Concept Analysis and Singular Value Decomposition 357-368. [Crossref]

138. Jan M. Czarnecki, Adrian J. Shepherd. Metabolic Pathway Mining 139-158. [Crossref]

139. Jean-Charles Lamirel, Pascal Cuxac, Kafil Hajlaoui. A Novel Approach to Feature Selection Based on Quality Estimation Metrics 121-140. [Crossref]

140. Safaa I. Hajeer, Rasha M. Ismail, Nagwa L. Badr, Mohamed Fahmy Tolba. A New Stemming Algorithm for Efficient Information Retrieval Systems and Web Search Engines 117-135. [Crossref]

141. Arpan Sen, Shrestha Ghosh, Debottam Kundu, Debleena Sarkar, Jaya Sil. Study of Engineered Features and Learning Features in Machine Learning - A Case Study in Document Classification 161-172. [Crossref]

142. Gregory J. Trevors, Krista R. Muis, Reinhard Pekrun, Gale M. Sinatra, Marloes M.L. Muijselaar. 2017. Exploring the relations between epistemic beliefs, emotions, and learning from texts. *Contemporary Educational Psychology* **48**, 116-132. [Crossref]

143. Moumita Basu, Anurag Roy, Kripabandhu Ghosh, Somprakash Bandyopadhyay, Saptarshi Ghosh. A Novel Word Embedding Based Stemming Approach for Microblog Retrieval During Disasters 589-597. [Crossref]

144. Colin Layfield, Joel Azzopardi, Chris Staff. Experiments with Document Retrieval from Small Text Collections Using Latent Semantic Analysis or Term Similarity with Query Coordination and Automatic Relevance Feedback 25-36. [Crossref]

145. Andreas Schoknecht, Nicolai Fischer, Andreas Oberweis. Process Model Search Using Latent Semantic Analysis 283-295. [Crossref]

146. Paul Nulty. Semantic/Content Analysis/Natural Language Processing 1-5. [Crossref]

147. Giacomo Domeniconi, Konstantinos Semertzidis, Gianluca Moro, Vanessa Lopez, Spyros Kotoulas, Elizabeth M. Daly. Identifying Conversational Message Threads by Integrating Classification and Data Clustering 25-46. [Crossref]

148. Arthur Huang, David Ebert, Parker Rider. You Are What You Tweet: A New Hybrid Model for Sentiment Analysis 403-416. [Crossref]

149. Amit Rathee, Jitender Kumar Chhabra. Software Remodularization by Estimating Structural and Conceptual Relations Among Classes and Using Hierarchical Clustering 94-106. [Crossref]

150. Akira Murakami, Paul Thompson, Susan Hunston, Dominik Vajn. 2017. 'What is this corpus about?': using topic modelling to explore a specialised corpus. *Corpora* **12**:2, 243. [Crossref]

151. Alba Núñez-Reyes, Esaú Villatoro-Tello, Gabriela Ramírez-de-la-Rosa, Christian Sánchez-Sánchez. A Compact Representation for Cross-Domain Short Text Clustering 16-26. [Crossref]

152. Manoj Bhat, Klym Shumaiev, Andreas Biesdorf, Uwe Hohenstein, Florian Matthes. Automatic Extraction of Design Decisions from Issue Management Systems: A Machine Learning Based Approach 138-154. [Crossref]

153. Giovanny Quiazúa, Carlos Rojas, Javier Ordoñez, Darwin Martinez, César Enciso-Olivera, Francisco Gómez. Computer Assisted Assignment of ICD Codes for Primary Admission Diagnostic in ICUs 211-223. [Crossref]

154. Clay Posey, Uzma Raja, Robert E. Crossler, A. J. Burns. 2017. Taking stock of organisations' protection of privacy: categorising and assessing threats to personally identifiable information in the USA. *European Journal of Information Systems* . [Crossref]

155. Sebastian Bank. 2017. Assessing the typology of person portmanteaus. *Morphology* . [Crossref]

156. Fedelucio Narducci, Cataldo Musto, Marco de Gemmis, Pasquale Lops, Giovanni Semeraro. 2017. TV-Program Retrieval and Classification: A Comparison of Approaches based on Machine Learning. *Information Systems Frontiers* . [Crossref]

157. Eun-kyung KIM, Key-Sun CHOI. 2017. Entity Summarization Based on Entity Grouping in Multilingual Projected Entity Space. *IEICE Transactions on Information and Systems* **E100.D**:9, 2138-2146. [Crossref]

158. Dimitrios Effrosynidis, Symeon Symeonidis, Avi Arampatzis. A Comparison of Pre-processing Techniques for Twitter Sentiment Analysis 394-406. [Crossref]

159. Emmanouil Marakakis, Haridimos Kondylakis, Papakonstantinou Aris. APANTISIS: A Greek Question-Answering System for Knowledge-Base Exploration 501-510. [Crossref]

160. Andrei M. Butnaru, Radu Tudor Ionescu. 2017. From Image to Text Classification: A Novel Approach based on Clustering Word Embeddings. *Procedia Computer Science* **112**, 1783-1792. [Crossref]

161. Yu. N. Orlov, Yu. A. Parfenova. Structure Analysis of an Explanatory Dictionary Ontological Graph 408-419. [Crossref]

162. Jon-Patrick Allem, Jagannathan Ramanujam, Kristina Lerman, Kar-Hai Chu, Tess Boley Cruz, Jennifer B Unger. 2017. Identifying Sentiment of Hookah-Related Posts on Twitter. *JMIR Public Health and Surveillance* **3**:4, e74. [Crossref]

163. Sandra Geisler, Christoph Quix, Rihan Hai, Sanchit Alekh. An Integrated Ontology-Based Approach for Patent Classification in Medical Engineering 38-52. [Crossref]

164. Omar ElShaweesh, Farookh Khadeer Hussain, Haiyan Lu, Malak Al-Hassan, Sadegh Kharazmi. Personalized Web Search Based on Ontological User Profile in Transportation Domain 239-248. [Crossref]

165. Nenad Ukić, Josip Maras, Ljiljana Šerić. 2016. The influence of cyclomatic complexity distribution on the understandability of xtUML models. *Software Quality Journal* . [Crossref]

166. Wael M.S. Yafoo. 2016. Model of Textual Data Linking and Clustering in Relational Databases. *Research Journal of Information Technology* **9**:1, 7-17. [Crossref]

167. Diana Maynard, Kalina Bontcheva, Isabelle Augenstein. 2016. Natural Language Processing for the Semantic Web. *Synthesis Lectures on the Semantic Web: Theory and Technology* **6**:2, 1-194. [Crossref]

168. Nan Wang, Shanwu Sun, Dantong OuYang. 2016. Business Process Modeling Abstraction Based on Semi-Supervised Clustering Analysis. *Business & Information Systems Engineering* . [Crossref]

169. Ulrike Kappler, Karen Davenport, Scott Beatson, Alla Lapidus, Chongle Pan, Cliff Han, Maria del Carmen Montero-Calasanz, Miriam Land, Loren Hauser, Manfred Rohde, Markus Göker, Natalia Ivanova, Tanja Woyke, Hans-Peter Klenk, Nikos C. Kyrpides. 2016. Complete genome sequence of the haloalkaliphilic, obligately chemolithoautotrophic thiosulfate and sulfide-oxidizing γ-proteobacterium Thioalkalimicrobium cyclicum type strain ALM 1 (DSM 14477T). *Standards in Genomic Sciences* **11**:1. . [Crossref]

170. Haijun Zhang, Tommy W. S. Chow, Q. M. Jonathan Wu. 2016. Organizing Books and Authors by Multilayer SOM. *IEEE Transactions on Neural Networks and Learning Systems* **27**:12, 2537-2550. [Crossref]

171. Abdulaziz Shehab, Mohamed Elhoseny, Aboul Ella Hassanien. A hybrid scheme for Automated Essay Grading based on LVQ and NLP techniques 65-70. [Crossref]

172. Melanie Erkens, Daniel Bodemer, H. Ulrich Hoppe. 2016. Improving collaborative learning in the classroom: Text mining based grouping and representing. *International Journal of Computer-Supported Collaborative Learning* **11**:4, 387-415. [Crossref]

173. Zhangjie Fu, Xinle Wu, Chaowen Guan, Xingming Sun, Kui Ren. 2016. Toward Efficient Multi-Keyword Fuzzy Search Over Encrypted Outsourced Data With Accuracy Improvement. *IEEE Transactions on Information Forensics and Security* **11**:12, 2706-2716. [Crossref]

174. Robert Karam, Ruchir Puri, Swarup Bhunia. 2016. Energy-Efficient Adaptive Hardware Accelerator for Text Mining Application Kernels. *IEEE Transactions on Very Large Scale Integration (VLSI) Systems* **24**:12, 3526-3537. [Crossref]

175. Fang-Yu Chao, Jia Xu, Chia-Wen Lin. Mining user interests from social media by fusing textual and visual features 1-8. [Crossref]

176. Albin Petit, Thomas Cerqueus, Antoine Boutet, Sonia Ben Mokhtar, David Coquil, Lionel Brunie, Harald Kosch. 2016. SimAttack: private web search under fire. *Journal of Internet Services and Applications* **7**:1. . [Crossref]

177. Luke Miratrix, Robin Ackerman. 2016. Conducting sparse feature selection on arbitrarily long phrases in text corpora with a focus on interpretability. *Statistical Analysis and Data Mining: The ASA Data Science Journal* **9**:6, 435-460. [Crossref]

178. Marcos Mourino-Garcia, Roberto Perez-Rodriguez, Luis Anido-Rifon, Miguel Gomez-Carballa. Bag-of-Concepts Document Representation for Bayesian Text Classification 281-288. [Crossref]

179. Majid Darabi, Nasseh Tabrizi. An Ontology-Based Framework to Model User Interests 398-403. [Crossref]

180. Vinaya B. Savadekar, Manoj E. Patil. Improved recommendation system with review analysis 79-82. [Crossref]

181. T. Liu, H. Wang, Y. He. 2016. Intelligent knowledge recommending approach for new product development based on workflow context matching. *Concurrent Engineering* **24**:4, 318-329. [Crossref]

182. Tzu-Tsung Wong, Chao-Rui Liu. 2016. An efficient parameter estimation method for generalized Dirichlet priors in naïve Bayesian classifiers with multinomial models. *Pattern Recognition* **60**, 62-71. [Crossref]

183. Kewen Chen, Zuping Zhang, Jun Long, Hao Zhang. 2016. Turning from TF-IDF to TF-IGM for term weighting in text classification. *Expert Systems with Applications* **66**, 245-260. [Crossref]

184. Lei Qi, Chuanhai Zhang, Adisak Sukul, Wallapak Tavanapong, David A. M. Peterson. Automated Coding of Political Video Ads for Political Science Research 7-13. [Crossref]

185. Rupali R. Gangarde. Pattern co-occurrence matrix to reduce the low frequency problem and effective pattern discovery 7-11. [Crossref]

186. In-Su Kang. 2016. Using Collective Citing Sentences to Recognize Cited Text in Computational Linguistics Articles. *Journal of the Korea Society of Computer and Information* **21**:11, 85-91. [Crossref]

187. Jyh-Jian Sheu, Yin-Kai Chen, Ko-Tsung Chu, Jih-Hsin Tang, Wei-Pang Yang. 2016. An intelligent three-phase spam filtering method based on decision tree data mining. *Security and Communication Networks* **9**:17, 4013-4026. [Crossref]

188. Fabon Dzogang, Thomas Lansdall-Welfare, Nello Cristianini. 2016. Discovering Periodic Patterns in Historical News. *PLOS ONE* **11**:11, e0165736. [Crossref]

189. Divya Padmanabhan, Satyanath Bhat, Shirish Shevade, Y. Narahari. Topic Model Based Multi-Label Classification 996-1003. [Crossref]

190. . Big Data Suite for Market Prediction and Reducing Complexity Using Bloom Filter 281-302. [Crossref]

191. . Insight and Intelligence 51-108. [Crossref]

192. Fatema Nafa, Javed I. Khan, Salem Othman, Amal Babour. Discovering Bloom Taxonomic Relationships between Knowledge Units Using Semantic Graph Triangularity Mining 224-233. [Crossref]

193. Shaowei Wang, David Lo. 2016. AmaLgam+: Composing Rich Information Sources for Accurate Bug Localization. *Journal of Software: Evolution and Process* **28**:10, 921-942. [Crossref]

194. Junbum Cha, Jeongwoo Kim, Sanghyun Park. GRiD: Gathering rich data from PubMed using one-class SVM 004325-004331. [Crossref]

195. Tanmay Basu, C. A. Murthy. 2016. A supervised term selection technique for effective text categorization. *International Journal of Machine Learning and Cybernetics* **7**:5, 877-892. [Crossref]

196. Ling Feng, Bin Liu, Zhiyong Peng. 2016. A patent hotspot discovery method. *Wuhan University Journal of Natural Sciences* **21**:5, 369-375. [Crossref]

197. Mohammad Hadi Bokaei, Hossein Sameti, Yang Liu. 2016. Summarizing Meeting Transcripts Based on Functional Segmentation. *IEEE/ACM Transactions on Audio, Speech, and Language Processing* **24**:10, 1831-1841. [Crossref]

198. Lucelene Lopes, Paulo Fernandes, Renata Vieira. ExATO - High Quality Term Extraction for Portuguese and English 540-545. [Crossref]

199. Tse-Hsun Chen, Stephen W. Thomas, Ahmed E. Hassan. 2016. A survey on the use of topic models when mining software repositories. *Empirical Software Engineering* **21**:5, 1843-1919. [Crossref]

200. Nasser Alsaedi, Pete Burnap, Omer Rana. Sensing Real-World Events Using Social Media Data and a Classification-Clustering Framework 216-223. [Crossref]

201. Deepshikha Patel, Prashant Kumar Singh. Kids safe search classification model 1-7. [Crossref]

202. Bogdan Dit. Configuring and Assembling Information Retrieval Based Solutions for Software Engineering Tasks 641-646. [Crossref]

203. David C. Anastasiu, George Karypis. Efficient Identification of Tanimoto Nearest Neighbors 156-165. [Crossref]

204. Amal Babour, Javed I. Khan, Fatema Nafa. Deepening Prose Comprehension by Incremental Free Text Conceptual Graph Mining and Knowledge 208-215. [Crossref]

205. Fabio Palomba. Alternative Sources of Information for Code Smell Detection: Postcards from Far Away 636-640. [Crossref]

206. . Bibliography 245-273. [Crossref]

207. Jasmeet Singh, Vishal Gupta. 2016. Text Stemming. *ACM Computing Surveys* **49**:3, 1-46. [Crossref]

208. Robert Leaman, Zhiyong Lu. 2016. TaggerOne: joint named entity recognition and normalization with semi-Markov Models. *Bioinformatics* **32**:18, 2839-2846. [Crossref]

209. Yogesh Kumar Meena, Dinesh Gopalani. 2016. Efficient Voting-Based Extractive Automatic Text Summarization Using Prominent Feature Set. *IETE Journal of Research* **62**:5, 581-590. [Crossref]

210. Jiansong Zhang, Nora M. El-Gohary. 2016. Extending Building Information Models Semiautomatically Using Semantic Natural Language Processing Techniques. *Journal of Computing in Civil Engineering* **30**:5, C4016004. [Crossref]

211. Hossein Soleimani, David J. Miller. 2016. ATD: Anomalous Topic Discovery in High Dimensional Discrete Data. *IEEE Transactions on Knowledge and Data Engineering* **28**:9, 2267-2280. [Crossref]

212. Andrew Bimba, Norisma Idris, Norazlina Khamis, Nurul Fazmidar Mohd Noor. 2016. Stemming Hausa text: using affix-stripping rules and reference look-up. *Language Resources and Evaluation* **50**:3, 687-703. [Crossref]

213. Matthijs Meire, Michel Ballings, Dirk Van den Poel. 2016. The added value of auxiliary data in sentiment analysis of Facebook posts. *Decision Support Systems* **89**, 98-112. [Crossref]

214. Zhangjie Fu, Kui Ren, Jiangang Shu, Xingming Sun, Fengxiao Huang. 2016. Enabling Personalized Search over Encrypted Outsourced Data with Efficiency Improvement. *IEEE Transactions on Parallel and Distributed Systems* **27**:9, 2546-2559. [Crossref]

215. Rebecca R. Carter, Jiayang Sun, Robin L. P. Jump. 2016. A Survey and Analysis of the American Public's Perceptions and Knowledge About Antibiotic Resistance. *Open Forum Infectious Diseases* **3**:3, ofw112. [Crossref]

216. Thorsten Merten, Matus Falis, Paul Hubner, Thomas Quirchmayr, Simone Bursner, Barbara Paech. Software Feature Request Detection in Issue Tracking Systems 166-175. [Crossref]

217. Felipe N. Flores, Viviane P. Moreira. 2016. Assessing the impact of Stemming Accuracy on Information Retrieval – A multilingual perspective. *Information Processing & Management* **52**:5, 840-854. [Crossref]

218. Dhiren A. Audich, Rozita Dara, Blair Nonnecke. Extracting keyword and keyphrase from online privacy policies 127-132. [Crossref]

219. Eliza D. Evans. 2016. Measuring Interdisciplinarity Using Text. *Socius: Sociological Research for a Dynamic World* **2**, 237802311665414. [Crossref]

220. Jun Miao, Jimmy Xiangji Huang, Jiashu Zhao. 2016. TopPRF. *ACM Transactions on Information Systems* **34**:4, 1-36. [Crossref]

221. Ming-Feng Tsai, Chuan-Ju Wang, Po-Chuan Chien. 2016. Discovering Finance Keywords via Continuous-Space Language Models. *ACM Transactions on Management Information Systems* **7**:3, 1-17. [Crossref]

222. Jordi Sancho. 2016. Learning Opportunities for Mass Collaboration Projects Through Learning Analytics: a Case Study. *IEEE Revista Iberoamericana de Tecnologias del Aprendizaje* **11**:3, 148-158. [Crossref]

223. Daniel Zoller, Stephan Doerfel, Robert Jäschke, Gerd Stumme, Andreas Hotho. 2016. Posted, visited, exported: Altmetrics in the social tagging system BibSonomy. *Journal of Informetrics* **10**:3, 732-749. [Crossref]

224. Zhilong Zhen, Haijuan Wang, Yan Xing, Lixin Han. Text feature selection approach by means of class difference 1582-1586. [Crossref]

225. Pushpalatha K, Ananthanarayana V S. 2016. Feature pattern based representation of multimedia documents for efficient knowledge discovery. *Multimedia Tools and Applications* **75**:15, 9461-9487. [Crossref]

226. Antreas Xenaros, Panagiotis Karampelas, Ioanna Lekea. Profiling individuals based on email analysis and ego networks: A visualization technique 1262-1269. [Crossref]

227. Jingfei Li, Peng Zhang, Dawei Song, Yuexian Hou. 2016. An adaptive contextual quantum language model. *Physica A: Statistical Mechanics and its Applications* **456**, 51-67. [Crossref]

228. Luo Yuansheng. Multi-lingual text clustering method using bilingual semantic correspondence analysis 1216-1221. [Crossref]

229. Phoebe Liu, Dylan F. Glas, Takayuki Kanda, Hiroshi Ishiguro. 2016. Data-Driven HRI: Learning Social Behaviors by Example From Human–Human Interaction. *IEEE Transactions on Robotics* **32**:4, 988-1008. [Crossref]

230. Mohamed Watfa, Nikiforos Rota, Henna Wadhwani, Nsimba Kupessa. An innovative mobile application for M-learning 453-458. [Crossref]

231. Mostafa A. Alksher, Azreen Azman, Razali Yaakob, Rabiah Abdul Kadir, Abdulmajid Mohamed, Eissa M. Alshari. A review of methods for mining idea from text 88-93. [Crossref]

232. Matthew C. Nowlin. 2016. Modeling Issue Definitions Using Quantitative Text Analysis. *Policy Studies Journal* **44**:3, 309-331. [Crossref]

233. Katalin Orosz, Illés J. Farkas, Péter Pollner. 2016. Quantifying the changing role of past publications. *Scientometrics* **108**:2, 829-853. [Crossref]

234. Lea Vega, Andres Mendez-Vazquez. Dynamic Neural Networks for Text Classification 6-11. [Crossref]

235. Kamil Topal, Mehmet Koyuturk, Gultekin Ozsoyoglu. Emotion -and area-driven topic shift analysis in social media discussions 510-518. [Crossref]

236. Ruhui Shen, Jialiang Shen, Yuhong Li, Haohan Wang. Predicting usefulness of Yelp reviews with localized linear regression models 189-192. [Crossref]

237. SeungJin Lim. An Algorithm for Detection of Chiastic Structures in Text Databases 263-269. [Crossref]

238. Titipat Achakulvisut, Daniel E. Acuna, Tulakan Ruangrong, Konrad Kording. 2016. Science Concierge: A Fast Content-Based Recommendation System for Scientific Publications. *PLOS ONE* **11**:7, e0158423. [Crossref]

239. Wei Wei, Gao Cong, Chunyan Miao, Feida Zhu, Guohui Li. 2016. Learning to Find Topic Experts in Twitter via Different Relations. *IEEE Transactions on Knowledge and Data Engineering* **28**:7, 1764-1778. [Crossref]

240. Sara Elena Garza Villarreal, Satu Elisa Schaeffer. 2016. Local bilateral clustering for identifying research topics and groups from bibliographical data. *Knowledge and Information Systems* **48**:1, 179-199. [Crossref]

241. Hyun Kyung Kim, Minyoung Kim. 2016. Model-induced term-weighting schemes for text classification. *Applied Intelligence* **45**:1, 30-43. [Crossref]

242. Charles H. P. Ferreira, Debora M. R. de Medeiros, Fabiana Santana. FCFilter: Feature selection based on clustering and genetic algorithms 2106-2113. [Crossref]

243. Lishuang Li, Shanshan Liu, Meiyue Qin, Yiwen Wang, Degen Huang. 2016. Extracting Biomedical Event with Dual Decomposition Integrating Word Embeddings. *IEEE/ACM Transactions on Computational Biology and Bioinformatics* **13**:4, 669-677. [Crossref]

244. Shansong Yang, Weiming Lu, Zhanjiang Zhang, Baogang Wei, Wenjia An. 2016. Amplifying scientific paper's abstract by leveraging data-weighted reconstruction. *Information Processing & Management* **52**:4, 698-719. [Crossref]

245. Minyoung Kim. 2016. Robust Algorithms for Combining Multiple Term Weighting Vectors for Document Classification. *The International Journal of Fuzzy Logic and Intelligent Systems* **16**:2, 81-86. [Crossref]

246. Ivan Candela, Gabriele Bavota, Barbara Russo, Rocco Oliveto. 2016. Using Cohesion and Coupling for Software Remodularization. *ACM Transactions on Software Engineering and Methodology* **25**:3, 1-28. [Crossref]

247. Marwa Naili, Anja Habacha Chaibi, Henda Hajjami Ben Ghezala. 2016. Exogenous approach to improve topic segmentation. *International Journal of Intelligent Computing and Cybernetics* **9**:2, 165-178. [Abstract] [Full Text] [PDF]

248. Ryan Light, jimi adams. 2016. Knowledge in motion: the evolution of HIV/AIDS research. *Scientometrics* **107**:3, 1227-1248. [Crossref]

249. Yu Zhao, Shaohua Wang, Ying Zou, Joanna Ng, Tinny Ng. Mining User Intents to Compose Services for End-Users 348-355. [Crossref]

250. Stuart McIlroy, Nasir Ali, Hammad Khalid, Ahmed E. Hassan. 2016. Analyzing and automatically labelling the types of user issues that are raised in mobile app reviews. *Empirical Software Engineering* **21**:3, 1067-1106. [Crossref]

251. Anna Knyazeva, Oleg Kolobov, Igor Turchanovsky. An example of empirical approach for bibliographic record linkage 1-6. [Crossref]

252. Richard Khoury, Francesca Sapsford. 2016. Latin word stemming using Wiktionary:. *Digital Scholarship in the Humanities* **31**:2, 368-373. [Crossref]

253. Andreas Kanavos, Christos Makris, Yannis Plegas, Evangelos Theodoridis. 2016. Ranking Web Search Results Exploiting Wikipedia. *International Journal on Artificial Intelligence Tools* **25**:03, 1650018. [Crossref]

254. Bekir Parlak, Alper Kursat Uysal. The impact of feature selection on medical document classification 1-5. [Crossref]

255. Peter Klimek, Aleksandar S. Jovanovic, Rainer Egloff, Reto Schneider. 2016. Successful fish go with the flow: citation impact prediction based on centrality measures for term–document networks. *Scientometrics* **107**:3, 1265-1282. [Crossref]

256. Jia Zhang, Runyu Shi, Weiyi Wang, Shenggu Lu, Yuanchen Bai, Qihao Bao, Tsengdar J. Lee, Kiran Nagaraja, Nimish Radia. A Bloom Filter-Powered Technique Supporting Scalable Semantic Service Discovery in Service Networks 81-90. [Crossref]

257. Dominic Duval, François Pétry. 2016. L'analyse automatisée du ton médiatique : construction et utilisation de la version française du Lexicoder Sentiment Dictionary. *Canadian Journal of Political Science* **49**:02, 197-220. [Crossref]

258. Evi Yulianti, Sharin Huspi, Mark Sanderson. 2016. Tweet-biased summarization. *Journal of the Association for Information Science and Technology* **67**:6, 1289-1300. [Crossref]

259. Fabio Palomba, Annibale Panichella, Andrea De Lucia, Rocco Oliveto, Andy Zaidman. A textual-based technique for Smell Detection 1-10. [Crossref]

260. Stefan Evert, Paul Greiner, João Filipe Baigger, Bastian Lang. 2016. A distributional approach to open questions in market research. *Computers in Industry* **78**, 16-28. [Crossref]

261. Johann Hauswald, Michael A. Laurenzano, Yunqi Zhang, Cheng Li, Austin Rovinski, Arjun Khurana, Ronald G. Dreslinski, Trevor Mudge, Vinicius Petrucci, Lingjia Tang, Jason Mars. 2016. Sirius Implications for Future Warehouse-Scale Computers. *IEEE Micro* **36**:3, 42-53. [Crossref]

262. Shaden Smith, George Karypis. A Medium-Grained Algorithm for Sparse Tensor Factorization 902-911. [Crossref]

263. Mohamad Nizam Kassim, Mohd Aizaini Maarof, Anazida Zainal, Amirudin Abdul Wahab. Word stemming challenges in Malay texts: A literature review 1-6. [Crossref]

264. Hanchen Jiang, Maoshan Qiang, Peng Lin. 2016. A topic modeling based bibliometric exploration of hydropower research. *Renewable and Sustainable Energy Reviews* **57**, 226-237. [Crossref]

265. Maria Lin, David W. Cheung. 2016. An automatic approach for tagging Web services using machine learning techniques1. *Web Intelligence* **14**:2, 99-118. [Crossref]

266. Johann Hauswald, Trevor Mudge, Vinicius Petrucci, Lingjia Tang, Jason Mars, Michael A. Laurenzano, Yunqi Zhang, Hailong Yang, Yiping Kang, Cheng Li, Austin Rovinski, Arjun Khurana, Ronald G. Dreslinski. 2016. Designing Future Warehouse-Scale Computers for Sirius, an End-to-End Voice and Vision Personal Assistant. *ACM Transactions on Computer Systems* **34**:1, 1-32. [Crossref]

267. Michael Siering, Jascha-Alexander Koch, Amit V. Deokar. 2016. Detecting Fraudulent Behavior on Crowdfunding Platforms: The Role of Linguistic and Content-Based Cues in Static and Dynamic Contexts. *Journal of Management Information Systems* **33**:2, 421-455. [Crossref]

268. Chetana T. Baviskar, Sandip S. Patil. Improvement of data object's membership by using Fuzzy K-Means clustering approach 139-147. [Crossref]

269. Nai-Wen Chi, Ken-Yu Lin, Nora El-Gohary, Shang-Hsien Hsieh. 2016. Evaluating the strength of text classification categories for supporting construction field inspection. *Automation in Construction* **64**, 78-88. [Crossref]

270. Wiem Chebil, Lina Fatima Soualmia, Mohamed Nazih Omri, Stéfan Jacques Darmoni. 2016. Indexing biomedical documents with a possibilistic network. *Journal of the Association for Information Science and Technology* **67**:4, 928-941. [Crossref]

271. Halil Kilicoglu, Dina Demner-Fushman. 2016. Bio-SCoRes: A Smorgasbord Architecture for Coreference Resolution in Biomedical Text. *PLOS ONE* **11**:3, e0148538. [Crossref]

272. Zhe Liu, Bernard J. Jansen. 2016. Understanding and Predicting Question Subjectivity in Social Question and Answering. *IEEE Transactions on Computational Social Systems* **3**:1, 32-41. [Crossref]

273. Bhawani Selvaretnam, Mohammed Belkhatir. 2016. A linguistically driven framework for query expansion via grammatical constituent highlighting and role-based concept weighting. *Information Processing & Management* **52**:2, 174-192. [Crossref]

274. Arash Heidarian, Michael J. Dinneen. A Hybrid Geometric Approach for Measuring Similarity Level Among Documents and Document Clustering 142-151. [Crossref]

275. Sameen Maruf, Kashif Javed, Haroon A. Babri. 2016. Improving Text Classification Performance with Random Forests-Based Feature Selection. *Arabian Journal for Science and Engineering* **41**:3, 951-964. [Crossref]

276. Annibale Panichella, Bogdan Dit, Rocco Oliveto, Massimiliano Di Penta, Denys Poshyvanyk, Andrea De Lucia. Parameterizing and Assembling IR-Based Solutions for SE Tasks Using Genetic Algorithms 314-325. [Crossref]

277. Rong Hu, Brian Mac Namee, Sarah Jane Delany. 2016. Active learning for text classification with reusability. *Expert Systems with Applications* **45**, 438-449. [Crossref]

278. Abeed Sarker, Karen O'Connor, Rachel Ginn, Matthew Scotch, Karen Smith, Dan Malone, Graciela Gonzalez. 2016. Social Media Mining for Toxicovigilance: Automatic Monitoring of Prescription Medication Abuse from Twitter. *Drug Safety* **39**:3, 231-240. [Crossref]

279. Edward Kai Fung Dang, Robert W.P. Luk, James Allan. 2016. A context-dependent relevance model. *Journal of the Association for Information Science and Technology* **67**:3, 582-593. [Crossref]

280. Gabriele Bavota. Mining Unstructured Data in Software Repositories: Current and Future Trends 1-12. [Crossref]

281. Froila Helixia Dsouza, Ananthanarayana V.S.. Document classification with a weighted frequency pattern tree algorithm 29-34. [Crossref]

282. Alfredo Kalaitzis, Maria Ivanova Gorinova, Yoad Lewenberg, Yoram Bachrach, Michael Fagan, Dean Carignan, Nitin Gautam. Predicting Gaming Related Properties from Twitter Profiles 28-35. [Crossref]

283. Dragan Ivanovi?, Yuh-Shan Ho. 2016. Highly cited articles in the Information Science and Library Science category in Social Science Citation Index: A bibliometric analysis. *Journal of Librarianship and Information Science* **48**:1, 36-46. [Crossref]

284. Iman Amini, David Martinez, Xiaodong Li, Mark Sanderson. 2016. Improving patient record search: A meta-data based approach. *Information Processing & Management* **52**:2, 258-272. [Crossref]

285. Wei Wei, ZhaoYan Ming, Liqiang Nie, Guohui Li, Jianjun Li, Feida Zhu, Tianfeng Shang, Changyin Luo. 2016. Exploring heterogeneous features for query-focused summarization of categorized community answers. *Information Sciences* **330**, 403-423. [Crossref]

286. Sumit Bhatia, Prakhar Biyani, Prasenjit Mitra. 2016. Identifying the role of individual user messages in an online discussion and its use in thread retrieval. *Journal of the Association for Information Science and Technology* **67**:2, 276-288. [Crossref]

287. Archana Chougule, Vijay Kumar Jha, Debajyoti Mukhopadhyay. Ontology Based System for Pests and Disease Management of Grapes in India 133-138. [Crossref]

288. Abeed Sarker, Diego Mollá, Cecile Paris. 2016. Query-oriented evidence extraction to support evidence-based medicine practice. *Journal of Biomedical Informatics* **59**, 169-184. [Crossref]

289. Bilel Elayeb, Ibrahim Bounhas. 2016. Arabic Cross-Language Information Retrieval. *ACM Transactions on Asian and Low-Resource Language Information Processing* **15**:3, 1-44. [Crossref]

290. James C. Brau, James Cicon, Grant McQueen. 2016. Soft Strategic Information and IPO Underpricing. *Journal of Behavioral Finance* **17**:1, 1-17. [Crossref]

291. S. Wagner. Natural language processing is no free lunch 175-179. [Crossref]

292. Yong Yang, Ying Li, Yang Yue, Zhonghai Wu, Wenlong Shao. CUT: A Combined Approach for Tag Recommendation in Software Information Sites 599-612. [Crossref]

293. Benjamin Guthier, Ralf Dörner, Hector P. Martinez. Affective Computing in Games 402-441. [Crossref]

294. Atif Khan, Naomie Salim, Haleem Farman. Clustered genetic semantic graph approach for multi-document abstractive summarization 63-70. [Crossref]

295. Mohamad Nizam Kassim, Mohd Aizaini Maarof, Anazida Zainal, Amirudin Abdul Wahab. Enhanced Rules Application Order to Stem Affixation, Reduplication and Compounding Words in Malay Texts 71-85. [Crossref]

296. Yin-Lai Yeong, Tien-Ping Tan, Siti Khaotijah Mohammad. 2016. Using Dictionary and Lemmatizer to Improve Low Resource English-Malay Statistical Machine Translation System. *Procedia Computer Science* **81**, 243-249. [Crossref]

297. Josiah Wang, Katja Markert, Mark Everingham. Harvesting Training Images for Fine-Grained Object Categories Using Visual Descriptions 549-560. [Crossref]

298. Praveen Krishnan, C. V. Jawahar. Matching Handwritten Document Images 766-782. [Crossref]

299. João Guerreiro, Paulo Rita, Duarte Trigueiros. 2016. A Text Mining-Based Review of Cause-Related Marketing Literature. *Journal of Business Ethics* **139**:1, 111. [Crossref]

300. Miroslav Smatana, Peter Butka. Extraction of keyphrases from single document based on hierarchical concepts 93-98. [Crossref]

301. Giacomo Domeniconi, Gianluca Moro, Andrea Pagliarani, Roberto Pasolini. Cross-Domain Sentiment Classification via Polarity-Driven State Transitions in a Markov Model 118-138. [Crossref]

302. Archana Chougule, Vijay Kumar Jha, Debajyoti Mukhopadhyay. Using IoT for integrated pest management 17-22. [Crossref]

303. David Kaczynski, Lisa Gandy, Gongzhu Hu. Innovations in News Media: Crisis Classification System 125-138. [Crossref]

304. Azad Dehghan, Tom Liptrot, Daniel Tibble, Matthew Barker-Hewitt, Goran Nenadic. Identification of Occupation Mentions in Clinical Narratives 359-365. [Crossref]

305. Agnese Augello, Alfredo Cuzzocrea, Giovanni Pilato, Carmelo Spiccia, Giorgio Vassallo. An Innovative Similarity Measure for Sentence Plagiarism Detection 552-566. [Crossref]

306. Ali Jahanian. Design Mining Color Semantics 15-55. [Crossref]

307. Chris D. Paice. Stemming 1-5. [Crossref]

308. Rosie Jones, Fuchun Peng. Web Search Query Rewriting 1-4. [Crossref]

309. Eiman Aeiad, Farid Meziane. Validating Learning Outcomes of an E-Learning System Using NLP 292-300. [Crossref]

310. Rogelio Nazar. 2016. Distributional analysis applied to terminology extraction. *Terminology* **22**:2, 141-170. [Crossref]

311. Anastasia Giachanou, Morgan Harvey, Fabio Crestani. Topic-Specific Stylistic Variations for Opinion Retrieval on Twitter 466-478. [Crossref]

312. R. W. Grundmeier, A. J. Masino, T. C. Casper, J. M. Dean, J. Bell, R. Enriquez, S. Deakyne, J. M. Chamberlain, E. R. Alpern. 2016. Identification of Long Bone Fractures in Radiology Reports Using Natural Language Processing to support Healthcare Quality Improvement. *Applied Clinical Informatics* **7**:4, 1051-1068. [Crossref]

313. Daniela Pohl, Abdelhamid Bouchachia, Hermann Hellwagner. 2016. Online indexing and clustering of social media data for emergency management. *Neurocomputing* **172**, 168-179. [Crossref]

314. Mounia Haddoud, Aïcha Mokhtari, Thierry Lecroq, Saïd Abdeddaïm. Supervised Term Weights for Biomedical Text Classification: Improvements in Nearest Centroid Computation 98-113. [Crossref]

315. Maxim Yakovlev, Ekaterina Chernyak. Using Annotated Suffix Tree Suffix tree Similarity Similarity Measure for Text Summarisation 103-112. [Crossref]

316. Jan-Otto Jandl. Information Processing in Securitized Real Estate Markets -- How Newspaper Content and Online Search Behavior Help Explaining Market Movements 1092-1101. [Crossref]

317. Ananthi Sheshasaayee, V. R. Angela Deepa. A Conceptual Model for Acquisition of Morphological Features of Highly Agglutinative Tamil Language Using Unsupervised Approach 499-507. [Crossref]

318. Wael M. S. Yafooz, Siti Z. Z. Abidin, Nasiroh Omar, Shadi Hilles. Interactive Big Data Visualization Model Based on Hot Issues (Online News Articles) 89-99. [Crossref]

319. Tuhin Sharma, Durga Toshniwal. A Generalized Framework for Quantifying Trust of Social Media Text Documents 692-713. [Crossref]

320. Yllias Chali, Mohsin Uddin. Multi-document Summarization Based on Atomic Semantic Events and Their Temporal Relationships 366-377. [Crossref]

321. Eman M. Bahgat, Sherine Rady, Walaa Gad. An E-mail Filtering Approach Using Classification Techniques 321-331. [Crossref]

322. Ilias Gialampoukidis, Stefanos Vrochidis, Ioannis Kompatsiaris. A Hybrid Framework for News Clustering Based on the DBSCAN-Martingale and LDA 170-184. [Crossref]

323. Jinhyuk Yun, Sang Hoon Lee, Hawoong Jeong. 2016. Intellectual interchanges in the history of the massive online open-editing encyclopedia, Wikipedia. *Physical Review E* **93**:1. . [Crossref]

324. Blessing Mbipom, Susan Craw, Stewart Massie. Harnessing Background Knowledge for E-Learning Recommendation 3-17. [Crossref]

325. Alper Kursat Uysal. 2016. An improved global feature selection scheme for text classification. *Expert Systems with Applications* **43**, 82-92. [Crossref]

326. Shweta, Asif Ekbal, Sriparna Saha, Pushpak Bhattacharyya. A deep learning architecture for protein-protein Interaction Article identification 3128-3133. [Crossref]

327. Jagbir Kaur, Meenakshi Bansal. Multi-layered sentiment analytical model for product review mining 415-420. [Crossref]

328. Xiaojun Quan, Chunyu Kit. 2015. Towards non-monotonic sentence alignment. *Information Sciences* **323**, 34-47. [Crossref]

329. Guido Zuccon, Sankalp Khanna, Anthony Nguyen, Justin Boyle, Matthew Hamlet, Mark Cameron. 2015. Automatic detection of tweets reporting cases of influenza like illnesses in Australia. *Health Information Science and Systems* **3**:S1. . [Crossref]

330. Onder Coban, Baris Ozyer and Gulsah Tumuklu Ozyer. A Comparison of Similarity Metrics for Sentiment Analysis on Turkish Twitter Feeds 333-338. [Crossref]

331. Walaa Gad, Sherine Rady. Email filtering based on supervised learning and mutual information feature selection 147-152. [Crossref]

332. Kabil Boukhari, Mohamed Nazih Omri. SAID: A new stemmer algorithm to indexing unstructured Document 59-63. [Crossref]

333. Anastasia Giachanou, Michail Salampasis, Georgios Paltoglou. 2015. Multilayer source selection as a tool for supporting patent search and classification. *Information Retrieval Journal* **18**:6, 559-585. [Crossref]

334. Amina Madani, Omar Boussaid, Djamel Eddine Zegour. 2015. Real-time trending topics detection and description from Twitter content. *Social Network Analysis and Mining* **5**:1. . [Crossref]

335. Bobby G Sumpter, Rama K Vasudevan, Thomas Potok, Sergei V Kalinin. 2015. A bridge for accelerating materials by design. *npj Computational Materials* **1**:1. . [Crossref]

336. Vahid Behbood, Jie Lu, Guangquan Zhang, Witold Pedrycz. 2015. Multistep Fuzzy Bridged Refinement Domain Adaptation Algorithm and Its Application to Bank Failure Prediction. *IEEE Transactions on Fuzzy Systems* **23**:6, 1917-1935. [Crossref]

337. Spyridon Ntougias, Alla Lapidus, Alex Copeland, T. B. K. Reddy, Amrita Pati, Natalia N. Ivanova, Victor M. Markowitz, Hans-Peter Klenk, Tanja Woyke, Constantinos Fasseas, Nikos C. Kyrpides, Georgios I. Zervakis. 2015. High-quality permanent draft genome sequence of the extremely osmotolerant diphenol degrading bacterium Halotalea alkalilenta AW-7T, and emended description of the genus Halotalea. *Standards in Genomic Sciences* **10**:1. . [Crossref]

338. Nancy M. Rizk, Amr Ebada, Eman S. Nasr. Investigating mobile applications' requirements evolution through sentiment analysis of users' reviews 123-130. [Crossref]

339. Pietro Di Lena, Giacomo Domeniconi, Luciano Margara, Gianluca Moro. 2015. GOTA: GO term annotation of biomedical literature. *BMC Bioinformatics* **16**:1. . [Crossref]

340. Jean-Charles Lamirel, Pascal Cuxac, Aneesh Sreevallabh Chivukula, Kafil Hajlaoui. 2015. Optimizing text classification through efficient feature selection based on quality metric. *Journal of Intelligent Information Systems* **45**:3, 379-396. [Crossref]

341. WeePheng Goh, Xiaohui Tao, Ji Zhang, Jianming Yong. A Study of Drug Interaction for Personalised Decision Support in Dental Clinics 88-91. [Crossref]

342. Kenric W. Hammond, Alon Y. Ben-Ari, Ryan J. Laundry, Edward J. Boyko, Matthew H. Samore. 2015. The Feasibility of Using Large-Scale Text Mining to Detect Adverse Childhood Experiences in a VA-Treated Population. *Journal of Traumatic Stress* **28**:6, 505-514. [Crossref]

343. Yun Zhang, David Lo, Xin Xia, Bowen Xu, Jianling Sun, Shanping Li. Combining Software Metrics and Text Features for Vulnerable File Prediction 40-49. [Crossref]

344. Keisuke Yano, Akihiko Matsuo. Labeling Feature-Oriented Software Clusters for Software Visualization Application 354-361. [Crossref]

345. Khaled Albishre, Mubarak Albathan, Yuefeng Li. Effective 20 Newsgroups Dataset Cleaning 98-101. [Crossref]

346. Dingding Wang, Mitsunori Ogihara, Carlos Gallo, Juan A. Villamar, Justin D. Smith, Wouter Vermeer, Gracelyn Cruden, Nanette Benbow, C. Hendricks Brown. 2015. Automatic classification of communication logs into implementation stages via text analysis. *Implementation Science* **11**:1. . [Crossref]

347. R.J. Prathibha, M.C. Padma. Design of rule based lemmatizer for Kannada inflectional words 264-269. [Crossref]

348. Bowen Xu, David Lo, Xin Xia, Ashish Sureka, Shanping Li. EFSPredictor: Predicting Configuration Bugs with Ensemble Feature Selection 206-213. [Crossref]

349. Sherine Rady. A business intelligent technique for sentiment estimation by management sectors 370-376. [Crossref]

350. Hyunuk Kim, Sang-Jin Ahn, Woo-Sung Jung. 2015. A Study on Technology Forecasting based on Co-occurrence Network of Keyword in Multidisciplinary Journals. *Journal of the Korean Operations Research and Management Science Society* **40**:4, 49-63. [Crossref]

351. Hossein Taghi-Zadeh, Mohammad Hadi Sadreddini, Mohammad Hasan Diyanati, Amir Hossein Rasekh. 2015. A New Hybrid Stemming Method for Persian Language. *Digital Scholarship in the Humanities* fqv053. [Crossref]

352. Vasileios Papapanagiotou, Christos Diou, Anastasios Delopoulos. 2015. Improving Concept-Based Image Retrieval with Training Weights Computed from Tags. *ACM Transactions on Multimedia Computing, Communications, and Applications* **12**:2, 1-22. [Crossref]

353. Seyedeh Homa Alizadeh, N. R. Arghami, Majid Vafaei Jahan. A recommender system based on trust and semantics in collaborative systems using a new measure of association 530-535. [Crossref]

354. Daniel Ansari. Sentiment Polarity Classification Using Structural Features 1270-1273. [Crossref]

355. Chayan Chakrabarti, George F. Luger. 2015. Artificial conversations for customer service chatter bots: Architecture, algorithms, and evaluation metrics. *Expert Systems with Applications* **42**:20, 6878-6897. [Crossref]

356. Yasuhiro Yamamoto, Daisuke Miyamoto, Masaya Nakayama. Text-Mining Approach for Estimating Vulnerability Score 67-73. [Crossref]

357. Sheng-hua Zhong, Yan Liu, Bin Li, Jing Long. 2015. Query-oriented unsupervised multi-document summarization via deep learning model. *Expert Systems with Applications* **42**:21, 8146-8155. [Crossref]

358. Fei Lv, Hongyu Zhang, Jian-guang Lou, Shaowei Wang, Dongmei Zhang, Jianjun Zhao. CodeHow: Effective Code Search Based on API Understanding and Extended Boolean Model (E) 260-270. [Crossref]

359. Gianni Costa, Riccardo Ortale. Fully-Automatic XML Clustering by Structure-Constrained Phrases 146-153. [Crossref]

360. Cheng-Ying Liu, Ming-Syan Chen, Chi-Yao Tseng. 2015. IncreSTS: Towards Real-Time Incremental Short Text Summarization on Comment Streams from Social Network Services. *IEEE Transactions on Knowledge and Data Engineering* **27**:11, 2986-3000. [Crossref]

361. Le Nguyen Hoai Nam, Ho Bao Quoc. A Combined Approach for Filter Feature Selection in Document Classification 317-324. [Crossref]

362. B. P. Pande, Pawan Tamta, H. S. Dhami. 2015. A simple algorithm for the problem of suffix stripping. *International Journal of Applied Linguistics* **25**:3, 315-328. [Crossref]

363. Soroush Vosoughi, Deb Roy. A Human-Machine Collaborative System for Identifying Rumors on Twitter 47-50. [Crossref]

364. Saksham Singhal, Vikram Pudi. Dispersion Based Similarity for Mining Similar Papers in Citation Network 524-531. [Crossref]

365. Ji-Hyun Lee, Wei-Chang Yeh, Mei-Chi Chuang. 2015. Web page classification based on a simplified swarm optimization. *Applied Mathematics and Computation* **270**, 13-24. [Crossref]

366. Mostefai Abdelkader, Ricardo Perez-Castillo, Malki Mimoun. Using text mining to locate concepts in execution traces 1-6. [Crossref]

367. . References 445-458. [Crossref]

368. Marjori N. M. Klinczak, Celso A. A. Kaestner. A study on topics identification on Twitter using clustering algorithms 1-6. [Crossref]

369. Nicolas Boissel-Dallier, Fréderick Benaben, Jean-Pierre Lorré, Hervé Pingaud. 2015. Mediation information system engineering based on hybrid service composition mechanism. *Journal of Systems and Software* **108**, 39-59. [Crossref]

370. Filippo Galgani, Paul Compton, Achim Hoffmann. 2015. LEXA: Building knowledge bases for automatic legal citation classification. *Expert Systems with Applications* **42**:17-18, 6391-6407. [Crossref]

371. Edgar Kalkowski, Bernhard Sick. Using Ontology-Based Similarity Measures to Find Training Data for Problems with Sparse Data 1693-1699. [Crossref]

372. Manoj Kumar, D K Yadav, Vijay Kumar Gupta. Frequent term based text document clustering: A new approach 11-15. [Crossref]

373. Taemin Jo, Jee-Hyong Lee. 2015. Latent Keyphrase Extraction Using Deep Belief Networks. *The International Journal of Fuzzy Logic and Intelligent Systems* **15**:3, 153-158. [Crossref]

374. Christopher J. Fariss, Fridolin J. Linder, Zachary M. Jones, Charles D. Crabtree, Megan A. Biek, Ana-Sophia M. Ross, Taranamol Kaur, Michael Tsai. 2015. Human Rights Texts: Converting Human Rights Primary Source Documents into Data. *PLOS ONE* **10**:9, e0138935. [Crossref]

375. Karolina Sylwester, Matthew Purver. 2015. Twitter Language Use Reflects Psychological Differences between Democrats and Republicans. *PLOS ONE* **10**:9, e0137422. [Crossref]

376. S. Poireiton Meitei, Bipul Syam Purkayastha, H. Mamata Devi. Development of a Manipuri stemmer: A hybrid approach 128-131. [Crossref]

377. Jeff Keith, Chris Westbury, James Goldman. 2015. Performance impact of stop lists and morphological decomposition on word–word corpus-based semantic space models. *Behavior Research Methods* **47**:3, 666-684. [Crossref]

378. SAAD Y SAIT, AKSHAY BHANDARI, SHREYA KHARE, CYRIAC JAMES, HEMA A MURTHY. 2015. Multi-level anomaly detection: Relevance of big data analytics in networks. *Sadhana* **40**:6, 1737-1767. [Crossref]

379. Qingliang Miao, Ruiyu Fang, Yao Meng. 2015. Extracting and Integrating Nutrition Related Linked Data. *International Journal of Semantic Computing* **09**:03, 353-372. [Crossref]

380. Fabio Palomba, Mario Linares-Vasquez, Gabriele Bavota, Rocco Oliveto, Massimiliano Di Penta, Denys Poshyvanyk, Andrea De Lucia. User reviews matter! Tracking crowdsourced reviews to support evolution of successful apps 291-300. [Crossref]

381. Gillian J. Greene, Bernd Fischer. Interactive tag cloud visualization of software version control repositories 56-65. [Crossref]

382. Muslim Chochlov, Michael English, Jim Buckley. Using changeset descriptions as a data source to assist feature location 51-60. [Crossref]

383. Amina Madani, Omar Boussaid, Djamel Eddine Zegour. 2015. New Information in Trending Topics of Tweets by Labelled Clusters. *Journal of Information & Knowledge Management* **14**:03, 1550019. [Crossref]

384. Chris Staff, Joel Azzopardi, Colin Layfield, Daniel Mercieca. Search Results Clustering without External Resources 276-280. [Crossref]

385. Arie Croitoru, N. Wayant, A. Crooks, J. Radzikowski, A. Stefanidis. 2015. Linking cyber and physical spaces through community detection and clustering in social media feeds. *Computers, Environment and Urban Systems* **53**, 47-64. [Crossref]

386. Yun Zhang, David Lo, Xin Xia, Jian-Ling Sun. 2015. Multi-Factor Duplicate Question Detection in Stack Overflow. *Journal of Computer Science and Technology* **30**:5, 981-997. [Crossref]

387. Erel Joffe, Emily J Pettigrew, Jorge R Herskovic, Charles F Bearden, Elmer V Bernstam. 2015. Expert guided natural language processing using one-class classification. *Journal of the American Medical Informatics Association* **22**:5, 962-966. [Crossref]

388. Andrew Karl, James Wisnowski, W. Heath Rushing. 2015. A practical guide to text mining with topic extraction. *Wiley Interdisciplinary Reviews: Computational Statistics* **7**:5, 326-340. [Crossref]

389. Getachew Mamo Wegari, Massimo Melucci, Solomon Teferra. Suffix sequences based morphological segmentation for Afaan Oromo 1-6. [Crossref]

390. Takahiro Komamizu, Toshiyuki Amagasa, Hiroyuki Kitagawa. 2015. Facet-value extraction scheme from textual contents in XML data. *International Journal of Web Information Systems* **11**:3, 270-290. [Abstract] [Full Text] [PDF]

391. Giuseppe Rizzo, Federico Tomassetti, Antonio Vetrò, Luca Ardito, Marco Torchiano, Maurizio Morisio, Raphaël Troncy. 2015. Semantic Enrichment for Recommendation of Primary Studies in a Systematic Literature Review. *Digital Scholarship in the Humanities* fqv031. [Crossref]

392. Obada Alhabashneh, Rahat Iqbal, Faiyaz Doctor, Saad Amin. Adaptive information retrieval system based on fuzzy profiling 1-8. [Crossref]

393. Robert Karam, Ruchir Puri, Swaroop Ghosh, Swarup Bhunia. 2015. Emerging Trends in Design and Applications of Memory-Based Computing and Content-Addressable Memories. *Proceedings of the IEEE* **103**:8, 1311-1330. [Crossref]

394. Siddharth Biswal, Zarina Nip, Valdery Moura Junior, Matt T. Bianchi, Eric S. Rosenthal, M. Brandon Westover. Automated information extraction from free-text EEG reports 6804-6807. [Crossref]

395. Waddah Waheeb, Rozaida Ghazali, Mustafa Mat Deris. Content-based SMS spam filtering based on the Scaled Conjugate Gradient backpropagation algorithm 675-680. [Crossref]

396. Christopher Scaffidi. What Training is Needed by Practicing Engineers Who Create Cyberphysical Systems? 298-305. [Crossref]

397. Yijun Yu, Virginia N.L. Franqueira, Thein Than Tun, Roel J. Wieringa, Bashar Nuseibeh. 2015. Automated analysis of security requirements through risk-based argumentation. *Journal of Systems and Software* **106**, 102-116. [Crossref]

398. Long Ma, Yanqing Zhang, Raj Sunderraman, Peter T. Fox, Angela R. Laird, Jessica A. Turner, Matthew D. Turner. Hybrid feature selection methods for online biomedical publication classification 1-8. [Crossref]

399. Zhiyuan Zhang, Hengpan Xu, Weigang Huo. Topical text network construction based on seed word augmentation 1470-1474. [Crossref]

400. Eleonora D'Andrea, Pietro Ducange, Beatrice Lazzerini, Francesco Marcelloni. 2015. Real-Time Detection of Traffic From Twitter Stream Analysis. *IEEE Transactions on Intelligent Transportation Systems* **16**:4, 2269-2283. [Crossref]

401. Tanmay Basu, C.A. Murthy. 2015. A similarity assessment technique for effective grouping of documents. *Information Sciences* **311**, 149-162. [Crossref]

402. Kasun Amarasinghe, Milos Manic, Ryan Hruska. Optimal stop word selection for text mining in critical infrastructure domain 1-6. [Crossref]

403. Theodora Tsikrika, George Kalpakis, Stefanos Vrochidis, Ioannis Kompatsiaris, Iraklis Paraskakis, Isaak Kavasidis, Jonathan Middleton, Una Williamson. A Framework for the Discovery, Analysis, and Retrieval of Multimedia Homemade Explosives Information on the Web 601-610. [Crossref]

404. Saprativa Bhattacharjee, Anirban Das, Ujjwal Bhattacharya, Swapan K. Parui, Sudipta Roy. Sentiment analysis using cosine similarity measure 27-32. [Crossref]

405. Kamal Sarkar, Khushbu Saraf, Avishikta Ghosh. Improving graph based multidocument text summarization using an enhanced sentence similarity measure 359-365. [Crossref]

406. Juan Manuel Rodriguez, Alejandro Zunino, Cristian Mateos, Felix Oscar Segura, Emmanuel Rodriguez. Improving REST Service Discovery with Unsupervised Learning Techniques 97-104. [Crossref]

407. Dong Ren, Long Ma, Yanqing Zhang, Raj Sunderraman, Peter T. Fox, Angela R. Laird, Jessica A. Turner, Matthew D. Turner. Online biomedical publication classification using Multi-Instance Multi-Label algorithms with feature reduction 234-241. [Crossref]

408. Ruben Costa, Celson Lima. 2015. Document Clustering Using an Ontology-Based Vector Space Model. *International Journal of Information Retrieval Research* **5**:3, 39-60. [Crossref]

409. Wei Zong, Feng Wu, Lap-Keung Chu, Domenic Sculli. 2015. A discriminative and semantic feature selection method for text categorization. *International Journal of Production Economics* **165**, 215-222. [Crossref]

410. Yiguo Pu, Jinqiao Shi, Xiaojun Chen, Li Guo, Tingwen Liu. Towards misdirected email detection based on multi-attributes 796-802. [Crossref]

411. Saif M. Mohammad, Xiaodan Zhu, Svetlana Kiritchenko, Joel Martin. 2015. Sentiment, emotion, purpose, and style in electoral tweets. *Information Processing & Management* **51**:4, 480-499. [Crossref]

412. Chih-Fong Tsai, Shih-Wen Ke, Kenneth McGarry, Ming-Yi Lin. 2015. LocalContent: a personal scientific document retrieval system. *The Electronic Library* **33**:3, 373-385. [Abstract] [Full Text] [PDF]

413. Juliana Tarossi Pollettini, Hugo Cesar Pessotti, Antonio Pazin Filho, Evandro Eduardo Seron Ruiz, Mario Sergio Adolfi Junior. Applying Natural Language Processing, Information Retrieval and Machine Learning to Decision Support in Medical Coordination in an Emergency Medicine Context 316-319. [Crossref]

414. Law Sheng Xun, Swapna Gottipati, Venky Shankararaman. Text-mining approach for verifying alignment of information systems curriculum with industry skills 1-6. [Crossref]

415. Xiaocao Hu, Zhiyong Feng, Keman Huang, Shizhan Chen. Classification Based Parameter Association for Non-redundant Annotation 688-695. [Crossref]

416. A.M. Supraja, S. Archana, S. Suvetha, T.V. Geetha. Patent search and trend analysis 501-506. [Crossref]

417. Samad Paydar, Mohsen Kahani. 2015. A semantic web enabled approach to reuse functional requirements models in web engineering. *Automated Software Engineering* **22**:2, 241-288. [Crossref]

418. Kazunari Sugiyama, Min-Yen Kan. 2015. A comprehensive evaluation of scholarly paper recommendation using potential citation papers. *International Journal on Digital Libraries* **16**:2, 91-109. [Crossref]

419. Abeed Sarker, Diego Mollá, Cécile Paris. 2015. Automatic evidence quality prediction to support evidence-based decision making. *Artificial Intelligence in Medicine* **64**:2, 89-103. [Crossref]

420. Yehia Elshater, Khalid Elgazzar, Patrick Martin. goDiscovery: Web Service Discovery Made Efficient 711-716. [Crossref]

421. Nikitas N. Karanikolas. 2015. Supervised learning for building stemmers. *Journal of Information Science* **41**:3, 315-328. [Crossref]

422. Edward Kai FUNG Dang, Robert Wing Pong Luk, James Allan. 2015. Fast-Forward Index Methods for Pseudo-Relevance Feedback Retrieval. *ACM Transactions on Information Systems* **33**:4, 1-33. [Crossref]

423. Artemy Kolchinsky, Anália Lourenço, Heng-Yi Wu, Lang Li, Luis M. Rocha. 2015. Extraction of Pharmacokinetic Evidence of Drug–Drug Interactions from the Literature. *PLOS ONE* **10**:5, e0122199. [Crossref]

424. Matilde Trevisani, Arjuna Tuzzi. 2015. A portrait of JASA: the History of Statistics through analysis of keyword counts in an early scientific journal. *Quality & Quantity* **49**:3, 1287-1304. [Crossref]

425. Tom Chao Zhou, Michael Rung-Tsong Lyu, Irwin King, Jie Lou. 2015. Learning to suggest questions in social media. *Knowledge and Information Systems* **43**:2, 389-416. [Crossref]

426. Tien-Duy B. Le, Mario Linares-Vasquez, David Lo, Denys Poshyvanyk. RCLinker: Automated Linking of Issue Reports and Commits Leveraging Rich Contextual Information 36-47. [Crossref]

427. Jingxuan Li, Haijun Zhang, Bin Luo, Yan Li. Mining Tags from Flickr User Comments Using a Hybrid Ranking Model 198-204. [Crossref]

428. Ardiles Sinaga, Adiwijaya, Hertog Nugroho. Development of word-based text compression algorithm for Indonesian language document 450-454. [Crossref]

429. Lorena Leal Bando, Falk Scholer, Andrew Turpin. 2015. Query-biased summary generation assisted by query expansion. *Journal of the Association for Information Science and Technology* **66**:5, 961-979. [Crossref]

430. Yuncheng Jiang, Xiaopei Zhang, Yong Tang, Ruihua Nie. 2015. Feature-based approaches to semantic similarity assessment of concepts using Wikipedia. *Information Processing & Management* **51**:3, 215-234. [Crossref]

431. Yavuz Selim Bozan, Onder Coban, Gulsah Tumuklu Ozyer, Baris Ozyer. SMS spam filtering based on text classification and expert system 2345-2348. [Crossref]

432. Taemin Cho, Jee-Hyong Lee. 2015. Latent Keyphrase Extraction Using LDA Model. *Journal of Korean Institute of Intelligent Systems* **25**:2, 180-185. [Crossref]

433. F. Hjorth, R. Klemmensen, S. Hobolt, M. E. Hansen, P. Kurrild-Klitgaard. 2015. Computers, coders, and voters: Comparing automated methods for estimating party positions. *Research & Politics* **2**:2. . [Crossref]

434. Yongwen Huang, Zuhua Jiang, Chengneng He, Jianfeng Liu, Bo Song, Lijun Liu. 2015. A semantic-based visualised wiki system (SVWkS) for lesson-learned knowledge reuse situated in product design. *International Journal of Production Research* **53**:8, 2524-2541. [Crossref]

435. Yi Liu, Yushun Fan, Keman Huang, Wei Tan. 2015. Failure analysis and tolerance strategies in web service ecosystems. *Concurrency and Computation: Practice and Experience* **27**:5, 1355-1374. [Crossref]

436. Qin Jin, Chengxin Li, Shizhe Chen, Huimin Wu. Speech emotion recognition with acoustic and lexical features 4749-4753. [Crossref]

437. Guymon R. Hall, Kazem Taghva. Using the Web 1T 5-Gram Database for Attribute Selection in Formal Concept Analysis to Correct Overstemmed Clusters 651-654. [Crossref]

438. Wei Song, Yingying Qiao, Soon Cheol Park, Xuezhong Qian. 2015. A hybrid evolutionary computation approach with its application for optimizing text document clustering. *Expert Systems with Applications* **42**:5, 2517-2524. [Crossref]

439. Helge Holzmann, Nina Tahmasebi, Thomas Risse. 2015. Named entity evolution recognition on the Blogosphere. *International Journal on Digital Libraries* **15**:2-4, 209-235. [Crossref]

440. Johann Hauswald, Lingjia Tang, Jason Mars, Michael A. Laurenzano, Yunqi Zhang, Cheng Li, Austin Rovinski, Arjun Khurana, Ronald G. Dreslinski, Trevor Mudge, Vinicius Petrucci. 2015. Sirius. *ACM SIGPLAN Notices* **50**:4, 223-238. [Crossref]

441. Johann Hauswald, Lingjia Tang, Jason Mars, Michael A. Laurenzano, Yunqi Zhang, Cheng Li, Austin Rovinski, Arjun Khurana, Ronald G. Dreslinski, Trevor Mudge, Vinicius Petrucci. 2015. Sirius. *ACM SIGARCH Computer Architecture News* **43**:1, 223-238. [Crossref]

442. Sahar Changuel, Nicolas Labroche, Bernadette Bouchon-Meunier. 2015. Resources Sequencing Using Automatic Prerequisite--Outcome Annotation. *ACM Transactions on Intelligent Systems and Technology* **6**:1, 1-30. [Crossref]

443. Adrian-Gabriel Chifu, Florentina Hristea, Josiane Mothe, Marius Popescu. 2015. Word sense discrimination in information retrieval: A spectral clustering-based approach. *Information Processing & Management* **51**:2, 16-31. [Crossref]

444. Daniel Abril, Guillermo Navarro-Arribas, Vicenç Torra. 2015. Spherical microaggregation: Anonymizing sparse vector spaces. *Computers & Security* **49**, 28-44. [Crossref]

445. Lijun Wang, Ming Dong. 2015. Exemplar-based low-rank matrix decomposition for data clustering. *Data Mining and Knowledge Discovery* **29**:2, 324-357. [Crossref]

446. Hyunseok Park, Janghyeok Yoon. 2015. A chance discovery-based approach for new product–service system (PSS) concepts. *Service Business* **9**:1, 115-135. [Crossref]

447. Dave Binkley, Dawn Lawrie, Christopher Uehlinger, Daniel Heinz. 2015. Enabling improved IR-based feature location. *Journal of Systems and Software* **101**, 30-42. [Crossref]

448. Deba P. Mandal, Dipankar Kundu, Saptaditya Maiti. Finding experts in community question answering services: A theme based query likelihood language approach 423-427. [Crossref]

449. MASOUD NAROUEI, MANSOUR AHMADI, ASHKAN SAMI. 2015. SePaS: Word sense disambiguation by sequential patterns in sentences. *Natural Language Engineering* **21**:02, 251-269. [Crossref]

450. LiGuo Huang, Vincent Ng, Isaac Persing, Mingrui Chen, Zeheng Li, Ruili Geng, Jeff Tian. 2015. AutoODC: Automated generation of orthogonal defect classifications. *Automated Software Engineering* **22**:1, 3-46. [Crossref]

451. Kazunari Sugiyama, Min-Yen Kan. 2015. "Towards higher relevance and serendipity in scholarly paper recommendation" by Kazunari Sugiyama and Min-Yen Kan with Martin Vesely as coordinator. *ACM SIGWEB Newsletter* :Winter, 1-16. [Crossref]

452. Jinhyuk Yun, Pan-Jun Kim, Hawoong Jeong. 2015. Anatomy of Scientific Evolution. *PLOS ONE* **10**:2, e0117388. [Crossref]

453. Matthias Zwick. 2015. Automated curation of gene name normalization results using the Konstanz information miner. *Journal of Biomedical Informatics* **53**, 58-64. [Crossref]

454. Antonio Jimeno Yepes, Rafael Berlanga. 2015. Knowledge based word-concept model estimation and refinement for biomedical text mining. *Journal of Biomedical Informatics* **53**, 300-307. [Crossref]

455. Pranpreya Sriwannawit, Ulf Sandström. 2015. Large-scale bibliometric review of diffusion research. *Scientometrics* **102**:2, 1615-1645. [Crossref]

456. Surafel Lemma Abebe, Paolo Tonella. 2015. Extraction of domain concepts from the source code. *Science of Computer Programming* **98**, 680-706. [Crossref]

457. Andreas Kanavos, Christos Makris, Evangelos Theodoridis. 2015. Topic Categorization of Biomedical Abstracts. *International Journal on Artificial Intelligence Tools* **24**:01, 1540004. [Crossref]

458. Xiaodan Zhu, Peter Turney, Daniel Lemire, André Vellino. 2015. Measuring academic influence: Not all citations are equal. *Journal of the Association for Information Science and Technology* **66**:2, 408-427. [Crossref]

459. Anirban Kundu, Vipul Jain, Sameer Kumar, Charu Chandra. 2015. A journey from normative to behavioral operations in supply chain management: A review using Latent Semantic Analysis. *Expert Systems with Applications* **42**:2, 796-809. [Crossref]

460. Vaishali Gupta, Nisheeth Joshi, Iti Mathur. Design & development of rule based inflectional and derivational Urdu stemmer ‘Usal’ 7-12. [Crossref]

461. Abeed Sarker, Graciela Gonzalez. 2015. Portable automatic text classification for adverse drug reaction detection via multi-corpus training. *Journal of Biomedical Informatics* **53**, 196-207. [Crossref]

462. Tope Omitola, Sebastián A. Ríos, John G. Breslin. 2015. Social Semantic Web Mining. *Synthesis Lectures on the Semantic Web: Theory and Technology* **5**:1, 1-154. [Crossref]

463. Hong Wang, Anqi Liu, Jing Wang, Brian D. Ziebart, Clement T. Yu, Warren Shen. Context Retrieval for Web Tables 251-260. [Crossref]

464. R. Romero, E. L. Iglesias, L. Borrajo. 2015. A Linear-RBF Multikernel SVM to Classify Big Text Corpora. *BioMed Research International* **2015**, 1-14. [Crossref]

465. Laura Moreno, Gabriele Bavota, Sonia Haiduc, Massimiliano Di Penta, Rocco Oliveto, Barbara Russo, Andrian Marcus. Query-based configuration of text retrieval solutions for software engineering tasks 567-578. [Crossref]

466. Tien-Duy B. Le, Richard J. Oentaryo, David Lo. Information retrieval and spectrum based bug localization: better together 579-590. [Crossref]

467. Thomas Riedel, Stefan Spring, Anne Fiebig, Carmen Scheuner, Jörn Petersen, Markus Göker, Hans-Peter Klenk. 2015. Genome sequence of the Roseovarius mucosus type strain (DSM 17069T), a bacteriochlorophyll a-containing representative of the marine Roseobacter group isolated from the dinoflagellate Alexandrium ostenfeldii. *Standards in Genomic Sciences* **10**:1, 17. [Crossref]

468. Marcos Antonio Mouriño García, Roberto Pérez Rodríguez, Luis E. Anido Rifón. 2015. Biomedical literature classification using encyclopedic knowledge: a Wikipedia-based bag-of-concepts approach. *PeerJ* **3**, e1279. [Crossref]

469. Debasis Ganguly, Gareth J.F. Jones. Partially Labeled Supervised Topic Models for RetrievingSimilar Questions in CQA Forums 161-170. [Crossref]

470. Massimiliano Ruocco, Heri Ramampiaro. 2015. Geo-temporal distribution of tag terms for event-related image retrieval. *Information Processing & Management* **51**:1, 92-110. [Crossref]

471. Robert Leaman, Chih-Hsuan Wei, Zhiyong Lu. 2015. tmChem: a high performance approach for chemical named entity recognition and normalization. *Journal of Cheminformatics* **7**:Suppl 1, S3. [Crossref]

472. Debasis Ganguly, Santanu Pal, Gareth J. F. Jones. DCU@FIRE-2014 80-85. [Crossref]

473. Qiuchi Li, Jingfei Li, Peng Zhang, Dawei Song. Modeling Multi-query Retrieval Tasks Using Density Matrix Transformation 871-874. [Crossref]

474. Xiaonan Ji, Po-Yin Yen. 2015. Using MEDLINE Elemental Similarity to Assist in the Article Screening Process for Systematic Reviews. *JMIR Medical Informatics* **3**:3, e28. [Crossref]

475. Shuo Xu, Xin An, Lijun Zhu, Yunliang Zhang, Haodong Zhang. 2015. A CRF-based system for recognizing chemical entity mentions (CEMs) in biomedical literature. *Journal of Cheminformatics* **7**:Suppl 1, S11. [Crossref]

476. Ananthi Sheshasaayee, V. R. Angela Deepa. An unsupervised approach for morphological segmentation of highly agglutinative Tamil language 739-743. [Crossref]

477. Edwin D. Simpson, Matteo Venanzi, Steven Reece, Pushmeet Kohli, John Guiver, Stephen J. Roberts, Nicholas R. Jennings. Language Understanding in the Wild 992-1002. [Crossref]

478. Bhanu C. Vattikonda, Santhosh Kodipaka, Hongyan Zhou, Vacha Dave, Saikat Guha, Alex C. Snoeren. Interpreting Advertiser Intent in Sponsored Search 2177-2185. [Crossref]

479. Filippo Galgani, Paul Compton, Achim Hoffmann. 2015. Summarization based on bi-directional citation analysis. *Information Processing & Management* **51**:1, 1-24. [Crossref]

480. Edgar Gonzàlez, Jordi Turmo. 2015. Unsupervised ensemble minority clustering. *Machine Learning* **98**:1-2, 217-268. [Crossref]

481. Vijay Kotu, Bala Deshpande. Text Mining 275-303. [Crossref]

482. Jose Gustavo S. Paiva, William Robson Schwartz, Helio Pedrini, Rosane Minghim. 2015. An Approach to Supporting Incremental Visual Data Classification. *IEEE Transactions on Visualization and Computer Graphics* **21**:1, 4-17. [Crossref]

483. Samad Paydar, Mohsen Kahani. 2015. A semi-automated approach to adapt activity diagrams for new use cases. *Information and Software Technology* **57**, 543-570. [Crossref]

484. Doris Hoogeveen, Karin M. Verspoor, Timothy Baldwin. CQADupStack 1-8. [Crossref]

485. Zhaohui Wu, Shuiguang Deng, Jian Wu. Complex Service Computing 267-315. [Crossref]

486. Hsiang-Fu Yu, Cho-Jui Hsieh, Hyokun Yun, S.V.N. Vishwanathan, Inderjit S. Dhillon. A Scalable Asynchronous Distributed Algorithm for Topic Modeling 1340-1350. [Crossref]

487. JianGuo Wang, Joshua Zhexue Huang, Dingming Wu. 2015. Recommending High Utility Queries via Query-Reformulation Graph. *Mathematical Problems in Engineering* **2015**, 1-14. [Crossref]

488. Manisha Chawla, Mounika Mesa, Krishna P. Miyapuram. Graph Clustering for Large-Scale Text-Mining of Brain Imaging Studies 163-168. [Crossref]

489. Ugo Erra, Sabrina Senatore, Fernando Minnella, Giuseppe Caggianese. 2015. Approximate TF–IDF based on topic extraction from massive message stream using the GPU. *Information Sciences* **292**, 143-161. [Crossref]

490. Alon Ben-Ari, Kenric Hammond. Text Mining the EMR for Modeling and Predicting Suicidal Behavior among US Veterans of the 1991 Persian Gulf War 3168-3175. [Crossref]

491. Dominik Siemon, Linda Eckardt, Susanne Robra-Bissantz. Tracking Down the Negative Group Creativity Effects with the Help of an Artificial Intelligence-Like Support System 236-243. [Crossref]

492. Federica Bisio, Claudia Meda, Rodolfo Zunino, Roberto Surlinelli, Eugenio Scillia, Augusto Ottaviano. Real-time monitoring of Twitter traffic by using semantic networks 966-969. [Crossref]

493. Markus Christen, Thomas Niederberger, Thomas Ott, Suleiman Aryobsei, Reto Hofstetter. 2015. Micro-text classification between small and big data. *Nonlinear Theory and Its Applications, IEICE* **6**:4, 556-569. [Crossref]

494. Rishiraj Saha Roy, Rahul Katare, Niloy Ganguly, Srivatsan Laxman, Monojit Choudhury. 2015. Discovering and understanding word level user intent in Web search queries. *Web Semantics: Science, Services and Agents on the World Wide Web* **30**, 22-38. [Crossref]

495. Deovrat Kakde, Arin Chaudhuri. Leveraging unstructured data to detect emerging reliability issues 1-5. [Crossref]

496. Johann Hauswald, Lingjia Tang, Jason Mars, Michael A. Laurenzano, Yunqi Zhang, Cheng Li, Austin Rovinski, Arjun Khurana, Ronald G. Dreslinski, Trevor Mudge, Vinicius Petrucci. Sirius 223-238. [Crossref]

497. Niladri Chatterjee, Pramod Kumar Sahoo. 2015. Random Indexing and Modified Random Indexing based approach for extractive text summarization. *Computer Speech & Language* **29**:1, 32-44. [Crossref]

498. Nicolas Prollochs, Stefan Feuerriegel, Dirk Neumann. Enhancing Sentiment Analysis of Financial News by Detecting Negation Scopes 959-968. [Crossref]

499. Sven Amann, Stefanie Beyer, Katja Kevic, Harald Gall. Software Mining Studies: Goals, Approaches, Artifacts, and Replicability 121-158. [Crossref]

500. Wanying Chiu, Kun Lu. 2015. Random walk on co-word network: Ranking terms using structural features. *Proceedings of the Association for Information Science and Technology* **52**:1, 1-11. [Crossref]

501. Majid Rastegar-Mojarad, Zhan Ye, Daniel Wall, Narayana Murali, Simon Lin. 2015. Collecting and Analyzing Patient Experiences of Health Care From Social Media. *JMIR Research Protocols* **4**:3, e78. [Crossref]

502. Dmitry Bondarchuk, Galina Timofeeva. Vector space model based on semantic relatedness 020005. [Crossref]

503. Zhe Zhao, Paul Resnick, Qiaozhu Mei. Enquiring Minds 1395-1405. [Crossref]

504. Tomáš Brychcín, Miloslav Konopík. 2015. HPS: High precision stemmer. *Information Processing & Management* **51**:1, 68-91. [Crossref]

505. Rami Ayadi, Mohsen Maraoui, Mounir Zrigui. LDA and LSI as a Dimensionality Reduction Method in Arabic Document Classification 491-502. [Crossref]

506. Steffen Eger. Designing and Comparing G2P-Type Lemmatizers for a Morphology-Rich Language 27-40. [Crossref]

507. Wei Li, Gareth J. F. Jones. Enhancing Medical Information Retrieval by Exploiting a Content-Based Recommender Method 142-153. [Crossref]

508. Brian Dickinson, Michael Ganger, Wei Hu. 2015. Dimensionality Reduction of Distributed Vector Word Representations and Emoticon Stemming for Sentiment Analysis. *Journal of Data Analysis and Information Processing* **03**:04, 153-162. [Crossref]

509. John Wheal, Yanyan Yang. 2015. CSRecommender: A Cloud Service Searching and Recommendation System. *Journal of Computer and Communications* **03**:06, 65-73. [Crossref]

510. Fatemeh Ghiyafeh Davoodi, Yllias Chali. Semi-extractive Multi-document Summarization via Submodular Functions 96-110. [Crossref]

511. Sanghoon Lee, Seung-Jin Moon. 2014. Automatic Document Summary Technique Using Fuzzy Theory. *KIPS Transactions on Software and Data Engineering* **3**:12, 531-536. [Crossref]

512. Tarique Anwar, Muhammad Abulaish. 2014. A social graph based text mining framework for chat log investigation. *Digital Investigation* **11**:4, 349-362. [Crossref]

513. Roman Lukyanenko, Jeffrey Parsons, Yolanda F. Wiersma. 2014. The IQ of the Crowd: Understanding and Improving Information Quality in Structured User-Generated Content. *Information Systems Research* **25**:4, 669-689. [Crossref]

514. Tanmay Basu, C. A. Murthy. 2014. Towards enriching the quality of k-nearest neighbor rule for document classification. *International Journal of Machine Learning and Cybernetics* **5**:6, 897-905. [Crossref]

515. SriJyothsna Yeleswarapu, Aditya Rao, Thomas Joseph, Vangala Govindakrishnan Saipradeep, Rajgopal Srinivasan. 2014. A pipeline to extract drug-adverse event pairs from multiple data sources. *BMC Medical Informatics and Decision Making* **14**:1. . [Crossref]

516. Ehab ElGindy, Alia Abdelmoty. 2014. Capturing Place Semantics on the GeoSocial Web. *Journal on Data Semantics* **3**:4, 207-223. [Crossref]

517. Taemin Cho, Hana Cho, Jaedong Lee, Jee-Hyong Lee. Latent keyphrase generation by combining contextually similar primitive words 600-604. [Crossref]

518. Mohsen JafariAsbagh, Emilio Ferrara, Onur Varol, Filippo Menczer, Alessandro Flammini. 2014. Clustering memes in social media streams. *Social Network Analysis and Mining* **4**:1. . [Crossref]

519. Nishu Garg, P Nikhitha, B. K. Tripathy. Image retrieval using latent feature learning by deep architecture 1-4. [Crossref]

520. Muhammad Imran Rafique, Mehdi Hassan. Utilizing distinct terms for proximity and phrases in the document for better information retrieval 100-105. [Crossref]

521. Giovanni Acampora, Georgina Cosma. A hybrid computational intelligence approach for efficiently evaluating customer sentiments in E-commerce reviews 73-80. [Crossref]

522. Petra Heck, Andy Zaidman. 2014. Horizontal traceability for just-in-time requirements: the case for open source feature requests. *Journal of Software: Evolution and Process* **26**:12, 1280-1296. [Crossref]

523. Meera Sharma, Punam Bedi, V. B. Singh. 2014. An empirical evaluation of cross project priority prediction. *International Journal of System Assurance Engineering and Management* **5**:4, 651-663. [Crossref]

524. A. A. G. A. K Perera, R. P. E. T Jayarathne, B.Y Thilantha, S. I. G Kalupahana, P. S. Haddela, A. Kirupananda, E. A. T. D. Edirisinghe. Ads-In Site: Location based advertising framework with social network analyzer 116-123. [Crossref]

525. Fei Xie, Xindong Wu, Xingquan Zhu. Document-Specific Keyphrase Extraction Using Sequential Patterns with Wildcards 1055-1060. [Crossref]

526. Nicole L. Beebe, Lishu Liu. 2014. Clustering digital forensic string search output. *Digital Investigation* **11**:4, 314-322. [Crossref]

527. Tingting Liang, Liang Chen, Haochao Ying, Jian Wu. Co-Clustering WSDL Documents to Bootstrap Service Discovery 215-222. [Crossref]

528. Nikoletta K. Bassiou, Constantine L. Kotropoulos. 2014. Online PLSA: Batch Updating Techniques Including Out-of-Vocabulary Words. *IEEE Transactions on Neural Networks and Learning Systems* **25**:11, 1953-1966. [Crossref]

529. Shan Jiang, Hsinchun Chen, Jay F. Nunamaker, David Zimbra. 2014. Analyzing firm-specific social media and market: A stakeholder-based event analysis framework. *Decision Support Systems* **67**, 30-39. [Crossref]

530. Jehoshua Eliashberg, Sam K. Hui, Z. John Zhang. 2014. Assessing Box Office Performance Using Movie Scripts: A Kernel-Based Approach. *IEEE Transactions on Knowledge and Data Engineering* **26**:11, 2639-2648. [Crossref]

531. Timm O. Sprenger, Andranik Tumasjan, Philipp G. Sandner, Isabell M. Welpe. 2014. Tweets and Trades: the Information Content of Stock Microblogs. *European Financial Management* **20**:5, 926-957. [Crossref]

532. Soumaya Medini, Venera Arnaoudova, Massimiliano Di Penta, Giuliano Antoniol, Yann-Gaël Guéhéneuc, Paolo Tonella. 2014. SCAN: an approach to label and relate execution trace segments. *Journal of Software: Evolution and Process* **26**:11, 962-995. [Crossref]

533. Tien-Duy B. Le, Ferdian Thung, David Lo. Predicting Effectiveness of IR-Based Bug Localization Techniques 335-345. [Crossref]

534. Gianni Costa, Riccardo Ortale. XML Document Co-clustering via Non-negative Matrix Tri-factorization 607-614. [Crossref]

535. Taras Lehinevych, Nikolaos Kokkinis-Ntrenis, Giorgos Siantikos, A. Seza Dogruoz, Theodoros Giannakopoulos, Stasinos Konstantopoulos. Discovering Similarities for Content-Based Recommendation and Browsing in Multimedia Collections 237-243. [Crossref]

536. Sanghoon Lee, Seung-Jin Moon. 2014. HBase-based Automatic Summary System using Twitter Trending Topics. *Journal of Internet Computing and Services* **15**:5, 63-72. [Crossref]

537. Diana de la Iglesia, Miguel García-Remesal, Alberto Anguita, Miguel Muñoz-Mármol, Casimir Kulikowski, Víctor Maojo. 2014. A Machine Learning Approach to Identify Clinical Trials Involving Nanodrugs and Nanodevices from ClinicalTrials.gov. *PLoS ONE* **9**:10, e110331. [Crossref]

538. Navanath Saharia, Utpal Sharma, Jugal Kalita. 2014. Stemming resource-poor Indian languages. *ACM Transactions on Asian Language Information Processing* **13**:3, 1-26. [Crossref]

539. Asieh Ahani, Karl Wiegand, Umut Orhan, Murat Akcakaya, Mohammad Moghadamfalahi, Hooman Nezamfar, Rupal Patel, Deniz Erdogmus. 2014. RSVP IconMessenger: icon-based brain-interfaced alternative and augmentative communication. *Brain-Computer Interfaces* **1**:3-4, 192-203. [Crossref]

540. Yindalon Aphinyanaphongs, Lawrence D. Fu, Zhiguo Li, Eric R. Peskin, Efstratios Efstathiadis, Constantin F. Aliferis, Alexander Statnikov. 2014. A comprehensive empirical comparison of modern supervised classification and feature selection methods for text categorization. *Journal of the Association for Information Science and Technology* **65**:10, 1964-1987. [Crossref]

541. Andrea De Lucia, Massimiliano Di Penta, Rocco Oliveto, Annibale Panichella, Sebastiano Panichella. 2014. Labeling source code with information retrieval methods: an empirical study. *Empirical Software Engineering* **19**:5, 1383-1420. [Crossref]

542. Dima Badawi, Hakan Altınçay. 2014. A novel framework for termset selection and weighting in binary text classification. *Engineering Applications of Artificial Intelligence* **35**, 38-53. [Crossref]

543. Shailesh S. Kulkarni, Uday M. Apte, Nicholas E. Evangelopoulos. 2014. The Use of Latent Semantic Analysis in Operations Management Research. *Decision Sciences* **45**:5, 971-994. [Crossref]

544. Isabel Segura-Bedmar, Paloma Martínez, María Herrero-Zazo. 2014. Lessons learnt from the DDIExtraction-2013 Shared Task. *Journal of Biomedical Informatics* **51**, 152-164. [Crossref]

545. Dumidu Wijayasekara, Milos Manic, Miles McQueen. Vulnerability identification and classification via text mining bug databases 3612-3618. [Crossref]

546. . Bibliography 309-342. [Crossref]

547. Maria Eskevich, Gareth J.F. Jones. 2014. Exploring speech retrieval from meetings using the AMI corpus. *Computer Speech & Language* **28**:5, 1021-1044. [Crossref]

548. Deng-Yiv Chiu, Ya-Chen Pan. 2014. Topic knowledge map and knowledge structure constructions with genetic algorithm, information retrieval, and multi-dimension scaling method. *Knowledge-Based Systems* **67**, 412-428. [Crossref]

549. Wladmir C. Brandão, Rodrygo L. T. Santos, Nivio Ziviani, Edleno S. de Moura, Altigran S. da Silva. 2014. Learning to expand queries using entities. *Journal of the Association for Information Science and Technology* **65**:9, 1870-1883. [Crossref]

550. Shaowei Wang, David Lo, Julia Lawall. Compositional Vector Space Models for Improved Bug Localization 171-180. [Crossref]

551. L. J. Dijkstra, A. V. Yakushev, P. A. C. Duijn, A. V. Boukhanovsky, P. M. A. Sloot. 2014. Inference of the Russian drug community from one of the largest social networks in the Russian Federation. *Quality & Quantity* **48**:5, 2739-2755. [Crossref]

552. Timm O. Sprenger, Philipp G. Sandner, Andranik Tumasjan, Isabell M. Welpe. 2014. News or Noise? Using Twitter to Identify and Understand Company-specific News Flow. *Journal of Business Finance & Accounting* **41**:7-8, 791-830. [Crossref]

553. Alan De Renzis, Martin Garriga, Andres Flores, Alejandra Cechich, Alejandro Zunino. Semantic-structural assessment scheme for integrability in service-oriented applications 1-11. [Crossref]

554. Hammad Haleem, Pankaj Kumar Sharma, M. M. Sufyan Beg. Novel frequent sequential patterns based probabilistic model for effective classification of web documents 361-371. [Crossref]

555. M. Mirhosseini, M. Mashinchi, H. Nezamabadi-pour. 2014. Improving n-Similarity Problem by Genetic Algorithm and Its Application in Text Document Resemblance. *Fuzzy Information and Engineering* **6**:3, 263-278. [Crossref]

556. Bincy Thomas, Dhanya K. A, Vinod P. Exploration of robust features for multiclass emotion classification 1704-1709. [Crossref]

557. Le Wang, Gang Hua, Jianru Xue, Zhanning Gao, Nanning Zheng. 2014. Joint Segmentation and Recognition of Categorized Objects From Noisy Web Image Collection. *IEEE Transactions on Image Processing* **23**:9, 4070-4086. [Crossref]

558. Ben J Marafino, Jason M Davies, Naomi S Bardach, Mitzi L Dean, R Adams Dudley. 2014. N-gram support vector machines for scalable procedure and diagnosis classification, with applications to clinical free text data from the intensive care unit. *Journal of the American Medical Informatics Association* **21**:5, 871-875. [Crossref]

559. Vincenzo Gervasi, Didar Zowghi. Supporting traceability through affinity mining 143-152. [Crossref]

560. Atsushi Keyaki, Jun Miyazaki, Kenji Hatano. An Expansion Method of XML Element Retrieval Techniques into Web Documents 853-858. [Crossref]

561. Thomas Hansmann, Peter Niemeyer. Big Data - Characterizing an Emerging Research Field Using Topic Models 43-51. [Crossref]

562. Pavlina Fragkou. 2014. Text Segmentation for Language Identification in Greek Forums. *Procedia - Social and Behavioral Sciences* **147**, 160-166. [Crossref]

563. Nikitas N. Karanikolas. 2014. A Methodology for Building Simple but Robust Stemmers without Language Knowledge: Stemmer Configuration. *Procedia - Social and Behavioral Sciences* **147**, 370-375. [Crossref]

564. Michael Symonds, Peter Bruza, Guido Zuccon, Bevan Koopman, Laurianne Sitbon, Ian Turner. 2014. Automatic query expansion: A structural linguistic perspective. *Journal of the Association for Information Science and Technology* **65**:8, 1577-1596. [Crossref]

565. Phoebe Liu, Dylan F. Glas, Takayuki Kanda, Hiroshi Ishiguro, Norihiro Hagita. How to train your robot - teaching service robots to reproduce human social behavior 961-968. [Crossref]

566. Boliang Sun, Min Tang, Guohui Li. Sparse Online Co-regularization Using Conjugate Functions 3666-3671. [Crossref]

567. Mubarak Albathan, Yuefeng Li, Yue Xu. Using Extended Random Set to Find Specific Patterns 30-37. [Crossref]

568. Rupali Gangarde, V.L. Kolhe. Effective pattern discovery by cleaning patterns with pattern co-occurrence matrix and PDCS deploying approach 119-124. [Crossref]

569. Md Abul Bashar, Yuefeng Li, Yan Shen, Mubarak Albathan. Interpreting Discovered Patterns in Terms of Ontology Concepts 432-437. [Crossref]

570. Maria Lin, David W. Cheung. Automatic Tagging Web Services Using Machine Learning Techniques 258-265. [Crossref]

571. . Bibliography 277-287. [Crossref]

572. Giorgos Margaritis, Stergios V. Anastasiadis. 2014. Incremental Text Indexing for Fast Disk-Based Search. *ACM Transactions on the Web* **8**:3, 1-31. [Crossref]

573. Hsiao Wen-Feng, Chang Te-Min, Thomas Erwin. 2014. Extracting bibliographical data for PDF documents with HMM and external resources. *Program* **48**:3, 293-313. [Abstract] [Full Text] [PDF]

574. Pavel Pecina, Ondřej Dušek, Lorraine Goeuriot, Jan Hajič, Jaroslava Hlaváčová, Gareth J.F. Jones, Liadh Kelly, Johannes Leveling, David Mareček, Michal Novák, Martin Popel, Rudolf Rosa, Aleš Tamchyna, Zdeňka Urešová. 2014. Adaptation of machine translation for multilingual information retrieval in the medical domain. *Artificial Intelligence in Medicine* **61**:3, 165-185. [Crossref]

575. Wei Song, Jiu Zhen Liang, Soon Cheol Park. 2014. Fuzzy control GA with a novel hybrid semantic similarity strategy for text clustering. *Information Sciences* **273**, 156-170. [Crossref]

576. Hakan Altınçay, Zafer Erenel. 2014. Ternary encoding based feature extraction for binary text classification. *Applied Intelligence* **41**:1, 310-326. [Crossref]

577. Triss Ashton, Nicholas Evangelopoulos, Victor Prybutok. 2014. Extending monitoring methods to textual data: a research agenda. *Quality & Quantity* **48**:4, 2277-2294. [Crossref]

578. Dnyanesh G. Rajpathak, Satnam Singh. 2014. An Ontology-Based Text Mining Method to Develop D-Matrix From Unstructured Text. *IEEE Transactions on Systems, Man, and Cybernetics: Systems* **44**:7, 966-977. [Crossref]

579. Filippo Galgani, Paul Compton, Achim Hoffmann. 2014. HAUSS: Incrementally building a summarizer combining multiple techniques. *International Journal of Human-Computer Studies* **72**:7, 584-605. [Crossref]

580. Sangwoo Kang, Harksoo Kim, Hyun-Kyu Kang, Jungyun Seo. 2014. Lightweight Morphological Analysis Model for Smart Home Applications Based on Natural Language Interfaces. *International Journal of Distributed Sensor Networks* **10**:6, 570634. [Crossref]

581. Georgia M. Kapitsaki. Annotating Web Service Sections with Combined Classification 622-629. [Crossref]

582. R. Santodomingo, S. Rohjans, M. Uslar, J.A. Rodríguez-Mondéjar, M.A. Sanz-Bobi. 2014. Ontology matching system for future energy smart grids. *Engineering Applications of Artificial Intelligence* **32**, 242-257. [Crossref]

583. Thiago Nunes, Daniel Schwabe. Building Distant Supervised Relation Extractors 44-51. [Crossref]

584. Yu Liu, Xiaobing Sun, Xiangyue Liu, Yun Li. Supporting program comprehension with program summarization 363-368. [Crossref]

585. Semih Ergin, Sahin Isik. The investigation on the effect of feature vector dimension for spam email detection with a new framework 1-4. [Crossref]

586. Dongqing Zhu, Stephen Wu, Ben Carterette, Hongfang Liu. 2014. Using large clinical corpora for query expansion in text-based cohort identification. *Journal of Biomedical Informatics* **49**, 275-281. [Crossref]

587. Cheng Hua Li, Laurence T. Yang, Man Lin. 2014. Parallel Training of An Improved Neural Network for Text Categorization. *International Journal of Parallel Programming* **42**:3, 505-523. [Crossref]

588. ###. 2014. A Manchu speller: With a practical introduction to the natural language processing of minority languages. *ALTAI HAKPO* **null**:24, 39-67. [Crossref]

589. Marie Katsurai, Takahiro Ogawa, Miki Haseyama. 2014. A Cross-Modal Approach for Extracting Semantic Relationships Between Concepts Using Tagged Images. *IEEE Transactions on Multimedia* **16**:4, 1059-1074. [Crossref]

590. G. Tambouratzis. 2014. Optimizing word segmentation tasks using ant colony metaheuristics. *Literary and Linguistic Computing* **29**:2, 234-254. [Crossref]

591. Sven S. Groth, Michael Siering, Peter Gomber. 2014. How to enable automated trading engines to cope with news-related liquidity shocks? Extracting signals from unstructured data. *Decision Support Systems* **62**, 32-42. [Crossref]

592. Ryan Light. 2014. From Words to Networks and Back. *Social Currents* **1**:2, 111-129. [Crossref]

593. Apostolos Gerasoulis. Web Search Engines:: Practice and Experience 1-24. [Crossref]

594. Thomas L. Ngo-Ye, Atish P. Sinha. 2014. The influence of reviewer engagement characteristics on online review helpfulness: A text regression model. *Decision Support Systems* **61**, 47-58. [Crossref]

595. Cheng Zheng, Weiming Shen, Hamada H. Ghenniwa. A study of intents resolving for service discovery 649-654. [Crossref]

596. Mayy M. Al-Tahrawi. 2014. The Significance Of Low Frequent Terms in Text Classification. *International Journal of Intelligent Systems* **29**:5, 389-406. [Crossref]

597. Venera Arnaoudova, Laleh M. Eshkevari, Massimiliano Di Penta, Rocco Oliveto, Giuliano Antoniol, Yann-Gael Gueheneuc. 2014. REPENT: Analyzing the Nature of Identifier Renamings. *IEEE Transactions on Software Engineering* **40**:5, 502-532. [Crossref]

598. Min Xu, Jinqiao Wang, Xiangjian He, Jesse S. Jin, Suhuai Luo, Hanqing Lu. 2014. A three-level framework for affective content analysis and its case studies. *Multimedia Tools and Applications* **70**:2, 757-779. [Crossref]

599. Tingting Mu, Makoto Miwa, Junichi Tsujii, Sophia Ananiadou. 2014. DISCOVERING ROBUST EMBEDDINGS IN (DIS)SIMILARITY SPACE FOR HIGH-DIMENSIONAL LINGUISTIC FEATURES. *Computational Intelligence* **30**:2, 285-315. [Crossref]

600. Xiaogang Han, Wei Wei, Chunyan Miao, Jian-Ping Mei, Hengjie Song. 2014. Context-Aware Personal Information Retrieval From Multiple Social Networks. *IEEE Computational Intelligence Magazine* **9**:2, 18-28. [Crossref]

601. Denis L. Nkweteyim. Data structures for information retrieval 1-8. [Crossref]

602. Arezki Hammache, Mohand Boughanem, Rachid Ahmed-Ouamer. 2014. Combining compound and single terms under language model framework. *Knowledge and Information Systems* **39**:2, 329-349. [Crossref]

603. Konstantinos Pliakos, Constantine Kotropoulos. Personalized music tagging using ranking on hypergraphs 628-631. [Crossref]

604. Yuan Xu, Chunhong Zhang, Yang Ji. 2014. An Upper-Ontology-Based Approach for Automatic Construction of IOT Ontology. *International Journal of Distributed Sensor Networks* **10**:4, 594782. [Crossref]

605. Qiwei He, Cees A.W. Glas, Michal Kosinski, David J. Stillwell, Bernard P. Veldkamp. 2014. Predicting self-monitoring skills using textual posts on Facebook. *Computers in Human Behavior* **33**, 69-78. [Crossref]

606. Liang Chen, Jian Wu, Zibin Zheng, Michael R. Lyu, Zhaohui Wu. 2014. Modeling and exploiting tag relevance for Web service mining. *Knowledge and Information Systems* **39**:1, 153-173. [Crossref]

607. P Pabitha, M. Mohana, S. Suganthi, B. Sivanandhini. Automatic Question Generation system 1-5. [Crossref]

608. V. D'Orazio, S. T. Landis, G. Palmer, P. Schrodt. 2014. Separating the Wheat from the Chaff: Applications of Automated Document Classification Using Support Vector Machines. *Political Analysis* **22**:2, 224-242. [Crossref]

609. Arash Joorabchi, Abdulhussain E. Mahdi. 2014. Towards linking libraries and Wikipedia: automatic subject indexing of library records with Wikipedia concepts. *Journal of Information Science* **40**:2, 211-221. [Crossref]

610. Kun Lu, Margaret E.I. Kipp. 2014. Understanding the retrieval effectiveness of collaborative tags and author keywords in different retrieval environments: An experimental study on medical collections. *Journal of the Association for Information Science and Technology* **65**:3, 483-500. [Crossref]

611. George E. Noel, Gilbert L. Peterson. 2014. Applicability of Latent Dirichlet Allocation to multi-disk search. *Digital Investigation* **11**:1, 43-56. [Crossref]

612. Hsin-Chang Yang, Chung-Hong Lee. 2014. Detecting tag spams for social bookmarking Websites using a text mining approach. *International Journal of Information Technology & Decision Making* **13**:02, 387-406. [Crossref]

613. Martin Wiesner, Daniel Pfeifer. 2014. Health Recommender Systems: Concepts, Requirements, Technical Basics and Challenges. *International Journal of Environmental Research and Public Health* **11**:3, 2580-2607. [Crossref]

614. Juan C. Rendon-Miranda, Julia Y. Arana-Llanes, Juan G. Gonzalez-Serna, Nimrod Gonzalez-Franco. Automatic Classification of Scientific Papers in PDF for Populating Ontologies 319-320. [Crossref]

615. Lea-Rachel Kosnik. 2014. Determinants of contract completeness: An environmental regulatory application. *International Review of Law and Economics* **37**, 198-208. [Crossref]

616. J.J. García Adeva, J.M. Pikatza Atxa, M. Ubeda Carrillo, E. Ansuategi Zengotitabengoa. 2014. Automatic text classification to support systematic reviews in medicine. *Expert Systems with Applications* **41**:4, 1498-1508. [Crossref]

617. Zhen Wang, Yuan-Hai Shao, Tie-Ru Wu. 2014. Proximal parametric-margin support vector classifier and its applications. *Neural Computing and Applications* **24**:3-4, 755-764. [Crossref]

618. Sabrina de Azevedo Silveira, Raquel Cardoso de Melo-Minardi, Carlos Henrique da Silveira, Marcelo Matos Santoro, Wagner Meira Jr. 2014. ENZYMAP: Exploiting Protein Annotation for Modeling and Predicting EC Number Changes in UniProt/Swiss-Prot. *PLoS ONE* **9**:2, e89162. [Crossref]

619. Soon Chong Johnson Lim, Shilong Wang, Ying Liu. 2014. Discovering contextual tags from product review using semantic relatedness. *Journal of Industrial and Production Engineering* **31**:2, 108-118. [Crossref]

620. Hyun-Joong Kim, Sungzoon Cho, Pilsung Kang. 2014. KR-WordRank : An Unsupervised Korean Word Extraction Method Based on WordRank. *Journal of Korean Institute of Industrial Engineers* **40**:1, 18-33. [Crossref]

621. Gabriele Bavota, Andrea De Lucia, Rocco Oliveto, Genoveffa Tortora. 2014. Enhancing software artefact traceability recovery processes with link count information. *Information and Software Technology* **56**:2, 163-182. [Crossref]

622. Shreya Khare, Akshay Bhandari, Hema A. Murthy. URL classification using non negative matrix factorization 1-6. [Crossref]

623. Annamaria Conti, Olgert Denas, Fabiana Visentin. 2014. Knowledge Specialization in Ph.D. Student Groups. *IEEE Transactions on Engineering Management* **61**:1, 52-67. [Crossref]

624. Fakhri Hasanzadeh, Mehrdad Jalali, Majid Vafaei Jahan. Detecting communities in social networks by techniques of clustering and analysis of communications 1-5. [Crossref]

625. Llanos Tobarra, Antonio Robles-Gómez, Salvador Ros, Roberto Hernández, Agustín C. Caminero. 2014. Analyzing the students' behavior and relevant topics in virtual learning communities. *Computers in Human Behavior* **31**, 659-669. [Crossref]

626. Tao Wang, Huaimin Wang, Gang Yin, Charles X. Ling, Xiao Li, Peng Zou. 2014. Tag recommendation for open source software. *Frontiers of Computer Science* **8**:1, 69-82. [Crossref]

627. Abhijit Debbarma, BS Purkayastha, Paritosh Bhattacharya. Stemmer for resource scarce language using string similarity measure 96-98. [Crossref]

628. Raymond Y.K. Lau, Yunqing Xia, Yunming Ye. 2014. A Probabilistic Generative Model for Mining Cybercriminal Networks from Online Social Media. *IEEE Computational Intelligence Magazine* **9**:1, 31-43. [Crossref]

629. Yuan Tian, David Lo, Julia Lawall. Automated construction of a software-specific word similarity database 44-53. [Crossref]

630. Seunghee Kim, Jinwook Choi. 2014. An SVM-based high-quality article classifier for systematic reviews. *Journal of Biomedical Informatics* **47**, 153-159. [Crossref]

631. Jikitsha Sheth, Bankim Patel. Dhiya: A stemmer for morphological level analysis of Gujarati language 151-154. [Crossref]

632. Tanja Bekhuis, Eugene Tseytlin, Kevin J. Mitchell, Dina Demner-Fushman. 2014. Feature Engineering and a Proposed Decision-Support System for Systematic Reviewers of Medical Evidence. *PLoS ONE* **9**:1, e86277. [Crossref]

633. R. Alexander Bentley, Alberto Acerbi, Paul Ormerod, Vasileios Lampos. 2014. Books Average Previous Decade of Economic Misery. *PLoS ONE* **9**:1, e83147. [Crossref]

634. Timothy Clark, Mike Wright, Zilia Iskoujina, Philip Garnett. 2014. JMS at 50: Trends over Time. *Journal of Management Studies* **51**:1, 19-37. [Crossref]

635. . References 201-221. [Crossref]

636. Esra Saraç, Selma Ayşe Özel. 2014. An Ant Colony Optimization Based Feature Selection for Web Page Classification. *The Scientific World Journal* **2014**, 1-16. [Crossref]

637. Tomer Geva, Jacob Zahavi. 2014. Empirical evaluation of an automated intraday stock recommendation system incorporating both market data and textual news. *Decision Support Systems* **57**, 212-223. [Crossref]

638. Jun-Ming Chen, Meng-Chang Chen, Yeali S. Sun. 2014. A tag based learning approach to knowledge acquisition for constructing prior knowledge and enhancing student reading comprehension. *Computers & Education* **70**, 256-268. [Crossref]

639. Daswin De Silva, Frada Burstein. 2014. An Intelligent Content Discovery Technique for Health Portal Content Management. *JMIR Medical Informatics* **2**:1, e7. [Crossref]

640. A. R. Rivas, E. L. Iglesias, L. Borrajo. 2014. Study of Query Expansion Techniques and Their Application in the Biomedical Information Retrieval. *The Scientific World Journal* **2014**, 1-10. [Crossref]

641. Jan Žižka, Arnošt Svoboda, František Dařena. 2014. Selecting text entries using a few positive samples and similarity ranking. *Acta Universitatis Agriculturae et Silviculturae Mendelianae Brunensis* **59**:4, 399-408. [Crossref]

642. Y. Raimond, C. Lowis, R. Hodgson, D. Tinley. 2014. Automated Metadata Enrichment of Large Speech Radio Archives. *SMPTE Motion Imaging Journal* **123**:1, 35-41. [Crossref]

643. Tzu-Tsung Wong. 2014. Generalized Dirichlet priors for Naïve Bayesian classifiers with multinomial models in document classification. *Data Mining and Knowledge Discovery* **28**:1, 123-144. [Crossref]

644. Elad Yom-Tov, Diana Borsa, Ingemar J Cox, Rachel A McKendry. 2014. Detecting Disease Outbreaks in Mass Gatherings Using Internet Data. *Journal of Medical Internet Research* **16**:6, e154. [Crossref]

645. Yinqing Xu, Tianyi Lin, Wai Lam, Zirui Zhou, Hong Cheng, Anthony Man-Cho So. Latent Aspect Mining via Exploring Sparsity and Intrinsic Information 879-888. [Crossref]

646. Andreas Holzinger, Johannes Schantl, Miriam Schroettner, Christin Seifert, Karin Verspoor. Biomedical Text Mining: State-of-the-Art, Open Problems and Future Challenges 271-300. [Crossref]

647. Takahiro Komamizu, Toshiyuki Amagasa, Hiroyuki Kitagawa. Extracting Facets from Textual Contents for Faceted Search over XML Data 420-429. [Crossref]

648. Alper Kursat Uysal, Serkan Gunal. 2014. The impact of preprocessing on text classification. *Information Processing & Management* **50**:1, 104-112. [Crossref]

649. Jian Wu, Liang Chen, Zibin Zheng, Michael R. Lyu, Zhaohui Wu. 2014. Clustering Web services to facilitate service discovery. *Knowledge and Information Systems* **38**:1, 207-229. [Crossref]

650. Todd Bodnar, Victoria C. Barclay, Nilam Ram, Conrad S. Tucker, Marcel Salathé. On the ground validation of online diagnosis with Twitter and medical records 651-656. [Crossref]

651. Andrew Trotman, Antti Puurula, Blake Burgess. Improvements to BM25 and Language Models Examined 58-65. [Crossref]

652. Johannes Schanda, Mark Sanderson, Paul Clough. Examining New Event Detection 26-33. [Crossref]

653. Ángel Felices-Lago, Pedro Ureña Gómez-Moreno. FunGramKB term extractor: A tool for building terminological ontologies
from specialised corpora 251-270. [Crossref]

654. Mingjie Qian, Chengxiang Zhai. Unsupervised Feature Selection for Multi-View Clustering on Text-Image Web News Data 1963-1966. [Crossref]

655. Ywen Huang, Zhua Jiang, Xde Xiang, Cneng He, Jfeng Liu, Ying Huang. An Inner-Enterprise Wiki System (IWkS) Integrated with an Expert Finding Mechanism for Lesson-Learned Knowledge Accumulation in Product Design 198-208. [Crossref]

656. Stefano Ferilli, Floriana Esposito, Domenico Grieco. 2014. Automatic Learning of Linguistic Resources for Stopword Removal and Stemming from Text. *Procedia Computer Science* **38**, 116-123. [Crossref]

657. Jingbo Xia, Alex Chengyu Fang, Xing Zhang. 2014. A Novel Feature Selection Strategy for Enhanced Biomedical Event Extraction Using the Turku System. *BioMed Research International* **2014**, 1-12. [Crossref]

658. Ea-Ee Jan, Kuan-Yu Chen, Tsuyoshi Ide. A Probabilistic Concept Annotation for IT Service Desk Tickets 21-23. [Crossref]

659. Ywen HUANG, Zhua JIANG. 2014. A Semantic-Based Topic Knowledge Map System (STKMS) for Lesson-Learned Documents Reuse in Product Design. *IEICE Transactions on Information and Systems* **E97.D**:5, 1049-1057. [Crossref]

660. Onur Varol, Filippo Menczer. Connecting dream networks across cultures 1267-1272. [Crossref]

661. Wojciech Gryc, Karo Moilanen. Leveraging Textual Sentiment Analysis with Social Network Modelling: Sentiment Analysis of Political Blogs in the 2008 U.S. Presidential Election 47-70. [Crossref]

662. Michael Wiegand, Manfred Klenner, Dietrich Klakow. 2013. Bootstrapping polarity classifiers with rule-based classification. *Language Resources and Evaluation* **47**:4, 1049-1088. [Crossref]

663. Zhonghua Li, Bingjun Zhang, Yi Yu, Jialie Shen, Ye Wang. 2013. Query-Document-Dependent Fusion: A Case Study of Multimodal Music Retrieval. *IEEE Transactions on Multimedia* **15**:8, 1830-1842. [Crossref]

664. Gianni Costa, Riccardo Ortale. A Latent Semantic Approach to XML Clustering by Content and Structure Based on Non-negative Matrix Factorization 179-184. [Crossref]

665. A. Kundu, V. Jain. On development of supplier segmentation ontology using latent semantic analysis for supplier knowledge management in supply chain 1007-1011. [Crossref]

666. Mohammed N. Al-Kabi. Towards improving Khoja rule-based Arabic stemmer 1-6. [Crossref]

667. Anwar Ali Yahya, Addin Osman, Ahmed Abdu Alattab. Educational data mining: A case study of teacher's classroom questions 92-97. [Crossref]

668. Daniela Pohl, Abdelhamid Bouchachia, Hermann Hellwagner. Online Processing of Social Media Data for Emergency Management 408-413. [Crossref]

669. Lawrence D. Fu, Yindalon Aphinyanaphongs, Constantin F. Aliferis. 2013. Computer models for identifying instrumental citations in the biomedical literature. *Scientometrics* **97**:3, 871-882. [Crossref]

670. Jenomi De Silva, Prasanna S. Haddela. A term weighting method for identifying emotions from text content 381-386. [Crossref]

671. Shehan R. Gunarathne, Jenomi De Silva, E. M. C. P Ekanayake, I. Samaradiwakara, P. S. Haddela, P. A Fernando. Intellemo: A mobile instant messaging application with intelligent emotion identification 627-632. [Crossref]

672. Prajitha U, Sreejith C, P.C. Reghu Raj. LALITHA: A light weight Malayalam stemmer using suffix stripping method 244-248. [Crossref]

673. P Jaganathan, S Jaiganesh. An improved K-means algorithm combined with Particle Swarm Optimization approach for efficient web document clustering 772-776. [Crossref]

674. Dave Debarr, Hao Sun, Harry Wechsler. Adversarial Spam Detection Using the Randomized Hough Transform-Support Vector Machine 299-304. [Crossref]

675. Elena Hensinger, Ilias Flaounas, Nello Cristianini. 2013. Modelling and predicting news popularity. *Pattern Analysis and Applications* **16**:4, 623-635. [Crossref]

676. Sanghoon Lee, Sunny Shakya, Raj Sunderraman, Saeid Belkasim. Real Time Micro-blog Summarization Based on Hadoop/HBase 46-49. [Crossref]

677. Gianni Costa, Riccardo Ortale. Developments in Partitioning XML Documents by Content and Structure Based on Combining Multiple Clusterings 477-482. [Crossref]

678. FILIPPO SCIARRONE. 2013. AN EXTENSION OF THE Q DIVERSITY METRIC FOR INFORMATION PROCESSING IN MULTIPLE CLASSIFIER SYSTEMS: A FIELD EVALUATION. *International Journal of Wavelets, Multiresolution and Information Processing* **11**:06, 1350049. [Crossref]

679. Matt Taddy. 2013. Measuring Political Sentiment on Twitter: Factor Optimal Design for Multinomial Inverse Regression. *Technometrics* **55**:4, 415-425. [Crossref]

680. Anwar Ali Yahya, Addin Osman, Ahmad Taleb, Ahmed Abdu Alattab. 2013. Analyzing the Cognitive Level of Classroom Questions Using Machine Learning Techniques. *Procedia - Social and Behavioral Sciences* **97**, 587-595. [Crossref]

681. Alan S. Abrahams, Jian Jiao, Weiguo Fan, G. Alan Wang, Zhongju Zhang. 2013. What's buzzing in the blizzard of buzz? Automotive component isolation in social media postings. *Decision Support Systems* **55**:4, 871-882. [Crossref]

682. Mohammad-hadi Zahedi, Mohsen Kahani. 2013. SREC: Discourse-level semantic relation extraction from text. *Neural Computing and Applications* **23**:6, 1573-1582. [Crossref]

683. Jean-Charles Lamirel. Enhancing Classification Accuracy with the Help of Feature Maximization Metric 569-574. [Crossref]

684. M.G. Armentano, D. Godoy, A.A. Amandi. 2013. Followee recommendation based on text analysis of micro-blogging activity. *Information Systems* **38**:8, 1116-1127. [Crossref]

685. . Linguistic Features 217-229. [Crossref]

686. Jalel Akaichi, Zeineb Dhouioui, Maria Jose Lopez-Huertas Perez. Text mining facebook status updates for sentiment classification 640-645. [Crossref]

687. Alberto Acerbi, Vasileios Lampos, R. Alexander Bentley. Robustness of emotion extraction from 20th century English books 1-8. [Crossref]

688. Marlies Rybnicek, Rainer Poisel, Simon Tjoa. Facebook Watchdog: A Research Agenda for Detecting Online Grooming and Bullying Activities 2854-2859. [Crossref]

689. Stephen W. Thomas, Meiyappan Nagappan, Dorothea Blostein, Ahmed E. Hassan. 2013. The Impact of Classifier Configuration and Classifier Combination on Bug Localization. *IEEE Transactions on Software Engineering* **39**:10, 1427-1443. [Crossref]

690. PAVLINA FRAGKOU. 2013. INFORMATION EXTRACTION VERSUS TEXT SEGMENTATION FOR WEB CONTENT MINING. *International Journal of Software Engineering and Knowledge Engineering* **23**:08, 1109-1137. [Crossref]

691. Ursula Torres-Parejo, Jesús R. Campaña, Miguel Delgado, M. Amparo Vila. 2013. MTCIR: A multi-term tag cloud information retrieval system. *Expert Systems with Applications* **40**:14, 5448-5455. [Crossref]

692. C. Gorg, Zhicheng Liu, Jaeyeon Kihm, Jaegul Choo, Haesun Park, J. Stasko. 2013. Combining Computational Analyses and Interactive Visualization for Document Exploration and Sensemaking in Jigsaw. *IEEE Transactions on Visualization and Computer Graphics* **19**:10, 1646-1663. [Crossref]

693. Behnam Taraghi, Martin Grossegger, Martin Ebner, Andreas Holzinger. 2013. Web analytics of user path tracing and a novel algorithm for generating recommendations in Open Journal Systems. *Online Information Review* **37**:5, 672-691. [Abstract] [Full Text] [PDF]

694. Ryan Burns, André Skupin. 2013. Towards Qualitative Geovisual Analytics: A Case Study Involving Places, People, and Mediated Experience. *Cartographica: The International Journal for Geographic Information and Geovisualization* **48**:3, 157-176. [Crossref]

695. Vaishali Gupta, Nisheeth Joshi, Iti Mathur. Rule based stemmer in Urdu 129-132. [Crossref]

696. Pooja Mudgil, A.K. Sharma, Pooja Gupta. An Improved Indexing Mechanism to Index Web Documents 460-464. [Crossref]

697. Marco Helbich, Julian Hagenauer, Michael Leitner, Ricky Edwards. 2013. Exploration of unstructured narrative crime reports: an unsupervised neural network and point pattern analysis approach. *Cartography and Geographic Information Science* **40**:4, 326-336. [Crossref]

698. Radu Machedon, William Rand, Yogesh Joshi. Automatic Crowdsourcing-Based Classification of Marketing Messaging on Twitter 975-978. [Crossref]

699. Martin Weber, Maarten H. Lamers. Finding News in a Haystack - Event Based News Clustering with Social Media Based Ranking 321-326. [Crossref]

700. Jalel Akaichi. Social Networks' Facebook' Statutes Updates Mining for Sentiment Classification 886-891. [Crossref]

701. Matt Taddy. 2013. Multinomial Inverse Regression for Text Analysis. *Journal of the American Statistical Association* **108**:503, 755-770. [Crossref]

702. Y. Jahnavi, Y. Radhika. Hot topic extraction based on frequency, position, scattering and topical weight for time sliced news documents 1-6. [Crossref]

703. Esa Rantanen, Christopher Claeys, Daniel Roder, William Moroney. 2013. Archival Human Factors Jobs Database as a Tool for Tracking Trends in Skills and Knowledge Expectations in the Labor Market. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* **57**:1, 457-461. [Crossref]

704. Rasia Naseem, Sheena Anees, K. Muneer, K. Syed Farook. Near Duplicate Web Page Detection with Analytic Feature Weighting 324-327. [Crossref]

705. S. Sarasvady, P. Vijayakumar, Daisy Jocabs. Mining and computing phrase weight in texts 371-374. [Crossref]

706. Arkaitz Zubiaga, Victor Fresno, Raquel Martinez, Alberto Perez Garcia-Plaza. 2013. Harnessing Folksonomies to Produce a Social Classification of Resources. *IEEE Transactions on Knowledge and Data Engineering* **25**:8, 1801-1813. [Crossref]

707. Bipin Upadhyaya, Ran Tang, Ying Zou. 2013. An approach for mining service composition patterns from execution logs. *Journal of Software: Evolution and Process* **25**:8, 841-870. [Crossref]

708. Roland Schäfer, Felix Bildhauer. 2013. Web Corpus Construction. *Synthesis Lectures on Human Language Technologies* **6**:4, 1-145. [Crossref]

709. Myunggwon Hwang, Do-Heon Jeong, Jinhyung Kim, Sa-Kwang Song, Hanmin Jung, Juhyun Shin, Pankoo Kim. 2013. A term normalization method for efficient knowledge acquisition through text processing. *Multimedia Tools and Applications* **65**:1, 75-91. [Crossref]

710. Fekade Getahun, Richard Chbeir. 2013. RSS query algebra: Towards a better news management. *Information Sciences* **237**, 313-342. [Crossref]

711. Xiaocao Hu, Zhiyong Feng, Shizhan Chen. Analyzing Distribution of Implicit Semantic Information in Web Services 415-420. [Crossref]

712. J. Grimmer, B. M. Stewart. 2013. Text as Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts. *Political Analysis* **21**:3, 267-297. [Crossref]

713. Vahid Behbood, Jie Lu, Guangquan Zhang. Text categorization by fuzzy domain adaptation 1-7. [Crossref]

714. Giovanni Capobianco, Andrea De Lucia, Rocco Oliveto, Annibale Panichella, Sebastiano Panichella. 2013. Improving IR-based traceability recovery via noun-based indexing of software artifacts. *Journal of Software: Evolution and Process* **25**:7, 743-762. [Crossref]

715. Junyan Peng, Patrick P. K. Chan. Revised Naive Bayes classifier for combating the focus attack in spam filtering 610-614. [Crossref]

716. Aaron K. Massey, Jacob Eisenstein, Annie I. Anton, Peter P. Swire. Automated text mining for requirements analysis of policy documents 4-13. [Crossref]

717. Daniela Pohl, Abdelhamid Bouchachia, Hermann Hellwagner. 2013. Supporting Crisis Management via Detection of Sub-Events in Social Networks. *International Journal of Information Systems for Crisis Response and Management* **5**:3, 20-36. [Crossref]

718. Michael Hagenau, Michael Liebmann, Dirk Neumann. 2013. Automated news reading: Stock price prediction based on financial news using context-capturing features. *Decision Support Systems* **55**:3, 685-697. [Crossref]

719. Xiaocao Hu, Zhiyong Feng, Shizhan Chen. Augmenting Semantics of Web Services Based on Public Open Ontology 304-311. [Crossref]

720. Percy Liang, Michael I. Jordan, Dan Klein. 2013. Learning Dependency-Based Compositional Semantics. *Computational Linguistics* **39**:2, 389-446. [Crossref]

721. Yueh-Min Huang, Yong-Ming Huang, Chien-Hung Liu, Chin-Chung Tsai. 2013. Applying social tagging to manage cognitive load in a Web 2.0 self-learning environment. *Interactive Learning Environments* **21**:3, 273-289. [Crossref]

722. Samah Jamal Fodeh, Cynthia Brandt, Thai Binh Luong, Ali Haddad, Martin Schultz, Terrence Murphy, Michael Krauthammer. 2013. Complementary ensemble clustering of biomedical data. *Journal of Biomedical Informatics* **46**:3, 436-443. [Crossref]

723. Dumidu Wijayasekara, Milos Manic, Miles McQueen. Information gain based dimensionality selection for classifying text documents 440-445. [Crossref]

724. Hu Zongzhen, Zhu Weina, Liyue, Du Xiaojuan, Yan fan. A Fuzzy Approach to Clustering of Text Documents Based on MapReduce 666-669. [Crossref]

725. Arash Joorabchi, Abdulhussain E. Mahdi. 2013. Automatic keyphrase annotation of scientific documents using Wikipedia and genetic algorithms. *Journal of Information Science* **39**:3, 410-426. [Crossref]

726. Maryam Bahojb Imani, Mohammad Reza Keyvanpour, Reza Azmi. 2013. A NOVEL EMBEDDED FEATURE SELECTION METHOD: A COMPARATIVE STUDY IN THE APPLICATION OF TEXT CATEGORIZATION. *Applied Artificial Intelligence* **27**:5, 408-427. [Crossref]

727. Ignazio Infantino, Giovanni Pilato, Riccardo Rizzo, Filippo Vella. 2013. Humanoid Introspection: A Practical Approach. *International Journal of Advanced Robotic Systems* **10**:5, 246. [Crossref]

728. Constantin Houy, Khulan Sainbuyan, Peter Fettke, Peter Loos. Towards automated analysis of fads and trends in information systems research: Concept, implementation and exemplary application in the context of business process management research 1-11. [Crossref]

729. Diana Diaz, Gabriele Bavota, Andrian Marcus, Rocco Oliveto, Silvia Takahashi, Andrea De Lucia. Using code ownership to improve IR-based Traceability Link Recovery 123-132. [Crossref]

730. P. Butka, J. Pocsova. Hybrid approach for visualization of documents clusters using GHSOM and sammon projection 337-342. [Crossref]

731. Gabriele Bavota, Andrea De Lucia, Rocco Oliveto, Annibale Panichella, Fabio Ricci, Genoveffa Tortora. The role of artefact corpus in LSI-based traceability recovery 83-89. [Crossref]

732. Antonis Theodoridis, Constantine Kotropoulos, Yannis Panagakis. Music recommendation using hypergraphs and group sparsity 56-60. [Crossref]

733. Zafer Erenel, Hakan Altınçay. 2013. Improving the precision-recall trade-off in undersampling-based binary text categorization using unanimity rule. *Neural Computing and Applications* **22**:S1, 83-100. [Crossref]

734. Xiaodan Zhu, Colin Cherry, Gerald Penn. 2013. A Graph-Partitioning Framework for Aligning Hierarchical Topic Structures to Presentations. *IEEE Transactions on Audio, Speech, and Language Processing* **21**:5, 1102-1112. [Crossref]

735. Bunyamin Sisman, Avinash C. Kak. Assisting code search with automatic Query Reformulation for bug localization 309-318. [Crossref]

736. Kenneth Cosh. Text mining wikipedia to discover alternative destinations 43-48. [Crossref]

737. Shady Shehata, Fakhri Karray, Mohamed S. Kamel. 2013. An efficient concept-based retrieval model for enhancing text retrieval quality. *Knowledge and Information Systems* **35**:2, 411-434. [Crossref]

738. Felix Weninger, Claudia Wagner, Martin Wollmer, Bjorn Schuller, Louis-Philippe Morency. Speaker trait characterization in web videos: Uniting speech, language, and facial features 3647-3651. [Crossref]

739. Zoe Y. Zhuang, Carla L. Wilkin, Andrzej Ceglowski. 2013. A framework for an intelligent decision support system: A case in pathology test ordering. *Decision Support Systems* **55**:2, 476-487. [Crossref]

740. John Slankas, Laurie Williams. Automated extraction of non-functional requirements in available documentation 9-16. [Crossref]

741. Nasir Ali, Yann-Gael Gueheneuc, Giuliano Antoniol. 2013. Trustrace: Mining Software Repositories to Improve the Accuracy of Requirement Traceability Links. *IEEE Transactions on Software Engineering* **39**:5, 725-741. [Crossref]

742. Ioannis Karydis, Katia Lida Kermanidis, Spyros Sioutas, Lazaros Iliadis. 2013. Comparing content and context based similarity for musical data. *Neurocomputing* **107**, 69-76. [Crossref]

743. K. Bretonnel Cohen, Lawrence E. Hunter. 2013. Chapter 16: Text Mining for Translational Bioinformatics. *PLoS Computational Biology* **9**:4, e1003044. [Crossref]

744. Andrea De Lucia, Massimiliano Di Penta, Rocco Oliveto, Annibale Panichella, Sebastiano Panichella. 2013. Applying a smoothing filter to improve IR-based traceability recovery processes: An empirical investigation. *Information and Software Technology* **55**:4, 741-754. [Crossref]

745. Francesco Camastra, Angelo Ciaramella, Antonino Staiano. 2013. Machine learning and soft computing for ICT security: an overview of current trends. *Journal of Ambient Intelligence and Humanized Computing* **4**:2, 235-247. [Crossref]

746. Maha Amami, Rim Faiz, Aymen Elkhlifi. A classification approach to extract biological events 1-6. [Crossref]

747. Amy J. C. Trappey, Charles V. Trappey, Tzu-An Chiang, Yi-Hsuan Huang. 2013. Ontology-based neural network for patent knowledge management in design collaboration. *International Journal of Production Research* **51**:7, 1992-2005. [Crossref]

748. Meliha Yetisgen-Yildiz, Martin L. Gunn, Fei Xia, Thomas H. Payne. 2013. A text processing pipeline to extract recommendations from radiology reports. *Journal of Biomedical Informatics* **46**:2, 354-362. [Crossref]

749. Lisheng Zhang, Yan Tang. Keywords Extraction method based on deleting actor index 167-170. [Crossref]

750. Houqiang Li, Yang Wang, Tao Mei, Jingdong Wang, Shipeng Li. 2013. Interactive Multimodal Visual Search on Mobile Device. *IEEE Transactions on Multimedia* **15**:3, 594-607. [Crossref]

751. Bogdan Dit, Meghan Revelle, Denys Poshyvanyk. 2013. Integrating information retrieval, execution and link analysis algorithms to improve feature location in software. *Empirical Software Engineering* **18**:2, 277-309. [Crossref]

752. Alberto Acerbi, Vasileios Lampos, Philip Garnett, R. Alexander Bentley. 2013. The Expression of Emotions in 20th Century Books. *PLoS ONE* **8**:3, e59030. [Crossref]

753. Zied Kechaou, Ali Wali, Mohamed Ben Ammar, Hichem Karray, Adel M. Alimi. 2013. A novel system for video news' sentiment analysis. *Journal of Systems and Information Technology* **15**:1, 24-44. [Abstract] [Full Text] [PDF]

754. Wei You, Dominique Fontaine, Jean-Paul Barthès. 2013. An automatic keyphrase extraction system for scientific documents. *Knowledge and Information Systems* **34**:3, 691-724. [Crossref]

755. Debarchana (Debs) Ghosh, Rajarshi Guha. 2013. What are we 'tweeting' about obesity? Mapping tweets with topic modeling and Geographic Information System. *Cartography and Geographic Information Science* **40**:2, 90-102. [Crossref]

756. A. Chavali, A. D. Kulkarni. A fuzzy self-constructing algorithm for feature reduction 35-40. [Crossref]

757. ASMAA M. EL-SAID, ALI I. ELDESOKY, HESHAM A. ARAFAT. 2013. AN EFFICIENT OBJECT ORIENTED TEXT ANALYSIS (OOTA) APPROACH TO CONSTRUCT STATIC STRUCTURE WITH DYNAMIC BEHAVIOR. *International Journal of Information Acquisition* **09**:01, 1350006. [Crossref]

758. . Bibliography 225-266. [Crossref]

759. . Logistic Regression and Text Classification 59-84. [Crossref]

760. . Bibliography 227-241. [Crossref]

761. Alexandre Trilla, Francesc Alias. 2013. Sentence-Based Sentiment Analysis for Expressive Text-to-Speech. *IEEE Transactions on Audio, Speech, and Language Processing* **21**:2, 223-233. [Crossref]

762. Uzma Raja. 2013. All complaints are not created equal: text analysis of open source software defect reports. *Empirical Software Engineering* **18**:1, 117-138. [Crossref]

763. Yonghee Shin, Laurie Williams. 2013. Can traditional fault prediction models be used for vulnerability prediction?. *Empirical Software Engineering* **18**:1, 25-59. [Crossref]

764. Lejun Gong, Ronggen Yang, Xiao Sun. 2013. BRES: EXTRACTING MULTICLASS BIOMEDICAL RELATIONS WITH SEMANTIC NETWORK. *Biomedical Engineering: Applications, Basis and Communications* **25**:01, 1350009. [Crossref]

765. Rodrigo Moraes, João Francisco Valiati, Wilson P. Gavião Neto. 2013. Document-level sentiment classification: An empirical comparison between SVM and ANN. *Expert Systems with Applications* **40**:2, 621-633. [Crossref]

766. Ilias Flaounas, Omar Ali, Thomas Lansdall-Welfare, Tijl De Bie, Nick Mosdell, Justin Lewis, Nello Cristianini. 2013. RESEARCH METHODS IN THE AGE OF DIGITAL JOURNALISM. *Digital Journalism* **1**:1, 102-116. [Crossref]

767. Michael Scharkow. 2013. Thematic content analysis using supervised machine learning: An empirical evaluation using German online news. *Quality & Quantity* **47**:2, 761-773. [Crossref]

768. Xin Jin, Jiebo Luo, Jie Yu, Gang Wang, D. Joshi, Jiawei Han. 2013. Reinforced Similarity Integration in Image-Rich Information Networks. *IEEE Transactions on Knowledge and Data Engineering* **25**:2, 448-460. [Crossref]

769. Rahul Singh, Ya-Wen Hsu, Naureen Moon. 2013. Multiple perspective interactive search: a paradigm for exploratory search and information retrieval on the web. *Multimedia Tools and Applications* **62**:2, 507-543. [Crossref]

770. Michael Hagenau, Matthias Hauser, Michael Liebmann, Dirk Neumann. Reading All the News at the Same Time: Predicting Mid-term Stock Price Developments Based on News Momentum 1279-1288. [Crossref]

771. Boliang Sun, Guohui Li, Li Jia, Kuihua Huang. 2013. Online Coregularization for Multiview Semisupervised Learning. *The Scientific World Journal* **2013**, 1-15. [Crossref]

772. Andrian Marcus, Sonia Haiduc. Text Retrieval Approaches for Concept Location in Source Code 126-158. [Crossref]

773. Anika Oellrich, Christoph Grabmüller, Dietrich Rebholz-Schuhmann. 2013. Automatically transforming pre- to post-composed phenotypes: EQ-lising HPO and MP. *Journal of Biomedical Semantics* **4**:1, 29. [Crossref]

774. Michael Gowanlock, Rich Gazan. 2013. Assessing researcher interdisciplinarity: a case study of the University of Hawaii NASA Astrobiology Institute. *Scientometrics* **94**:1, 133-161. [Crossref]

775. Jordan Boyd-Graber, Kimberly Glasgow, Jackie Sauter Zajac. 2013. Spoiler alert: Machine learning approaches to detect social media posts with revelatory information. *Proceedings of the American Society for Information Science and Technology* **50**:1, 1-9. [Crossref]

776. Mitra Akasereh. 2013. A quantitative evaluation of query expansion in domain specific information retrieval. *Proceedings of the American Society for Information Science and Technology* **50**:1, 1-7. [Crossref]

777. Hadi Fanaee Tork, Mehran Yazdi. 2013. A Semantic VSM-Based Recommender System. *International Journal of Computer Theory and Engineering* 331-336. [Crossref]

778. Stuti Karol, Veenu Mangat. 2013. Evaluation of text document clustering approach based on particle swarm optimization. *Open Computer Science* **3**:2. . [Crossref]

779. Muhammad Naveed Anwar. 2013. Mining and analysis of audiology data to find significant factors associated with tinnitus masker. *SpringerPlus* **2**:1, 595. [Crossref]

780. Davide Falessi, Giovanni Cantone, Gerardo Canfora. 2013. Empirical Principles and an Industrial Case Study in Retrieving Equivalent Requirements via Natural Language Processing Techniques. *IEEE Transactions on Software Engineering* **39**:1, 18-44. [Crossref]

781. Mayy M. Al-Tahrawi. 2013. The Role of Rare Terms in Enhancing the Performance of Polynomial Networks Based Text Categorization. *Journal of Intelligent Learning Systems and Applications* **05**:02, 84-89. [Crossref]

782. Mita K. Dalal, Mukesh A. Zaveri. 2013. Automatic Classification of Unstructured Blog Text. *Journal of Intelligent Learning Systems and Applications* **05**:02, 108-114. [Crossref]

783. Daniela Godoy. 2012. One-class support vector machines for personalized tag-based resource classification in social bookmarking systems. *Concurrency and Computation: Practice and Experience* **24**:17, 2193-2206. [Crossref]

784. Juliana T. Pollettini, Sylvia R. G. Panico, Julio C. Daneluzzi, Renato Tinós, José A. Baranauskas, Alessandra A. Macedo. 2012. Using Machine Learning Classifiers to Assist Healthcare-Related Decisions: Classification of Electronic Patient Records. *Journal of Medical Systems* **36**:6, 3861-3874. [Crossref]

785. Thierry Paquet, Laurent Heutte, Guillaume Koch, Clément Chatelain. 2012. A categorization system for handwritten documents. *International Journal on Document Analysis and Recognition (IJDAR)* **15**:4, 315-330. [Crossref]

786. Mohamad Khairul Annuar B Zainudin, Riaza Mohd Rias. M-Hadith: Retrieving Malay Haditli text in a mobile application 60-63. [Crossref]

787. Jacques Lonchamp. 2012. Computational analysis and mapping of ijCSCL content. *International Journal of Computer-Supported Collaborative Learning* **7**:4, 475-497. [Crossref]

788. Tanmay Basu, C.A. Murthy. Effective Text Classification by a Supervised Feature Selection Approach 918-925. [Crossref]

789. Nicolas Leveau, Sandra Jhean-Larose, Guy Denhière, Ba-Linh Nguyen. 2012. Validating an interlingual metanorm for emotional analysis of texts. *Behavior Research Methods* **44**:4, 1007-1014. [Crossref]

790. Damir Vandic, Jan-Willem van Dam, Flavius Frasincar. 2012. A semantic-based approach for searching and browsing tag spaces. *Decision Support Systems* **54**:1, 644-654. [Crossref]

791. Stefanos Vrochidis, Anastasia Moumtzidou, Ioannis Kompatsiaris. 2012. Concept-based patent image retrieval. *World Patent Information* **34**:4, 292-303. [Crossref]

792. Chia-Hui Chang, Jhih-Ming Chen. Automatic Extraction of Blog Post from Diverse Blog Pages 129-136. [Crossref]

793. Alan S. Abrahams, Jian Jiao, G. Alan Wang, Weiguo Fan. 2012. Vehicle defect discovery from social media. *Decision Support Systems* **54**:1, 87-97. [Crossref]

794. Dirk Thorleuchter, Dirk Van den Poel. 2012. Improved multilevel security with latent semantic indexing. *Expert Systems with Applications* **39**:18, 13462-13471. [Crossref]

795. Qingliang Miao, Shu Zhang, Yao Meng, Hao Yu. Polarity Analysis for Food and Disease Relationships 188-195. [Crossref]

796. Weifeng Pan, Shan Li. 2012. Tag Ontology Automatic Building for Semantic Searching of Services: a Case Study on Mashup Services. *Journal of Computers* **7**:12. . [Crossref]

797. Axel Gerstmair, Philipp Daumke, Kai Simon, Mathias Langer, Elmar Kotter. 2012. Intelligent image retrieval based on radiology reports. *European Radiology* **22**:12, 2750-2758. [Crossref]

798. Alper Kursat Uysal, Serkan Gunal. 2012. A novel probabilistic feature selection method for text classification. *Knowledge-Based Systems* **36**, 226-235. [Crossref]

799. Ian Kaplan, Andrew Rosenberg. Analysis of speech transcripts to predict winners of U.S. Presidential and Vice-Presidential debates 449-454. [Crossref]

800. Yen-Liang Chen, Yuan-Che Chang. 2012. A three-phase method for patent classification. *Information Processing & Management* **48**:6, 1017-1030. [Crossref]

801. Colin Layfield. With LSA Size DOES Matter 127-131. [Crossref]

802. Krunoslav Zubrinic, Damir Kalpic, Mario Milicevic. 2012. The automatic creation of concept maps from documents written using morphologically rich languages. *Expert Systems with Applications* **39**:16, 12709-12718. [Crossref]

803. Chayan Chakrabarti, George F. Luger. A semantic architecture for artificial conversations 21-26. [Crossref]

804. V. Viswanathan, K. Ilango. 2012. Ranking semantic relationships between two entities using personalization in context specification. *Information Sciences* **207**, 35-49. [Crossref]

805. Meera Sharma, Punam Bedi, K.K. Chaturvedi, V.B. Singh. Predicting the priority of a reported bug using machine learning techniques and cross project validation 539-545. [Crossref]

806. Shrihari A. Hudli, Aditi A. Hudli, Anand V. Hudli. Identifying online opinion leaders using K-means clustering 416-419. [Crossref]

807. Yunfei Yi, Lijun Liu, Cheng Hua Li, Wei Song, Shuai Liu. Machine Learning Algorithms with Co-occurrence Based Term Association for Text Mining 958-962. [Crossref]

808. Yu Jing, Wang Junli, Zhao Xiaodong. An Ontology Term Extracting Method Based on Latent Dirichlet Allocation 366-369. [Crossref]

809. Soumaya Medini, Giuliano Antoniol, Yann-Gael Gueheneuc, Massimiliano Di Penta, Paolo Tonella. SCAN: An Approach to Label and Relate Execution Trace Segments 135-144. [Crossref]

810. Haydemar Nunez, Esmeralda Ramos. Automatic classification of academic documents using text mining techniques 1-7. [Crossref]

811. Guixiang Ma, Youli Qu. A local LDA based method for Latent Aspect Rating Analysis on reviews 2240-2245. [Crossref]

812. Methanias Colaco Junior, Janisson Gois de Souza, Carlos Alberto Goncalves. An Approach to Improve the Empathy of Text Interactions in Collaborative Systems 11-15. [Crossref]

813. Zafer Erenel, Hakan Altınçay. 2012. Nonlinear transformation of term frequencies for term weighting in text categorization. *Engineering Applications of Artificial Intelligence* **25**:7, 1505-1514. [Crossref]

814. Efthimia Mavridou, George Hassapis, Dionisis D. Kehagias, Dimitrios Tzovaras. Semantic Categorization of Web Services Based on Feature Space Transformation 162-167. [Crossref]

815. Zhen Li, Derrick Tate, Christopher Lane, Christopher Adams. 2012. A framework for automatic TRIZ level of invention estimation of patents using natural language processing, knowledge-transfer and patent citation metrics. *Computer-Aided Design* **44**:10, 987-1010. [Crossref]

816. Shuguang Wang, Milos Hauskrecht. Keyword annotation of biomedicai documents with graph-based similarity methods 1-4. [Crossref]

817. Gonenc Ercan, Ilyas Cicekli. 2012. Keyphrase extraction through query performance prediction. *Journal of Information Science* **38**:5, 476-488. [Crossref]

818. Ulrike Kappler, Karen Davenport, Scott Beatson, Susan Lucas, Alla Lapidus, Alex Copeland, Kerrie W. Berry, Tijana Glavina Del Rio, Nancy Hammon, Eileen Dalin, Hope Tice, Sam Pitluck, Paul Richardson, David Bruce, Lynne A. Goodwin, Cliff Han, Roxanne Tapia, John C. Detter, Yun-juan Chang, Cynthia D. Jeffries, Miriam Land, Loren Hauser, Nikos C. Kyrpides, Markus Göker, Natalia Ivanova, Hans-Peter Klenk, Tanja Woyke. 2012. Complete genome sequence of the facultatively chemolithoautotrophic and methylotrophic alpha Proteobacterium Starkeya novella type strain (ATCC 8093T). *Standards in Genomic Sciences* **7**:1, 44-58. [Crossref]

819. David Buffoni, Sabrina Tollari, Patrick Gallinari. 2012. A Learning to Rank framework applied to text-image retrieval. *Multimedia Tools and Applications* **60**:1, 161-180. [Crossref]

820. Alessio Ferrari, Stefania Gnesi. Using collective intelligence to detect pragmatic ambiguities 191-200. [Crossref]

821. Jae Hwa Lee, Aviv Segev. 2012. Knowledge maps for e-learning. *Computers & Education* **59**:2, 353-364. [Crossref]

822. Hong-Jie Dai, Chun-Yu Chen, Chi-Yang Wu, Po-Ting Lai, Richard Tzong-Han Tsai, Wen-Lian Hsu. 2012. Coreference resolution of medical concepts in discharge summaries by exploiting contextual information. *Journal of the American Medical Informatics Association* **19**:5, 888-896. [Crossref]

823. Claudia Wagner, Vera Liao, Peter Pirolli, Les Nelson, Markus Strohmaier. It's Not in Their Tweets: Modeling Topical Expertise of Twitter Users 91-100. [Crossref]

824. K. K. Chaturvedi, V. B. Singh. Determining Bug severity using machine learning techniques 1-6. [Crossref]

825. Claudio Biancalana, Fabio Gasparetti, Alessandro Micarelli, Giuseppe Sansonetti. Enhancing Query Expansion through Folksonomies and Semantic Classes 611-616. [Crossref]

826. Kathleen M. Carley, Michael W. Bigrigg, Boubacar Diallo. 2012. Data-to-model: a mixed initiative approach for rapid ethnographic assessment. *Computational and Mathematical Organization Theory* **18**:3, 300-327. [Crossref]

827. Emily Hill, Shivani Rao, Avinash Kak. On the Use of Stemming for Concern Location and Bug Localization in Java 184-193. [Crossref]

828. P.S.M. Nizer, J.C. Nievola. 2012. Predicting published news effect in the Brazilian stock market. *Expert Systems with Applications* **39**:12, 10674-10680. [Crossref]

829. Willem Robert van Hage, Marieke van Erp, Véronique Malaisé. 2012. Linked Open Piracy: A Story about e-Science, Linked Data, and Statistics. *Journal on Data Semantics* **1**:3, 187-201. [Crossref]

830. John A. Bullinaria, Joseph P. Levy. 2012. Extracting semantic representations from word co-occurrence statistics: stop-lists, stemming, and SVD. *Behavior Research Methods* **44**:3, 890-907. [Crossref]

831. Philips K. Prasetyo, David Lo, Palakorn Achananuparp, Yuan Tian, Ee-Peng Lim. Automatic classification of software related microblogs 596-599. [Crossref]

832. Hossein Mohammadhassanzadeh, Hamid Reza Shahriari. Using user similarity to infer trust values in social networks regardless of direct ratings 66-72. [Crossref]

833. Qiwei He, Bernard P. Veldkamp, Theo de Vries. 2012. Screening for posttraumatic stress disorder using verbal features in self narratives: A text mining approach. *Psychiatry Research* **198**:3, 441-447. [Crossref]

834. Ralph Ewerth, Khalid Ballafkir, Markus Muhling, Dominik Seiler, Bernd Freisleben. 2012. Long-Term Incremental Web-Supervised Learning of Visual Concepts via Random Savannas. *IEEE Transactions on Multimedia* **14**:4, 1008-1020. [Crossref]

835. Nicolás Tourné, Daniela Godoy. 2012. Evaluating tag filtering techniques for web resource classification in folksonomies. *Expert Systems with Applications* **39**:10, 9723-9729. [Crossref]

836. Gang Kou, Chunwei Lou. 2012. Multiple factor hierarchical clustering algorithm for large scale web page and search engine clickstream data. *Annals of Operations Research* **197**:1, 123-134. [Crossref]

837. Christian Gschwend, Ross S. Purves. 2012. Exploring Geomorphometry through User Generated Content: Comparing an Unsupervised Geomorphometric Classification with Terms Attached to Georeferenced Images in Great Britain. *Transactions in GIS* **16**:4, 499-522. [Crossref]

838. Alberto Pérez García-Plaza, Arkaitz Zubiaga, Víctor Fresno, Raquel Martínez. 2012. Reorganizing clouds: A study on tag clustering and evaluation. *Expert Systems with Applications* **39**:10, 9483-9493. [Crossref]

839. Lan Huang, David Milne, Eibe Frank, Ian H. Witten. 2012. Learning a concept-based document similarity measure. *Journal of the American Society for Information Science and Technology* **63**:8, 1593-1608. [Crossref]

840. Stephen Davies, Stacey Aylor Seal, Jesse Hatfield. 2012. Cinefile: A Category-Based Analytic Browser. *Computer* **45**:8, 62-69. [Crossref]

841. Jacques Savoy, Olena Zubaryeva. 2012. Simple and efficient classification scheme based on specific vocabulary. *Computational Management Science* **9**:3, 401-415. [Crossref]

842. Iain Anderson, Olga Chertkov, Amy Chen, Elizabeth Saunders, Alla Lapidus, Matt Nolan, Susan Lucas, Nancy Hammon, Shweta Deshpande, Jan-Fang Cheng, Cliff Han, Roxanne Tapia, Lynne A. Goodwin, Sam Pitluck, Konstantinos Liolios, Ioanna Pagani, Natalia Ivanova, Natalia Mikhailova, Amrita Pati, Krishna Palaniappan, Miriam Land, Chongle Pan, Manfred Rohde, Rüdiger Pukall, Markus Göker, John C. Detter, Tanja Woyke, James Bristow, Jonathan A. Eisen, Victor Markowitz, Philip Hugenholtz, Nikos C. Kyrpides, Hans-Peter Klenk, Konstantinos Mavromatis. 2012. Complete genome sequence of the moderately thermophilic mineral-sulfide-oxidizing firmicute Sulfobacillus acidophilus type strain (NALT). *Standards in Genomic Sciences* **6**:3, 293-303. [Crossref]

843. Zhongqiang Chen, Mema Roussopoulos, Zhanyan Liang, Yuan Zhang, Zhongrong Chen, Alex Delis. 2012. Malware characteristics and threats on the internet ecosystem. *Journal of Systems and Software* **85**:7, 1650-1672. [Crossref]

844. Adriano Veloso, Anderson A. Ferreira, Marcos André Gonçalves, Alberto H.F. Laender, Wagner Meira. 2012. Cost-effective on-demand associative author name disambiguation. *Information Processing & Management* **48**:4, 680-697. [Crossref]

845. F. Murtagh. 2012. Ultrametric model of mind, II: Application to text content analysis. *P-Adic Numbers, Ultrametric Analysis, and Applications* **4**:3, 207-221. [Crossref]

846. Giovanni Pilato, Riccardo Rizzo, Filippo Vella, Ignazio Infantino. Human-Robot Interaction Based on Introspective Capability 461-468. [Crossref]

847. Roger Blake, Ramakrishna Ayyagari. 2012. Analyzing Information Systems Security Research to Find Key Topics, Trends, and Opportunities. *Journal of Information Privacy and Security* **8**:3, 37-67. [Crossref]

848. Mostafa M. Hassan, Fakhri Karray, Mohamed S. Kamel. Automatic Document Topic Identification using Wikipedia Hierarchical Ontology 237-242. [Crossref]

849. Alper Kursat Uysal, Serkan Gunal, Semih Ergin, Efnan Sora Gunal. A novel framework for SMS spam filtering 1-4. [Crossref]

850. Magdalini Eirinaki, Shamita Pisal, Japinder Singh. 2012. Feature-based opinion mining and ranking. *Journal of Computer and System Sciences* **78**:4, 1175-1184. [Crossref]

851. Sung-Jeon Song, Young-Mee Chung. 2012. A Study on Improving the Performance of Document Classification Using the Context of Terms. *Journal of the Korean Society for information Management* **29**:2, 205-224. [Crossref]

852. Jonathon Read, Erik Velldal, Lilja Ovrelid. 2012. Topic Classification for Suicidology. *Journal of Computing Science and Engineering* **6**:2, 143-150. [Crossref]

853. P. Natarajan, Shuang Wu, S. Vitaladevuni, Xiaodan Zhuang, S. Tsakalidis, Unsang Park, R. Prasad, P. Natarajan. Multimodal feature fusion for robust event detection in web videos 1298-1305. [Crossref]

854. B. Sisman, A. C. Kak. Incorporating version histories in Information Retrieval based bug localization 50-59. [Crossref]

855. Zheng Li. An On-demand Service Discovery Approach Based on Mined Domain Knowledge 393-396. [Crossref]

856. Andrea De Lucia, Massimiliano Di Penta, Rocco Oliveto, Annibale Panichella, Sebastiano Panichella. Using IR methods for labeling source code artifacts: Is it worthwhile? 193-202. [Crossref]

857. Lijun Wang, Manjeet Rege, Ming Dong, Yongsheng Ding. 2012. Low-Rank Kernel Matrix Factorization for Large-Scale Evolutionary Clustering. *IEEE Transactions on Knowledge and Data Engineering* **24**:6, 1036-1050. [Crossref]

858. Andrei P. Kirilenko, Svetlana O. Stepchenkova. 2012. Climate change discourse in mass media: application of computer-assisted content analysis. *Journal of Environmental Studies and Sciences* **2**:2, 178-191. [Crossref]

859. Dumidu Wijayasekara, Milos Manic, Jason L. Wright, Miles McQueen. Mining Bug Databases for Unidentified Software Vulnerabilities 89-96. [Crossref]

860. Daniela Pohl, Abdelhamid Bouchachia, Hermann Hellwagner. Supporting Crisis Management via Sub-event Detection in Social Networks 373-378. [Crossref]

861. Cheolhyun Jeong, Kwangsoo Kim. Technology relationship analysis using problem and solution similarities 516-521. [Crossref]

862. Richard Mason, Brian McInnis, Siddhartha Dalal. Machine learning for the automatic identification of terrorist incidents in worldwide news media 84-89. [Crossref]

863. Sanjiv R. Das. News analytics: Framework, techniques, and metrics 41-71. [Crossref]

864. Hans-Peter Klenk, Megan Lu, Susan Lucas, Alla Lapidus, Alex Copeland, Sam Pitluck, Lynne A. Goodwin, Cliff Han, Roxanne Tapia, Evelyne-Marie Brambilla, Gabriele Pötter, Miriam Land, Natalia Ivanova, Manfred Rohde, Markus Göker, John C. Detter, Wen-Jun Li, Nikos C. Kyrpides, Tanja Woyke. 2012. Genome sequence of the ocean sediment bacterium Saccharomonospora marina type strain (XMU15T). *Standards in Genomic Sciences* **6**:2, 265-275. [Crossref]

865. Markus Göker, Elisabeth Saunders, Alla Lapidus, Matt Nolan, Susan Lucas, Nancy Hammon, Shweta Deshpande, Jan-Fang Cheng, Cliff Han, Roxanne Tapia, Lynne A. Goodwin, Sam Pitluck, Konstantinos Liolios, Konstantinos Mavromatis, Ioanna Pagani, Natalia Ivanova, Natalia Mikhailova, Amrita Pati, Amy Chen, Krishna Palaniappan, Miriam Land, Yun-juan Chang, Cynthia D. Jeffries, Evelyne-Marie Brambilla, Manfred Rohde, Stefan Spring, John C. Detter, Tanja Woyke, James Bristow, Jonathan A. Eisen, Victor Markowitz, Philip Hugenholtz, Nikos C. Kyrpides, Hans-Peter Klenk. 2012. Genome sequence of the moderately thermophilic, amino-acid-degrading and sulfur-reducing bacterium Thermovirga lienii type strain (Cas60314T). *Standards in Genomic Sciences* **6**:2, 230-239. [Crossref]

866. Alex Copeland, Xiaojing Zhang, Monica Misra, Alla Lapidus, Matt Nolan, Susan Lucas, Shweta Deshpande, Jan-Fang Cheng, Roxanne Tapia, Lynne A. Goodwin, Sam Pitluck, Konstantinos Liolios, Ioanna Pagani, Natalia Ivanova, Natalia Mikhailova, Amrita Pati, Amy Chen, Krishna Palaniappan, Miriam Land, Loren Hauser, Chongle Pan, Cynthia D. Jeffries, John C. Detter, Evelyne-Marie Brambilla, Manfred Rohde, Olivier D. Ngatchou Djao, Markus Göker, Johannes Sikorski, Brian J. Tindall, Tanja Woyke, James Bristow, Jonathan A. Eisen, Victor Markowitz, Philip Hugenholtz, Nikos C. Kyrpides, Hans-Peter Klenk, Konstantinos Mavromatis. 2012. Complete genome sequence of the aquatic bacterium Runella slithyformis type strain (LSU 4T). *Standards in Genomic Sciences* **6**:2, 145-154. [Crossref]

867. Marcel Huntemann, Hazuki Teshima, Alla Lapidus, Matt Nolan, Susan Lucas, Nancy Hammon, Shweta Deshpande, Jan-Fang Cheng, Roxanne Tapia, Lynne A. Goodwin, Sam Pitluck, Konstantinos Liolios, Ioanna Pagani, Natalia Ivanova, Konstantinos Mavromatis, Natalia Mikhailova, Amrita Pati, Amy Chen, Krishna Palaniappan, Miriam Land, Loren Hauser, Chongle Pan, Evelyne-Marie Brambilla, Manfred Rohde, Stefan Spring, Markus Göker, John C. Detter, James Bristow, Jonathan A. Eisen, Victor Markowitz, Philip Hugenholtz, Nikos C. Kyrpides, Hans-Peter Klenk, Tanja Woyke. 2012. Complete genome sequence of the facultatively anaerobic, appendaged bacterium Muricauda ruestringensis type strain (B1T). *Standards in Genomic Sciences* **6**:2, 185-193. [Crossref]

868. Iain Anderson, Elizabeth Saunders, Alla Lapidus, Matt Nolan, Susan Lucas, Hope Tice, Tijana Glavina Del Rio, Jan-Fang Cheng, Cliff Han, Roxanne Tapia, Lynne A. Goodwin, Sam Pitluck, Konstantinos Liolios, Konstantinos Mavromatis, Ioanna Pagani, Natalia Ivanova, Natalia Mikhailova, Amrita Pati, Amy Chen, Krishna Palaniappan, Miriam Land, Loren Hauser, Cynthia D. Jeffries, Yun-juan Chang, Evelyne-Marie Brambilla, Manfred Rohde, Stefan Spring, Markus Göker, John C. Detter, Tanja Woyke, James Bristow, Jonathan A. Eisen, Victor Markowitz, Philip Hugenholtz, Nikos C. Kyrpides, Hans-Peter Klenk. 2012. Complete genome sequence of the thermophilic sulfate-reducing ocean bacterium Thermodesulfatator indicus type strain (CIR29812T). *Standards in Genomic Sciences* **6**:2, 155-164. [Crossref]

869. Iain Anderson, Brittany Held, Alla Lapidus, Matt Nolan, Susan Lucas, Hope Tice, Tijana Glavina Del Rio, Jan-Fang Cheng, Cliff Han, Roxanne Tapia, Lynne A. Goodwin, Sam Pitluck, Konstantinos Liolios, Konstantinos Mavromatis, Ioanna Pagani, Natalia Ivanova, Natalia Mikhailova, Amrita Pati, Amy Chen, Krishna Palaniappan, Miriam Land, Evelyne-Marie Brambilla, Manfred Rohde, Stefan

Spring, Markus Göker, John C. Detter, Tanja Woyke, James Bristow, Jonathan A. Eisen, Victor Markowitz, Philip Hugenholtz, Hans-Peter Klenk, Nikos C. Kyrpides. 2012. Genome sequence of the homoacetogenic bacterium Holophaga foetida type strain (TMBS4T). *Standards in Genomic Sciences* **6**:2, 174-184. [Crossref]

870. Konstantinos Mavromatis, Olga Chertkov, Alla Lapidus, Matt Nolan, Susan Lucas, Hope Tice, Tijana Glavina Del Rio, Jan-Fang Cheng, Cliff Han, Roxanne Tapia, David Bruce, Lynne A. Goodwin, Sam Pitluck, Marcel Huntemann, Konstantinos Liolios, Ioanna Pagani, Natalia Ivanova, Natalia Mikhailova, Amrita Pati, Amy Chen, Krishna Palaniappan, Miriam Land, Evelyne-Marie Brambilla, Manfred Rohde, Stefan Spring, Markus Göker, John C. Detter, James Bristow, Jonathan A. Eisen, Victor Markowitz, Philip Hugenholtz, Nikos C. Kyrpides, Hans-Peter Klenk, Tanja Woyke. 2012. Permanent draft genome sequence of the gliding predator Saprospira grandis strain Sa g1 (= HR1). *Standards in Genomic Sciences* **6**:2, 210-219. [Crossref]

871. Birte Abt, Cliff Han, Carmen Scheuner, Megan Lu, Alla Lapidus, Matt Nolan, Susan Lucas, Nancy Hammon, Shweta Deshpande, Jan-Fang Cheng, Roxanne Tapia, Lynne A. Goodwin, Sam Pitluck, Konstantinos Liolios, Ioanna Pagani, Natalia Ivanova, Konstantinos Mavromatis, Natalia Mikhailova, Marcel Huntemann, Amrita Pati, Amy Chen, Krishna Palaniappan, Miriam Land, Loren Hauser, Evelyne-Marie Brambilla, Manfred Rohde, Stefan Spring, Sabine Gronow, Markus Göker, Tanja Woyke, James Bristow, Jonathan A. Eisen, Victor Markowitz, Philip Hugenholtz, Nikos C. Kyrpides, Hans-Peter Klenk, John C. Detter. 2012. Complete genome sequence of the termite hindgut bacterium Spirochaeta coccoides type strain (SPN1T), reclassification in the genus Sphaerochaeta as Sphaerochaeta coccoides comb. nov. and emendations of the family Spirochaetaceae and the genus Sphaerochaeta. *Standards in Genomic Sciences* **6**:2, 194-209. [Crossref]

872. M. Sanderson, W. B. Croft. 2012. The History of Information Retrieval Research. *Proceedings of the IEEE* **100**:Special Centennial Issue, 1444-1451. [Crossref]

873. Takaya Yoshida, Hayato Ohwada. Using abstracts to classify life medicine papers 2272-2275. [Crossref]

874. Jung Song Lee, Soon Cheol Park. Document Clustering Using Multi-Objective Genetic Algorithms on MATLAB Distributed Computing 1-6. [Crossref]

875. Jer Lang Hong. OntoLabel - data labeling for deep web using WordNet 1320-1324. [Crossref]

876. Alex Copeland, Ahmet Zeytun, Montri Yassawong, Matt Nolan, Susan Lucas, Nancy Hammon, Shweta Deshpande, Jan-Fang Cheng, Cliff Han, Roxanne Tapia, Lynne A. Goodwin, Sam Pitluck, Konstantinos Mavromatis, Konstantinos Liolios, Ioanna Pagani, Natalia Ivanova, Natalia Mikhailova, Amrita Pati, Amy Chen, Krishna Palaniappan, Miriam Land, Loren Hauser, Cynthia D. Jeffries, Evelyne-Marie Brambilla, Manfred Rohde, Johannes Sikorski, Rüdiger Pukall, Markus Göker, John C. Detter, Tanja Woyke, James Bristow, Jonathan A. Eisen, Victor Markowitz, Philip Hugenholtz, Nikos C. Kyrpides, Hans-Peter Klenk, Alla Lapidus. 2012. Complete genome sequence of the orange-red pigmented, radioresistant Deinococcus proteolyticus type strain (MRPT). *Standards in Genomic Sciences* **6**:2, 240-250. [Crossref]

877. David Sánchez, Antonio Moreno, Luis Del Vasto-Terrientes. 2012. Learning relation axioms from text: An automatic Web-based approach. *Expert Systems with Applications* **39**:5, 5792-5805. [Crossref]

878. P. Swathy Priyadharsini, A. Natarajan, A. M. Natarajan. Optimization of bloom filter using evolutionary algorithm 86-90. [Crossref]

879. Geir Solskinnsbakk, Jon Atle Gulla, Veronika Haderlein, Per Myrseth, Olga Cerrato. 2012. Quality of hierarchies in ontologies and folksonomies. *Data & Knowledge Engineering* **74**, 13-25. [Crossref]

880. Yongli Liu, Qianqian Guo, Lishen Yang, Yingying Li. Research on incremental clustering 2803-2806. [Crossref]

881. Fethi Ferjani, Samir Elloumi, Ali Jaoua, Sadok Ben Yahia, Sahar Ismail, Sheikha Ravan. 2012. Formal context coverage based on isolated labels: An efficient solution for text feature extraction. *Information Sciences* **188**, 198-214. [Crossref]

882. Alper Kursat Uysal, Serkan Gunal, Semih Ergin, Efnan Sora Gunal. Detection of SMS spam messages on mobile phones 1-4. [Crossref]

883. Nouman Azam, JingTao Yao. 2012. Comparison of term frequency and document frequency based feature selection metrics in text categorization. *Expert Systems with Applications* **39**:5, 4760-4768. [Crossref]

884. Goksel Biricik. Comparing the impacts of dimension reduction methods that use class labels on text classification 1-4. [Crossref]

885. Gang Li, Fei Liu. 2012. Application of a clustering method on sentiment analysis. *Journal of Information Science* **38**:2, 127-139. [Crossref]

886. Wei-Feng PAN, Bing LI, Bo SHAO, Peng HE. 2012. Service Classification and Recommendation Based on Software Networks. *Chinese Journal of Computers* **34**:12, 2355-2369. [Crossref]

887. Cheng-Shiun Tasi, Yong-Ming Huang, Chien-Hung Liu, Yueh-Min Huang. 2012. Applying VSM and LCS to develop an integrated text retrieval mechanism. *Expert Systems with Applications* **39**:4, 3974-3982. [Crossref]

888. Jukka Perkiö, Antti Tuominen, Taneli Vähäkangas, Petri Myllymäki. 2012. Image similarity: from syntax to weak semantics. *Multimedia Tools and Applications* **57**:1, 5-27. [Crossref]

889. Ronald J. Mann, Marian Underweiser. 2012. A New Look at Patent Quality: Relating Patent Prosecution to Validity. *Journal of Empirical Legal Studies* **9**:1, 1-32. [Crossref]

890. Maha Amami, Aymen Elkhlifi, Rim Faiz. BioEv: A system for learning biological event extraction 1-5. [Crossref]

891. Weizhong Zhao, Qing He, Huifang Ma, Zhongzhi Shi. 2012. Effective semi-supervised document clustering via active learning with instance-level constraints. *Knowledge and Information Systems* **30**:3, 569-587. [Crossref]

892. Suleyman Cetintas, Luo Si. 2012. Effective query generation and postprocessing strategies for prior art patent search. *Journal of the American Society for Information Science and Technology* **63**:3, 512-527. [Crossref]

893. Krishnendu Ghosh, K. Sreenivasa Rao. Subword based approach for grapheme-to-phoneme conversion in Bengali text-to-speech synthesis system 1-5. [Crossref]

894. Michael Niemann, Melanie Siebenhaar, Stefan Schulte, Ralf Steinmetz. 2012. Comparison and retrieval of process models using related cluster pairs. *Computers in Industry* **63**:2, 168-180. [Crossref]

895. Yuan PING, Ya-jian ZHOU, Chao XUE, Yi-xian YANG. 2012. Efficient representation of text with multiple perspectives. *The Journal of China Universities of Posts and Telecommunications* **19**:1, 101-111. [Crossref]

896. Haijun Zhang, Tommy W.S. Chow. 2012. A multi-level matching method with hybrid similarity for document retrieval. *Expert Systems with Applications* **39**:3, 2710-2719. [Crossref]

897. Marta Indulska, Dirk S Hovorka, Jan Recker. 2012. Quantitative approaches to content analysis: identifying conceptual drift across publication outlets. *European Journal of Information Systems* **21**:1, 49-69. [Crossref]

898. Sampo Pyysalo, Tomoko Ohta, Rafal Rak, Dan Sullivan, Chunhong Mao, Chunxia Wang, Bruno Sobral, Jun'ichi Tsujii, Sophia Ananiadou. 2012. Overview of the ID, EPI and REL tasks of BioNLP Shared Task 2011. *BMC Bioinformatics* **13**:Suppl 11, S2. [Crossref]

899. Michael Siering. "Boom" or "Ruin"--Does It Make a Difference? Using Text Mining and Sentiment Analysis to Support Intraday Investment Decisions 1050-1059. [Crossref]

900. Cheng Hua Li, Ju Cheng Yang, Soon Cheol Park. 2012. Text categorization algorithms using semantic approaches, corpus-based thesaurus and WordNet. *Expert Systems with Applications* **39**:1, 765-772. [Crossref]

901. Jari Björne, Filip Ginter, Tapio Salakoski. 2012. University of Turku in the BioNLP'11 Shared Task. *BMC Bioinformatics* **13**:Suppl 11, S4. [Crossref]

902. Paul F. Cook, Kai R. Larsen, Teresa J. Sakraida, Leli Pedro. 2012. A Novel Quantitative Approach to Concept Analysis. *Nursing Research* **61**:5, 369-378. [Crossref]

903. Majid Soltani Moghaddam, Morteza Nasiri, Alireza Zarea, Sajjad Sepehrinia. 2012. Teaching Pronunciation: The Lost Ring of the Chain. *Journal of Language Teaching and Research* **3**:1. . [Crossref]

904. Hisham Al-Mubaid, Sandeep Gungu. 2012. A Learning-Based Approach for Biomedical Word Sense Disambiguation. *The Scientific World Journal* **2012**, 1-8. [Crossref]

905. Roger Blake, Ganesan Shankaranarayanan. 2012. Discovering Data and Information Quality Research Insights Gained through Latent Semantic Analysis. *International Journal of Business Intelligence Research* **3**:1, 1-16. [Crossref]

906. Nicholas Evangelopoulos, Xiaoni Zhang, Victor R Prybutok. 2012. Latent Semantic Analysis: five methodological recommendations. *European Journal of Information Systems* **21**:1, 70-86. [Crossref]

907. Sofie Van Landeghem, Jari Björne, Thomas Abeel, Bernard De Baets, Tapio Salakoski, Yves Van de Peer. 2012. Semantically linking molecular entities in literature through entity relationships. *BMC Bioinformatics* **13**:Suppl 11, S6. [Crossref]

908. Hakan Altınçay, Zafer Erenel. 2012. Using the absolute difference of term occurrence probabilities in binary text categorization. *Applied Intelligence* **36**:1, 148-160. [Crossref]

909. Pierre de Keyser. Techniques applied in automatic indexing of text material 65-78. [Crossref]

910. Bidyut Das. 2012. Extracting Collocations from Bengali Text Corpus. *Procedia Technology* **4**, 325-329. [Crossref]

911. Miguel Pragosa, Vitor Basto-Fernandes. 2012. Semantic Integration For Food Sales Analysis. *Procedia Technology* **5**, 581-588. [Crossref]

912. Adrian-Gabriel Chifu, Radu-Tudor Ionescu. 2012. Word sense disambiguation to improve precision for ambiguous queries. *Open Computer Science* **2**:4. . [Crossref]

913. Vivek Anandan Ramachandran, Ilango Krishnamurthi. An Iterative Stemmer for Tamil Language 197-205. [Crossref]

914. . Conceptual Foundations of Text Mining and Preprocessing Steps 43-51. [Crossref]

915. Jan Czarnecki, Irene Nobeli, Adrian M Smith, Adrian J Shepherd. 2012. A text-mining system for extracting metabolic reactions from full-text articles. *BMC Bioinformatics* **13**:1, 172. [Crossref]

916. Muhammad N Anwar, Michael P Oakes. 2012. Data mining of audiology patient records: factors influencing the choice of hearing aid type. *BMC Medical Informatics and Decision Making* **12**:Suppl 1, S6. [Crossref]

917. Michael Hagenau, Michael Liebmann, Markus Hedwig, Dirk Neumann. Automated News Reading: Stock Price Prediction Based on Financial News Using Context-Specific Features 1040-1049. [Crossref]

918. Christopher J Mungall, Carlo Torniai, Georgios V Gkoutos, Suzanna E Lewis, Melissa A Haendel. 2012. Uberon, an integrative multi-species anatomy ontology. *Genome Biology* **13**:1, R5. [Crossref]

919. Robert M. Patton, Wade McNair, Christopher T. Symons, Jim N. Treadwell, Thomas E. Potok. A Text Analysis Approach to Motivate Knowledge Sharing via Microsoft SharePoint 3670-3678. [Crossref]

920. Zhe Liu, Bernard J. Jansen. 2012. Almighty Twitter, what are people asking for?. *Proceedings of the American Society for Information Science and Technology* **49**:1, 1-10. [Crossref]

921. Ning Zhong, Yuefeng Li, Sheng-Tang Wu. 2012. Effective Pattern Discovery for Text Mining. *IEEE Transactions on Knowledge and Data Engineering* **24**:1, 30-44. [Crossref]

922. Mita K. Dalal, Mukesh A. Zaveri. Automatic Text Classification of sports blog data 219-222. [Crossref]

923. Seunghee Kim, Jinwook Choi. 2012. Improving the Performance of Text Categorization Models used for the Selection of High Quality Articles. *Healthcare Informatics Research* **18**:1, 18. [Crossref]

924. Matjaž Juršič, Borut Sluban, Bojan Cestnik, Miha Grčar, Nada Lavrač. Bridging Concept Identification for Constructing Information Networks from Text Documents 66-90. [Crossref]

925. Teemu Hynönen, Sébastien Mahler, Hannu Toivonen. Discovery of Novel Term Associations in a Document Collection 91-103. [Crossref]

926. Daniela Godoy. Comparing One-Class Classification Algorithms for Finding Interesting Resources in Social Bookmarking Systems 88-103. [Crossref]

927. M. Gashler, D. Ventura, T. Martinez. 2011. Manifold Learning by Graduated Optimization. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)* **41**:6, 1458-1470. [Crossref]

928. Jangwon Seo, W. Bruce Croft, David A. Smith. 2011. Online community search using conversational structures. *Information Retrieval* **14**:6, 547-571. [Crossref]

929. E. Emadzadeh, S. Jonnalagadda, G. Gonzalez. Evaluating Distributional Semantic and Feature Selection for Extracting Relationships from Biological Text 66-71. [Crossref]

930. Shiliang Sun, Qingjiu Zhang. 2011. Multiple-View Multiple-Learner Semi-Supervised Learning. *Neural Processing Letters* **34**:3, 229-240. [Crossref]

931. Juan D. Velásquez, Luis E. Dujovne, Gaston L'Huillier. 2011. Extracting significant Website Key Objects: A Semantic Web mining approach. *Engineering Applications of Artificial Intelligence* **24**:8, 1532-1541. [Crossref]

932. Ladda Suanmali, Naomie Salim, Mohammed Salem Binwahlan. Fuzzy Genetic Semantic Based Text Summarization 1184-1191. [Crossref]

933. Dengping Wei, Ting Wang, Ji Wang, Abraham Bernstein. 2011. SAWSDL-iMatcher: A customizable and effective Semantic Web Service matchmaker. *Web Semantics: Science, Services and Agents on the World Wide Web* **9**:4, 402-417. [Crossref]

934. MARCO A. ALVAREZ, CHANGHUI YAN. 2011. A GRAPH-BASED SEMANTIC SIMILARITY MEASURE FOR THE GENE ONTOLOGY. *Journal of Bioinformatics and Computational Biology* **09**:06, 681-695. [Crossref]

935. Rong Hu, Wanchun Dou, Xiaoqing Frank Liu, Jianxun Liu. Personalized Searching for Web Service Using User Interests 156-163. [Crossref]

936. Meghan Revelle, Malcom Gethers, Denys Poshyvanyk. 2011. Using structural and textual information to capture feature coupling in object-oriented software. *Empirical Software Engineering* **16**:6, 773-811. [Crossref]

937. A. Zelaia, I. Alegria, O. Arregi, B. Sierra. 2011. A multiclass/multilabel document categorization system: Combining multiple classifiers in a reduced dimension. *Applied Soft Computing* **11**:8, 4981-4990. [Crossref]

938. Xin Jin, Scott Spangler, Ying Chen, Keke Cai, Rui Ma, Li Zhang, Xian Wu, Jiawei Han. Patent Maintenance Recommendation with Patent Information Network Model 280-289. [Crossref]

939. Brent Wenerstrom, Mehmed Kantardzic. 2011. ReClose: web page summarization combining summary techniques. *International Journal of Web Information Systems* **7**:4, 333-359. [Abstract] [Full Text] [PDF]

940. Aixin Sun, Ying Liu, Ee-Peng Lim. 2011. Web classification of conceptual entities using co-training. *Expert Systems with Applications* **38**:12, 14367-14375. [Crossref]

941. Muhammad Usman Ghani Khan, Lei Zhang, Yoshihiko Gotoh. Human Focused Video Description 1480-1487. [Crossref]

942. Björn Schuller, Anton Batliner, Stefan Steidl, Dino Seppi. 2011. Recognising realistic emotions and affect in speech: State of the art and lessons learnt from the first challenge. *Speech Communication* **53**:9-10, 1062-1087. [Crossref]

943. Hsin-Chang Yang, Chung-Hong Lee. Mining tag semantics for social tag recommendation 760-766. [Crossref]

944. Jari Björne, Juho Heimonen, Filip Ginter, Antti Airola, Tapio Pahikkala, Tapio Salakoski. 2011. EXTRACTING CONTEXTUALIZED COMPLEX BIOLOGICAL EVENTS WITH RICH GRAPH-BASED FEATURE SETS. *Computational Intelligence* **27**:4, 541-557. [Crossref]

945. Lyle Ungar, Stephen Leibholz, Carole Chaski. IntentFinder: A system for discovering significant information implicit in large, heterogeneous document collections and computationally mapping social networks and command nodes 219-223. [Crossref]

946. LiGuo Huang, Vincent Ng, Isaac Persing, Ruili Geng, Xu Bai, Jeff Tian. AutoODC: Automated generation of Orthogonal Defect Classifications 412-415. [Crossref]

947. Mayank Saini, Dharmendar Sharma. Tuning topical queries for effective information retrieval 1-5. [Crossref]

948. Janghyeok Yoon, Kwangsoo Kim. 2011. An automated method for identifying TRIZ evolution trends from patents. *Expert Systems with Applications* **38**:12, 15540-15548. [Crossref]

949. Gianni Costa, Riccardo Ortale, Ettore Ritacco. Effective XML Classification Using Content and Structural Information via Rule Learning 102-109. [Crossref]

950. Roger Blake, Oscar Gutierrez. 2011. A semantic analysis approach for assessing professionalism using free-form text entered online. *Computers in Human Behavior* **27**:6, 2249-2262. [Crossref]

951. Sofus A. Macskassy. 2011. Contextual linking behavior of bloggers: leveraging text mining to enable topic-based analysis. *Social Network Analysis and Mining* **1**:4, 355-375. [Crossref]

952. Sofie Van Landeghem, Bernard De Baets, Yves Van de Peer, Yvan Saeys. 2011. HIGH-PRECISION BIO-MOLECULAR EVENT EXTRACTION FROM TEXT USING PARALLEL BINARY CLASSIFIERS. *Computational Intelligence* **27**:4, 645-664. [Crossref]

953. Manuel Baena-Garcia, Jose M. Carmona-Cejudo, Gladys Castillo, Rafael Morales-Bueno. TF-SIDF: Term frequency, sketched inverse document frequency 1044-1049. [Crossref]

954. Nicole Lang Beebe, Jan Guynes Clark, Glenn B. Dietrich, Myung S. Ko, Daijin Ko. 2011. Post-retrieval search hit clustering to improve information retrieval effectiveness: Two digital forensics case studies. *Decision Support Systems* **51**:4, 732-744. [Crossref]

955. Georg Groh, Christoph Fuchs. 2011. Multi-modal social networks for modeling scientific fields. *Scientometrics* **89**:2, 569-590. [Crossref]

956. Carlos C. Rojas, Robert M. Patton, Barbara G. Beckerman. 2011. Characterizing Mammography Reports for Health Analytics. *Journal of Medical Systems* **35**:5, 1197-1210. [Crossref]

957. Shiro Usui, Nilton Liuji Kamiji, Tatsuki Taniguchi, Naonori Ueda. 2011. RAST: finding related documents based on triplet similarity. *Neural Computing and Applications* **20**:7, 993-999. [Crossref]

958. Isabel Segura-Bedmar, Paloma Martínez, Cesar de Pablo-Sánchez. 2011. Using a shallow linguistic kernel for drug–drug interaction extraction. *Journal of Biomedical Informatics* **44**:5, 789-804. [Crossref]

959. Vivek Kumar Singh, Nisha Tiwari, Shekhar Garg. Document Clustering Using K-Means, Heuristic K-Means and Fuzzy C-Means 297-301. [Crossref]

960. Xiaohui Cui, Jesse St. Charles, Thomas E. Potok. The GPU Enhanced Parallel Computing for Large Scale Data Clustering 220-225. [Crossref]

961. N. Bassiou, C. Kotropoulos. 2011. RPLSA: A novel updating scheme for Probabilistic Latent Semantic Analysis. *Computer Speech & Language* **25**:4, 741-760. [Crossref]

962. S. Cetintas, Luo Si, H. Aagard, K. Bowen, M. Cordova-Sanchez. 2011. Microblogging in a Classroom: Classifying Students' Relevant and Irrelevant Questions in a Microblogging-Supported Classroom. *IEEE Transactions on Learning Technologies* **4**:4, 292-300. [Crossref]

963. Harun Uğuz. 2011. A two-stage feature selection method for text categorization by using information gain, principal component analysis and genetic algorithm. *Knowledge-Based Systems* **24**:7, 1024-1032. [Crossref]

964. Sayyid Mohammad Hoseini Khozani, Hosein Bayat. Specialization of keyword extraction approach to Persian texts 112-116. [Crossref]

965. K. Saravanakumar, K. Deepa. Alternate Query Construction Agent for Improving Web Search Result Using WordNet 117-120. [Crossref]

966. Laura Dabbish, Ben Towne, Jana Diesner, James Herbsleb. 2011. Construction of association networks from communication in teams working on complex projects. *Statistical Analysis and Data Mining* **4**:5, 547-563. [Crossref]

967. Juan Martinez-Romo, Lourdes Araujo, Javier Borge-Holthoefer, Alex Arenas, José A. Capitán, José A. Cuesta. 2011. Disentangling categorical relationships through a graph of co-occurrences. *Physical Review E* **84**:4. . [Crossref]

968. Haijun Zhang, Gang Liu, T. W. S. Chow, Wenyin Liu. 2011. Textual and Visual Content-Based Anti-Phishing: A Bayesian Approach. *IEEE Transactions on Neural Networks* **22**:10, 1532-1546. [Crossref]

969. Zhilong Zhen, Haijuan Wang, Lixin Han, Zhan Shi. Categorical Document Frequency Based Feature Selection for Text Categorization 65-68. [Crossref]

970. Hakim Sultanov, Jane Huffman Hayes, Wei-Keat Kong. 2011. Application of swarm techniques to requirements tracing. *Requirements Engineering* **16**:3, 209-226. [Crossref]

971. Jung-Tae Lee, Jangwon Seo, Jiwoon Jeon, Hae-Chang Rim. 2011. Sentence-based relevance flow analysis for high accuracy retrieval. *Journal of the American Society for Information Science and Technology* **62**:9, 1666-1675. [Crossref]

972. Hsin-Chang Yang, Han-Wei Hsiao, Chung-Hong Lee. 2011. Multilingual document mining and navigation using self-organizing maps. *Information Processing & Management* **47**:5, 647-666. [Crossref]

973. Reihaneh Rabbany Khorasgani, Eleni Stroulia, Osmar R. Zaiane. Web service matching for RESTful web services 115-124. [Crossref]

974. Xin Hou, S.K. Ong, A.Y.C. Nee, X.T. Zhang, W.J. Liu. 2011. GRAONTO: A graph-based approach for automatic construction of domain ontology. *Expert Systems with Applications* **38**:9, 11958-11975. [Crossref]

975. Chun-Ling Chen, Frank S. C. Tseng, Tyne Liang. 2011. An integration of fuzzy association rules and WordNet for document clustering. *Knowledge and Information Systems* **28**:3, 687-708. [Crossref]

976. Rahul Chougule, Dnyanesh Rajpathak, Pulak Bandyopadhyay. 2011. An integrated framework for effective service and repair in the automotive domain: An application of association mining and case-based-reasoning. *Computers in Industry* **62**:7, 742-754. [Crossref]

977. Pekka Abrahamsson, Ilenia Fronza, Raimund Moser, Jelena Vlasenko, Witold Pedrycz. Predicting Development Effort from User Stories 400-403. [Crossref]

978. Malcom Gethers, Rocco Oliveto, Denys Poshyvanyk, Andrea De Lucia. On integrating orthogonal information retrieval methods to improve traceability recovery 133-142. [Crossref]

979. Jian Wang, Jia Zhang, Patrick C.K. Hung, Zheng Li, Jianxiao Liu, Keqing He. Leveraging Fragmental Semantic Data to Enhance Services Discovery 687-694. [Crossref]

980. Andrew Wiese, Valerie Ho, Emily Hill. A comparison of stemmers on source code identifiers for software search 496-499. [Crossref]

981. Dirk Thorleuchter, Dirk Van den Poel. Companies website optimising concerning consumer's searching for new products 40-43. [Crossref]

982. Mihaela Dinsoreanu, Bogdan Neacsa, Catalin Cupse, Rodica Potolea. A layered approach for enabling context-sensitive content 19-26. [Crossref]

983. David G. Gordon, Travis D. Breaux. Comparing requirements from multiple jurisdictions 43-49. [Crossref]

984. Ahmad Hawalah, Maria Fasli. A Multi-agent System Using Ontological User Profiles for Dynamic User Modelling 430-437. [Crossref]

985. Zhao Rongying, Chen Rui. Visual analysis on the research of cross-language information retrieval 32-35. [Crossref]

986. Dirk Thorleuchter, Dirk Van den Poel. Semantic technology classification — A defence and security case study 36-39. [Crossref]

987. Hanze Liu, Orland Hoeber. A Luhn-Inspired Vector Re-weighting Approach for Improving Personalized Web Search 301-305. [Crossref]

988. Giuliano Armano, Alessandro Giuliani, Eloisa Vargiu. Studying the Impact of Text Summarization on Contextual Advertising 172-176. [Crossref]

989. Ran Chen, Hongfei Lin, Zhihao Yang. 2011. Passage retrieval based hidden knowledge discovery from biomedical literature. *Expert Systems with Applications* **38**:8, 9958-9964. [Crossref]

990. Lijun Wang, Ming Dong. On the clustering of large-scale data: A matrix-based approach 139-144. [Crossref]

991. Zonghu Wang, Zhijing Liu, Donghui Chen, Kai Tang. A new partitioning based algorithm for document clustering 1741-1745. [Crossref]

992. Robert M. Patton, Carlos C. Rojas, Barbara G. Beckerman, Thomas E. Potok. A Computational Framework for Search, Discovery, and Trending of Patient Health in Radiology Reports 104-111. [Crossref]

993. Marc-Andre Orthmann, Martin Friedrich, Uwe Kirschenmann, Katja Niemann, Maren Scheffel, Hans-Christian Schmitz, Martin Wolpers. Usage-based Clustering of Learning Objects for Recommendation 553-557. [Crossref]

994. Patrick A. De Mazière, Marc M. Van Hulle. 2011. A clustering study of a 7000 EU document inventory using MDS and SOM. *Expert Systems with Applications* **38**:7, 8835-8849. [Crossref]

995. Krzysztof Jedrzejewski, Mikolaj Morzy. Opinion Mining and Social Networks: A Promising Match 599-604. [Crossref]

996. Maria Vargas-Vera, Yesica Castellanos, Miltiadis D. Lytras. 2011. CONQUIRO: A cluster-based meta-search engine. *Computers in Human Behavior* **27**:4, 1303-1309. [Crossref]

997. Jia Zhang, Ravi Madduri, Wei Tan, Kevin Deichl, John Alexander, Ian Foster. Toward Semantics Empowered Biomedical Web Services 371-378. [Crossref]

998. Keerthiram Murugesan, Jun Zhang. Hybrid Bisect K-Means Clustering Algorithm 216-219. [Crossref]

999. M. M. Brut, F. Sedes, S. D. Dumitrescu. 2011. A Semantic-Oriented Approach for Organizing and Developing Annotation for E-Learning. *IEEE Transactions on Learning Technologies* **4**:3, 239-248. [Crossref]

1000. Neal Lewis, Daniel Gruhl, Hui Yang. Dependency Parsing for Extracting Family History 237-242. [Crossref]

1001. Tengku Mohd T. Sembok, Zainab Abu Bakar, Fatimah Ahmad. Experiments in malay information retrieval 1-5. [Crossref]

1002. Hsin-Chang Yang, Chung-Hong Lee. Post-Level Spam Detection for Social Bookmarking Web Sites 180-185. [Crossref]

1003. Malihe Danesh, Mahmoud Naghibzadeh, Ahad Harati. Fuzzy discrete correlation for document clustering 59-65. [Crossref]

1004. Syed Abdullah Fadzli, Rossitza Setchi. Ontology-based indexing of annotated images using semantic DNA and vector space model 40-47. [Crossref]

1005. Jihn-Chang Jehng, Shihchieh Chou, Chin-Yi Cheng, Jia-Sheng Heh. An Evaluation of the Formal Concept Analysis-Based Document Vector on Document Clustering 207-210. [Crossref]

1006. N. F. Nabila, A. Mamat, M. A. Azmi-Murad, N. Mustapha. Enriching non-taxonomic relations extracted from domain texts 99-105. [Crossref]

1007. Andrea De Lucia, Massimiliano Di Penta, Rocco Oliveto, Annibale Panichella, Sebastiano Panichella. Improving IR-based Traceability Recovery Using Smoothing Filters 21-30. [Crossref]

1008. Simone Paolo Ponzetto, Michael Strube. 2011. Taxonomy induction based on a collaboratively built knowledge repository. *Artificial Intelligence* **175**:9-10, 1737-1756. [Crossref]

1009. Bogdan Dit, Latifa Guerrouj, Denys Poshyvanyk, Giuliano Antoniol. Can Better Identifier Splitting Techniques Help Feature Location? 11-20. [Crossref]

1010. Alaa Abi-Haidar, Luis M. Rocha. 2011. Collective classification of textual documents by guided self-organization in T-Cell cross-regulation dynamics. *Evolutionary Intelligence* **4**:2, 69-80. [Crossref]

1011. Derek Hao Hu, Vincent Wenchen Zheng, Qiang Yang. 2011. Cross-domain activity recognition via transfer learning. *Pervasive and Mobile Computing* **7**:3, 344-358. [Crossref]

1012. Somayye Estahbanati, Reza Javidan. A new stemmer for Farsi language 25-29. [Crossref]

1013. Giuseppe Scanniello, Andrian Marcus. Clustering Support for Static Concept Location in Source Code 1-10. [Crossref]

1014. Yingcai Wu, Thomas Provan, Furu Wei, Shixia Liu, Kwan-Liu Ma. 2011. Semantic-Preserving Word Clouds by Seam Carving. *Computer Graphics Forum* **30**:3, 741-750. [Crossref]

1015. Guillaume Cabanac. 2011. Accuracy of inter-researcher similarity measures based on topical and social clues. *Scientometrics* **87**:3, 597-620. [Crossref]

1016. Georgios Mamakis, Athanasios G. Malamos, J. Andrew Ware. 2011. An alternative approach for statistical single-label document classification of newspaper articles. *Journal of Information Science* **37**:3, 293-303. [Crossref]

1017. Noah A. Smith. 2011. Linguistic Structure Prediction. *Synthesis Lectures on Human Language Technologies* **4**:2, 1-274. [Crossref]

1018. Xiaojun Quan, Liu Wenyin, Bite Qiu. 2011. Term Weighting Schemes for Question Categorization. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **33**:5, 1009-1021. [Crossref]

1019. Lei Shi, Xinming Ma, Lei Xi, Qiguo Duan, Jingying Zhao. 2011. Rough set and ensemble learning based semi-supervised algorithm for text classification. *Expert Systems with Applications* **38**:5, 6300-6306. [Crossref]

1020. M. Arun Kumar, M. Gopal. 2011. Reduced one-against-all method for multiclass SVM classification. *Expert Systems with Applications* . [Crossref]

1021. Paul-Armand Verhaegen, Joris D'hondt, Dennis Vandevenne, Simon Dewulf, Joost R. Duflou. 2011. Identifying candidates for design-by-analogy. *Computers in Industry* **62**:4, 446-459. [Crossref]

1022. Amrita Pati, Birte Abt, Hazuki Teshima, Matt Nolan, Alla Lapidus, Susan Lucas, Nancy Hammon, Shweta Deshpande, Jan-Fang Cheng, Roxane Tapia, Cliff Han, Lynne Goodwin, Sam Pitluck, Konstantinos Liolios, Ioanna Pagani, Konstantinos Mavromatis, Galina Ovchinikova, Amy Chen, Krishna Palaniappan, Miriam Land, Loren Hauser, Cynthia D. Jeffries, John C. Detter, Evelyne-Marie Brambilla, K. Palani Kannan, Manfred Rohde, Stefan Spring, Markus Göker, Tanja Woyke, James Bristow, Jonathan A. Eisen, Victor Markowitz, Philip Hugenholtz, Nikos C. Kyrpides, Hans-Peter Klenk, Natalia Ivanova. 2011. Complete genome sequence of Cellulophaga lytica type strain (LIM-21T). *Standards in Genomic Sciences* **4**:2, 221-232. [Crossref]

1023. Ioanna Pagani, Olga Chertkov, Alla Lapidus, Susan Lucas, Tijana Glavina Del Rio, Hope Tice, Alex Copeland, Jan-Fang Cheng, Matt Nolan, Elizabeth Saunders, Sam Pitluck, Brittany Held, Lynne Goodwin, Konstantinos Liolios, Galina Ovchinikova, Natalia Ivanova, Konstantinos Mavromatis, Amrita Pati, Amy Chen, Krishna Palaniappan, Miriam Land, Loren Hauser, Cynthia D. Jeffries, John C. Detter, Cliff Han, Roxanne Tapia, Olivier D. Ngatchou-Djao, Manfred Rohde, Markus Göker, Stefan Spring, Johannes Sikorski, Tanja Woyke, Jim Bristow, Jonathan A. Eisen, Victor Markowitz, Philip Hugenholtz, Hans-Peter Klenk, Nikos C. Kyrpides. 2011. Complete genome sequence of Marivirga tractuosa type strain (H-43T). *Standards in Genomic Sciences* **4**:2, 154-162. [Crossref]

1024. Konstantinos Mavromatis, Megan Lu, Monica Misra, Alla Lapidus, Matt Nolan, Susan Lucas, Nancy Hammon, Shweta Deshpande, Jan-Fang Cheng, Roxane Tapia, Cliff Han, Lynne Goodwin, Sam Pitluck, Konstantinos Liolios, Ioanna Pagani, Natalia Ivanova, Natalia Mikhailova, Amrita Pati, Amy Chen, Krishna Palaniappan, Miriam Land, Loren Hauser, Cynthia D. Jeffries, John C. Detter, Evelyne-Marie Brambilla, Manfred Rohde, Markus Göker, Sabine Gronow, Tanja Woyke, James Bristow, Jonathan A. Eisen, Victor Markowitz, Philip Hugenholtz, Hans-Peter Klenk, Nikos C. Kyrpides. 2011. Complete genome sequence of Riemerella anatipestifer type strain (ATCC 11845T). *Standards in Genomic Sciences* **4**:2, 144-153. [Crossref]

1025. Reinhard Wirth, Olga Chertkov, Brittany Held, Alla Lapidus, Matt Nolan, Susan Lucas, Nancy Hammon, Shweta Deshpande, Jan-Fang Cheng, Roxane Tapia, Cliff Han, Lynne Goodwin, Sam Pitluck, Konstantinos Liolios, Pagani Ioanna, Natalia Ivanova, Konstantinos Mavromatis, Natalia Mikhailova, Amrita Pati, Amy Chen, Krishna Palaniappan, Miriam Land, Loren Hauser, Yun-Juan Chang, Cynthia D. Jeffries, Yvonne Bilek, Thomas Hader, Manfred Rohde, Stefan Spring, Johannes Sikorski, Markus Göker, Tanja Woyke, James Bristow, Jonathan A. Eisen, Victor Markowitz, Philip Hugenholtz, Nikos C. Kyrpides, Hans-Peter Klenk. 2011. Complete genome sequence

of Desulfurococcus mucosus type strain (O7/1T). *Standards in Genomic Sciences* **4**:2, 173-182. [Crossref]

1026. M.A. Yunus, R. Zainuddin, N. Abdullah. Semantic speech query via stemmer for Quran documents results 17-21. [Crossref]

1027. Vera Hollink, Theodora Tsikrika, Arjen P. de Vries. 2011. Semantic search log analysis: A method and a study on professional image search. *Journal of the American Society for Information Science and Technology* **62**:4, 691-713. [Crossref]

1028. Nadia Lachetar, Halima Bahi. Application of an ant colony algorithm for text indexing 1-6. [Crossref]

1029. Selma Ayşe Özel. 2011. A Web page classification system based on a genetic algorithm using tagged-terms as features. *Expert Systems with Applications* **38**:4, 3407-3415. [Crossref]

1030. Remco Dijkman, Marlon Dumas, Boudewijn van Dongen, Reina Käärik, Jan Mendling. 2011. Similarity of business process models: Metrics and evaluation. *Information Systems* **36**:2, 498-516. [Crossref]

1031. Yao-Tsung Chen, Meng Chang Chen. 2011. Using chi-square statistics to measure similarities for text categorization. *Expert Systems with Applications* **38**:4, 3085-3090. [Crossref]

1032. Sahar Changuel, Nicolas Labroche. Distinguishing defined concepts from prerequisite concepts in learning resources 22-29. [Crossref]

1033. T J Jankun-Kelly, David Wilson, Andrew S Stamps, Josh Franck, Jeffery Carver, J Edward Swan. 2011. Visual Analysis for Textual Relationships in Digital Forensic Evidence. *Information Visualization* **10**:2, 134-144. [Crossref]

1034. Jesús Vilares, Manuel Vilares, Juan Otero. 2011. Managing misspelled queries in IR applications. *Information Processing & Management* **47**:2, 263-286. [Crossref]

1035. Rung-Ching Chen, Cho-Tscan Bau, Chun-Ju Yeh. 2011. Merging domain ontologies based on the WordNet system and Fuzzy Formal Concept Analysis techniques. *Applied Soft Computing* **11**:2, 1908-1923. [Crossref]

1036. Argyris Kalogeratos, Aristidis Likas. 2011. Document clustering using synthetic cluster prototypes. *Data & Knowledge Engineering* **70**:3, 284-306. [Crossref]

1037. Ahmed Lamkanfi, Serge Demeyer, Quinten David Soetens, Tim Verdonck. Comparing Mining Algorithms for Predicting the Severity of a Reported Bug 249-258. [Crossref]

1038. M.C. Naldi, R.J.G.B. Campello, E.R. Hruschka, A.C.P.L.F. Carvalho. 2011. Efficiency issues of evolutionary k-means. *Applied Soft Computing* **11**:2, 1938-1952. [Crossref]

1039. Andrea De Lucia, Massimiliano Di Penta, Rocco Oliveto. 2011. Improving Source Code Lexicon via Traceability and Information Retrieval. *IEEE Transactions on Software Engineering* **37**:2, 205-227. [Crossref]

1040. Teng-Kai Fan, Chia-Hui Chang. 2011. Blogger-Centric Contextual Advertising. *Expert Systems with Applications* **38**:3, 1777-1788. [Crossref]

1041. Sven S. Groth, Jan Muntermann. 2011. An intraday market risk management approach based on textual analysis. *Decision Support Systems* **50**:4, 680-691. [Crossref]

1042. Johanna Strandell, Ola Caster, Andrew Bate, Niklas Norén, I. Ralph Edwards. 2011. Reporting Patterns Indicative of Adverse Drug Interactions. *Drug Safety* **34**:3, 253-266. [Crossref]

1043. Natalia Ivanova, Christine Rohde, Christine Munk, Matt Nolan, Susan Lucas, Tijana Glavina Del Rio, Hope Tice, Shweta Deshpande, Jan-Fang Cheng, Roxane Tapia, Cliff Han, Lynne Goodwin, Sam Pitluck, Konstantinos Liolios, Konstantinos Mavromatis, Natalia Mikhailova, Amrita Pati, Amy

Chen, Krishna Palaniappan, Miriam Land, Loren Hauser, Yun-Juan Chang, Cynthia D. Jeffries, Evelyne Brambilla, Manfred Rohde, Markus Göker, Brian J. Tindall, Tanja Woyke, James Bristow, Jonathan A. Eisen, Victor Markowitz, Philip Hugenholtz, Nikos C. Kyrpides, Hans-Peter Klenk, Alla Lapidus. 2011. Complete genome sequence of Truepera radiovictrix type strain (RQ-24T). *Standards in Genomic Sciences* **4**:1, 91-99. [Crossref]

1044. Birte Abt, Megan Lu, Monica Misra, Cliff Han, Matt Nolan, Susan Lucas, Nancy Hammon, Shweta Deshpande, Jan-Fang Cheng, Roxane Tapia, Lynne Goodwin, Sam Pitluck, Konstantinos Liolios, Ioanna Pagani, Natalia Ivanova, Konstantinos Mavromatis, Galina Ovchinikova, Amrita Pati, Amy Chen, Krishna Palaniappan, Miriam Land, Loren Hauser, Yun-Juan Chang, Cynthia D. Jeffries, John C. Detter, Evelyne Brambilla, Manfred Rohde, Brian J. Tindall, Markus Göker, Tanja Woyke, James Bristow, Jonathan A. Eisen, Victor Markowitz, Philip Hugenholtz, Nikos C. Kyrpides, Hans-Peter Klenk, Alla Lapidus. 2011. Complete genome sequence of Cellulophaga algicola type strain (IC166T). *Standards in Genomic Sciences* **4**:1, 72-80. [Crossref]

1045. Markus Göker, David Cleland, Elizabeth Saunders, Alla Lapidus, Matt Nolan, Susan Lucas, Nancy Hammon, Shweta Deshpande, Jan-Fang Cheng, Roxane Tapia, Cliff Han, Lynne Goodwin, Sam Pitluck, Konstantinos Liolios, Ioanna Pagani, Natalia Ivanova, Konstantinos Mavromatis, Amrita Pati, Amy Chen, Krishna Palaniappan, Miriam Land, Loren Hauser, Yun-Juan Chang, Cynthia D. Jeffries, John C. Detter, Brian Beck, Tanja Woyke, James Bristow, Jonathan A. Eisen, Victor Markowitz, Philip Hugenholtz, Nikos C. Kyrpides, Hans-Peter Klenk. 2011. Complete genome sequence of Isosphaera pallida type strain (IS1BT). *Standards in Genomic Sciences* **4**:1, 63-71. [Crossref]

1046. Sabine Gronow, Christine Munk, Alla Lapidus, Matt Nolan, Susan Lucas, Nancy Hammon, Shweta Deshpande, Jan-Fang Cheng, Roxane Tapia, Cliff Han, Lynne Goodwin, Sam Pitluck, Konstantinos Liolios, Natalia Ivanova, Konstantinos Mavromatis, Natalia Mikhailova, Amrita Pati, Amy Chen, Krishna Palaniappan, Miriam Land, Loren Hauser, Yun-Juan Chang, Cynthia D. Jeffries, Evelyne Brambilla, Manfred Rohde, Markus Göker, John C. Detter, Tanja Woyke, James Bristow, Jonathan A. Eisen, Victor Markowitz, Philip Hugenholtz, Nikos C. Kyrpides, Hans-Peter Klenk. 2011. Complete genome sequence of Paludibacter propionicigenes type strain (WB4T). *Standards in Genomic Sciences* **4**:1, 36-44. [Crossref]

1047. Ioanna Pagani, Alla Lapidus, Matt Nolan, Susan Lucas, Nancy Hammon, Shweta Deshpande, Jan-Fang Cheng, Olga Chertkov, Karen Davenport, Roxane Tapia, Cliff Han, Lynne Goodwin, Sam Pitluck, Konstantinos Liolios, Konstantinos Mavromatis, Natalia Ivanova, Natalia Mikhailova, Amrita Pati, Amy Chen, Krishna Palaniappan, Miriam Land, Loren Hauser, Yun-Juan Chang, Cynthia D. Jeffries, John C. Detter, Evelyne Brambilla, K. Palani Kannan, Olivier D. Ngatchou Djao, Manfred Rohde, Rüdiger Pukall, Stefan Spring, Markus Göker, Johannes Sikorski, Tanja Woyke, James Bristow, Jonathan A. Eisen, Victor Markowitz, Philip Hugenholtz, Nikos C. Kyrpides, Hans-Peter Klenk. 2011. Complete genome sequence of Desulfobulbus propionicus type strain (1pr3T). *Standards in Genomic Sciences* **4**:1, 100-110. [Crossref]

1048. Mike Thelwall, Kevan Buckley, Georgios Paltoglou. 2011. Sentiment in Twitter events. *Journal of the American Society for Information Science and Technology* **62**:2, 406-418. [Crossref]

1049. Chitrita Chaudhuri, Atal Chaudhuri. Detection of Verbatim or Partial Duplication from Multiple Source Documents Using Data Mining Techniques and Case-Based Reasoning Methodologies 129-132. [Crossref]

1050. SEBASTIÁN A. RÍOS, JUAN D. VELÁSQUEZ. 2011. FINDING REPRESENTATIVE WEB PAGES BASED ON A SOM AND A REVERSE CLUSTER ANALYSIS. *International Journal on Artificial Intelligence Tools* **20**:01, 93-118. [Crossref]

1051. Sheng-Hui Dai, Gang-Yong Lin, Hua-Qing Zhou. 2011. A Novel Method for Speech Data Mining. *Journal of Software* **6**:1. . [Crossref]

1052. Vaclav Beĺak, Marcel Karnstedt, Conor Hayes. 2011. Life-Cycles and Mutual E_ects of Scientific Communities. *Procedia - Social and Behavioral Sciences* **22**, 37-48. [Crossref]

1053. Daqing He, Dan Wu. 2011. Enhancing query translation with relevance feedback in translingual information retrieval. *Information Processing & Management* **47**:1, 1-17. [Crossref]

1054. Somayyeh Estahbanati, Reza Javidan, Mehdi Nikkhah. 2011. A New Multi-Phase Algorithm for Stemming in Farsi Language Based on Morphology. *International Journal of Computer Theory and Engineering* 623-627. [Crossref]

1055. Kwan Yi. 2011. An empirical study on the automatic resolution of semantic ambiguity in social tags. *Proceedings of the American Society for Information Science and Technology* **48**:1, 1-10. [Crossref]

1056. R. Subhashini, V.Jawahar Senthil Kumar. 2011. A Roadmap to Integrate Document Clustering in Information Retrieval. *International Journal of Information Retrieval Research* **1**:1, 31-44. [Crossref]

1057. Jörg Waitelonis, Nadine Ludwig, Harald Sack. Use What You Have: Yovisto Video Search Engine Takes a Semantic Turn 173-185. [Crossref]

1058. Danesh Irani, Marco Balduzzi, Davide Balzarotti, Engin Kirda, Calton Pu. Reverse Social Engineering Attacks in Online Social Networks 55-74. [Crossref]

1059. Stefanos Vrochidis, Ioannis Kompatsiaris, Ioannis Patras. 2011. Utilizing Implicit User Feedback to Improve Interactive Video Retrieval. *Advances in Multimedia* **2011**, 1-18. [Crossref]

1060. Mari-Sanna Paukkeri, Ilkka Kivimäki, Santosh Tirunagari, Erkki Oja, Timo Honkela. Effect of Dimensionality Reduction on Different Distance Measures in Document Clustering 167-176. [Crossref]

1061. Gareth J. F. Jones, Marguerite Fuller, Eamonn Newman, Ying Zhang. Hybrid and Interactive Domain-Specific Translation for Multilingual Access to Digital Libraries 72-91. [Crossref]

1062. Björn Schuller, Tobias Knaup. Learning and Knowledge-Based Sentiment Analysis in Movie Review Key Excerpts 448-472. [Crossref]

1063. Raymond Y. K. Lau, Wenping Zhang, Yunqing Xia, Dawei Song. Multi-facets Quality Assessment of Online Opinionated Expressions 212-225. [Crossref]

1064. Catherine Blake. 2011. Text mining. *Annual Review of Information Science and Technology* **45**:1, 121-155. [Crossref]

1065. Sujeevan Aseervatham, Anestis Antoniadis, Eric Gaussier, Michel Burlet, Yves Denneulin. 2011. A sparse version of the ridge logistic regression for large-scale text categorization. *Pattern Recognition Letters* **32**:2, 101-106. [Crossref]

1066. Sa-kwang Song, Heung-seon Oh, Sung Hyon Myaeng, Sung-pil Choi, Hong-woo Chun, Yun-soo Choi, Chang-hoo Jeong. Procedural Knowledge Extraction on MEDLINE Abstracts 345-354. [Crossref]

1067. Anil Turel, Fazli Can. A New Approach to Search Result Clustering and Labeling 283-292. [Crossref]

1068. Mohamed Salah Hamdi. 2011. SOMSE: A semantic map based meta-search engine for the purpose of web information customization. *Applied Soft Computing* **11**:1, 1310-1321. [Crossref]

1069. Fatih Oguz, Wallace C. Koehler. 2011. Document constancy and persistence: A study of Web pages in library and information science domain. *Proceedings of the American Society for Information Science and Technology* **48**:1, 1-9. [Crossref]

1070. Chidansh Amitkumar Bhatt, Mohan S. Kankanhalli. 2011. Multimedia data mining: state of the art and challenges. *Multimedia Tools and Applications* **51**:1, 35-76. [Crossref]

1071. Nikoletta Bassiou, Constantine Kotropoulos. 2011. Long distance bigram models applied to word clustering. *Pattern Recognition* **44**:1, 145-158. [Crossref]

1072. Uta Christoph, Daniel Götten, Karl-Heinz Krempels, Christoph Terwelp. Efficient Literature Research Based on Semantic Tagnets: Implemented and Evaluated for a German Text-Corpus 162-172. [Crossref]

1073. Mohd Amin MohdYunus, Roziati Zainuddin, Noorhidawati Abdullah. A Framework of Hybrid Semantic Speech Query via Stemmer for Quran Documents Results 91-101. [Crossref]

1074. S Das, P Mitra. A rule-based approach of stemming for inflectional and derivational words in Bengali 134-136. [Crossref]

1075. Simone Teufel, Min-Yen Kan. Robust Argumentative Zoning for Sensemaking in Scholarly Documents 154-170. [Crossref]

1076. Benoît Sagot, Géraldine Walther. Non-canonical Inflection: Data, Formalisation and Complexity Measures 23-45. [Crossref]

1077. Gareth J. F. Jones, Daragh Byrne, Mark Hughes, Noel E. O'Connor, Andrew Salway. Automated Annotation of Landmark Images Using Community Contributed Datasets and Web Resources 111-126. [Crossref]

1078. Elena Hensinger, Ilias Flaounas, Nello Cristianini. Learning Readers' News Preferences with Support Vector Machines 322-331. [Crossref]

1079. Soumaya Medini, Philippe Galinier, Massimiliano Di Penta, Yann-Gaël Guéhéneuc, Giuliano Antoniol. A Fast Algorithm to Locate Concepts in Execution Traces 252-266. [Crossref]

1080. Martin Gropp, Elmar Nöth, Korbinian Riedhammer. A Novel Lecture Browsing System Using Ranked Key Phrases and StreamGraphs 17-24. [Crossref]

1081. Zbigniew Suraj, Piotr Grochowalski. RoSetOn: The Open Project for Ontology of Rough Sets and Related Fields 414-419. [Crossref]

1082. Günther Schatter, Andreas Eiselt. An Enhanced Concept of a Digital Radio Incorporating a Multimodal Interface and Searchable Spoken Content 291-304. [Crossref]

1083. Paul-Armand Verhaegen, Jef Peeters, Dennis Vandevenne, Simon Dewulf, Joost R. Duflou. 2011. Effectiveness of the PAnDA ideation tool. *Procedia Engineering* **9**, 63-76. [Crossref]

1084. Naoaki Okazaki, Jun'ichi Tsujii. 2011. A Simple and Fast Algorithm for Approximate String Matching with Set Similarity. *Journal of Natural Language Processing* **18**:2, 89-117. [Crossref]

1085. Vundavalli Srinivasarao, Prasad Pingali, Vasudeva Varma. Effective Term Weighting in ALT Text Prediction for Web Image Retrieval 237-244. [Crossref]

1086. Yanping Lu, Shengrui Wang, Shaozi Li, Changle Zhou. 2011. Particle swarm optimizer for variable weighting in clustering high-dimensional data. *Machine Learning* **82**:1, 43-70. [Crossref]

1087. Shiliang Sun. Multi-view Laplacian Support Vector Machines 209-222. [Crossref]

1088. Mohammad Sadegh Rasooli, Omid Kashefi, Behrouz Minaei-Bidgoli. Extracting Parallel Paragraphs and Sentences from English-Persian Translated Documents 574-583. [Crossref]

1089. Liang Chen, Liukai Hu, Zibin Zheng, Jian Wu, Jianwei Yin, Ying Li, Shuiguang Deng. WTCluster: Utilizing Tags for Web Services Clustering 204-218. [Crossref]

1090. Ujwala Bharambe, Archana Kale. Landscape of Web Search Results Clustering Algorithms 95-107. [Crossref]

1091. Debasis Ganguly, Johannes Leveling, Gareth J. F. Jones. Query Expansion for Language Modeling Using Sentence Similarities 62-77. [Crossref]

1092. Peyman Nasirifard, Conor Hayes. Tadvise: A Twitter Assistant Based on Twitter Lists 153-160. [Crossref]

1093. Antonio Fernández, Josval Díaz, Yoan Gutiérrez, Rafael Muñoz. An Unsupervised Method to Improve Spanish Stemmer 221-224. [Crossref]

1094. Xiang-Fei Jia, David Alexander, Vaughn Wood, Andrew Trotman. University of Otago at INEX 2010 250-268. [Crossref]

1095. Georgina Ramírez. UPF at INEX 2010: Towards Query-Type Based Focused Retrieval 206-218. [Crossref]

1096. Roque Enrique López, Dennis Barreda, Javier Tejada, Ernesto Cuadros. MFSRank: An Unsupervised Method to Extract Keyphrases Using Semantic Information 338-344. [Crossref]

1097. Marco A. Palomino, Tom Wuytack. Unsupervised Extraction of Keywords from News Archives 544-555. [Crossref]

1098. J.R. Duflou, P.-A. Verhaegen. 2011. Systematic innovation through patent based product aspect analysis. *CIRP Annals* **60**:1, 203-206. [Crossref]

1099. Yong-Hun Lee, Sang-Bum Lee. 2010. A Research on Enhancement of Text Categorization Performance by using Okapi BM25 Word Weight Method. *Journal of the Korea Academia-Industrial cooperation Society* **11**:12, 5089-5096. [Crossref]

1100. Robert M. Patton, Barbara G. Beckerman, Thomas E. Potok. Analysis and Classification of Mammography Reports Using Maximum Variation Sampling 112-131. [Crossref]

1101. Ea-Ee Jan, Brian Kingsbury. Rapid and inexpensive development of speech action classifiers for natural language call routing systems 348-353. [Crossref]

1102. Christian Senger, Jens Kaltschmidt, Simon P.W. Schmitt, Markus G. Pruszydlo, Walter E. Haefeli. 2010. Misspellings in drug information system queries: Characteristics of drug name spelling errors and strategies for their prevention. *International Journal of Medical Informatics* **79**:12, 832-839. [Crossref]

1103. Byeong-Chul Kang, Zee-Won Sur, Chulhwan Park, Man-gi Cho. 2010. Document clustering of MEDLINE abstracts based on non-negative matrix factorization using local confidence assessment. *BioChip Journal* **4**:4, 336-349. [Crossref]

1104. M A Yunus, R Zainuddin, N Abdullah. Visualizing Quran documents results by stemming semantic speech query 209-213. [Crossref]

1105. E.K.F. Dang, R.W.P. Luk, J. Allan, K.S. Ho, S.C.F. Chan, K.F.L. Chung, D.L. Lee. 2010. A new context-dependent term weight computed by boost and discount using relevance information. *Journal of the American Society for Information Science and Technology* **61**:12, 2514-2530. [Crossref]

1106. Dieter Pfoser, Alexandros Efentakis, Thanasis Hadzilacos, Sophia Karagiorgou, Giorgos Vasiliou. 2010. Language-Independent Access to History Textbooks: Utilizing GIS Functionality with Spatiotemporal and Thematic Metadata. *Transactions in GIS* **14**:6, 773-790. [Crossref]

1107. Besiki Stvilia, Corinne Jörgensen. 2010. Member activities and quality of tags in a collection of historical photographs in Flickr. *Journal of the American Society for Information Science and Technology* **61**:12, 2477-2489. [Crossref]

1108. P. Warintarawej, A. Laurent, P. Pompidor, B. Laurent. Classification of brand names based on n-grams 12-17. [Crossref]

1109. Guillaume Cabanac, Max Chevalier, Claude Chrisment, Christine Julien. 2010. Organization of digital resources as an original facet for exploring the quiescent information capital of a community. *International Journal on Digital Libraries* **11**:4, 239-261. [Crossref]

1110. Elaheh Rahimtoroghi, Hesham Faili, Azadeh Shakery. A structural rule-based stemmer for Persian 574-578. [Crossref]

1111. D Jayasinghe, S Hettiarachchi, S Abeywickrama, C Ketteepearachchi, D Alahakoon, S Matharage, U Gunasinghe. txtKnot 239-243. [Crossref]

1112. Shajith Ikbal, Karthik Visweswariah. Call transcript segmentation using word cooccurrence model 372-377. [Crossref]

1113. M A Yunus, R Zainuddin, N Abdullah. Semantic query with stemmer for Quran documents results 40-44. [Crossref]

1114. Franco Nazareno, Kyung-Hee Lee, Wan-Sup Cho. Mining molecular interactions from scientific literature using cloud computing 864-865. [Crossref]

1115. Deborah De Felice, Giovanni Giuffrida, Giuseppe Giura, Calogero Zarba. 2010. La descrizione dei reati di criminalità organizzata e di tipo mafioso nel testo delle sentenze. *Quaderni di Sociologia* :54, 57-80. [Crossref]

1116. Jahiruddin, Muhammad Abulaish, Lipika Dey. 2010. A concept-driven biomedical knowledge extraction and visualization framework for conceptualization of text corpora. *Journal of Biomedical Informatics* **43**:6, 1020-1035. [Crossref]

1117. Nikolaos Nanas, Manolis Vavalis, Anne De Roeck. 2010. Words, antibodies and their interactions. *Swarm Intelligence* **4**:4, 275-300. [Crossref]

1118. FENG LI. 2010. The Information Content of Forward-Looking Statements in Corporate Filings-A Naïve Bayesian Machine Learning Approach. *Journal of Accounting Research* **48**:5, 1049-1102. [Crossref]

1119. Viorel Milea, Nurfadhlina Mohd Sharef, Rui Jorge Almeida, Uzay Kaymak, Flavius Frasincar. Prediction of the MSCI EURO index based on fuzzy grammar fragments extracted from European Central Bank statements 231-236. [Crossref]

1120. Abdulhussain E. Mahdi, Arash Joorabchi. 2010. A citation-based approach to automatic topical indexing of scientific literature. *Journal of Information Science* **36**:6, 798-811. [Crossref]

1121. Maxim Mozgovoy, Tuomo Kakkonen, Georgina Cosma. 2010. Automatic Student Plagiarism Detection: Future Perspectives. *Journal of Educational Computing Research* **43**:4, 511-531. [Crossref]

1122. Junjie Wu, Hua Yuan, Hui Xiong, Guoqing Chen. 2010. Validation of overlapping clustering: A random clustering perspective. *Information Sciences* **180**:22, 4353-4369. [Crossref]

1123. Hao Wu, Jun He, Yijian Pei. 2010. Scientific impact at the topic level: A case study in computational linguistics. *Journal of the American Society for Information Science and Technology* **61**:11, 2274-2287. [Crossref]

1124. Yuan Ping, Ya-jian Zhou, Yi-xian Yang, Wei-ping Peng. A novel term weighting scheme with distributional coefficient for text categorization with support vector machine 182-185. [Crossref]

1125. Salha Alzahrani, Naomie Salim, Chow Kok Kent, Mohammed Salem Binwahlan, Ladda Suanmali. The development of cross-language plagiarism detection tool utilising fuzzy swarm-based summarisation 86-90. [Crossref]

1126. Irene Diaz-Valenzuela, Maria J. Martin-Bautista, M. Amparo Vila. An automatic data mining authority control system: A first approach 569-574. [Crossref]

1127. Krzysztof Pancerz, Piotr Grochowalski. Matching ontological subgraphs to concepts: A preliminary rough set approach 1394-1399. [Crossref]

1128. Stephan Neuhaus, Thomas Zimmermann. Security Trend Analysis with CVE Topic Models 111-120. [Crossref]

1129. Yaw-Huei Chen, Sin-Sian Li, Yu-Ta Chen. Extracting Topics Information from Conference Web Pages Using Page Segmentation and SVM 270-277. [Crossref]

1130. Te-Min Chang, Wen-Feng Hsiao, Cheng-Wei Wu. Incremental learning of aspect model on streaming documents 360-365. [Crossref]

1131. Jun-Ming Chen, Meng-Chang Chen, Yeali S. Sun. 2010. A novel approach for enhancing student reading comprehension and assisting teacher assessment of literacy. *Computers & Education* **55**:3, 1367-1382. [Crossref]

1132. Ljiljana Dolamic, Jacques Savoy. 2010. Retrieval effectiveness of machine translated queries. *Journal of the American Society for Information Science and Technology* **61**:11, 2266-2273. [Crossref]

1133. Lawrence D. Fu, Constantin F. Aliferis. 2010. Using content-based and bibliometric features for machine learning models to predict citation counts in the biomedical literature. *Scientometrics* **85**:1, 257-270. [Crossref]

1134. A Arslan, O Yilmazel. Quality benchmarking relational databases and Lucene in the TREC4 adhoc task environment 365-372. [Crossref]

1135. Farkhund Iqbal, Hamad Binsalleeh, Benjamin C.M. Fung, Mourad Debbabi. 2010. Mining writeprints from anonymous e-mails for forensic investigation. *Digital Investigation* **7**:1-2, 56-64. [Crossref]

1136. Liping Jing, Michael K. Ng, Joshua Z. Huang. 2010. Knowledge-based vector space model for text clustering. *Knowledge and Information Systems* **25**:1, 35-55. [Crossref]

1137. Rekha Baghel, Renu Dhir. Text document clustering based on frequent concepts 366-371. [Crossref]

1138. P Maciolek, G Dobrowolski. Is shallow semantic analysis really that shallow? A study on improving text classification performance 455-460. [Crossref]

1139. Fabricio J. Barth. Using Social Network Analysis and Hierarchical Clustering to Identify Groups in Wikis 8-11. [Crossref]

1140. Shady Shehata, Fakhri Karray, Mohamed Kamel. 2010. An Efficient Concept-Based Mining Model for Enhancing Text Clustering. *IEEE Transactions on Knowledge and Data Engineering* **22**:10, 1360-1371. [Crossref]

1141. Shiliang Sun, David R. Hardoon. 2010. Active learning with extremely sparse labeled examples. *Neurocomputing* **73**:16-18, 2980-2988. [Crossref]

1142. Dirk Thorleuchter, Dirk Van den Poel, Anita Prinzie. 2010. Mining ideas from textual information. *Expert Systems with Applications* **37**:10, 7182-7188. [Crossref]

1143. Bo Jin, Xinghua Lu. 2010. Identifying informative subsets of the Gene Ontology with information bottleneck methods. *Bioinformatics* **26**:19, 2445-2451. [Crossref]

1144. S. Van Landeghem, T. Abeel, Y. Saeys, Y. Van de Peer. 2010. Discriminative and informative features for biomolecular text mining with ensemble feature selection. *Bioinformatics* **26**:18, i554-i560. [Crossref]

1145. JungWon Yoon, Brian O'Connor. 2010. Engineering an image-browsing environment: re-purposing existing denotative descriptors. *Journal of Documentation* **66**:5, 750-774. [Abstract] [Full Text] [PDF]

1146. Hakim Sultanov, Jane Huffman Hayes. Application of Swarm Techniques to Requirements Engineering: Requirements Tracing 211-220. [Crossref]

1147. Ricardo G. Cota, Anderson A. Ferreira, Cristiano Nascimento, Marcos André Gonçalves, Alberto H. F. Laender. 2010. An unsupervised heuristic-based hierarchical method for name disambiguation in bibliographic citations. *Journal of the American Society for Information Science and Technology* **61**:9, 1853-1870. [Crossref]

1148. Senthil Karthikeyan Sundaram, Jane Huffman Hayes, Alex Dekhtyar, E. Ashlee Holbrook. 2010. Assessing traceability of software engineering artifacts. *Requirements Engineering* **15**:3, 313-335. [Crossref]

1149. Wendy Wang, Drew Hwang. 2010. Abstraction Assistant: An automatic text abstraction system. *Journal of the American Society for Information Science and Technology* **61**:9, 1790-1799. [Crossref]

1150. J. Slivka, A. Kovacevic, Z. Konjovic. Co-training based algorithm for datasets without the natural feature split 279-284. [Crossref]

1151. M. Kestemont, W. Daelemans, G. De Pauw. 2010. Weigh your words--memory-based lemmatization for Middle Dutch. *Literary and Linguistic Computing* **25**:3, 287-301. [Crossref]

1152. Ritesh Srivastava, M. P. S. Bhatia, Hemant Kr Srivastava, C. P. Sahu. Exploiting grammatical dependencies for fine-grained opinion mining 768-775. [Crossref]

1153. Fatemeh Asadi, Giuliano Antoniol, Yann-Gael Gueheneuc. Concept Location with Genetic Algorithms: A Comparison of Four Distributed Architectures 153-162. [Crossref]

1154. D. Thorleuchter, D. Van den Poel, A. Prinzie. 2010. A compared R&D-based and patent-based cross impact analysis for identifying relationships between technologies. *Technological Forecasting and Social Change* **77**:7, 1037-1050. [Crossref]

1155. Levent Özgür, Tunga Güngör. 2010. Text classification with the support of pruned dependency patterns. *Pattern Recognition Letters* **31**:12, 1598-1607. [Crossref]

1156. Addison Y.S. Su, Stephen J.H. Yang, Wu-Yuin Hwang, Jia Zhang. 2010. A Web 2.0-based collaborative annotation system for enhancing knowledge sharing in collaborative learning environments. *Computers & Education* **55**:2, 752-766. [Crossref]

1157. Barbara Hammer, Alexander Hasenfuss. 2010. Topographic Mapping of Large Dissimilarity Data Sets. *Neural Computation* **22**:9, 2229-2284. [Crossref]

1158. Kashish Sukhija. Zycox: File analyzer on the web V1-548-V1-551. [Crossref]

1159. Liang Feng, Ming-Hui Qiu, Yu-Xuan Wang, Qiao-Liang Xiang, Yin-Fei Yang, Kai Liu. 2010. A fast divisive clustering algorithm using an improved discrete particle swarm optimizer. *Pattern Recognition Letters* **31**:11, 1216-1225. [Crossref]

1160. Vladimír Bartík. Text-Based Web Page Classification with Use of Visual Information 416-420. [Crossref]

1161. Gang Luo. On search guide phrase compilation for recommending home medical products 2167-2171. [Crossref]

1162. Sofus A. Macskassy. Leveraging Contextual Information to Explore Posting and Linking Behaviors of Bloggers 64-71. [Crossref]

1163. Hakan Altınçay, Zafer Erenel. 2010. Analytical evaluation of term weighting schemes for text categorization. *Pattern Recognition Letters* **31**:11, 1310-1323. [Crossref]

1164. Yuen-Hsien Tseng, Chun-Yen Chang, Shu-Nu Chang Rundgren, Carl-Johan Rundgren. 2010. Mining concept maps from news stories for measuring civic scientific literacy in media. *Computers & Education* **55**:1, 165-177. [Crossref]

1165. M. Arun Kumar, M. Gopal. 2010. A comparison study on multiple binary-class SVM methods for unilabel text categorization. *Pattern Recognition Letters* **31**:11, 1437-1444. [Crossref]

1166. R. Subhashini, V. Jawahar Senthil Kumar. Evaluating the Performance of Similarity Measures Used in Document Clustering and Information Retrieval 27-31. [Crossref]

1167. Ken Q Pu, Russell Cheung. Tag grid: Supporting collaborative and fuzzy multidimensional queries of tagged datasets 364-367. [Crossref]

1168. Jagan Sankaranarayanan, Hanan Samet. Images in News 3240-3243. [Crossref]

1169. Bassiou Nikoletta, Kotropoulos Constantine. Word Clustering Using PLSA Enhanced with Long Distance Bigrams 4226-4229. [Crossref]

1170. Wei-Chung Hu, Dung-Feng Yu, Hewijin Christine Jiau. A FAQ Finding Process in Open Source Project Forums 259-264. [Crossref]

1171. Ya Gao, Shiliang Sun. An empirical evaluation of linear and nonlinear kernels for text classification using Support Vector Machines 1502-1505. [Crossref]

1172. Theodoros Giannakopoulos, Aggelos Pikrakis, Sergios Theodoridis. A Multimodal Approach to Violence Detection in Video Sharing Sites 3244-3247. [Crossref]

1173. Shady Shehata, Fakhri Karray, Mohamed S. Kamel. 2010. AN EFFICIENT MODEL FOR ENHANCING TEXT CATEGORIZATION USING SENTENCE SEMANTICS. *Computational Intelligence* **26**:3, 215-231. [Crossref]

1174. Nikitas N. Karanikolas, Christos Skourlas. 2010. A parametric methodology for text classification. *Journal of Information Science* **36**:4, 421-442. [Crossref]

1175. Anna Sidorova, Oyku Isik. 2010. Business process research: a cross-disciplinary review. *Business Process Management Journal* **16**:4, 566-597. [Abstract] [Full Text] [PDF]

1176. Nobel Khandaker, Leen-Kiat Soh. 2010. ClassroomWiki: A Collaborative Wiki for Instructional Use with Multiagent Group Formation. *IEEE Transactions on Learning Technologies* **3**:3, 190-202. [Crossref]

1177. Artemy Kolchinsky, Alaa Abi-Haidar, Jasleen Kaur, Ahmed Abdeen Hamed, Luis M Rocha. 2010. Classification of Protein-Protein Interaction Full-Text Documents Using Text and Citation Network Features. *IEEE/ACM Transactions on Computational Biology and Bioinformatics* **7**:3, 400-411. [Crossref]

1178. Man Lan, Jian Su. 2010. Empirical Investigations into Full-Text Protein Interaction Article Categorization Task (ACT) in the BioCreative II.5 Challenge. *IEEE/ACM Transactions on Computational Biology and Bioinformatics* **7**:3, 421-427. [Crossref]

1179. Natsu Ishii, Asako Koike, Yasunori Yamamoto, Toshihisa Takagi. 2010. Figure classification in biomedical literature to elucidate disease mechanisms, based on pathways. *Artificial Intelligence in Medicine* **49**:3, 135-143. [Crossref]

1180. Xiao-Bo Jin, Cheng-Lin Liu, Xinwen Hou. 2010. Regularized margin-based conditional log-likelihood loss for prototype learning. *Pattern Recognition* **43**:7, 2428-2438. [Crossref]

1181. Eduardo Akira Kinto, Emilio Del-Moral-Hernandez. SVM - Neighbor based candidate working set selection applied on text-categorization 1-7. [Crossref]

1182. Nikolaos Nanas, Stefanos Kodovas, Manolis Vavalis. Revisiting Evolutionary Information Filtering 1-8. [Crossref]

1183. Mohand Boughanem, Ihab Mallak, Henri Prade. A new factor for computing the relevance of a document to a query 1-6. [Crossref]

1184. Khalid Elgazzar, Ahmed E. Hassan, Patrick Martin. Clustering WSDL Documents to Bootstrap the Discovery of Web Services 147-154. [Crossref]

1185. Olle Persson. 2010. Identifying research themes with weighted direct citation links. *Journal of Informetrics* **4**:3, 415-422. [Crossref]

1186. Sebastián Peña Saldarriaga, Christian Viard-Gaudin, Emmanuel Morin. 2010. Impact of online handwriting recognition performance on text categorization. *International Journal on Document Analysis and Recognition (IJDAR)* **13**:2, 159-171. [Crossref]

1187. Haijun Zhang, Tommy W.S. Chow, M.K.M. Rahman. 2010. A novel dual wing harmonium model aided by 2-D wavelet transform subbands for document data mining. *Expert Systems with Applications* **37**:6, 4403-4412. [Crossref]

1188. Ying Liu, Han Tong Loh. Domain concept handling in automated text categorization 1543-1549. [Crossref]

1189. Ola Amayri, Nizar Bouguila. 2010. A study of spam filtering using support vector machines. *Artificial Intelligence Review* **34**:1, 73-108. [Crossref]

1190. Tao Gu, Shaxun Chen, Xianping Tao, Jian Lu. 2010. An unsupervised approach to activity recognition and segmentation based on object-use fingerprints. *Data & Knowledge Engineering* **69**:6, 533-544. [Crossref]

1191. Rocco Oliveto, Malcom Gethers, Denys Poshyvanyk, Andrea De Lucia. On the Equivalence of Information Retrieval Methods for Automated Traceability Link Recovery 68-71. [Crossref]

1192. Vasileios Lampos, Nello Cristianini. Tracking the flu pandemic by monitoring the social web 411-416. [Crossref]

1193. Charles-Edmond Bichot. 2010. Co-clustering Documents and Words by Minimizing the Normalized Cut Objective Function. *Journal of Mathematical Modelling and Algorithms* **9**:2, 131-147. [Crossref]

1194. Stefanos Vrochidis, Anastasia Moumtzidou, Paul King, Anastasios Dimou, Vasileios Mezaris, Ioannis Kompatsiaris. VERGE: A video interactive retrieval engine 1-6. [Crossref]

1195. Teng-Kai Fan, Chia-Hui Chang. 2010. Sentiment-oriented contextual advertising. *Knowledge and Information Systems* **23**:3, 321-344. [Crossref]

1196. Catarina Silva, Uros Lotric, Bernardete Ribeiro, Andrej Dobnikar. 2010. Distributed Text Classification With an Ensemble Kernel-Based Learning Approach. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)* **40**:3, 287-297. [Crossref]

1197. Shikui Wei, Yao Zhao, Zhenfeng Zhu, Nan Liu. 2010. A Cooperative Learning Scheme for Interactive Video Search. *Journal of Signal Processing Systems* **59**:2, 189-199. [Crossref]

1198. Antonina Dattolo, Felice Ferrara, Carlo Tasso. The role of tags for recommendation: A survey 548-555. [Crossref]

1199. Ahmed Lamkanfi, Serge Demeyer, Emanuel Giger, Bart Goethals. Predicting the severity of a reported bug 1-10. [Crossref]

1200. V. Mascardi, A. Locoro, P. Rosso. 2010. Automatic Ontology Matching via Upper Ontologies: A Systematic Evaluation. *IEEE Transactions on Knowledge and Data Engineering* **22**:5, 609-623. [Crossref]

1201. Tzu-An Chiang, Amy J. C. Trappey, Chun-Yi Wu, Charles V. Trappey. Applying BPANN and hierarchical ontology to develop a methodology for binary knowledge document classification and content analysis 263-268. [Crossref]

1202. Chiara Franzoni, Christopher L. Simpkins, Baoli Li, Ashwin Ram. 2010. Using content analysis to investigate the research paths chosen by scientists over time. *Scientometrics* **83**:1, 321-335. [Crossref]

1203. Agustin Casamayor, Daniela Godoy, Marcelo Campo. 2010. Identification of non-functional requirements in textual specifications: A semi-supervised learning approach. *Information and Software Technology* **52**:4, 436-445. [Crossref]

1204. Bader Aljaber, Nicola Stokes, James Bailey, Jian Pei. 2010. Document clustering of scientific texts using citation contexts. *Information Retrieval* **13**:2, 101-131. [Crossref]

1205. Keith N. Hampton. 2010. Internet Use and the Concentration of Disadvantage: Glocalization and the Urban Underclass. *American Behavioral Scientist* **53**:8, 1111-1132. [Crossref]

1206. Fekade Getahun Taddesse, Joe Tekli, Richard Chbeir, Marco Viviani, Kokou Yetongnon. 2010. Semantic-based Merging of RSS Items. *World Wide Web* **13**:1-2, 169-207. [Crossref]

1207. S. Decherchi, S. Ridella, R. Zunino, P. Gastaldo, D. Anguita. 2010. Using Unsupervised Analysis to Constrain Generalization Bounds for Support Vector Classifiers. *IEEE Transactions on Neural Networks* **21**:3, 424-438. [Crossref]

1208. Weiwei Cui, Yingcai Wu, Shixia Liu, Furu Wei, Michelle X. Zhou, Huamin Qu. Context preserving dynamic word cloud visualization 121-128. [Crossref]

1209. Chen-Huei Chou, Atish P. Sinha, Huimin Zhao. 2010. Commercial Internet filters: Perils and opportunities. *Decision Support Systems* **48**:4, 521-530. [Crossref]

1210. Chun-Ling Chen, Frank S.C. Tseng, Tyne Liang. 2010. Mining fuzzy frequent itemsets for hierarchical document clustering. *Information Processing & Management* **46**:2, 193-211. [Crossref]

1211. Angel R. Martinez. 2010. Bigram proximity matrix. *Wiley Interdisciplinary Reviews: Computational Statistics* **2**:2, 251-255. [Crossref]

1212. Wei Song, Soon Cheol Park. 2010. Latent semantic analysis for vector space expansion and fuzzy logic-based genetic clustering. *Knowledge and Information Systems* **22**:3, 347-369. [Crossref]

1213. F Asadi, M Di Penta, G Antoniol, Yann-Gaël Guéhéneuc. A Heuristic-Based Approach to Identify Concepts in Execution Traces 31-40. [Crossref]

1214. Junjie Wu, Hui Xiong, Jian Chen. 2010. COG: local decomposition for rare class analysis. *Data Mining and Knowledge Discovery* **20**:2, 191-220. [Crossref]

1215. Micah Weinberg. 2010. Measuring Governors' Political Orientations Using Words as Data. *State Politics & Policy Quarterly* **10**:1, 96-109. [Crossref]

1216. Zacharias Detorakis, George Tambouratzis. 2010. Applying a sectioned genetic algorithm to word segmentation. *Pattern Analysis and Applications* **13**:1, 93-104. [Crossref]

1217. Alexandros Karatzoglou, Ingo Feinerer. 2010. Kernel-based machine learning for fast text mining in R. *Computational Statistics & Data Analysis* **54**:2, 290-297. [Crossref]

1218. Tom Magerman, Bart Van Looy, Xiaoyan Song. 2010. Exploring the feasibility and accuracy of Latent Semantic Analysis based text mining techniques to detect similarity between patent documents and scientific publications. *Scientometrics* **82**:2, 289-306. [Crossref]

1219. Paulito Palmes, Hung Keng Pung, Tao Gu, Wenwei Xue, Shaxun Chen. 2010. Object relevance weight pattern mining for activity recognition and segmentation. *Pervasive and Mobile Computing* **6**:1, 43-57. [Crossref]

1220. Courtney D. Corley, Diane J. Cook, Armin R. Mikler, Karan P. Singh. 2010. Text and Structural Data Mining of Influenza Mentions in Web and Social Media. *International Journal of Environmental Research and Public Health* **7**:2, 596-615. [Crossref]

1221. Kevin M. Quinn, Burt L. Monroe, Michael Colaresi, Michael H. Crespin, Dragomir R. Radev. 2010. How to Analyze Political Attention with Minimal Assumptions and Costs. *American Journal of Political Science* **54**:1, 209-228. [Crossref]

1222. . References 453-472. [Crossref]

1223. Daniel J. Hopkins, Gary King. 2010. A Method of Automated Nonparametric Content Analysis for Social Science. *American Journal of Political Science* **54**:1, 229-247. [Crossref]

1224. M. Arun Kumar, M. Gopal. An Investigation on Linear SVM and its Variants for Text Categorization 27-31. [Crossref]

1225. Musa Mammadov, John Yearwood, Lei Zhao. A New Supervised Term Ranking Method for Text Categorization 102-111. [Crossref]

1226. D. Thorleuchter, D. Van den Poel, A. Prinzie. Extracting Consumers Needs for New Products - A Web Mining Approach 440-443. [Crossref]

1227. Rizauddin Saian, Ku Ruhana Ku-Mahamud. Comparison of Attribute Selection Methods for Web Texts Categorization 115-118. [Crossref]

1228. Bao H. Do, Andrew Wu, Sandip Biswal, Aya Kamaya, Daniel L. Rubin. 2010. Informatics in Radiology: RADTF: A Semantic Search–enabled, Natural Language Processor–generated Radiology Teaching File. *RadioGraphics* **30**:7, 2039. [Crossref]

1229. Tingyu Liu, Zhonghua Ni, Lei Jiao, Xiaojun Liu. The stem-based vector space model for automatic resource allocation in workflow V1-631-V1-634. [Crossref]

1230. Sven S. Groth, Jan Muntermann. Discovering Intraday Market Risk Exposures in Unstructured Data Sources: The Case of Corporate Disclosures 1-10. [Crossref]

1231. Hiroya Takamura, Manabu Okumura. 2010. Text Summarization Model based on Facility Location Problem. *Transactions of the Japanese Society for Artificial Intelligence* **25**, 174-182. [Crossref]

1232. . References 597-606. [Crossref]

1233. Hazra Imran, Aditi Sharan. Improving Effectiveness of Query Expansion Using Information Theoretic Approach 1-11. [Crossref]

1234. Robert J. Hendley, Barry Wilkins, Russell Beale. 2010. Organix. *International Journal of Creative Interfaces and Computer Graphics* **1**:1, 40-53. [Crossref]

1235. Alex Kohn, François Bry, Alexander Manta. 2010. Semantic Search on Unstructured Data. *International Journal on Semantic Web and Information Systems* **6**:2, 17-35. [Crossref]

1236. Youngik Yang, Donald Gilbert, Sun Kim. 2010. Annotation confidence score for genome annotation: a genome comparison approach. *Bioinformatics* **26**:1, 22-29. [Crossref]

1237. Justin Grimmer. 2010. A Bayesian Hierarchical Topic Model for Political Texts: Measuring Expressed Agendas in Senate Press Releases. *Political Analysis* **18**:01, 1-35. [Crossref]

1238. Geir Solskinnsbakk, Jon Atle Gulla. A Hybrid Approach to Constructing Tag Hierarchies 975-982. [Crossref]

1239. Anestis Gkanogiannis, Theodore Kalamboukis. A Perceptron-Like Linear Supervised Algorithm for Text Classification 86-97. [Crossref]

1240. Ying Zhao. Topic-Constrained Hierarchical Clustering for Document Datasets 181-192. [Crossref]

1241. Dionisis D. Kehagias, Dimitrios Tzovaras, Efthimia Mavridou, Kostantinos Kalogirou, Martin Becker. Implementing an Open Reference Architecture Based on Web Service Mining for the Integration of Distributed Applications and Multi-Agent Systems 162-177. [Crossref]

1242. Ingo Feinerer. Analysis and Algorithms for Stemming Inversion 290-299. [Crossref]

1243. Sergio Martínez, David Sánchez, Aida Valls. Ontology-Based Anonymization of Categorical Values 243-254. [Crossref]

1244. Sonia Ordoñez-Salinas, Alexander Gelbukh. Information Retrieval with a Simplified Conceptual Graph-Like Representation 92-104. [Crossref]

1245. Ljiljana Dolamic, Jacques Savoy. 2009. Indexing and searching strategies for the Russian language. *Journal of the American Society for Information Science and Technology* **60**:12, 2540-2547. [Crossref]

1246. I-Chin Wu, Duen-Ren Liu, Pei-Cheng Chang. 2009. Learning dynamic information needs: A collaborative topic variation inspection approach. *Journal of the American Society for Information Science and Technology* **60**:12, 2430-2451. [Crossref]

1247. A. Sharma, R. Dhir. A Wordsets based document clustering algorithm for large datasets 1-7. [Crossref]

1248. Pu Wang, Carlotta Domeniconi. Towards a Universal Text Classifier: Transfer Learning Using Encyclopedic Knowledge 435-440. [Crossref]

1249. Yi-Ning Tu, Jia-Lang Seng. 2009. Research intelligence involving information retrieval − An example of conferences and journals. *Expert Systems with Applications* **36**:10, 12151-12166. [Crossref]

1250. Feng Yang, Lei Liu. A Flexible Approach for Ontology Matching 1-4. [Crossref]

1251. Tommy W.S. Chow, Haijun Zhang, M.K.M. Rahman. 2009. A new document representation using term frequency and vectorized graph connectionists with application to document retrieval. *Expert Systems with Applications* **36**:10, 12023-12035. [Crossref]

1252. Yanhong Liang, Runhua Tan, Chaoyang Wang, Zhiguang Li. Computer-aided classification of patents oriented to TRIZ 2389-2393. [Crossref]

1253. Xianyang Qu, Hailong Sun, Xiang Li, Xudong Liu, Wei Lin. WSSM: A WordNet-Based Web Services Similarity Mining Mechanism 339-345. [Crossref]

1254. Hua Jiang, Ping Li, Xin Hu, Shuyan Wang. An improved method of term weighting for text classification 294-298. [Crossref]

1255. Yanhua Chen, Lijun Wang, Ming Dong, Jing Hua. 2009. Exemplar-based Visualization of Large Document Corpus (InfoVis2009-1115). *IEEE Transactions on Visualization and Computer Graphics* **15**:6, 1161-1168. [Crossref]

1256. Haijun Zhang, Tommy W.S. Chow, M.K.M. Rahman. 2009. A new dual wing harmonium model for document retrieval. *Pattern Recognition* **42**:11, 2950-2960. [Crossref]

1257. M. Bhavani, K. Thammi, M. Shashi. A rough set based approach to detect plagiarism 1-8. [Crossref]

1258. Allen P. Nikora, Galen Balcom. Automated Identification of LTL Patterns in Natural Language Requirements 185-194. [Crossref]

1259. Biju Issac, Wendy J. Jap. Implementing spam detection using Bayesian and Porter Stemmer keyword stripping approaches 1-5. [Crossref]

1260. Wei Wang, Bo Yu. 2009. Text categorization based on combination of modified back propagation neural network and latent semantic analysis. *Neural Computing and Applications* **18**:8, 875-881. [Crossref]

1261. Jing Chen, Jianfeng Wu. Improved algorithm for keywords extraction from documents without corpus 2339-2341. [Crossref]

1262. Wei Song, Shi Tong Wang, Cheng Hua Li. 2009. Parametric and nonparametric evolutionary computing with a content-based feature selection approach for parallel categorization. *Expert Systems with Applications* **36**:9, 11934-11943. [Crossref]

1263. Alain Lifchitz, Sandra Jhean-Larose, Guy Denhière. 2009. Effect of tuned parameters on an LSA multiple choice questions answering model. *Behavior Research Methods* **41**:4, 1201-1209. [Crossref]

1264. Ljiljana Dolamic, Jacques Savoy. 2009. Indexing and stemming approaches for the Czech language. *Information Processing & Management* **45**:6, 714-720. [Crossref]

1265. Chun-Wei Tsai, Ko-Wei Huango, Ming-Chao Chiang, Chu-Sing Yang. A fast tree-based search algorithm for cluster search engine 1603-1608. [Crossref]

1266. Erik Boiy, Marie-Francine Moens. 2009. A machine learning approach to sentiment analysis in multilingual Web texts. *Information Retrieval* **12**:5, 526-558. [Crossref]

1267. Sotiris Batsakis, Euripides G.M. Petrakis, Evangelos Milios. 2009. Improving the performance of focused web crawlers. *Data & Knowledge Engineering* **68**:10, 1001-1013. [Crossref]

1268. Cheng Hua Li, Wei Song, Soon Cheol Park. 2009. An automatically constructed thesaurus for neural network based document categorization. *Expert Systems with Applications* **36**:8, 10969-10975. [Crossref]

1269. Ira Goldstein, Özlem Uzuner. 2009. Specializing for predicting obesity and its co-morbidities. *Journal of Biomedical Informatics* **42**:5, 873-886. [Crossref]

1270. Sukanya Manna, Zoltan Petres, Tom Gedeon. Tensor term indexing: An application of HOSVD for document summarization 135-141. [Crossref]

1271. Marco A. Palomino, Michael P. Oakes, Tom Wuytack. An evaluation of concept suggestion strategies for professional multimedia archives 273-279. [Crossref]

1272. Priscille Durville, Fabien Gandon. Filling the Gap between Web 2.0 Technologies and Natural Language Processing Pipelines with Semantic Web 109-112. [Crossref]

1273. J. Ignacio Serrano, M. Dolores del Castillo, Ángel Iglesias, Jesús Oliva. 2009. Assessing aspects of reading by a connectionist model. *Neurocomputing* **72**:16-18, 3659-3668. [Crossref]

1274. Ahmet Uyar. 2009. Google stemming mechanisms. *Journal of Information Science* **35**:5, 499-514. [Crossref]

1275. Henryk Fukś, Mark Krzemiński. 2009. Topological structure of dictionary graphs. *Journal of Physics A: Mathematical and Theoretical* **42**:37, 375101. [Crossref]

1276. Shanfeng Zhu, Ichigaku Takigawa, Jia Zeng, Hiroshi Mamitsuka. 2009. Field independent probabilistic model for clustering multi-field documents. *Information Processing & Management* **45**:5, 555-570. [Crossref]

1277. Yunus Emre Esin, Ozgur Alan, Ferda Nur Alpaslan. Improvement on corpus-based word similarity using vector space models 280-285. [Crossref]

1278. Remco Dijkman, Marlon Dumas, Luciano Garcia-Banuelos, Reina Kaarik. Aligning Business Process Models 45-53. [Crossref]

1279. Cecile Levasseur, Uwe F. Mayer, Ken Kreutz-Delgado. Classifying non-gaussian and mixed data sets in their natural parameter space 1-6. [Crossref]

1280. Vincenzo Moscato, Antonio Picariello, Fabio Persia, Antonio Penta. A System for Automatic Image Categorization 624-629. [Crossref]

1281. Goksel Biricik, Banu Diri, Ahmet Coskun Sonmez. A new method for attribute extraction with application on text classification 1-4. [Crossref]

1282. Matthew G. Snover, Nitin Madnani, Bonnie Dorr, Richard Schwartz. 2009. TER-Plus: paraphrase, semantic, and alignment enhancements to Translation Edit Rate. *Machine Translation* **23**:2-3, 117-127. [Crossref]

1283. Domenico Beneventano, Sonia Bergamaschi, Serena Sorrentino. Extending WordNet with compound nouns for semi-automatic annotation in data integration systems 1-8. [Crossref]

1284. Cem Aksoy, Ahmet Bugdayci, Tunay Gur, Ibrahim Uysal, Fazli Can. Semantic argument frequency-based multi-document summarization 460-464. [Crossref]

1285. Francisco C. Pereira, Ana Alves, João Oliveirinha, Assaf Biderman. Perspectives on Semantics of the Place from Online Resources 215-220. [Crossref]

1286. Billy Wong, Chunyu Kit. 2009. ATEC: automatic evaluation of machine translation via word choice and word order. *Machine Translation* **23**:2-3, 141-155. [Crossref]

1287. T.W.S. Chow, M.K.M. Rahman. 2009. Multilayer SOM With Tree-Structured Data for Efficient Document Retrieval and Plagiarism Detection. *IEEE Transactions on Neural Networks* **20**:9, 1385-1402. [Crossref]

1288. Zafer Erenel, Hakan Altincay, Ekrem Varoglu. A symmetric term weighting scheme for text categorization based on term occurrence probabilities 1-4. [Crossref]

1289. Miguel Á. García-Cumbreras, José M. Perea-Ortega, Manuel García-Vega, L. Alfonso Ureña-López. 2009. Information retrieval with geographical references. Relevant documents filtering vs. query expansion. *Information Processing & Management* **45**:5, 605-614. [Crossref]

1290. F. Canan Pembe, Tunga Güngör. 2009. Structure-preserving and query-biased document summarisation for web searching. *Online Information Review* **33**:4, 696-719. [Abstract] [Full Text] [PDF]

1291. Zhiping Wang, Seongho Kim, Sara K. Quinney, Yingying Guo, Stephen D. Hall, Luis M. Rocha, Lang Li. 2009. Literature mining on pharmacokinetics numerical data: A feasibility study. *Journal of Biomedical Informatics* **42**:4, 726-735. [Crossref]

1292. Simone Stumpf, Vidya Rajaram, Lida Li, Weng-Keen Wong, Margaret Burnett, Thomas Dietterich, Erin Sullivan, Jonathan Herlocker. 2009. Interacting meaningfully with machine learning systems: Three experiments. *International Journal of Human-Computer Studies* **67**:8, 639-662. [Crossref]

1293. Chih-Hung Wu, I-Ching Fang, Chia-Hsiang Wu, Wei-Ting Lin, Chi-Hua Li. A novel parallel cross-validated support vector machine on patent classification system 633-641. [Crossref]

1294. Kun Yang, Zhongzhi Shi. Exploring Web Resources in Social Annotation Services in Hierarchy 182-187. [Crossref]

1295. Claire Fautsch, Jacques Savoy. 2009. Algorithmic stemmers or morphological analysis? An evaluation. *Journal of the American Society for Information Science and Technology* **60**:8, 1616-1624. [Crossref]

1296. G. Tambouratzis. 2009. Using an Ant Colony Metaheuristic to Optimize Automatic Word Segmentation for Ancient Greek. *IEEE Transactions on Evolutionary Computation* **13**:4, 742-753. [Crossref]

1297. Juan Manuel Barrios, Diego Diaz-Espinoza, Benjamin Bustos. Text-Based and Content-Based Image Retrieval on Flickr: DEMO 156-157. [Crossref]

1298. Dora Alvarez, Hugo Hidalgo. 2009. Document analysis and visualization with zero-inflated poisson. *Data Mining and Knowledge Discovery* **19**:1, 1-23. [Crossref]

1299. Aishan Wumaier, Zaokere Kadeer, Parida Tursun, Shengwei Tian. Maximum Entropy combined FSM stemming method for Uyghur 51-55. [Crossref]

1300. M. Vazquez, P. Carmona-Saez, R. Nogales-Cadenas, M. Chagoyen, F. Tirado, J. M. Carazo, A. Pascual-Montano. 2009. SENT: semantic features in text. *Nucleic Acids Research* **37**:Web Server, W153-W159. [Crossref]

1301. Farag Ahmed, Andreas Nürnberger. 2009. Evaluation of n-gram conflation approaches for Arabic text retrieval. *Journal of the American Society for Information Science and Technology* **60**:7, 1448-1465. [Crossref]

1302. Zhixing Huang, Yan Yan, Yuhui Qiu, Shuqiong Qiao. Exploring Emergent Semantic Communities from DBLP Bibliography Database 219-224. [Crossref]

1303. Allen P. Nikora, Galen Balcom. Improving the Accuracy of Space Mission Software Anomaly Frequency Estimates 402-409. [Crossref]

1304. Deng-Yiv Chiu, Ya-Chen Pan, Shih-Yi Yu. 2009. An automated knowledge structure construction approach: Applying information retrieval and genetic algorithm to journal of Expert Systems with Applications. *Expert Systems with Applications* **36**:5, 9438-9447. [Crossref]

1305. Heejung Kim, Jae Yun Lee. 2009. Archiving research trends in LIS domain using profiling analysis. *Scientometrics* **80**:1, 75-90. [Crossref]

1306. Badam-Osor Khaltar, Atsushi Fujii. 2009. A lemmatization method for Mongolian and its application to indexing for information retrieval. *Information Processing & Management* **45**:4, 438-451. [Crossref]

1307. K. Verspoor, D. Dvorkin, K. B. Cohen, L. Hunter. 2009. Ontology quality assurance through analysis of term transformations. *Bioinformatics* **25**:12, i77-i84. [Crossref]

1308. Junjie Wu, Hui Xiong, Jian Chen. 2009. Towards understanding hierarchical clustering: A data distribution perspective. *Neurocomputing* **72**:10-12, 2319-2330. [Crossref]

1309. Hui Xiong, Michael Steinbach, Arifin Ruslim, Vipin Kumar. 2009. Characterizing pattern preserving clustering. *Knowledge and Information Systems* **19**:3, 311-336. [Crossref]

1310. R.Y.K. Lau, Dawei Song, Yuefeng Li, T.C.H. Cheung, Jin-Xing Hao. 2009. Toward a Fuzzy Domain Ontology Extraction Method for Adaptive e-Learning. *IEEE Transactions on Knowledge and Data Engineering* **21**:6, 800-813. [Crossref]

1311. Fotis Lazarinis, Jesús Vilares, John Tait, Efthimis N. Efthimiadis. 2009. Current research issues and trends in non-English Web searching. *Information Retrieval* **12**:3, 230-250. [Crossref]

1312. Pu Wang, Jian Hu, Hua-Jun Zeng, Zheng Chen. 2009. Using Wikipedia knowledge to improve text classification. *Knowledge and Information Systems* **19**:3, 265-281. [Crossref]

1313. Peter Wittek, Chew Lim Tan. A kernel-based feature weighting for text classification 3373-3379. [Crossref]

1314. M. Arun Kumar, M. Gopal. 2009. Least squares twin support vector machines for pattern classification. *Expert Systems with Applications* **36**:4, 7535-7543. [Crossref]

1315. Supakpong Jinarat, Choochart Haruechaiyasak, Arnon Rungsawang. Improving web search result categorization using knowledge from web taxonomy 726-730. [Crossref]

1316. Cheng Hua Li, Soon Cheol Park. 2009. Combination of modified BPNN algorithms and an efficient feature selection method for text categorization. *Information Processing & Management* **45**:3, 329-340. [Crossref]

1317. K.M. Hammouda, M.S. Kamel. 2009. Hierarchically Distributed Peer-to-Peer Document Clustering and Cluster Summarization. *IEEE Transactions on Knowledge and Data Engineering* **21**:5, 681-698. [Crossref]

1318. Ramiz M. Aliguliyev. 2009. Clustering of document collection – A weighting approach. *Expert Systems with Applications* **36**:4, 7904-7916. [Crossref]

1319. Ramiz M. Aliguliyev. 2009. A new sentence similarity measure and sentence based extractive technique for automatic text summarization. *Expert Systems with Applications* **36**:4, 7764-7772. [Crossref]

1320. Giovanni Capobianco, Andrea De Lucia, Rocco Oliveto, Annibale Panichella, Sebastiano Panichella. On the role of the nouns in IR-based traceability recovery 148-157. [Crossref]

1321. M. Ángeles Serrano, Alessandro Flammini, Filippo Menczer. 2009. Modeling Statistical Properties of Written Text. *PLoS ONE* **4**:4, e5372. [Crossref]

1322. Om D. Deshmukh, Kundan Kandhway, Ashish Verma, Kartik Audhkhasi. Automatic evaluation of spoken english fluency 4829-4832. [Crossref]

1323. Hiroshi Ogura, Hiromi Amano, Masato Kondo. 2009. Feature selection with a measure of deviations from Poisson in text categorization. *Expert Systems with Applications* **36**:3, 6826-6832. [Crossref]

1324. Ioannis Katakis, Grigorios Tsoumakas, Evangelos Banos, Nick Bassiliades, Ioannis Vlahavas. 2009. An adaptive personalized news dissemination system. *Journal of Intelligent Information Systems* **32**:2, 191-212. [Crossref]

1325. Man Lan, Chew Lim Tan, Jian Su, Yue Lu. 2009. Supervised and Traditional Term Weighting Methods for Automatic Text Categorization. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **31**:4, 721-735. [Crossref]

1326. Hui Xiong, Junjie Wu, Jian Chen. 2009. K-Means Clustering Versus Validation Measures: A Data-Distribution Perspective. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)* **39**:2, 318-331. [Crossref]

1327. M.C. Díaz-Galiano, M.T Martín-Valdivia, L.A. Ureña-López. 2009. Query expansion with a medical ontology to improve a multimodal information retrieval system. *Computers in Biology and Medicine* **39**:4, 396-403. [Crossref]

1328. Hei-Chia Wang, Tian-Hsiang Huang. 2009. Prediction of EST Functional Relationships via Literature Mining With User-Specified Parameters. *IEEE Transactions on Biomedical Engineering* **56**:4, 969-977. [Crossref]

1329. SUJEEVAN ASEERVATHAM. 2009. A CONCEPT VECTOR SPACE MODEL FOR SEMANTIC KERNELS. *International Journal on Artificial Intelligence Tools* **18**:02, 239-272. [Crossref]

1330. Kimmo Kettunen. 2009. Reductive and generative approaches to management of morphological variation of keywords in monolingual information retrieval. *Journal of Documentation* **65**:2, 267-290. [Abstract] [Full Text] [PDF]

1331. Jiaming Zhan, Han Tong Loh, Ying Liu. 2009. Gather customer concerns from online product reviews – A text summarization approach. *Expert Systems with Applications* **36**:2, 2107-2115. [Crossref]

1332. Robert Jay Milewski, Venu Govindaraju, Anurag Bhardwaj. 2009. Automatic recognition of handwritten medical forms for search engines. *International Journal of Document Analysis and Recognition (IJDAR)* **11**:4, 203-218. [Crossref]

1333. Robert M. Patton, Barbara G. Beckerman, Thomas E. Potok. Learning cue phrase patterns from radiology reports using a genetic algorithm 1-4. [Crossref]

1334. Juan Cao, Tian Xia, Jintao Li, Yongdong Zhang, Sheng Tang. 2009. A density-based method for adaptive LDA model selection. *Neurocomputing* **72**:7-9, 1775-1781. [Crossref]

1335. Samhaa R. El-Beltagy, Ahmed Rafea. 2009. KP-Miner: A keyphrase extraction system for English and Arabic documents. *Information Systems* **34**:1, 132-144. [Crossref]

1336. Utku Irmak, Vadim von Brzeski, Reiner Kraft. Contextual Ranking of Keywords Using Click Data 457-468. [Crossref]

1337. Carolyn Holton. 2009. Identifying disgruntled employee systems fraud risk through text mining: A simple solution for a multi-billion dollar problem. *Decision Support Systems* **46**:4, 853-864. [Crossref]

1338. A. E. ELdesoky, M. Saleh, N.A. Sakr. Novel similarity measure for document clustering based on topic phrases 92-96. [Crossref]

1339. Cheng Hua Li, Soon Choel Park. 2009. An efficient document classification model using an improved back propagation neural network and singular value decomposition. *Expert Systems with Applications* **36**:2, 3208-3215. [Crossref]

1340. Duen-Ren Liu, Meng-Jung Shih, Churn-Jung Liau, Chin-Hui Lai. 2009. Mining the change of event trends for decision support in environmental scanning. *Expert Systems with Applications* **36**:2, 972-984. [Crossref]

1341. Yanping Lu, Shengrui Wang, Shaozi Li, Changle Zhou. Text Clustering via Particle Swarm Optimization 45-51. [Crossref]

1342. Luh Yen, Francois Fouss, Christine Decaestecker, Pascal Francq, Marco Saerens. 2009. Graph nodes clustering with the sigmoid commute-time kernel: A comparative study. *Data & Knowledge Engineering* **68**:3, 338-361. [Crossref]

1343. Bai Rujiang, Liao Junhua. A Novel Conception Based Texts Classification Method 30-34. [Crossref]

1344. Asako Koike, Toshihisa Takagi. Classifying Biomedical Figures Using Combination of Bag of Keypoints and Bag of Words 848-853. [Crossref]

1345. S Nefti, M Oussalah, Y Rezgui. 2009. A modified fuzzy clustering for documents retrieval: application to document categorization. *Journal of the Operational Research Society* **60**:3, 384-394. [Crossref]

1346. Robert Moskovitch, Yuval Shahar. 2009. Vaidurya: A multiple-ontology, concept-based, context-sensitive clinical-guideline search engine. *Journal of Biomedical Informatics* **42**:1, 11-21. [Crossref]

1347. Jiaming Zhan, Han Tong Loh, Ying Liu. 2009. On macro- and micro-level information in multiple documents and its influence on summarization. *International Journal of Information Management* **29**:1, 57-66. [Crossref]

1348. Nizar Habash, Bonnie Dorr, Christof Monz. 2009. Symbolic-to-statistical hybridization: extending generation-heavy machine translation. *Machine Translation* **23**:1, 23-63. [Crossref]

1349. Andrea De Lucia, Rocco Oliveto, Genoveffa Tortora. 2009. Assessing IR-based traceability recovery tools through controlled experiments. *Empirical Software Engineering* **14**:1, 57-92. [Crossref]

1350. Constantin Orăsan. 2009. Comparative Evaluation of Term-Weighting Methods for Automatic Summarization*. *Journal of Quantitative Linguistics* **16**:1, 67-95. [Crossref]

1351. Hsin-Chang Yang, Chung-Hong Lee, Ding-Wen Chen. 2009. A method for multilingual text mining and retrieval using growing hierarchical self-organizing maps. *Journal of Information Science* **35**:1, 3-23. [Crossref]

1352. Kwan Yi, Jamshid Beheshti. 2009. A hidden Markov model-based text classification of medical documents. *Journal of Information Science* **35**:1, 67-81. [Crossref]

1353. Christian Sternitzke, Isumo Bergmann. 2009. Similarity measures for document mapping: A comparative study on the level of an individual scientist. *Scientometrics* **78**:1, 113-130. [Crossref]

1354. Ya-Xi Chen, Rodrigo Santamaría, Andreas Butz, Roberto Therón. TagClusters: Semantic Aggregation of Collaborative Tags beyond TagClouds 56-67. [Crossref]

1355. Sebastián Peña Saldarriaga, Emmanuel Morin, Christian Viard-Gaudin. Using top n Recognition Candidates to Categorize On-line Handwritten Documents 881-885. [Crossref]

1356. Limin Li, Junjie Wu, Shiwei Zhu. Implication intensity: Randomized F-measure for cluster evaluation 510-515. [Crossref]

1357. Mikko Kurimo, Mathias Creutz, Ville Turunen. Morpho Challenge Evaluation by Information Retrieval Experiments 991-998. [Crossref]

1358. José Manuel Perea-Ortega, L. Alfonso Ureña-López, Manuel García-Vega, Miguel Angel García-Cumbreras. Using Query Reformulation and Keywords in the Geographic Information Retrieval Task 855-862. [Crossref]

1359. Miro Lehtonen, Antoine Doucet. Enhancing Keyword Search with a Keyphrase Index 65-70. [Crossref]

1360. Mark Kilfoil, Ali Ghorbani. 2009. SWAMI. *International Journal of Information Technology and Web Engineering* 4:3, 1-24. [Crossref]

1361. Kazunari Sugiyama, Manabu Okumura. 2009. Personal Name Disambiguation in Web Search Results Using a Semi-Supervised Clustering Approach. *Journal of Natural Language Processing* 16:5, 23-49. [Crossref]

1362. Ljiljana Dolamic, Claire Fautsch, Jacques Savoy. UniNE at CLEF 2008: TEL, and Persian IR 178-185. [Crossref]

1363. Mostafa Keikha, Ahmad Khonsari, Farhad Oroumchian. 2009. Rich document representation and classification: An analysis. *Knowledge-Based Systems* 22:1, 67-71. [Crossref]

1364. Tolga Berber, Adil Alpkocak. A boolean information retrieval system for mammography reports 1-6. [Crossref]

1365. Alberto Barrón-Cedeño, Paolo Rosso. On Automatic Plagiarism Detection Based on n-Grams Comparison 696-700. [Crossref]

1366. Houda Benbrahim, Max Bramer. Text and Hypertext Categorization 11-38. [Crossref]

1367. André Pinto Geraldo, Viviane P. Moreira, Marcos A. Gonçalves. On-Demand Associative Cross-Language Information Retrieval 165-173. [Crossref]

1368. Walaa K. Gad, Mohamed S. Kamel. New Semantic Similarity Based Model for Text Clustering Using Extended Gloss Overlaps 663-677. [Crossref]

1369. Bart Van Rompaey, Serge Demeyer. Establishing Traceability Links between Unit Test Cases and Units under Test 209-218. [Crossref]

1370. Michiko YASUKAWA, Hui Tian LIM, Hidetoshi YOKOO. 2009. Stemming Malay Text and Its Application in Automatic Text Categorization. *IEICE Transactions on Information and Systems* E92-D:12, 2351-2359. [Crossref]

1371. Mao-Lung Chen, Richi Nayak, Shlomo Geva. Link-the-Wiki: Performance Evaluation Based on Frequent Phrases 326-336. [Crossref]

1372. Yanhong Liang, Dequan Gan, Yingchun Guo, Peng Zhang. Computer-Aided Analysis of Patents for Product Technology Maturity Forecasting 295-303. [Crossref]

1373. Hsin-Min Lu, Nina WanHsin Huang, Zhu Zhang, Tsai-Jyh Chen. Identifying Firm-Specific Risk Statements in News Articles 42-53. [Crossref]

1374. Nikitas N. Karanikolas. Bootstrapping the Albanian Information Retrieval 231-235. [Crossref]

1375. Archana Kale, Ujwala Bharambe, M. SashiKumar. A New Suffix Tree Similarity Measure and Labeling for Web Search Results Clustering 856-861. [Crossref]

1376. Aishan Wumaier, Tuergen Yibulayin, Zaokere Kadeer, Shengwei Tian. Conditional Random Fields combined FSM stemming method for Uyghur 295-299. [Crossref]

1377. Christophe Moulin, Cécile Barat, Mathias Géry, Christophe Ducottet, Christine Largeron. UJM at ImageCLEFwiki 2008 779-786. [Crossref]

1378. Sukanya Manna, Zoltan Petres, Tom Gedeon. Significant term extraction by Higher Order SVD 63-68. [Crossref]

1379. Alberto Barrón-Cedeño, Paolo Rosso, José-Miguel Benedí. Reducing the Plagiarism Detection Search Space on the Basis of the Kullback-Leibler Distance 523-534. [Crossref]

1380. Orland Hoeber, Xue Dong Yang. 2009. HotMap: Supporting visual exploration of Web search results. *Journal of the American Society for Information Science and Technology* **60**:1, 90-110. [Crossref]

1381. Shanmugasundaram Hariharan, Rengaramanujam Srinivasan. 2009. Studies on Graph Based Approaches for Singleand Multi Document Summarizations. *International Journal of Computer Theory and Engineering* 519-526. [Crossref]

1382. Ljiljana Dolamic, Jacques Savoy. Persian Language, Is Stemming Efficient? 388-392. [Crossref]

1383. N.L. Bhamidipati, S.K. Pal. 2009. Comparing Scores Intended for Ranking. *IEEE Transactions on Knowledge and Data Engineering* **21**:1, 21-34. [Crossref]

1384. Walaa Khaled Gad, Mohamed Salem Kamel. PH-SSBM: Phrase Semantic Similarity Based Model for Document Clustering 197-200. [Crossref]

1385. Mani Arun Kumar, Madan Gopal. Text Categorization Using Fuzzy Proximal SVM and Distributional Clustering of Words 52-61. [Crossref]

1386. Chun-Wei Tsai, Shih-Pang Tseng, Yu-Sheng Yang, Ming-Chao Chiang, Chu-Sing Yang. An Ontology-Based Method for Integrating Heterogeneous Itembanks 919-924. [Crossref]

1387. Surafel Lemma Abebe, Sonia Haiduc, Andrian Marcus, Paolo Tonella, Giuliano Antoniol. Analyzing the Evolution of the Source Code Vocabulary 189-198. [Crossref]

1388. Thomas Gottron. Document Word Clouds: Visualising Web Documents as Tag Clouds to Aid Users in Relevance Decisions 94-105. [Crossref]

1389. José Manuel Perea-Ortega, Arturo Montejo-Ráez, Manuel Carlos Díaz-Galiano, María Teresa Martín-Valdivia, L. Alfonso Ureña-López. Using an Information Retrieval System for Video Classification 927-930. [Crossref]

1390. Syed Toufeeq Ahmed, Ruchi Bhindwale, Hasan Davulcu. Tracking terrorism news threads by extracting event signatures 182-184. [Crossref]

1391. Raghavendra Udupa, Saravanan K, Anton Bakalov, Abhijit Bhole. "They Are Out There, If You Know Where to Look": Mining Transliterations of OOV Query Terms for Cross-Language Information Retrieval 437-448. [Crossref]

1392. Christopher Collins, Fernanda B. Viegas, Martin Wattenberg. Parallel Tag Clouds to explore and analyze faceted text corpora 91-98. [Crossref]

1393. Chun-Ling Chen, Frank S. C. Tseng, Tyne Liang. An Integration of Fuzzy Association Rules and WordNet for Document Clustering 147-159. [Crossref]

1394. Gleb Skobeltsyn, Toan Luu, Ivana Podnar Žarko, Martin Rajman, Karl Aberer. 2009. Query-driven indexing for scalable peer-to-peer text retrieval. *Future Generation Computer Systems* **25**:1, 89-99. [Crossref]

1395. Jay Urbain, Nazli Goharian, Ophir Frieder. 2009. A dimensional retrieval model for integrating semantics and statistical evidence in context for genomics literature search. *Computers in Biology and Medicine* **39**:1, 61-68. [Crossref]

1396. Sangeetha Kutty, Tien Tran, Richi Nayak, Yuefeng Li. Clustering XML Documents Using Frequent Subtrees 436-445. [Crossref]

1397. André Pinto Geraldo, Viviane P. Moreira. UFRGS@CLEF2008: Using Association Rules for Cross-Language Information Retrieval 66-74. [Crossref]

1398. Sidath Harshanath Samarasinghe, Siu Cheung Hui. Mathematical Document Retrieval for Problem Solving 583-587. [Crossref]

1399. Nattakarn Ratprasartporn, Jonathan Po, Ali Cakmak, Sulieman Bani-Ahmad, Gultekin Ozsoyoglu. 2009. Context-based literature digital collection search. *The VLDB Journal* **18**:1, 277-301. [Crossref]

1400. Hsin-Chang Yang, Han-Wei Hsiao, Chung-Hong Lee. Automatic Construction of Multilingual Web Directory Using Self-Organizing Maps 1283-1288. [Crossref]

1401. Mikko Kurimo, Ville Turunen, Matti Varjokallio. Overview of Morpho Challenge 2008 951-966. [Crossref]

1402. J. Ignacio Serrano, M. Dolores del Castillo, A. Iglesias. 2009. Dealing with written language semantics by a connectionist model of cognitive reading. *Neurocomputing* **72**:4-6, 713-725. [Crossref]

1403. Chih-Ming Chen, Hahn-Ming Lee, Yu-Jung Chang. 2009. Two novel feature selection approaches for web page classification. *Expert Systems with Applications* **36**:1, 260-272. [Crossref]

1404. Joann M. Paul, Mwaffaq Otoom, Marc Somers, Sean Pieper, Michael J. Schulte. Chapter 4 The Emerging Landscape of Computer Performance Evaluation 235-280. [Crossref]

1405. Andrzej Siemiński. Using WordNet to Measure the Similarity of Link Texts 720-731. [Crossref]

1406. Stephan Bloehdorn, Andreas Hotho. Ontologies for Machine Learning 637-661. [Crossref]

1407. Philipp Cimiano, Alexander Mädche, Steffen Staab, Johanna Völker. Ontology Learning 245-267. [Crossref]

1408. Mathias G?ry, Christine Largeron, Christophe Moulin. UJM at INEX 2008 XML Mining Track 446-452. [Crossref]

1409. Agus T. Kwee, Flora S. Tsai, Wenyin Tang. Sentence-Level Novelty Detection in English and Malay 40-51. [Crossref]

1410. Rasim ALGULIEV, Ramiz ALIGULIYEV. 2009. Evolutionary Algorithm for Extractive Text Summarization. *Intelligent Information Management* **01**:02, 128-138. [Crossref]

1411. Shaunak Chatterjee, Sudeep Juvekar, Koushik Sen. SNIFF: A Search Engine for Java Using Free-Form Queries 385-400. [Crossref]

1412. Bruno Antunes, Tiago Guerreiro, Daniel Gonçalves. This Just In! Your Life in the Newspaper 279-292. [Crossref]

1413. Jianwei Cui, Pei Li, Hongyan Liu, Jun He, Xiaoyong Du. A Neighborhood Search Method for Link-Based Tag Clustering 91-103. [Crossref]

1414. Rayner Alfred. A Parallel Hierarchical Agglomerative Clustering Technique for Billingual Corpora Based on Reduced Terms with Automatic Weight Optimization 19-30. [Crossref]

1415. Fei Liu, Feifan Liu, Yang Liu. Automatic keyword extraction for the meeting corpus using supervised approach and bigram expansion 181-184. [Crossref]

1416. Geraldo Xexéo, Jano de Souza, Patrícia F. Castro, Wallace A. Pinheiro. Using Wavelets to Classify Documents 272-278. [Crossref]

1417. Leonhard Hennig, Winfried Umbrath, Robert Wetzker. An Ontology-Based Approach to Text Summarization 291-294. [Crossref]

1418. Sudhindra Shukla, Ashok Samal. 2008. Recognition and quality assessment of data charts in mixed-mode documents. *International Journal of Document Analysis and Recognition (IJDAR)* **11**:3, 111-126. [Crossref]

1419. YI PENG, GANG KOU, YONG SHI, ZHENGXIN CHEN. 2008. A DESCRIPTIVE FRAMEWORK FOR THE FIELD OF DATA MINING AND KNOWLEDGE DISCOVERY. *International Journal of Information Technology & Decision Making* **07**:04, 639-682. [Crossref]

1420. Maria Fernanda Moura, Bruno Magalhaes Nogueira, Merley da Silva Conrado, Fabiano Fernandes dos Santos, Solange Oliveira Rezende. Making good choices of non-redundant n-gramwords 64-71. [Crossref]

1421. D. Suendermann, J. Liscombe, K. Evanini, K. Dayanidhi, R. Pieraccini. C^5 125-128. [Crossref]

1422. Chongyang Shi, Zhendong Niu, Tao Wang, Xiaomei Xu. A Digital Assets Navigating System Based on FCA 356-360. [Crossref]

1423. Michael G. Noll, Christoph Meinel. The Metadata Triumvirate: Social Annotations, Anchor Texts and Search Queries 640-647. [Crossref]

1424. Cai-Nicolas Ziegler, Michal Skubacz, Maximilian Viermetz. Mining and Exploring Unstructured Customer Feedback Data Using Language Models and Treemap Visualizations 932-937. [Crossref]

1425. XiQuan Yang, DiNa Guo, XueYa Cao, JianYuan Zhou. Research on Ontology-Based Text Clustering 141-146. [Crossref]

1426. A. Li, M. Haraguchi, Y. Okubo. Implicit Groups of Web Pages as Constrained Top N Concepts 190-194. [Crossref]

1427. N. Ishii, A. Koike, Y. Yamamoto, T. Takagi. Figure Classification in Biomedical Literature towards Figure Mining 263-269. [Crossref]

1428. Yongli Liu, Yuanxin Ouyang, Hao Sheng, Zhang Xiong. An Incremental Algorithm for Clustering Search Results 112-117. [Crossref]

1429. A. Fyshe, Y. Liu, D. Szafron, R. Greiner, P. Lu. 2008. Improving subcellular localization prediction using text classification and the gene ontology. *Bioinformatics* **24**:21, 2512-2517. [Crossref]

1430. Alexandra Laflamme-Sanders, Mu Zhu. 2008. LAGO on the unit sphere. *Neural Networks* **21**:9, 1220-1223. [Crossref]

1431. Hsin-Chang Yang, Chung-Hong Lee. Multilingual Information Retrieval Using GHSOM 225-228. [Crossref]

1432. Peter Z. Yeh, Alex Kass. Capturing the Semantics of Online News Sources for Business Intelligence Applications 111-117. [Crossref]

1433. Wouter Souffriau, Pieter Vansteenwegen, Joris Vertommen, Greet Vanden Berghe, Dirk Van Oudheusden. 2008. A PERSONALIZED TOURIST TRIP DESIGN ALGORITHM FOR MOBILE TOURIST GUIDES. *Applied Artificial Intelligence* **22**:10, 964-985. [Crossref]

1434. Khaled Hammouda, Mohamed Kamel. 2008. Distributed collaborative Web document clustering using cluster keyphrase summaries. *Information Fusion* **9**:4, 465-480. [Crossref]

1435. Kendall E. Giles, Michael W. Trosset, David J. Marchette, Carey E. Priebe. 2008. Iterative Denoising. *Computational Statistics* **23**:4, 497-517. [Crossref]

1436. Renato Fernandes Corrêa, Teresa Bernarda Ludermir. 2008. A quickly trainable hybrid SOM-based document organization system. *Neurocomputing* **71**:16-18, 3353-3359. [Crossref]

1437. Sara E. Garza, Ramón F. Brena, Eduardo Ramírez. Topic Calculation and Clustering: An Application to Wikipedia 88-93. [Crossref]

1438. Antoine Naud, Shiro Usui. 2008. Exploration of a collection of documents in neuroscience and extraction of topics by clustering. *Neural Networks* **21**:8, 1205-1211. [Crossref]

1439. Xuelin Shi, Ying Zhao, Xiangjun Dong. Web Page Categorization Based on k-NN and SVM Hybrid Pattern Recognition Algorithm 523-527. [Crossref]

1440. Jay Urbain, Nazli Goharian, Ophir Frieder. 2008. Probabilistic passage models for semantic search of genomics literature. *Journal of the American Society for Information Science and Technology* **59**:12, 2008-2023. [Crossref]

1441. Abdelghani Bellaachia, Ghita Amor-Tijani. 2008. Proper nouns in English-Arabic cross language information retrieval. *Journal of the American Society for Information Science and Technology* **59**:12, 1925-1932. [Crossref]

1442. DIMITRIOS P. LYRAS, KYRIAKOS N. SGARBAS, NIKOLAOS D. FAKOTAKIS. 2008. APPLYING SIMILARITY MEASURES FOR AUTOMATIC LEMMATIZATION: A CASE STUDY FOR MODERN GREEK AND ENGLISH. *International Journal on Artificial Intelligence Tools* **17**:05, 1043-1064. [Crossref]

1443. Cheng-Yuan Liou, Jau-Chi Huang, Wen-Chie Yang. 2008. Modeling word perception using the Elman network. *Neurocomputing* **71**:16-18, 3150-3157. [Crossref]

1444. Yun Chen, Flora S. Tsai, Kap Luk Chan. 2008. Machine learning techniques for business blog search and mining. *Expert Systems with Applications* **35**:3, 581-590. [Crossref]

1445. Yun Zhang, Boqin Feng, Yewei Xue. A New Search Results Clustering Algorithm Based on Formal Concept Analysis 356-360. [Crossref]

1446. Heng-Yao Hsu, Chun-Wei Tsai, Ming-Chao Chiang, Chu-Sing Yang. Topic generation for web document summarization 3702-3707. [Crossref]

1447. H. Shatkay, F. Pan, A. Rzhetsky, W. J. Wilbur. 2008. Multi-dimensional classification of biomedical text: Toward automated, practical provision of high-utility text to diverse users. *Bioinformatics* **24**:18, 2086-2093. [Crossref]

1448. K. Bretonnel Cohen, Martha Palmer, Lawrence Hunter. 2008. Nominalization and Alternations in Biomedical Language. *PLoS ONE* **3**:9, e3158. [Crossref]

1449. B. Yu. 2008. An evaluation of text classification methods for literary study. *Literary and Linguistic Computing* **23**:3, 327-343. [Crossref]

1450. Yanhong Liang, Runhua Tan, Jianhong Ma. Patent analysis with text mining for TRIZ 1147-1151. [Crossref]

1451. Chaitali Gupta, Rajdeep Bhowmik, Madhusudhan Govindaraju. Web Services Operation and Parameter Matchmaking Based on Free-Form User Queries 57-64. [Crossref]

1452. Renato Fernandes Corrêa, Teresa Bernarda Ludermir. Improved Semantic Mapping and SOM Applied to Document Organization 284-289. [Crossref]

1453. Sebastián Peña Saldarriaga, Emmanuel Morin, Christian Viard-Gaudin. Categorization of On-Line Handwritten Documents 95-102. [Crossref]

1454. J. Šnajder, B. Dalbelo Bašić, M. Tadić. 2008. Automatic acquisition of inflectional lexica for morphological normalisation. *Information Processing & Management* **44**:5, 1720-1731. [Crossref]

1455. Chen-Huei Chou, Atish P. Sinha, Huimin Zhao. 2008. A text mining approach to Internet abuse detection. *Information Systems and e-Business Management* **6**:4, 419-439. [Crossref]

1456. Ji He. Improving feature representation of natural language gene functional annotations using automatic term expansion 173-179. [Crossref]

1457. Xiaohui Cui, Justin M. Beaver, Jesse St. Charles, Thomas E. Potok. Dimensionality reduction particle swarm algorithm for high dimensional clustering 1-6. [Crossref]

1458. Joris Vertommen, Frizo Janssens, Bart De Moor, Joost R. Duflou. 2008. Multiple-vector user profiles in support of knowledge sharing. *Information Sciences* **178**:17, 3333-3346. [Crossref]

1459. Chin Pang Cheng, Gloria T. Lau, Kincho H. Law, Jiayi Pan, Albert Jones. 2008. Regulation retrieval using industry specific taxonomies. *Artificial Intelligence and Law* **16**:3, 277-303. [Crossref]

1460. Xia Sun, Ziqiang Wang, Dexian Zhang. A Web Document Classification Method Based on Shuffled Frog Leaping Algorithm 205-208. [Crossref]

1461. I-Chin Wu, Duen-Ren Liu, Pei-Cheng Chang. 2008. Toward incorporating a task-stage identification technique into the long-term document support process. *Information Processing & Management* **44**:5, 1649-1672. [Crossref]

1462. Tim Menzies, Andrian Marcus. Automated severity assessment of software defect reports 346-355. [Crossref]

1463. Shi Yu, Steven Van Vooren, Leon-Charles Tranchevent, Bart De Moor, Yves Moreau. 2008. Comparison of vocabularies, representations and ranking algorithms for gene prioritization by text mining. *Bioinformatics* **24**:16, i119-i125. [Crossref]

1464. R. Dhanapal. 2008. An intelligent information retrieval agent. *Knowledge-Based Systems* **21**:6, 466-470. [Crossref]

1465. Mats Sjöberg, Jorma Laaksonen, Timo Honkela, Matti Pöllä. 2008. Inferring semantics from textual information in multimedia retrieval. *Neurocomputing* **71**:13-15, 2576-2586. [Crossref]

1466. Smita Nirkhi, Kapil Hande. Optimization of Context Disambiguation in Web Wearch Results 820-824. [Crossref]

1467. S. Repp, A. Gross, C. Meinel. 2008. Browsing within Lecture Videos Based on the Chain Index of Speech Transcription. *IEEE Transactions on Learning Technologies* **1**:3, 145-156. [Crossref]

1468. CAROLINE SPORLEDER, ALEX LASCARIDES. 2008. Using automatically labelled examples to classify rhetorical relations: an assessment. *Natural Language Engineering* **14**:03. . [Crossref]

1469. Jesús Vilares, Miguel A. Alonso, Manuel Vilares. 2008. Extraction of complex index terms in non-English IR: A shallow parsing based approach. *Information Processing & Management* **44**:4, 1517-1537. [Crossref]

1470. Eray Uluhan, Bertan Badur. Development of a framework for sub-topic discovery from the Web 878-888. [Crossref]

1471. Olga Vechtomova, Murat Karamuftuoglu. 2008. Lexical cohesion and term proximity in document ranking. *Information Processing & Management* **44**:4, 1485-1502. [Crossref]

1472. Abdelghani Bellaachia, Ghita Amor-Tijani. Proper nouns in English-Arabic Cross Language Information Retrieval 849-857. [Crossref]

1473. Leila Kosseim, Jamileh Yousefi. 2008. Improving the performance of question answering with semantically equivalent answer patterns. *Data & Knowledge Engineering* **66**:1, 53-67. [Crossref]

1474. S. Haiduc, A. Marcus. On the Use of Domain Terms in Source Code 113-122. [Crossref]

1475. Kazuaki Kishida. 2008. Prediction of performance of cross-language information retrieval using automatic evaluation of translation. *Library & Information Science Research* **30**:2, 138-144. [Crossref]

1476. Chih-Ping Wei, Hsueh-Ching Chen, Tsang-Hsiang Cheng. 2008. Effective spam filtering: A single-class learning and ensemble approach. *Decision Support Systems* **45**:3, 491-503. [Crossref]

1477. Frizo Janssens, Wolfgang Glänzel, Bart De Moor. 2008. A hybrid mapping of information science. *Scientometrics* **75**:3, 607-631. [Crossref]

1478. Michael Chau, Jialun Qin, Yilu Zhou, Chunju Tseng, Hsinchun Chen. 2008. SpidersRUs: Creating specialized search engines in multiple languages. *Decision Support Systems* **45**:3, 621-640. [Crossref]

1479. Manjeet Rege, Ming Dong, Farshad Fotouhi. 2008. Bipartite isoperimetric graph partitioning for data co-clustering. *Data Mining and Knowledge Discovery* **16**:3, 276-312. [Crossref]

1480. A. Markov, M. Last, A. Kandel. 2008. The hybrid representation model for web document classification. *International Journal of Intelligent Systems* **23**:6, 654-679. [Crossref]

1481. MIHAI LINTEAN, VASILE RUS. 2008. LARGE SCALE EXPERIMENTS WITH NAIVE BAYES AND DECISION TREES FOR FUNCTION TAGGING. *International Journal on Artificial Intelligence Tools* **17**:03, 483-499. [Crossref]

1482. Heejung Kim, Jae Yun Lee. 2008. Exploring the emerging intellectual structure of archival studies using text mining: 2001—2004. *Journal of Information Science* **34**:3, 356-369. [Crossref]

1483. Ming-Shin Lu, Yu-Chun Wang, Jen-Hsiang Lin, Chao-Lin Liu, Zhao-Ming Gao, Chun-Yen Chang. 2008. Supporting the Translation and Authoring of Test Items with Techniques of Natural Language Processing. *Journal of Advanced Computational Intelligence and Intelligent Informatics* **12**:3, 234-242. [Crossref]

1484. W.S. Wong, R.W.P. Luk, H.V. Leong, K.S. Ho, D.L. Lee. 2008. Re-examining the effects of adding relevance information in a relevance feedback environment. *Information Processing & Management* **44**:3, 1086-1116. [Crossref]

1485. Y. Zhang, B. Feng. 2008. Clustering Search Results Based on Formal Concept Analysis. *Information Technology Journal* **7**:5, 746-753. [Crossref]

1486. D. Brezeale, D.J. Cook. 2008. Automatic Video Classification: A Survey of the Literature. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)* **38**:3, 416-430. [Crossref]

1487. Olena Medelyan, Ian H. Witten. 2008. Domain-independent automatic keyphrase indexing with small training sets. *Journal of the American Society for Information Science and Technology* **59**:7, 1026-1040. [Crossref]

1488. Ricardo Gacitua, Pete Sawyer. Ensemble Methods for Ontology Learning - An Empirical Experiment to Evaluate Combinations of Concept Acquisition Techniques 328-333. [Crossref]

1489. Xiaojun Wan, Jianwu Yang, Jianguo Xiao. 2008. Towards a unified approach to document similarity search using manifold-ranking of blocks. *Information Processing & Management* **44**:3, 1032-1048. [Crossref]

1490. Nian-Shing Chen, Kinshuk, Chun-Wang Wei, Hong-Jhe Chen. 2008. Mining e-Learning domain concept map from academic articles. *Computers & Education* **50**:3, 1009-1021. [Crossref]

1491. Xiaojun Wan. 2008. Beyond topical similarity: a structural similarity measure for retrieving highly similar documents. *Knowledge and Information Systems* **15**:1, 55-73. [Crossref]

1492. Marco Crasso, Alejandro Zunino, Marcelo Campo. 2008. Easy web service discovery: A query-by-example approach. *Science of Computer Programming* **71**:2, 144-164. [Crossref]

1493. Wen-Feng Hsiao, Te-Min Chang. 2008. An incremental cluster-based approach to spam filtering. *Expert Systems with Applications* **34**:3, 1599-1608. [Crossref]

1494. David E. Losada, Leif Azzopardi. 2008. An analysis on document length retrieval trends in language modeling smoothing. *Information Retrieval* **11**:2, 109-138. [Crossref]

1495. Ricardo Gacitua, Pete Sawyer, Paul Rayson. 2008. A flexible framework to experiment with ontology learning techniques. *Knowledge-Based Systems* **21**:3, 192-199. [Crossref]

1496. Joanna Yi-Hang Pong, Ron Chi-Wai Kwok, Raymond Yiu-Keung Lau, Jin-Xing Hao, Percy Ching-Chi Wong. 2008. A comparative study of two automatic document classification methods in a library setting. *Journal of Information Science* **34**:2, 213-230. [Crossref]

1497. Orland Hoeber, Xue Dong Yang. 2008. Evaluating WordBars in exploratory Web search scenarios. *Information Processing & Management* **44**:2, 485-510. [Crossref]

1498. Dennis van Heijst, Rob Potharst, Michiel van Wezel. 2008. A support system for predicting eBay end prices. *Decision Support Systems* **44**:4, 970-982. [Crossref]

1499. MAYY M. AL-TAHRAWI, RAED ABU ZITAR. 2008. POLYNOMIAL NETWORKS VERSUS OTHER TECHNIQUES IN TEXT CATEGORIZATION. *International Journal of Pattern Recognition and Artificial Intelligence* **22**:02, 295-322. [Crossref]

1500. Shanmugasundaram Hariharan, Rengaramanujam Srinivasan. 2008. A Comparison of Similarity Measures for Text Documents. *Journal of Information & Knowledge Management* **07**:01, 1-8. [Crossref]

1501. Daniel Gon, Tiago Guerreiro, Joaquim A. Jorge. Understanding Stories about Personal Documents 23-28. [Crossref]

1502. El-Sayed M. El-Alfy, Fares S. Al-Qunaieer. A fuzzy similarity approach for automated spam filtering 544-550. [Crossref]

1503. Samir Abdou, Jacques Savoy. 2008. Searching in Medline: Query expansion and manual indexing evaluation. *Information Processing & Management* **44**:2, 781-789. [Crossref]

1504. TIANMING HU, CHEW LIM TAN, YONG TANG, SAM YUAN SUNG, HUI XIONG, CHAO QU. 2008. CO-CLUSTERING BIPARTITE WITH PATTERN PRESERVATION FOR TOPIC EXTRACTION. *International Journal on Artificial Intelligence Tools* **17**:01, 87-107. [Crossref]

1505. CATARINA SILVA, BERNARDETE RIBEIRO. 2008. TOWARDS EXPANDING RELEVANCE VECTOR MACHINES TO LARGE SCALE DATASETS. *International Journal of Neural Systems* **18**:01, 45-58. [Crossref]

1506. Dengya Zhu, Heinz Dreher. Personalized information retrieval in digital ecosystems 580-585. [Crossref]

1507. Xiaojun Wan. 2008. Using only cross-document relationships for both generic and topic-focused multi-document summarizations. *Information Retrieval* **11**:1, 25-49. [Crossref]

1508. Ronald N. Kostoff, Jeffrey L. Solka, Robert L. Rushenberg, Jeffrey A. Wyatt. 2008. Literature-related discovery (LRD): Water purification. *Technological Forecasting and Social Change* **75**:2, 256-275. [Crossref]

1509. Sargur Srihari, Jim Collins, Rohini Srihari, Harish Srinivasan, Shravya Shetty, Janina Brutt-Griffler. 2008. Automatic scoring of short handwritten essays in reading comprehension tests. *Artificial Intelligence* **172**:2-3, 300-324. [Crossref]

1510. Bang-Rae Lee, Woon-Dong Yeo, June-Young Lee, Chang-Hoan Lee, Oh-Jin Kwon, Yeong-Ho Moon. 2008. Development of the KnowledgeMatrix as an Informetric Analysis System. *The Journal of the Korea Contents Association* **8**:1, 68-74. [Crossref]

1511. Ming Lin, Zhu Zhang. 2008. Question-driven segmentation of lecture speech text: Towards intelligent e-learning systems. *Journal of the American Society for Information Science and Technology* **59**:2, 186-200. [Crossref]

1512. Zhang Yun, Feng Boqin, Liu Lianmeng. Robust Self-Splitting Competitive learning for data clustering 302-307. [Crossref]

1513. J.M. Santos, J.M. de Sa, L.A. Alexandre. 2008. LEGClust—A Clustering Algorithm Based on Layered Entropic Subgraphs. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **30**:1, 62-75. [Crossref]

1514. Jacques Savoy. 2008. Searching strategies for the Hungarian language. *Information Processing & Management* **44**:1, 310-324. [Crossref]

1515. A. Zouaq, R. Nkambou. 2008. Building Domain Ontologies from Text for Educational Purposes. *IEEE Transactions on Learning Technologies* **1**:1, 49-62. [Crossref]

1516. Michael Chau, Hsinchun Chen. 2008. A machine learning approach to web page filtering using content and structure analysis. *Decision Support Systems* **44**:2, 482-494. [Crossref]

1517. Yueh-Min Huang, Juei-Nan Chen, Yen-Hung Kuo, Yu-Lin Jeng. 2008. An intelligent human-expert forum system based on fuzzy information retrieval technique. *Expert Systems with Applications* **34**:1, 446-458. [Crossref]

1518. Duen-Ren Liu, I-Chin Wu. 2008. Collaborative relevance assessment for task-based knowledge support. *Decision Support Systems* **44**:2, 524-543. [Crossref]

1519. Sadegh Aliakbary, Mahdy Khayyamian, Hassan Abolhassani. Using Social Annotations for Search Results Clustering 976-980. [Crossref]

1520. George Shih-Wen Ke, Michael P. Oakes, Marco A. Palomino, Yan Xu. Comparison between SVM-Light, a search engine-based approach and the mediamill baselines for assigning concepts to video shot annotations 381-387. [Crossref]

1521. M. Malenica, T. Šmuc, J. Šnajder, B. Dalbelo Bašić. 2008. Language morphology offset: Text classification on a Croatian–English parallel corpus. *Information Processing & Management* **44**:1, 325-339. [Crossref]

1522. Weijian Ni, Yalou Huang. Extracting and organizing acronyms based on ranking 4542-4547. [Crossref]

1523. Anton Karl Ingason, Sigrún Helgadóttir, Hrafn Loftsson, Eiríkur Rögnvaldsson. A Mixed Method Lemmatization Algorithm Using a Hierarchy of Linguistic Identities (HOLI) 205-216. [Crossref]

1524. Hiroya Takamura, Manabu Okumura. 2008. Text Summarization Model based on Maximum Coverage Problem and its Variant. *Transactions of the Japanese Society for Artificial Intelligence* **23**, 505-513. [Crossref]

1525. Cesar Koirala, Khaled Rasheed. Comparison of the Effects of Morphological and Ontological Information on Text Categorization 783-786. [Crossref]

1526. Shigeaki Sakurai, Ryohei Orihara. Classification Model Learning for Bulletin Board Site Analysis Based on Unbalanced Textual Examples 494-501. [Crossref]

1527. Diogo Krejci, Jonice Oliveira, Jano M. de Souza. The Use of Visualization for Analysis and Recommendation on People Replacement on Virtual Communities and Teams in the Brazilian Scientific Scenario 265-275. [Crossref]

1528. Florian Boudin, Juan-Manuel Torres-Moreno, Patricia Velázquez-Morales. An Efficient Statistical Approach for Automatic Organic Chemistry Summarization 89-99. [Crossref]

1529. Roberto Cornacchia, Sándor Héman, Marcin Zukowski, Arjen P. de Vries, Peter Boncz. 2008. Flexible and efficient IR using array databases. *The VLDB Journal* **17**:1, 151-168. [Crossref]

1530. Keith Noto, Milton H. Saier, Charles Elkan. Learning to Find Relevant Biological Articles without Negative Training Examples 202-213. [Crossref]

1531. Eun-Kyung Chung. 2007. A Study on Information Resource Evaluation for Text Categorization. *Journal of the Korean Society for information Management* **24**:4, 305-321. [Crossref]

1532. Vladimir Volkovich, Jacob Kogan, Charles Nicholas. 2007. Building initial partitions through sampling techniques. *European Journal of Operational Research* **183**:3, 1097-1105. [Crossref]

1533. C. E. Crangle, J. M. Cherry, E. L. Hong, A. Zbyslaw. 2007. Mining experimental evidence of molecular function claims from the literature. *Bioinformatics* **23**:23, 3232-3240. [Crossref]

1534. MUGE ERDOGMUS, OSMAN UGUR SEZERMAN. 2007. APPLICATION OF AUTOMATIC MUTATION–GENE PAIR EXTRACTION TO DISEASES. *Journal of Bioinformatics and Computational Biology* **05**:06, 1261-1275. [Crossref]

1535. Cherie Courseault Trumbach, Dinah Payne. 2007. Identifying synonymous concepts in preparation for technology mining. *Journal of Information Science* **33**:6, 660-677. [Crossref]

1536. Ram Dayal Goyal. Knowledge Based Neural Network for Text Classification 542-542. [Crossref]

1537. Shady Shehata, Fakhri Karray, Mohamed Kamel. Enhancing Search Engine Quality Using Concept-based Text Retrieval 26-32. [Crossref]

1538. Mirco Speretta, Susan Gauch. Automatic Ontology Identification for Reuse 419-422. [Crossref]

1539. Hoda K. Mohamed. Automatic documents classification 33-37. [Crossref]

1540. Orland Hoeber. ExploringWeb Search Results by Visually Specifying Utility Functions 650-654. [Crossref]

1541. Robert Wetzker, Tansu Alpcan, Christian Bauckhage, Winfried Umbrath, Sahin Albayrak. An unsupervised hierarchical approach to document categorization 482-486. [Crossref]

1542. JoËl Plisson, Peter Ljubi¿, Igor Mozeti¿, Nada Lavra¿. 2007. An Ontology for Virtual Organization Breeding Environments. *IEEE Transactions on Systems, Man and Cybernetics, Part C (Applications and Reviews)* **37**:6, 1327-1341. [Crossref]

1543. David Zajic, Bonnie J. Dorr, Jimmy Lin, Richard Schwartz. 2007. Multi-candidate reduction: Sentence compression as a tool for document summarization tasks. *Information Processing & Management* **43**:6, 1549-1570. [Crossref]

1544. Cheng Hua Li, Soon Cheol Park. Artificial Neural Network for Document Classification Using Latent Semantic Indexing 17-21. [Crossref]

1545. Diego Reforgiato Recupero. 2007. A new unsupervised method for document clustering by using WordNet lexical and conceptual relations. *Information Retrieval* **10**:6, 563-579. [Crossref]

1546. Jacques Savoy. 2007. Searching strategies for the Bulgarian language. *Information Retrieval* **10**:6, 509-529. [Crossref]

1547. Jay Urbain, Nazli Goharian, Ophir Frieder. Combining Semantics, Context, and Statistical Evidence in Genomics Literature Search 1313-1317. [Crossref]

1548. Giuliano Antoniol, Yann-Gael Gueheneuc, Ettore Merlo, Paolo Tonella. Mining the Lexicon Used by Programmers during Sofware Evolution 14-23. [Crossref]

1549. Junjie Wu, Hui Xiong, Jian Chen, Wenjun Zhou. A Generalization of Proximity Functions for K-Means 361-370. [Crossref]

1550. Kilian Thiel, Fabian Dill, Tobias Kotter, Michael R. Berthold. Towards Visual Exploration of Topic Shifts 522-527. [Crossref]

1551. Ana M. Cuadros, Fernando V. Paulovich, Rosane Minghim, Guilherme P. Telles. Point Placement by Phylogenetic Trees and its Application to Visual Analysis of Document Collections 99-106. [Crossref]

1552. Jebari Chaker, Ounelli Habib. Genre Categorization of Web Pages 455-464. [Crossref]

1553. Sylvain Lamprier, Tassadit Amghar, Bernard Levrat, Frederic Saubion. Document Length Normalization by Statistical Regression 11-18. [Crossref]

1554. Pu Wang, Jian Hu, Hua-Jun Zeng, Lijun Chen, Zheng Chen. Improving Text Classification by Using Encyclopedia Knowledge 332-341. [Crossref]

1555. Selim Mimaroglu, Kuniaki Uehara. Bit Sequences and Biclustering of Text Documents 51-56. [Crossref]

1556. Ying Liu, Yongjing Lin. Supervised HITS Algorithm for MEDLINE Citation Ranking 1323-1327. [Crossref]

1557. Jing Jiang, ChengXiang Zhai. 2007. An empirical study of tokenization strategies for biomedical information retrieval. *Information Retrieval* **10**:4-5, 341-363. [Crossref]

1558. H.C. Wu, Robert W.P. Luk, K.F. Wong, K.L. Kwok. 2007. A retrospective study of a hybrid document-context based retrieval model. *Information Processing & Management* **43**:5, 1308-1331. [Crossref]

1559. Cui Zifeng, Xu Baowen, Zhang Weifeng, Jiang Dawei, Xu Junling. CLDA: Feature Selection for Text Categorization Based on Constrained LDA 702-712. [Crossref]

1560. Yuen-Hsien Tseng, Chi-Jen Lin, Yu-I Lin. 2007. Text mining techniques for patent analysis. *Information Processing & Management* **43**:5, 1216-1247. [Crossref]

1561. Xiaojun Wan. 2007. A novel document similarity measure based on earth mover's distance. *Information Sciences* **177**:18, 3718-3730. [Crossref]

1562. G. Almpanidis, C. Kotropoulos, I. Pitas. 2007. Combining text and link analysis for focused crawling—An application for vertical search engines. *Information Systems* **32**:6, 886-908. [Crossref]

1563. Xin Li, Dawei Hu, Huan Li, Tianyong Hao, Enhong Chen, Wenyin Liu. 2007. Automatic question answering from Web documents. *Wuhan University Journal of Natural Sciences* **12**:5, 875-880. [Crossref]

1564. Sven Meyerzu Eissen, Benno Stein. An MDA Approach to Implement Personal IR Tools 259-263. [Crossref]

1565. Y. Lin, W. Li, K. Chen, Y. Liu. 2007. A Document Clustering and Ranking System for Exploring MEDLINE Citations. *Journal of the American Medical Informatics Association* **14**:5, 651-661. [Crossref]

1566. Yukihiro Tsuboshita, Hiroshi Okamoto. 2007. Context-dependent retrieval of information by neural-network dynamics with continuous attractors. *Neural Networks* **20**:6, 705-713. [Crossref]

1567. Limin Zeng, Yong Zhang, Robin G. Qiu. Adaptive User Profiling in Enhancing RSS-based Information Services 1-5. [Crossref]

1568. Ross S. Purves, Paul Clough, Christopher B. Jones, Avi Arampatzis, Benedicte Bucher, David Finch, Gaihua Fu, Hideo Joho, Awase Khirni Syed, Subodh Vaid, Bisheng Yang. 2007. The design and implementation of SPIRIT: a spatially aware search engine for information retrieval on the Internet. *International Journal of Geographical Information Science* **21**:7, 717-745. [Crossref]

1569. Kuan-Yu Chen, Luesak Luesukprasert, Seng-cho T. Chou. 2007. Hot Topic Extraction Based on Timeline Analysis and Multidimensional Sentence Modeling. *IEEE Transactions on Knowledge and Data Engineering* **19**:8, 1016-1025. [Crossref]

1570. Shiro Usui, Antoine Naud, Naonori Ueda, Tatsuki Taniguchi. 3D-SE Viewer: A Text Mining Tool based on Bipartite Graph Visualization 1103-1108. [Crossref]

1571. Tze-Haw Huang, Mao Lin Huang. Visualization of Individual's Knowledge by Analyzing the Citation Networks 465-470. [Crossref]

1572. Johan Bollen, Michael L. Nelson, Gary Geisler, Raquel Araujo. 2007. Usage derived recommendations for a video digital library. *Journal of Network and Computer Applications* **30**:3, 1059-1083. [Crossref]

1573. Yiyang Zhang, Jianxin (Roger) Jiao. 2007. An associative classification-based recommendation system for personalization in B2C e-commerce applications. *Expert Systems with Applications* **33**:2, 357-367. [Crossref]

1574. David B. Bracewell, Fuji Ren, Katsutoshi Hisazumi, Zhi Teng, Yuichi Furose, Shingo Kuroiwa. READING: A Self Sufficient Internet News System with Applications in Information and Knowledge Mining 190-196. [Crossref]

1575. Andrew Kwok-Fai Lui, Siu Cheung Li, Sheung On Choy. An Evaluation of Automatic Text Categorization in Online Discussion Analysis 205-209. [Crossref]

1576. Mohamed Salah Hamdi. Semantic Map Based Web Search Result Visualization 222-227. [Crossref]

1577. Yang Hu, Chen Hang. A Dynamic Reconfigurable Adaptive Software Architecture for Federate in HLA-based Simulation 682-687. [Crossref]

1578. Chaitali Gupta, Rajdeep Bhowmik, Michael R. Head, Madhusudhan Govindaraju, Weiyi Meng. A Query-based System for Automatic Invocation of Web Services 759-766. [Crossref]

1579. Duen-Ren Liu, Chih-Kun Ke. 2007. Knowledge support for problem-solving in a production process: A hybrid of knowledge discovery and case-based reasoning. *Expert Systems with Applications* **33**:1, 147-161. [Crossref]

1580. Ronan Cummins, Colm O'Riordan. 2007. An axiomatic comparison of learned term-weighting schemes in information retrieval: clarifications and extensions. *Artificial Intelligence Review* **28**:1, 51-68. [Crossref]

1581. Chirayu Wongchokprasitti, Peter Brusilovsky. NewsMe: A Case Study for Adaptive News Systems with Open User Model 69-69. [Crossref]

1582. Giuliano Armano, Andrea Manconi, Eloisa Vargiu. 2007. A Multiagent System for Retrieving Bioinformatics Publications From Web Sources. *IEEE Transactions on Nanobioscience* **6**:2, 104-109. [Crossref]

1583. Esma Aimeur, Flavien Serge Mani Onana, Anita Saleman. HELP: A Recommender System to Locate Expertise in Organizational Memories 866-874. [Crossref]

1584. M.K.M. Rahman, Wang Pi Yang, Tommy W.S. Chow, Sitao Wu. 2007. A flexible multi-layer self-organizing map for generic processing of tree-structured data. *Pattern Recognition* **40**:5, 1406-1424. [Crossref]

1585. Victor Rodriguez, Frizo Janssens, Koenraad Debackere, Bart De Moor. 2007. Do material transfer agreements affect the choice of research agendas? The case of biotechnology in Belgium. *Scientometrics* **71**:2, 239-269. [Crossref]

1586. MichaeL A. Covington, Wim J. Riedel, Cati Brown, Congzhou He, Eric Morris, Sara Weinstein, James SempLe, John Brown. 2007. Does ketamine mimic aspects of schizophrenic speech?. *Journal of Psychopharmacology* **21**:3, 338-346. [Crossref]

1587. H. Xu, J.-W. Fan, G. Hripcsak, E. A. Mendonca, M. Markatou, C. Friedman. 2007. Gene symbol disambiguation using knowledge-based profiles. *Bioinformatics* **23**:8, 1015-1022. [Crossref]

1588. Wenlei Mao, Wesley W. Chu. 2007. The phrase-based vector space model for automatic retrieval of free-text medical documents. *Data & Knowledge Engineering* **61**:1, 76-92. [Crossref]

1589. S. Usui, P. Palmes, K. Nagata, T. Taniguchi, N. Ueda. 2007. Keyword extraction, ranking, and organization for the neuroinformatics platform. *Biosystems* **88**:3, 334-342. [Crossref]

1590. Zehra Cataltepe, Eser Aygun. An Improvement of Centroid-Based Classification Algorithm for Text Classification 952-956. [Crossref]

1591. Yasunori Yamamoto, Toshihisa Takagi. 2007. Biomedical knowledge navigation by literature clustering. *Journal of Biomedical Informatics* **40**:2, 114-130. [Crossref]

1592. Narayan L. Bhamidipati, Sankar K. Pal. 2007. Stemming via Distribution-Based Word Segregation for Classification and Retrieval. *IEEE Transactions on Systems, Man and Cybernetics, Part B (Cybernetics)* **37**:2, 350-360. [Crossref]

1593. E. MONTAÑÉS, J. R. QUEVEDO, E. F. COMBARRO, I. DÍAZ, J. RANILLA. 2007. A HYBRID FEATURE SELECTION METHOD FOR TEXT CATEGORIZATION. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems* **15**:02, 133-151. [Crossref]

1594. Hong Yu, Won Kim, Vasileios Hatzivassiloglou, W. John Wilbur. 2007. Using MEDLINE as a knowledge source for disambiguating abbreviations and acronyms in full-text biomedical journal articles. *Journal of Biomedical Informatics* **40**:2, 150-159. [Crossref]

1595. Adrian Kuhn, Stéphane Ducasse, Tudor Gîrba. 2007. Semantic clustering: Identifying topics in source code. *Information and Software Technology* **49**:3, 230-243. [Crossref]

1596. Larry Manevitz, Malik Yousef. 2007. One-class document classification via Neural Networks. *Neurocomputing* **70**:7-9, 1466-1481. [Crossref]

1597. Massimo Melucci, Nicola Orio. 2007. Design, implementation, and evaluation of a methodology for automatic stemmer generation. *Journal of the American Society for Information Science and Technology* **58**:5, 673-686. [Crossref]

1598. R. Moskovitch, S. B. Martins, E. Behiri, A. Weiss, Y. Shahar. 2007. A Comparative Evaluation of Full-text, Concept-based, and Context-sensitive Search. *Journal of the American Medical Informatics Association* **14**:2, 164-174. [Crossref]

1599. Michelangelo Ceci, Donato Malerba. 2007. Classifying web documents in a hierarchy of categories: a comprehensive study. *Journal of Intelligent Information Systems* **28**:1, 37-78. [Crossref]

1600. Bruno Cassiman, Patrick Glenisson, Bart Van Looy. 2007. Measuring industry-science links through inventor-author relations: A profiling methodology. *Scientometrics* **70**:2, 379-391. [Crossref]

1601. Mike Thelwall, Rudy Prabowo. 2007. Identifying and characterizing public science-related fears from RSS feeds. *Journal of the American Society for Information Science and Technology* **58**:3, 379-390. [Crossref]

1602. Francisco P. Romero, Jos? A. Olivas, Pablo J. Garc?s. 2007. FzMail: Using FIS-CRM for E-mail Classification. *Journal of Advanced Computational Intelligence and Intelligent Informatics* **11**:1, 40-50. [Crossref]

1603. Elizabeth Leeds Hohman, David J. Marchette. A Dynamic Graph Model for Analyzing Streaming News Documents 462-469. [Crossref]

1604. Wen-feng Hsiao, Te-ming Chang, Guo-hsin Hu. A Cluster-based Approach to Filtering Spam under Skewed Class Distributions 53-53. [Crossref]

1605. Rudolf Mayer, Angela Roiger, Andreas Rauber. Supporting information management in digital libraries with map-based interfaces 55-60. [Crossref]

1606. Raymond Lau, Jin Hao, Maolin Tang, Xujuan Zhou. Towards Context-Sensitive Domain Ontology Extraction 60-60. [Crossref]

1607. Kazuaki Kishida. 2007. Term disambiguation techniques based on target document collection for cross-language information retrieval: An empirical comparison of performance between techniques. *Information Processing & Management* **43**:1, 103-120. [Crossref]

1608. Philipp Daumke, Kornél Markü, Michael Poprat, Stefan Schulz, Rüdiger Klar. 2007. Biomedical information retrieval across languages. *Medical Informatics and the Internet in Medicine* **32**:2, 131-147. [Crossref]

1609. Lukasz Itert, Wtodzislaw Duch, John Pestian. Influence of a priori Knowledge on Medical Document Categorization 163-170. [Crossref]

1610. Yulian Zhang, Chunxia Yin, Fuyong Yuan. An Application of Improved PageRank in Focused Crawler 331-335. [Crossref]

1611. Jun Won Lee, Yiu-Kai Ng. Using Fuzzy-Word Correlation Factors to Compute Document Similarity Based on Phrase Matching 186-192. [Crossref]

1612. Hsin-Chang Yang, Ding-Wen Chen, Chung-Hong Lee. Mining Multilingual Texts using Growing Hierarchical Self-Organizing Maps 2263-2268. [Crossref]

1613. Anubha Jain, Sujoy Das. Hindi Stemmer @ FIRE-2013 1-3. [Crossref]

1614. Keelan Evanini, David Suendermann, Roberto Pieraccini. Call classification for automated troubleshooting on large corpora 207-212. [Crossref]

1615. Zoe Zhuang, Rasika Amarasiri, Leonid Churilov, Damminda Alahakoon, Ken Sikaris. Exploring the Clinical Notes of Pathology Ordering by Australian General Practitioners: a text mining perspective 136-136. [Crossref]

1616. Marko A. Rodriguez, Johan Bollen, Herbert Van de Sompel. 2007. Mapping the bid behavior of conference referees. *Journal of Informetrics* **1**:1, 68-82. [Crossref]

1617. Debasis Ganguly, Johannes Leveling, Gareth J. F. Jones. DCU@Morpheme Extraction Task of FIRE-2012 1-5. [Crossref]

1618. Raymond K. Pon, Alfonso F. Cardenas, David J. Buttler, Terence J. Critchlow. iScore: Measuring the Interestingness of Articles in a Limited User Environment 354-361. [Crossref]

1619. Jan Martinovic, Tomas Novosad, Vaclav Snasel. Vector model improvement using suffix trees 180-187. [Crossref]

1620. David Pinto, José-Miguel Benedí, Paolo Rosso. Clustering Narrow-Domain Short Texts by Using the Kullback-Leibler Distance 611-622. [Crossref]

1621. Florian Boudin, Juan Manuel Torres Moreno. NEO-CORTEX: A Performant User-Oriented Multi-Document Summarization System 551-562. [Crossref]

1622. N. Okazaki, S. Ananiadou. 2006. Building an abbreviation dictionary using a term recognition approach. *Bioinformatics* **22**:24, 3089-3095. [Crossref]

1623. Sheng-tang Wu, Yuefeng Li, Yue Xu. Deploying Approaches for Pattern Refinement in Text Mining 1157-1161. [Crossref]

1624. Rasika Amarasiri, Damminda Alahakoon, Malin Premarathne. The Effect of Random Weight Updation in Dynamic Self Organizing Maps 183-188. [Crossref]

1625. Ying Pan, Xuhua Ding. Anomaly Based Web Phishing Page Detection 381-392. [Crossref]

1626. Chih-Ming Chen, Hahn-Ming Lee, Chia-Chen Tan. 2006. An intelligent web-page classifier with fair feature-subset selection. *Engineering Applications of Artificial Intelligence* **19**:8, 967-978. [Crossref]

1627. Ana G. Maguitman, Filippo Menczer, Fulya Erdinc, Heather Roinestad, Alessandro Vespignani. 2006. Algorithmic Computation and Approximation of Semantic Similarity. *World Wide Web* **9**:4, 431-456. [Crossref]

1628. Daniel Crabtree, Peter Andreae, Xiaoying Gao. Query Directed Web Page Clustering 202-210. [Crossref]

1629. A. Georgakis, H. Li. 2006. User behavior modeling and content based speculative web page prefetching. *Data & Knowledge Engineering* **59**:3, 770-788. [Crossref]

1630. Frizo Janssens, Jacqueline Leta, Wolfgang Glänzel, Bart De Moor. 2006. Towards mapping library and information science. *Information Processing & Management* **42**:6, 1614-1642. [Crossref]

1631. James Caverlee, Ling Liu, Daniel Rocco. 2006. Discovering Interesting Relationships among Deep Web Databases: A Source-Biased Approach. *World Wide Web* **9**:4, 585-622. [Crossref]

1632. Orland Hoeber, Xue Yang. Interactive Web Information Retrieval Using WordBars 875-882. [Crossref]

1633. Eamonn Keogh, Li Wei, Xiaopeng Xi, Stefano Lonardi, Jin Shieh, Scott Sirowy. Intelligent Icons: Integrating Lite-Weight Data Mining and Visualization into GUI Operating Systems 912-916. [Crossref]

1634. B. F. Momin, P. J. Kulkarni, Amol Chaudhari. Web Document Clustering Using Document Index Graph 32-37. [Crossref]

1635. Joel Reed, Yu Jiao, Thomas Potok, Brian Klump, Mark Elmore, Ali Hurson. TF-ICF: A New Term Weighting Scheme for Clustering Dynamic Data Streams 258-263. [Crossref]

1636. David Milne, Olena Medelyan, Ian Witten. Mining Domain-Specific Thesauri from Wikipedia: A Case Study 442-448. [Crossref]

1637. Rashid Ali, M. M. Sufyan Beg. Aggregating Subjective and Objective Measures of Web Search Quality using Modified Shimura Technique 231-234. [Crossref]

1638. Shady Shehata, Fakhri Karray, Mohamed Kamel. Enhancing Text Clustering Using Concept-based Mining Model 1043-1048. [Crossref]

1639. Carmen Galvez. 2006. Aplicación de transductores de estado-finito a los procesos de unificación de términos. *Ciência da Informação* **35**:3, 67-74. [Crossref]

1640. Russell Beale, Robert J Hendley, Andy Pryke, Barry Wilkins. 2006. Nature-Inspired Visualisation of Similarity and Relationships in Human Systems and Behaviours. *Information Visualization* **5**:4, 260-270. [Crossref]

1641. J. I. Serrano, M. D. del Castillo. 2006. Evolutionary learning of document categories. *Information Retrieval* **10**:1, 69-83. [Crossref]

1642. David E. Losada, Alvaro Barreiro. 2006. Negations and document length in logical retrieval. *Information Systems* **31**:7, 610-620. [Crossref]

1643. Cheryl Aasheim, Gary J. Koehler. 2006. Scanning World Wide Web documents with the vector space model. *Decision Support Systems* **42**:2, 690-699. [Crossref]

1644. Samhaa El-beltagy. KP-Miner: A Simple System for Effective Keyphrase Extraction 1-5. [Crossref]

1645. Tianming Hu, Chao Qu, Chew Tan, Sam Sung, Wenjun Zhou. Preserving Patterns in Bipartite Graph Partitioning 489-496. [Crossref]

1646. Ramon F. Brena, Eduardo H. Ramirez. A Soft Semantic Web 1-8. [Crossref]

1647. Ehab Mohamed, Samhaa El-beltagy, Salwa El-Gamal. A Feature Reduction Technique for Improved Web Page Clustering 1-5. [Crossref]

1648. Ian Garcia, Yiu-kai Ng. Eliminating Redundant and Less-Informative RSS News Articles Based on Word Similarity and a Fuzzy Equivalence Relation 465-473. [Crossref]

1649. Renato Correa, Teresa Ludermir. A Hybrid SOM-Based Document Organization System 16-16. [Crossref]

1650. Aekkasit Chongsuntornsri, Ohm Sornil. An Automatic Thai Text Summarization Using Topic Sensitive PageRank 547-552. [Crossref]

1651. A. Vinciarelli, J.-M. Odobez. 2006. Application of Information Retrieval Technologies to Presentation Slides. *IEEE Transactions on Multimedia* **8**:5, 981-995. [Crossref]

1652. Chun-Wei Tsai, Ting-Wen Liang, Jiun-Huei Ho, Chu-Sing Yang, Ming-Chao Chiang. A Document Clustering Approach for Search Engines 1050-1055. [Crossref]

1653. Duangjai Thammasut, Ohm Sornil. A Graph-Based Information Retrieval System 743-748. [Crossref]

1654. Zane Reynolds, Horst Bunke, Mark Last, Abraham Kandel. 2006. Comparing representative selection strategies for dissimilarity representations. *International Journal of Intelligent Systems* **21**:10, 1093-1109. [Crossref]

1655. Katrin Kirchhoff, Dimitra Vergyri, Jeff Bilmes, Kevin Duh, Andreas Stolcke. 2006. Morphology-based language modeling for conversational Arabic speech recognition. *Computer Speech & Language* **20**:4, 589-608. [Crossref]

1656. Yen-Liang Chen, Jhong-Jhih Wei, Shin-Yi Wu, Ya-Han Hu. 2006. A similarity-based method for retrieving documents from the SCI/SSCI database. *Journal of Information Science* **32**:5, 449-464. [Crossref]

1657. Robert M. Losee. 2006. Is 1 noun worth 2 adjectives? Measuring relative feature utility. *Information Processing & Management* **42**:5, 1248-1259. [Crossref]

1658. Andrea Lucia, Rocco Oliveto, Paola Sgueglia. Incremental Approach and User Feedbacks: a Silver Bullet for Traceability Recovery 299-309. [Crossref]

1659. Olga Vechtomova, Murat Karamuftuoglu, Stephen E. Robertson. 2006. On document relevance and lexical cohesion between query terms. *Information Processing & Management* **42**:5, 1230-1247. [Crossref]

1660. B. Han, Z. Obradovic, Z.-Z. Hu, C. H. Wu, S. Vucetic. 2006. Substring selection for biomedical document classification. *Bioinformatics* **22**:17, 2136-2142. [Crossref]

1661. Cui Zifeng, Xu Baowen, Zhang Weifeng, Xu Junling. 2006. A new approach of feature selection for text categorization. *Wuhan University Journal of Natural Sciences* **11**:5, 1335-1339. [Crossref]

1662. M. Ceccato, M. Marin, K. Mens, L. Moonen, P. Tonella, T. Tourwé. 2006. Applying and combining three different aspect Mining Techniques. *Software Quality Journal* **14**:3, 209-231. [Crossref]

1663. X. Lu, B. Zheng, A. Velivelli, C. Zhai. 2006. Enhancing Text Categorization with Semantic-enriched Representation and Training Data Augmentation. *Journal of the American Medical Informatics Association* **13**:5, 526-535. [Crossref]

1664. Mark Goadrich, Louis Oliphant, Jude Shavlik. 2006. Gleaner: Creating ensembles of first-order clauses to improve recall-precision curves. *Machine Learning* **64**:1-3, 231-261. [Crossref]

1665. Andreas Nürnberger, Marcin Detyniecki. 2006. Externally growing self-organizing maps and its application to e-mail database visualization and exploration. *Applied Soft Computing* **6**:4, 357-371. [Crossref]

1666. Mike Thelwall, Liz Price. 2006. Language evolution and the spread of ideas on the Web: A procedure for identifying emergent hybrid word family members. *Journal of the American Society for Information Science and Technology* **57**:10, 1326-1337. [Crossref]

1667. Xiaohui Cui, Jinzhu Gao, Thomas E. Potok. 2006. A flocking based algorithm for document clustering analysis. *Journal of Systems Architecture* **52**:8-9, 505-515. [Crossref]

1668. Honglin Li, Aijuan Dong. Hierarchical Segmentation of Presentation Videos through Visual and Text Analysis 314-319. [Crossref]

1669. Zhiguo Gong, Leong Hou U., Chan Wa Cheang. 2006. Web image indexing by using associated texts. *Knowledge and Information Systems* **10**:2, 243-264. [Crossref]

1670. Peter Willett. 2006. The Porter stemming algorithm: then and now. *Program* **40**:3, 219-223. [Abstract] [Full Text] [PDF]

1671. Ernesto William De Luca, Andreas Nürnberger. 2006. Using clustering methods to improve ontology-based query term disambiguation. *International Journal of Intelligent Systems* **21**:7, 693-709. [Crossref]

1672. Brieuc Conan-Guez, Fabrice Rossi, Aïcha El Golli. 2006. Fast algorithm and implementation of dissimilarity self-organizing maps. *Neural Networks* **19**:6-7, 855-863. [Crossref]

1673. Chun-Yan Liang, Li Guo, Zhao-Jie Xia, Feng-Guang Nie, Xiao-Xia Li, Liang Su, Zhang-Yuan Yang. 2006. Dictionary-based text categorization of chemical web pages. *Information Processing & Management* **42**:4, 1017-1029. [Crossref]

1674. Y. Aphinyanaphongs, A. Statnikov, C. F. Aliferis. 2006. A Comparison of Citation Metrics to Machine Learning Filters for the Identification of High Quality MEDLINE Documents. *Journal of the American Medical Informatics Association* **13**:4, 446-455. [Crossref]

1675. Juan Jose Garcia Adeva, Nicholas L. Carroll, Rafael A. Calvo. Applying Plagiarism Detection to Engineering Education 722-731. [Crossref]

1676. Robert Moskovitch, Shiva Cohen-Kashi, Uzi Dror, Iftah Levy, Amit Maimon, Yuval Shahar. 2006. Multiple hierarchical classification of free-text clinical guidelines. *Artificial Intelligence in Medicine* **37**:3, 177-190. [Crossref]

1677. Ohm Sornil, Kornnika Gree-ut. An Automatic Text Summarization Approach using Content-Based and Graph-Based Characteristics 1-6. [Crossref]

1678. Ronan Cummins, Colm O'Riordan. 2006. Evolving local and global weighting schemes in information retrieval. *Information Retrieval* **9**:3, 311-330. [Crossref]

1679. Youjin Chang, Minkoo Kim, Vijay V. Raghavan. 2006. Construction of query concepts based on feature clustering of documents. *Information Retrieval* **9**:3, 231-248. [Crossref]

1680. Ching-Chi Hsu, Fan Wu. 2006. Topic-specific crawling on the Web with the measurements of the relevancy context graph. *Information Systems* **31**:4-5, 232-246. [Crossref]

1681. Daniela Godoy, Analía Amandi. 2006. Modeling user interests by conceptual clustering. *Information Systems* **31**:4-5, 247-265. [Crossref]

1682. Carmen Galvez, Félix de Moya-Anegón. 2006. An evaluation of conflation accuracy using finite-state transducers. *Journal of Documentation* **62**:3, 328-349. [Abstract] [Full Text] [PDF]

1683. Tuomas Talvensaari, Jorma Laurikkala, Kalervo Järvelin, Martti Juhola. 2006. A study on automatic creation of a comparable document collection in cross-language information retrieval. *Journal of Documentation* **62**:3, 372-387. [Abstract] [Full Text] [PDF]

1684. Amy Trappey, Charles Trappey, Burgess S. Kao. Automated Patent Document Summarization for R&D Intellectual Property Management 1-6. [Crossref]

1685. Adenike M. Lam-Adesina, Gareth J.F. Jones. 2006. Examining and improving the effectiveness of relevance feedback for retrieval of scanned text documents. *Information Processing & Management* **42**:3, 633-649. [Crossref]

1686. S. Srinivasan ., P. Thambidurai .. 2006. STANS Algorithm for Root Word Stemming. *Information Technology Journal* **5**:4, 685-688. [Crossref]

1687. H. Xiong, Gaurav Pandey, M. Steinbach, Vipin Kumar. 2006. Enhancing data analysis with noise removal. *IEEE Transactions on Knowledge and Data Engineering* **18**:3, 304-319. [Crossref]

1688. XAVIER LLORÀ, DAVID E. GOLDBERG, YUKIO OHSAWA, NAOHIRO MATSUMURA, YUICHI WASHIDA, HIROSHI TAMURA, MASATAKA YOSHIKAWA, MICHAEL WELGE, LORETTA AUVIL, DUANE SEARSMITH, KEI OHNISHI, CHEN-JU CHAO. 2006. INNOVATION AND CREATIVITY SUPPORT VIA CHANCE DISCOVERY, GENETIC ALGORITHMS, AND DATA MINING. *New Mathematics and Natural Computation* **02**:01, 85-100. [Crossref]

1689. HANI ABU-SALEM, PHILIP KA-FAI CHAN. 2006. English-Arabic Cross-Language Information Retrieval Based on Parallel Documents. *International Journal of Computer Processing of Languages* **19**:01, 21-37. [Crossref]

1690. Kieran J. White, Richard F. E. Sutcliffe. 2006. Applying incremental tree induction to retrieval from manuals and medical texts. *Journal of the American Society for Information Science and Technology* **57**:5, 588-600. [Crossref]

1691. Youjin Chang, Iadh Ounis, Minkoo Kim. 2006. Query reformulation using automatically generated query concepts from a document space. *Information Processing & Management* **42**:2, 453-468. [Crossref]

1692. H. S. Ng, A. Toukourou, L. Soibelman. 2006. Knowledge Discovery in a Facility Condition Assessment Database Using Text Clustering. *Journal of Infrastructure Systems* **12**:1, 50-59. [Crossref]

1693. Stephan Bloehdorn, Andreas Hotho. Boosting for Text Classification with Semantic Features 149-166. [Crossref]

1694. F. Feger, I. Koprinska. Co-training using RBF Nets and Different Feature Splits 1878-1885. [Crossref]

1695. Olga Vechtomova, Murat Karamuftuoglu. 2006. Elicitation and use of relevance feedback information. *Information Processing & Management* **42**:1, 191-206. [Crossref]

1696. Cheng Li, Soon Park. High Performance Text Categorization System Based on a Novel Neural Network Algorithm 21-21. [Crossref]

1697. A. Kiani-B, M.R. Akbarzadeh-T. Automatic Text Summarization Using Hybrid Fuzzy GA-GP 977-983. [Crossref]

1698. Mustapha Baziz, Mohand Boughanem, Henri Prade, Gabriella Pasi. Chapter 18 A fuzzy logic approach to information retrieval using an ontology-based representation of documents 363-377. [Crossref]

1699. Xin Zhou, Yunlong Xu, Gongming Chen, Zhigeng Pan. A New Wildcard Search Method for Digital Dictionary Based on Mobile Platform 699-704. [Crossref]

1700. S. Usui, P. Palmes, K. Nagata, T. Taniguchi, N. Ueda. Extracting Keywords from Research Abstracts for the Neuroinformatics Platform Index Tree 5045-5050. [Crossref]

1701. Tayebeh Mosavi Miangah. 2006. Automatic lemmatization of Persian words*. *Journal of Quantitative Linguistics* **13**:1, 1-15. [Crossref]

1702. Yi Peng, Gang Kou, Yong Shi, Zhengxin Chen. A Hybrid Strategy for Clustering Data Mining Documents 838-842. [Crossref]

1703. Tao Wang, Bipin Desai. Extracting Document Semantics for Semantic Header 1878-1883. [Crossref]

1704. G.S. Tomar, Shekhar Verma, Ashish Jha. Web Page Classification using Modified Na�ve Bayesian Approach 1-4. [Crossref]

1705. A. Scharl, A. Weichselbraun. An Ontology-Based Architecture for Tracking Information across Interactive Electronic Environments 160b-160b. [Crossref]

1706. H. Al-Mubaid, Ping Chen. Biomedical Term Disambiguation: An Application to Gene-Protein Name Disambiguation 606-612. [Crossref]

1707. Ting Yu, T. Jan, J. Debenham, S. Simoff. Classify Unexpected News Impacts to Stock Price by Incorporating Time Series Analysis into Support Vector Machine 2993-2998. [Crossref]

1708. Amruta Joshi, Rajeev Motwani. Keyword Generation for Search Engine Advertising 490-496. [Crossref]

1709. Carlos Tobar, Claudio Oliveira, Ricardo De Freitas. An Explanation Tool to Support Learning of Basic Programming 21-26. [Crossref]

1710. Mohand Boughanem, Yannick Loiseau, Henri Prade. Rank-Ordering Documents According to Their Relevance in Information Retrieval Using Refinements of Ordered-Weighted Aggregations 44-54. [Crossref]

1711. K. Lakshmi, S. Mukherjee. An Improved Feature Selection using Maximized Signal to Noise Ratio Technique for TC 541-546. [Crossref]

1712. Eugen Popovici, Gildas Ménier, Pierre-François Marteau. SIRIUS: A Lightweight XML Indexing and Approximate Search System at INEX 2005 321-335. [Crossref]

1713. EunKyung Chung, Samantha K. Hastings. 2006. A conception-based approach to automatic subject term assignment for scientific journal articles. *Proceedings of the American Society for Information Science and Technology* **43**:1, 1-21. [Crossref]

1714. Hui Yang, Minjie Zhang. 2006. Two-stage statistical language models for text database selection. *Information Retrieval* **9**:1, 5-31. [Crossref]

1715. C.D. Paice. Stemming 149-150. [Crossref]

1716. Kazunari Sugiyama, Kenji Hatano, Masatoshi Yoshikawa, Shunsuke Uemura. 2005. Improvement in TF-IDF scheme for Web pages based on the contents of their hyperlinked neighboring pages. *Systems and Computers in Japan* **36**:14, 56-68. [Crossref]

1717. A. Vinciarelli. 2005. Noisy text categorization. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **27**:12, 1882-1895. [Crossref]

1718. Christian R. Huyck, Viviane Orengo. 2005. Information Retrieval and Categorisation using a Cell Assembly Network. *Neural Computing and Applications* **14**:4, 282-289. [Crossref]

1719. Jing Chen, Qing Li, Weijia Jia. 2005. Automatically Generating an E-textbook on the Web. *World Wide Web* **8**:4, 377-394. [Crossref]

1720. Patrick Glenisson, Wolfgang Glänzel, Frizo Janssens, Bart De Moor. 2005. Combining full text and bibliometric information in mapping scientific disciplines. *Information Processing & Management* **41**:6, 1548-1572. [Crossref]

1721. Kevyn Collins-Thompson, Jamie Callan. 2005. Predicting reading difficulty with statistical language models. *Journal of the American Society for Information Science and Technology* **56**:13, 1448-1462. [Crossref]

1722. Ronan Cummins, Colm O'riordan. 2005. Evolving General Term-Weighting Schemes for Information Retrieval: Tests on Larger Collections. *Artificial Intelligence Review* **24**:3-4, 277-299. [Crossref]

1723. Matevz Pogacnik, Jurij Tasic, Marko Meza, Andrej Kosir. 2005. Personal Content Recommender Based on a Hierarchical User Model for the Selection of TV Programmes. *User Modeling and User-Adapted Interaction* **15**:5, 425-457. [Crossref]

1724. J.H. Hayes, A. Dekhtyar, S.K. Sundaram. 2005. Improving After-the-Fact Tracing and Mapping: Supporting Software Quality Predictions. *IEEE Software* **22**:6, 30-37. [Crossref]

1725. Naohiro Matsumura, Yukio Ohsawa, Mitsuru Ishizuka. 2005. Combination retrieval for creating knowledge from sparse document-collection. *Knowledge-Based Systems* **18**:7, 327-333. [Crossref]

1726. Alessandro Vinciarelli. 2005. Application of information retrieval techniques to single writer documents. *Pattern Recognition Letters* **26**:14, 2262-2271. [Crossref]

1727. Lee Begeja, H. Drucker, D. Gibbon, P. Haffner, Zhu Liu, B. Renger, B. Shahraray. 2005. Semantic data mining of short utterances. *IEEE Transactions on Speech and Audio Processing* **13**:5, 672-680. [Crossref]

1728. Jieping Ye, Qi Li, Hui Xiong, H. Park, R. Janardan, V. Kumar. 2005. IDR/QR: an incremental dimension reduction algorithm via QR decomposition. *IEEE Transactions on Knowledge and Data Engineering* **17**:9, 1208-1222. [Crossref]

1729. J. Caverlee, L. Liu. 2005. QA-Pagelet: data preparation techniques for large-scale data analysis of the deep Web. *IEEE Transactions on Knowledge and Data Engineering* **17**:9, 1247-1262. [Crossref]

1730. E.F. Combarro, E. Montanes, I. Diaz, J. Ranilla, R. Mones. 2005. Introducing a family of linear measures for feature selection in text categorization. *IEEE Transactions on Knowledge and Data Engineering* **17**:9, 1223-1232. [Crossref]

1731. Carmen Galvez, Félix de Moya-Anegón, Víctor H. Solana. 2005. Term conflation methods in information retrieval. *Journal of Documentation* **61**:4, 520-547. [Abstract] [Full Text] [PDF]

1732. Kimmo Kettunen, Tuomas Kunttu, Kalervo Järvelin. 2005. To stem or lemmatize a highly inflectional language in a probabilistic IR environment?. *Journal of Documentation* **61**:4, 476-496. [Abstract] [Full Text] [PDF]

1733. Johan Gustav Bellika, Gunnar Hartvigsen. 2005. The oncological nurse assistant: A web-based intelligent oncological nurse advisor. *International Journal of Medical Informatics* **74**:7-8, 587-595. [Crossref]

1734. Ronald N. Kostoff, Joel A. Block. 2005. Factor matrix text filtering and clustering. *Journal of the American Society for Information Science and Technology* **56**:9, 946-968. [Crossref]

1735. Bowen Hui, Eric Yu. 2005. Extracting conceptual relationships from specialized documents. *Data & Knowledge Engineering* **54**:1, 29-55. [Crossref]

1736. Xiao-Jun Wan, Yu-Xin Peng. 2005. A New Retrieval Model Based on TextTiling for Document Similarity Search. *Journal of Computer Science and Technology* **20**:4, 552-558. [Crossref]

1737. Jacques Savoy. 2005. Bibliographic database access using free-text and controlled vocabulary: an evaluation. *Information Processing & Management* **41**:4, 873-890. [Crossref]

1738. Carly W.Y. Wong, Robert W.P. Luk, Edward K.S. Ho. 2005. Discovering "title-like" terms. *Information Processing & Management* **41**:4, 789-800. [Crossref]

1739. D. Godoy, A. Amandi. 2005. User Profiling for Web Page Filtering. *IEEE Internet Computing* **9**:4, 56-64. [Crossref]

1740. Jieping Ye, Qi Li. 2005. A two-stage linear discriminant analysis via QR-decomposition. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **27**:6, 929-941. [Crossref]

1741. Adam Schenker, Mark Last, Abraham Kandel. 2005. Design and implementation of a web mining system for organizing search engine results. *International Journal of Intelligent Systems* **20**:6, 607-625. [Crossref]

1742. Liz Price, Mike Thelwall. 2005. The clustering power of low frequency words in academic Webs. *Journal of the American Society for Information Science and Technology* **56**:8, 883-888. [Crossref]

1743. Kazuaki Kishida. 2005. Technical issues of cross-language information retrieval: a review. *Information Processing & Management* **41**:3, 433-455. [Crossref]

1744. E. Montanes, I. Diaz, J. Ranilla, E.F. Combarro, J. Fernandez. 2005. Scoring and Selecting Terms for Text Categorization. *IEEE Intelligent Systems* **20**:3, 40-47. [Crossref]

1745. Michael Schulte-Mecklenbeck, Moritz Neun. 2005. WebDiP: A tool for information search experiments on the World-Wide Web. *Behavior Research Methods* **37**:2, 293-300. [Crossref]

1746. Gina-Anne Levow, Douglas W. Oard, Philip Resnik. 2005. Dictionary-based techniques for cross-language information retrieval. *Information Processing & Management* **41**:3, 523-547. [Crossref]

1747. Jinxi Xu, Ralph Weischedel. 2005. Empirical studies on the impact of lexical resources on CLIR performance. *Information Processing & Management* **41**:3, 475-487. [Crossref]

1748. Vassilis Plachouras, Iadh Ounis. 2005. Dempster-Shafer Theory for a Query-Biased Combination of Evidence on the Web. *Information Retrieval* **8**:2, 197-218. [Crossref]

1749. Hugh E. Williams, Justin Zobel. 2005. Searchable words on the Web. *International Journal on Digital Libraries* **5**:2, 99-105. [Crossref]

1750. Mike Thelwall. 2005. Text characteristics of English language university Web sites. *Journal of the American Society for Information Science and Technology* **56**:6, 609-619. [Crossref]

1751. Mark P. Sinka, David W. Corne. 2005. The BankSearch web document dataset: investigating unsupervised clustering and category similarity. *Journal of Network and Computer Applications* **28**:2, 129-146. [Crossref]

1752. Camille Roth, Paul Bourgine. 2005. Epistemic Communities: Description and Hierarchic Categorization. *Mathematical Population Studies* **12**:2, 107-130. [Crossref]

1753. Ying Zhao, George Karypis, Usama Fayyad. 2005. Hierarchical Clustering Algorithms for Document Datasets. *Data Mining and Knowledge Discovery* **10**:2, 141-168. [Crossref]

1754. Krishnaprasad Thirunarayan, Aaron Berkovich, Dan Z. Sokol. 2005. An information extraction approach to reorganizing and summarizing specifications. *Information and Software Technology* **47**:4, 215-232. [Crossref]

1755. A. Vailaya, P. Bluvas, R. Kincaid, A. Kuchinsky, M. Creech, A. Adler. 2005. An architecture for biological information extraction and representation. *Bioinformatics* **21**:4, 430-438. [Crossref]

1756. Rajiv Yerra, Yiu-Kai Ng. Detecting similar HTML documents using a fuzzy set information retrieval approach 693-699 Vol. 2. [Crossref]

1757. K. Taghva, R. Beckley, M. Sadeh. A stemming algorithm for the Farsi language 158-162 Vol. 1. [Crossref]

1758. Min-Yuh Day, Chun-Hung Lu, Jin-Tan David Yang, Guey-Fa Chiou, Chorng Shyong Ong, Wen-Lian Hsu. Designing an ontology-based intelligent tutoring agent with instant messaging 318-320. [Crossref]

1759. G. Potamias, L. Koumakis, V.S. Moustakis. Enhancing Web Based Services by Coupling Document Classification with User Profile 205-208. [Crossref]

1760. Xiao-Lin Wang, Cloete. Learning to classify email: a survey 5716-5719 Vol. 9. [Crossref]

1761. Richard T. Freeman, Hujun Yin. 2005. Tree view self-organisation of web content. *Neurocomputing* **63**, 415-446. [Crossref]

1762. P. Srinivasan, F. Menczer, G. Pant. 2005. A General Evaluation Framework for Topical Crawlers. *Information Retrieval* **8**:3, 417-447. [Crossref]

1763. X. Cui, T.E. Potok, P. Palathingal. Document clustering using particle swarm optimization 185-191. [Crossref]

1764. Michela Bacchin, Nicola Ferro, Massimo Melucci. 2005. A probabilistic model for stemmer generation. *Information Processing & Management* **41**:1, 121-137. [Crossref]

1765. . References 387-411. [Crossref]

1766. S. Agarwal, P. Agarwal. A Fuzzy Logic Approach to Search Results&#146; Personalization by Tracking User&#146;s Web Navigation Pattern and Psychology 318-325. [Crossref]

1767. S. Watanabe, T. Ito, T. Ozono, T. Shintani. A paper recommendation mechanism for the research support system Papits 71-80. [Crossref]

1768. N. Polavarapu, S.B. Navathe, R. Ramnarayanan, A. ul Haque, S. Sahay, Y. Liu. Investigation into biomedical literature classification using support vector machines 366-374. [Crossref]

1769. Andreas Girgensohn, John Adcock, Matthew Cooper, Lynn Wilcox. A Synergistic Approach to Efficient Interactive Video Retrieval 781-794. [Crossref]

1770. L. Smith, W.J. Wilbur. 2004. Retrieving definitional content for ontology development. *Computational Biology and Chemistry* **28**:5-6, 387-391. [Crossref]

1771. PAWEL BRZEMINSKI, WITOLD PEDRYCZ. 2004. TEXTUAL-BASED CLUSTERING OF WEB DOCUMENTS. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems* **12**:06, 715-743. [Crossref]

1772. Filippo Menczer. 2004. Lexical and semantic clustering by Web links. *Journal of the American Society for Information Science and Technology* **55**:14, 1261-1269. [Crossref]

1773. Kai R. Larsen, David E. Monarchi. 2004. 9. A Mathematical Approach to Categorization and Labeling of Qualitative Data: The Latent Categorization Method. *Sociological Methodology* **34**:1, 349-392. [Crossref]

1774. Sandro Araya, Mariano Silva, Richard Weber. 2004. A methodology for web usage mining and its application to target group identification. *Fuzzy Sets and Systems* **148**:1, 139-152. [Crossref]

1775. M.J. Martín-Bautista, D. Sánchez, J. Chamorro-Martínez, J.M. Serrano, M.A. Vila. 2004. Mining web documents to find additional query terms using fuzzy association rules. *Fuzzy Sets and Systems* **148**:1, 85-104. [Crossref]

1776. Khaled M. Hammouda, Mohamed S. Kamel. 2004. Document Similarity Using a Phrase Indexing Graph Model. *Knowledge and Information Systems* **6**:6, 710-727. [Crossref]

1777. K.M. Hammouda, M.S. Kamel. 2004. Efficient phrase-based document indexing for Web document clustering. *IEEE Transactions on Knowledge and Data Engineering* **16**:10, 1279-1296. [Crossref]

1778. Ian Witten. Text Mining . [Crossref]

1779. Daniela Godoy, Silvia Schiaffino, Analía Amandi. 2004. Interface agents personalizing Web-based tasks. *Cognitive Systems Research* **5**:3, 207-222. [Crossref]

1780. Jieping Ye, R. Janardan, C.H. Park, H. Park. 2004. An optimization criterion for generalized discriminant analysis on undersampled problems. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **26**:8, 982-994. [Crossref]

1781. C. Clifton, R. Cooley, J. Rennie. 2004. TopCat: data mining for topic identification in a text corpus. *IEEE Transactions on Knowledge and Data Engineering* **16**:8, 949-964. [Crossref]

1782. K. Shima, M. Todoriki, A. Suzuki. 2004. SVM-based feature selection of latent semantic features. *Pattern Recognition Letters* **25**:9, 1051-1057. [Crossref]

1783. L.A.F. Park, K. Ramamohanarao, M. Palaniswami. 2004. Fourier domain scoring: a novel document ranking method. *IEEE Transactions on Knowledge and Data Engineering* **16**:5, 529-539. [Crossref]

1784. Irene Díaz, José Ranilla, Elena Montañes, Javier Fernández, Elías F. Combarro. 2004. Improving performance of text categorization by combining filtering and support vector machines. *Journal of the American Society for Information Science and Technology* **55**:7, 579-592. [Crossref]

1785. A. Georgakis, C. Kotropoulos, A. Xafopoulos, I. Pitas. 2004. Marginal median SOM for document organization and retrieval. *Neural Networks* **17**:3, 365-377. [Crossref]

1786. Y. MATSUO, M. ISHIZUKA. 2004. KEYWORD EXTRACTION FROM A SINGLE DOCUMENT USING WORD CO-OCCURRENCE STATISTICAL INFORMATION. *International Journal on Artificial Intelligence Tools* **13**:01, 157-169. [Crossref]

1787. Zachary J. Mason. 2004. CorMet: A Computational, Corpus-Based Conventional Metaphor Extraction System. *Computational Linguistics* **30**:1, 23-44. [Crossref]

1788. Imad A. Al-Sughaiyer, Ibrahim A. Al-Kharashi. 2004. Arabic morphological analysis techniques: A comprehensive survey. *Journal of the American Society for Information Science and Technology* **55**:3, 189-213. [Crossref]

1789. Konstantinos G. Zerfiridis, Helen D. Karatza. 2004. Brute force web search for wireless devices using mobile agents. *Journal of Systems and Software* **69**:1-2, 195-206. [Crossref]

1790. Min Song, Il-Yeol Song, Xiaohua Hu. 2004. Designing and developing an automatic interactive keyphrase extraction system with Unified Modeling Language (UML). *Proceedings of the American Society for Information Science and Technology* **41**:1, 367-372. [Crossref]

1791. Lucia Helena Machado Rino, Marcelo Módolo. SuPor: An Environment for AS of Texts in Brazilian Portuguese 419-430. [Crossref]

1792. Nuanwan Soonthornphisaj, Boonserm Kijsirikul. 2004. Iterative cross-training: An algorithm for learning from unlabeled Web pages. *International Journal of Intelligent Systems* **19**:1-2, 131-147. [Crossref]

1793. Ming Lin, J.F. Nunamaker, M. Chau, Hsinchun Chen. Segmentation of lecture videos based on text: a method combining multiple linguistic features 9 pp.. [Crossref]

1794. Manabu Matsuyama, Satoshi Shintani, Takayuki Ito, Yusuke Hiraoka, Satoshi Watanabe. 2004. A Keyword Extraction Method for generating a User Profile in the Paper Collection and Sharing System MiDoc. *IEEJ Transactions on Electronics, Information and Systems* **124**:12, 2489-2494. [Crossref]

1795. K.H. Lee, B.H. Kang. A new framework for uncertainty sampling: exploiting uncertain and positive-certain examples in similarity-based text classification 474-479 Vol.2. [Crossref]

1796. D. Mallett, J. Elding, M.A. Nascimento. Information-content based sentence extraction for text summarization 214-218 Vol.2. [Crossref]

1797. A. Kehagias, A. Nicolaou, V. Petridis, P. Fragkou. 2004. Text segmentation by product partition models and dynamic programming. *Mathematical and Computer Modelling* **39**:2-3, 209-217. [Crossref]

1798. . References 147-148. [Crossref]

1799. Nega Alemayehu, Peter Willett. 2003. The effectiveness of stemming for information retrieval in Amharic. *Program* **37**:4, 254-259. [Abstract] [Full Text] [PDF]

1800. Hsinchun Chen, Haiyan Fan, Michael Chau, Daniel Zeng. 2003. Testing a Cancer Meta Spider. *International Journal of Human-Computer Studies* **59**:5, 755-776. [Crossref]

1801. Vicente P. Guerrero-Bote, Cristina López-Pujalte, Félix de Moya-Anegón, Victor Herrero-Solana. 2003. Comparison of neural models for document clustering. *International Journal of Approximate Reasoning* **34**:2-3, 287-305. [Crossref]

1802. Miguel E. Ruiz, Padmini Srinivasan. 2003. Hybrid hierarchical classifiers for categorization of medical documents. *Proceedings of the American Society for Information Science and Technology* **40**:1, 65-70. [Crossref]

1803. Fabio Crestani, Massimo Melucci. 2003. Automatic construction of hypertexts for self-referencing: the Hyper-TextBook project. *Information Systems* **28**:7, 769-790. [Crossref]

1804. H. Davalcu, S. Vadrevu, S. Nagarajan, I.V. Ramakrishnan. 2003. OntoMiner: bootstrapping and populating ontologies from domain-specific Web sites. *IEEE Intelligent Systems* **18**:5, 24-33. [Crossref]

1805. Cristina López-Pujalte, Vicente P. Guerrero-Bote, Félix de Moya-Anegón. 2003. Genetic algorithms in relevance feedback: a second test and new contributions. *Information Processing & Management* **39**:5, 669-687. [Crossref]

1806. Wessel Kraaij, Jian-Yun Nie, Michel Simard. 2003. Embedding Web-Based Statistical Translation Models in Cross-Language Information Retrieval. *Computational Linguistics* **29**:3, 381-419. [Crossref]

1807. Harold Thimbleby. 2003. Explaining code for publication. *Software: Practice and Experience* **33**:10, 975-1001. [Crossref]

1808. Suleiman H. Mustafa. 2003. A morphology-driven string matching approach to Arabic text searching. *Journal of Systems and Software* **67**:2, 77-87. [Crossref]

1809. Debbie A. Travers, Stephanie W. Haas. 2003. Using nurses' natural language entries to build a concept-oriented terminology for patients' chief complaints in the emergency department. *Journal of Biomedical Informatics* **36**:4-5, 260-270. [Crossref]

1810. Marius Paşca. 2003. Question-Driven Semantic Filters for Answer Retrieval. *International Journal of Pattern Recognition and Artificial Intelligence* **17**:05, 741-756. [Crossref]

1811. Ronald N. Kostoff. 2003. The practice and malpractice of stemming. *Journal of the American Society for Information Science and Technology* **54**:10, 984-985. [Crossref]

1812. Bong-Hyun Cho, Changki Lee, Gary Geunbae Lee. 2003. Exploring term dependences in probabilistic information retrieval model. *Information Processing & Management* **39**:4, 505-519. [Crossref]

1813. Naohiro Matsumura, Yukio Ohsawa, Mitsuru Ishizuka. 2003. PAI: Automatic indexing for extracting asserted keywords from a document. *New Generation Computing* **21**:1, 37-47. [Crossref]

1814. HOLGER BILLHARDT, DANIEL BORRAJO, VICTOR MAOJO. 2003. LEARNING RETRIEVAL EXPERT COMBINATIONS WITH GENETIC ALGORITHMS. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems* **11**:01, 87-113. [Crossref]

1815. T.A. Runkler, J.C. Bezdek. 2003. Web mining with relational clustering. *International Journal of Approximate Reasoning* **32**:2-3, 217-236. [Crossref]

1816. Cristina López-Pujalte, Vicente P. Guerrero-Bote, Félix de Moya-Anegón. 2003. Order-based fitness functions for genetic algorithms applied to relevance feedback. *Journal of the American Society for Information Science and Technology* **54**:2, 152-160. [Crossref]

1817. Lina Zhou, Dongsong Zhang. 2003. NLPIR: A theoretical framework for applying natural language processing to information retrieval. *Journal of the American Society for Information Science and Technology* **54**:2, 115-123. [Crossref]

1818. Eric J. Stierna, Neil C. Rowe. 2003. Applying information-retrieval methods to software reuse: a case study. *Information Processing & Management* **39**:1, 67-74. [Crossref]

1819. Makoto Nakashima, Keizo Sato, Yanhua Qu, Tetsuro Ito. 2003. Browsing-based conceptual information retrieval incorporating dictionary term relations, keyword association, and a user's interest. *Journal of the American Society for Information Science and Technology* **54**:1, 16-28. [Crossref]

1820. G. Pilato, S. Vitabile, G. Vassallo, V. Conti, F. Sorbello. A neural multi-agent based system for smart HTML pages retrieval 233-239. [Crossref]

1821. Elías F. Combarro, Elena Montañés, José Ranilla, Javier Fernández. A comparison of the performance of SVM and ARNI on Text Categorization with new filtering measures on an unbalanced collection 742-749. [Crossref]

1822. R. W. P. Luk, B. K. Y. T'sou, T. B. Y. Lai, O. O. Y. Kwong, F. C. Y. Chik, L. Y. L. Cheung. 2003. Bilingual legal document retrieval and management using XML. *Software: Practice and Experience* **33**:1, 41-59. [Crossref]

1823. . References 307-326. [Crossref]

1824. Jacques Savoy. 2003. Cross-language information retrieval: experiments based on CLEF 2000 corpora. *Information Processing & Management* **39**:1, 75-115. [Crossref]

1825. . References 489-498. [Crossref]

1826. U. Hahn, M. Honeck, S. Schulz. Subword-based text retrieval 8 pp.. [Crossref]

1827. Naohiro MATSUMURA, Yu KATO, Yukio OHSAWA, Mitsuru ISHIZUKA. 2003. Visualization of Discussion Structure for Discovering and Understanding Discussion Points. *Journal of Japan Society for Fuzzy Theory and Intelligent Informatics* **15**:5, 554-564. [Crossref]

1828. RADA F. MIHALCEA. 2002. GOING BEYOND EXPLICIT KNOWLEDGE FOR IMPROVED SEMANTIC BASED INFORMATION RETRIEVAL. *International Journal on Artificial Intelligence Tools* **11**:04, 553-586. [Crossref]

1829. Xiaofeng He, Hongyuan Zha, Chris H.Q. Ding, Horst D. Simon. 2002. Web document clustering using hyperlink structures. *Computational Statistics & Data Analysis* **41**:1, 19-45. [Crossref]

1830. F. Menczer, W.N. Street, A.E. Monge. 2002. Adaptive assistants for customized E-shopping. *IEEE Intelligent Systems* **17**:6, 12-19. [Crossref]

1831. Cristina López-Pujalte, Vicente P Guerrero Bote, Félix de Moya Anegón. 2002. A test of genetic algorithms in relevance feedback. *Information Processing & Management* **38**:6, 793-805. [Crossref]

1832. Polona Vilar. 2002. The influence of English on Slovenian library terminology. *Aslib Proceedings* **54**:5, 309-316. [Abstract] [Full Text] [PDF]

1833. Michael Molla, Peter Andreae, Jeremy Glasner, Frederick Blattner, Jude Shavlik. 2002. Interpreting microarray expression data using text annotating the genes. *Information Sciences* **146**:1-4, 75-88. [Crossref]

1834. Mikko Kurimo. 2002. Thematic indexing of spoken documents by using self-organizing maps. *Speech Communication* **38**:1-2, 29-45. [Crossref]

1835. Thorsten Kurz, Kilian Stoffel. 2002. Going beyond stemming: creating concept signatures of complex medical terms. *Knowledge-Based Systems* **15**:5-6, 309-313. [Crossref]

1836. Brian Fox, Christopher J Fox. 2002. Efficient stemmer generation. *Information Processing & Management* **38**:4, 547-558. [Crossref]

1837. N.C. Rowe. 2002. Marie-4: a high-recall, self-improving Web crawler that finds images using captions. *IEEE Intelligent Systems* **17**:4, 8-14. [Crossref]

1838. Wooju Kim, Larry Kerschberg, Anthony Scime. 2002. Learning for automatic personalization in a semantic taxonomy-based meta-search agent. *Electronic Commerce Research and Applications* **1**:2, 150-173. [Crossref]

1839. Harris Drucker, Behzad Shahrary, David C Gibbon. 2002. Support vector machines: relevance feedback and information retrieval. *Information Processing & Management* **38**:3, 305-323. [Crossref]

1840. Malcolm Slaney. Semantic-audio retrieval IV-4108-IV-4111. [Crossref]

1841. Fabio Crestani. 2002. Spoken query processing for interactive information retrieval. *Data & Knowledge Engineering* **41**:1, 105-124. [Crossref]

1842. RAYMOND Y. K. LAU. 2002. THE STATE OF THE ART IN ADAPTIVE INFORMATION AGENTS. *International Journal on Artificial Intelligence Tools* **11**:01, 19-61. [Crossref]

1843. Ángel Alonso, José R. Villar, Carmen Benavides, Isaías García, Francisco Rodríguez. 2002. AN APPROACH TO DOCUMENT ENGINEERING AT FUNDACIÓN HULLERA VASCO-LEONESA. *IFAC Proceedings Volumes* **35**:1, 371-376. [Crossref]

1844. Shoji Goto, Tadachika Ozono, Toramatsu Shintani. 2002. *Transactions of the Japanese Society for Artificial Intelligence* **17**, 285-292. [Crossref]

1845. Naohiro Matsumura, Yukio Ohsawa, Mitsuru Ishizuka. 2002. *Transactions of the Japanese Society for Artificial Intelligence* **17**, 398-406. [Crossref]

1846. Vicente P Guerrero Bote, Félix de Moya Anegón, Victor Herrero Solana. 2002. Document organization using Kohonen's algorithm. *Information Processing & Management* **38**:1, 79-89. [Crossref]

1847. Ian Ruthven, Mounia Lalmas, Keith van Rijsbergen. 2002. Combining and selecting characteristics of information use. *Journal of the American Society for Information Science and Technology* **53**:5, 378-396. [Crossref]

1848. Giorgio Brajnik, Stefano Mizzaro, Carlo Tasso, Fabio Venuti. 2002. Strategic help in user interfaces for information retrieval. *Journal of the American Society for Information Science and Technology* **53**:5, 343-358. [Crossref]

1849. M.G. Brown, J.T. Foote, G.J.F. Jones, K. Spärck Jones, S.J. Young. Open-Vocabulary Speech Indexing for Voice and Video Mail Retrieval 237-246. [Crossref]

1850. P.F. Wang, C. Mosley. Associative memory neural networks for information retrieval of text word pairs 2200-2203 vol.5. [Crossref]

1851. Holger Billhardt, Daniel Borrajo, Victor Maojo. 2002. A context vector model for information retrieval. *Journal of the American Society for Information Science and Technology* **53**:3, 236-249. [Crossref]

1852. Vicente P. Guerrero, Félix Moya-Anegón, Victor Herrero-Solana. 2002. Automatic extraction of relationships between terms by means of Kohonen's algorithm. *Library & Information Science Research* **24**:3, 235-250. [Crossref]

1853. Irene Diaz, Jorge Morato, Juan Lloréns. 2002. An algorithm for term conflation based on tree structures. *Journal of the American Society for Information Science and Technology* **53**:3, 199-208. [Crossref]

1854. Sheila Tejada, Craig A Knoblock, Steven Minton. 2001. Learning object identification rules for information integration. *Information Systems* **26**:8, 607-633. [Crossref]

1855. Mohammed Benkhalifa, Abdelhak Mouradi, Houssaine Bouyakhf. 2001. Integrating WordNet knowledge to supplement training data in semi-supervised agglomerative hierarchical clustering for text categorization. *International Journal of Intelligent Systems* **16**:8, 929-947. [Crossref]

1856. Mohamed T. Elhadi. 2001. Using statutes-based ir drive legal cbr. *Applied Artificial Intelligence* **15**:6, 587-600. [Crossref]

1857. Ari Pirkola. 2001. Morphological typology of languages for IR. *Journal of Documentation* **57**:3, 330-348. [Abstract] [PDF]

1858. J.P.A. Charlesworth, P.N. Garner. 2001. SpokenContent representation in MPEG-7. *IEEE Transactions on Circuits and Systems for Video Technology* **11**:6, 730-736. [Crossref]

1859. Kiduk Yang, Kelly L Maglaughlin, Gregory B Newby. 2001. Passage feedback with IRIS. *Information Processing & Management* **37**:3, 521-541. [Crossref]

1860. W. Lam, H. M. L. Meng, K. L. Wong, J. C. H. Yen. 2001. Using contextual analysis for news event detection. *International Journal of Intelligent Systems* **16**:4, 525-546. [Crossref]

1861. Z.Z. Nick, P. Themis. 2001. Web search using a genetic algorithm. *IEEE Internet Computing* **5**:2, 18-26. [Crossref]

1862. RAYMOND LAU, ARTHUR H. M. TER HOFSTEDE, PETER D. BRUZA. 2001. BELIEF REVISION FOR ADAPTIVE INFORMATION FILTERING AGENTS. *International Journal of Cooperative Information Systems* **10**:01n02, 57-79. [Crossref]

1863. Turid Hedlund, Ari Pirkola, Kalervo Järvelin. 2001. Aspects of Swedish morphology and semantics from the perspective of mono- and cross-language information retrieval. *Information Processing & Management* **37**:1, 147-161. [Crossref]

1864. Michael Shepherd, John F. (Jack) Duffy, Carolyn Watters, Nitin Gugle. 2001. The role of user profiles for news filtering. *Journal of the American Society for Information Science and Technology* **52**:2, 149-160. [Crossref]

1865. P. Nadkarni, R. Chen, C. Brandt. 2001. UMLS Concept Indexing for Production Databases: A Feasibility Study. *Journal of the American Medical Informatics Association* **8**:1, 80-91. [Crossref]

1866. Masayuki Okabe, Seiji Yamada. 2001. *Transactions of the Japanese Society for Artificial Intelligence* **16**, 139-146. [Crossref]

1867. Marco Meier, Peter Mertens. 2001. The Editorial Workbench ? Handling the Information Supply Chain of External Internet Data for Strategic Decision Support. *Journal of Decision Systems* **10**:2, 149-174. [Crossref]

1868. Raymond Lau, Arthur H.M. ter Hofstede, Peter D. Bruza. 2001. Maxi-Adjustment and Possibilistic Deduction for Adaptive Information Agents. *Journal of Applied Non-Classical Logics* **11**:1-2, 169-201. [Crossref]

1869. Vicente P. Guerrero, Félix de Moya Anegón. 2001. Reduction of the dimension of a document space using the fuzzified output of a Kohonen network. *Journal of the American Society for Information Science and Technology* **52**:14, 1234-1241. [Crossref]

1870. Marco Meier, Peter Mertens. 2000. Linking key figures and internet business news for personalized management information. *Journal of Systems and Information Technology* **4**:2, 13-32. [Abstract] [PDF]

1871. MOUNIA LALMAS. 2000. COMBINING DOCUMENT REPRESENTATIONS. *International Journal of Cooperative Information Systems* **09**:04, 427-447. [Crossref]

1872. Danny C.C. Poo, Tech-Kang Toh, Christopher S.G. Khoo. 2000. Enhancing online catalog searches with an electronic referencer. *Journal of Systems and Software* **55**:2, 203-219. [Crossref]

1873. Carlos G. Figuerola, Raquel Gómez, Eva López de San Román. 2000. Stemming and n-grams in Spanish: an evaluation of their impact on information retrieval. *Journal of Information Science* **26**:6, 461-467. [Crossref]

1874. K. Sparck Jones, S. Walker, S.E. Robertson. 2000. A probabilistic model of information retrieval: development and comparative experiments. *Information Processing & Management* **36**:6, 779-808. [Crossref]

1875. Kenney Ng, Victor W. Zue. 2000. Subword-based approaches for spoken document retrieval. *Speech Communication* **32**:3, 157-186. [Crossref]

1876. Steve Renals, Dave Abberley, David Kirby, Tony Robinson. 2000. Indexing and retrieval of broadcast news. *Speech Communication* **32**:1-2, 5-20. [Crossref]

1877. Pierre Jourlin, Sue E. Johnson, Karen Spärck Jones, Philip C. Woodland. 2000. Spoken document representations for probabilistic retrieval. *Speech Communication* **32**:1-2, 21-36. [Crossref]

1878. Stefan Schulz, Udo Hahn. 2000. Morpheme-based, cross-lingual indexing for medical document retrieval. *International Journal of Medical Informatics* **58-59**, 87-99. [Crossref]

1879. J. Makhoul, F. Kubala, T. Leek, Daben Liu, Long Nguyen, R. Schwartz, A. Srivastava. 2000. Speech and language technologies for audio indexing and retrieval. *Proceedings of the IEEE* **88**:8, 1338-1353. [Crossref]

1880. Antonio G�mez Skarmeta, Amine Bensaid, Nadia Tazi. 2000. Data mining for text categorization with semi-supervised agglomerative hierarchical clustering. *International Journal of Intelligent Systems* **15**:7, 633-646. [Crossref]

1881. William W. Cohen. 2000. WHIRL: A word-based information representation language. *Artificial Intelligence* **118**:1-2, 163-196. [Crossref]

1882. Robert Krovetz. 2000. Viewing morphology as an inference process. *Artificial Intelligence* **118**:1-2, 277-294. [Crossref]

1883. M.T Elhadi. 2000. Bankruptcy support system: taking advantage of information retrieval and case-based reasoning. *Expert Systems with Applications* **18**:3, 215-219. [Crossref]

1884. Charles L.A. Clarke, Gordon V. Cormack, Elizabeth A. Tudhope. 2000. Relevance ranking for one to three term queries. *Information Processing & Management* **36**:2, 291-311. [Crossref]

1885. ELAINE G. TOMS. 2000. Understanding and facilitating the browsing of electronic text. *International Journal of Human-Computer Studies* **52**:3, 423-452. [Crossref]

1886. A Klose, A Nürnberger, R Kruse, G Hartmann, M Richards. 2000. Interactive text retrieval based on document similarities. *Physics and Chemistry of the Earth, Part A: Solid Earth and Geodesy* **25**:8, 649-654. [Crossref]

1887. Michael J. Wise. 2000. Protein annotators' assistant: A novel application of information retrieval techniques. *Journal of the American Society for Information Science* **51**:12, 1131-1136. [Crossref]

1888. Mark Rorvig, Steven Fitzpatrick. 2000. Shape recovery: A visual method for evaluation of information retrieval experiments. *Journal of the American Society for Information Science* **51**:13, 1205-1210. [Crossref]

1889. Zainab Abu Bakar, Tengku Mohd T. Sembok, Mohammed Yusoff. 2000. An evaluation of retrieval effectiveness using spelling-correction and string-similarity matching methods on Malay texts. *Journal of the American Society for Information Science* **51**:8, 691-706. [Crossref]

1890. Daniel Boley, Maria Gini, Robert Gross, Eui-Hong (Sam) Han, Kyle Hastings, George Karypis, Vipin Kumar, Bamshad Mobasher, Jerome Moore. 1999. Partitioning-based clustering for Web document categorization. *Decision Support Systems* **27**:3, 329-341. [Crossref]

1891. John M. Prager. 1999. Linguini: Language Identification for Multilingual Documents. *Journal of Management Information Systems* **16**:3, 71-101. [Crossref]

1892. F. Macdonald, J. Miller. 1999. ASSIST—a tool to support software inspection. *Information and Software Technology* **41**:15, 1045-1057. [Crossref]

1893. Chia-Hui Chang, Ching-Chi Hsu. 1999. Enabling concept-based relevance feedback for information retrieval on the WWW. *IEEE Transactions on Knowledge and Data Engineering* **11**:4, 595-609. [Crossref]

1894. M.T. Elhadi, T. Vamos. An IR-CBR approach to legal indexing and retrieval in bankruptcy law 769-774. [Crossref]

1895. Mikko Kurimo. Indexing Audio Documents by using Latent Semantic Analysis and SOM 363-374. [Crossref]

1896. Jacques Savoy. 1999. A stemming procedure and stopword list for general French corpora. *Journal of the American Society for Information Science* **50**:10, 944-952. [Crossref]

1897. M. Siegler, M. Withrock. Improving the suitability of imperfect transcriptions for information retrieval from spoken documents 505-508 vol.1. [Crossref]

1898. Daniel Boley, Maria Gini, Kyle Hastings, Bamshad Mobasher, Jerry Moore. 1998. A client-side Web agent for document categorization. *Internet Research* **8**:5, 387-399. [Abstract] [Full Text] [PDF]

1899. Mounia Lalmas, Ian Ruthven. 1998. Representing and retrieving structured documents using the Dempster-Shafer theory of evidence: modelling and evaluation. *Journal of Documentation* **54**:5, 529-565. [Abstract] [PDF]

1900. Michael P Oakes, Malcolm J Taylor. 1998. Automated assistance in the formulation of search statements for bibliographic databases. *Information Processing & Management* **34**:6, 645-668. [Crossref]

1901. Fabio Crestani, Massimo Melucci. 1998. A case study of automatic authoring: From a textbook to a hyper-textbook. *Data & Knowledge Engineering* **27**:1, 1-30. [Crossref]

1902. Alexander M. Robertson, Peter Willett. 1998. Applications of n-grams in textual information systems. *Journal of Documentation* **54**:1, 48-67. [Abstract] [PDF]

1903. M. Dreger, S. Lohrum, H. Schweppe, C. D. Ziegler. Ariadne, an interactive navigation and search system for computer science information on the world-wide web 127-143. [Crossref]

1904. Christopher S. G. Khoo, Danny C. C. Poo, Teck-Kang Toh, Soon-Kah Liew, Anne N. M. Goh. E-Referencer: A Prototype Expert System Web Interface to Online Catalogs 315-333. [Crossref]

1905. Michael L. Best. 1997. An Ecology of Text: Using Text Retrieval to Study Alife on the Net. *Artificial Life* **3**:4, 261-287. [Crossref]

1906. Jacques Savoy. 1997. Statistical inference in retrieval effectiveness evaluation. *Information Processing & Management* **33**:4, 495-512. [Crossref]

1907. F.C. Berger, P. van Bommel. 1997. Augmenting a characterization network with semantic information. *Information Processing & Management* **33**:4, 453-479. [Crossref]

1908. Valery J. Frants, Jacob Shapiro, Vladimir G. Voiskunskii. Automatic Indexing of Documents 136-165. [Citation] [PDF] [PDF]

1909. Ulrich Pfeifer, Thomas Poersch, Norbert Fuhr. 1996. Retrieval effectiveness of proper name search methods. *Information Processing & Management* **32**:6, 667-679. [Crossref]

1910. K. Chen-Chuan Chang, H. Garcia-Molina, A. Paepcke. 1996. Boolean query mapping across heterogeneous information sources. *IEEE Transactions on Knowledge and Data Engineering* **8**:4, 515-521. [Crossref]

1911. Maristella Agosti, Fabio Crestani, Massimo Melucci. 1996. Design and implementation of a tool for the automatic construction of hypertexts for information retrieval. *Information Processing & Management* **32**:4, 459-476. [Crossref]

1912. K. Sparck Jones, G.J.F. Jones, J.T. Foote, S.J. Young. 1996. Experiments in spoken document retrieval. *Information Processing & Management* **32**:4, 399-417. [Crossref]

1913. Jacques Savoy. 1996. An extended vector-processing scheme for searching information in hypertext systems. *Information Processing & Management* **32**:2, 155-170. [Crossref]

1914. ROBYN SCHINKE, MARK GREENGRASS, ALEXANDER M. ROBERTSON, PETER WILLETT. 1996. A STEMMING ALGORITHM FOR LATIN TEXT DATABASES. *Journal of Documentation* **52**:2, 172-187. [Abstract] [PDF]

1915. W.B. Croft, J. Broglio, H. Fujii. Applications of multilingual text retrieval 98-107 vol.5. [Crossref]

1916. W.B. Croft. 1995. Effective text retrieval based on combining evidence from the corpus and users. *IEEE Expert* **10**:6, 59-63. [Crossref]

1917. T.Z. Kalamboukis. 1995. Suffix stripping with modern Greek. *Program* **29**:3, 313-321. [Abstract] [PDF]

1918. Steve Wade, Patrick Baekevelt. 1995. IR_Framework: An object oriented framework for developing information retrieval systems. *Program* **29**:1, 15-29. [Abstract] [PDF]

1919. H.P. Frei, D. Stieger. 1995. The use of semantic links in hypertext information retrieval. *Information Processing & Management* **31**:1, 1-13. [Crossref]

1920. Margaret Fieldhouse, Micheline Hancock-Beaulieu. 1994. The Changing Face of Okapi. *Library Review* **43**:4, 38-50. [Abstract] [Full Text] [PDF]

1921. J. K. Wu, Y. H. Ang, P. Lam, H. H. Loh, A. Desai Narasimhalu. 1994. Inference and retrieval of facial images. *Multimedia Systems* **2**:1, 1-14. [Crossref]

1922. Christopher S.G. Khoo, Danny C.C. Poo. 1994. An expert system approach to online catalog subject searching. *Information Processing & Management* **30**:2, 223-238. [Crossref]

1923. M.D. Dunlop, C.J. van Rijsbergen. 1993. Hypermedia and free text retrieval. *Information Processing & Management* **29**:3, 287-298. [Crossref]

1924. Eugene L. Margulis. 1993. Modelling documents with multiple poisson distributions. *Information Processing & Management* **29**:2, 215-227. [Crossref]

1925. BRIAN VICKERY, ALINA VICKERY. 1993. ONLINE SEARCH INTERFACE DESIGN. *Journal of Documentation* **49**:2, 103-107. [Abstract] [PDF]

1926. Georgina Verity. 1993. Relational database management systems and open systems used in the development of Oracle Libraries. *Program* **27**:1, 73-82. [Abstract] [PDF]

1927. Milam Aiken, Judith Carlisle. 1992. An automated idea consolidation tool for computer supported cooperative work. *Information & Management* **23**:6, 373-382. [Crossref]

1928. F. Cuna Ekmekcioglu, Alexander M. Robertson, Peter Willett. 1992. Effectiveness of query expansion in ranked-output document retrieval systems. *Journal of Information Science* **18**:2, 139-147. [Crossref]

1929. BRIAN VICKERY, ALINA VICKERY. 1992. AN APPLICATION OF LANGUAGE PROCESSING FOR A SEARCH INTERFACE. *Journal of Documentation* **48**:3, 255-275. [Abstract] [PDF]

1930. W. John Wilbur, Karl Sirotkin. 1992. The automatic identification of stop words. *Journal of Information Science* **18**:1, 45-55. [Crossref]

1931. William R. Hersh, Robert A. Greenes. 1990. SAPHIRE—An information retrieval system featuring concept matching, automatic indexing, probabilistic retrieval, and hierarchical relationships. *Computers and Biomedical Research* **23**:5, 410-425. [Crossref]

1932. Alan F. Smeaton. 1990. retriev: an information retrieval system implemented on top of a relational database. *Program* **24**:1, 21-32. [Abstract] [PDF]

1933. Xin Lu. 1990. Document retrieval: A structural approach. *Information Processing & Management* **26**:2, 209-218. [Crossref]

1934. Tengku M.T. Sembok, C.J. van Rijsbergen. 1990. SILOL: A simple logical-linguistic document retrieval system. *Information Processing & Management* **26**:1, 111-134. [Crossref]

1935. Stephen J. Wade, Peter Willett, David Bawden. 1989. SIBRIS: the Sandwich Interactive Browsing and Ranking Information System. *Journal of Information Science* **15**:4-5, 249-260. [Crossref]

1936. Alan Griffiths. 1989. SAGIS: a proposal for a Sardinian Geographical Information System and an assessment of alternative implementation strategies. *Journal of Information Science* **15**:4-5, 261-267. [Crossref]

1937. R.H. Thompson, W.B. Croft. 1989. Support for browsing in an intelligent text retrieval system. *International Journal of Man-Machine Studies* **30**:6, 639-668. [Crossref]

1938. Stephen Wade, Peter Willett, Bruce Robinson, Brian Vickery, Alina Vickery. 1988. A comparison of knowledge-based and statistically-based techniques for reference retrieval. *Online Review* **12**:2, 91-108. [Abstract] [PDF]

1939. Stephen J. Wade, Peter Willett. 1988. INSTRUCT: a teaching package for experimental methods in information retrieval. Part III. Browsing, clustering and query expansion. *Program* **22**:1, 44-61. [Abstract] [PDF]

1940. Stephen Walker. 1988. Improving subject access painlessly: recent work on the Okapi online catalogue projects. *Program* **22**:1, 21-31. [Abstract] [PDF]

1941. K.L. Kwok, William Kuan. 1988. Experiments with document components for indexing and retrieval. *Information Processing & Management* **24**:4, 405-417. [Crossref]

1942. Peter Willett. 1988. Improving subject retrieval in online catalogues. *Information Processing & Management* **24**:6, 718-719. [Crossref]

1943. Michael F. Lynch, P. Willett. 1987. Information retrieval research in the Department of Information Studies, University of Sheffield: 1965-1985. *Journal of Information Science* **13**:4, 221-234. [Crossref]

1944. ALINA VICKERY, HELEN BROOKS, BRUCE ROBINSON, BRIAN VICKERY. 1987. A REFERENCE AND REFERRAL SYSTEM USING EXPERT SYSTEM TECHNIQUES. *Journal of Documentation* **43**:1, 1-23. [Abstract] [PDF]

1945. A. Vickery, H.M. Brooks. 1987. Plexus-the expert system for referral. *Information Processing & Management* **23**:2, 99-117. [Crossref]

1946. Ian G Hendry, Peter Willett, Frances E. Wood. 1986. INSTRUCT: a teaching package for experimental methods in information retrieval. Part I. The users view. *Program* **20**:3, 245-263. [Abstract] [PDF]

1947. Peter Willett. 1985. An algorithm for the calculation of exact term discrimination values. *Information Processing & Management* **21**:3, 225-232. [Crossref]

1948. Mary E. Rowbottom, Peter Willett. 1982. The effect of subject matter on the automatic indexing of full text. *Journal of the American Society for Information Science* **33**:3, 139-141. [Crossref]

1949. Martin Lennon, David S. Peirce, Brian D. Tarry, Peter Willett. 1981. An evaluation of some conflation algorithms for information retrieval. *Information Scientist* **3**:4, 177-183. [Crossref]

1950. Giuliano Armano, Alessandro Giuliani. Novel Text Summarization Techniques for Contextual Advertising 185-204. [Crossref]

1951. D.C.C. Poo, Teck Kang Toh, C.S.G. Khoo, G. Hong. Development of an intelligent Web interface to online library catalog databases 64-71. [Crossref]

1952. A. Di Stefano, G. Pappalardo, C. Santoro, E. Tramontana. SHARK, a Multi-Agent System to Support Document Sharing and Promote Collaboration 86-93. [Crossref]

1953. Cecil Eng Huang Chua, Roger H. Chiang, Veda C. Storey. Improving Domain Searches through Customized Search Engines 1-22. [Crossref]

1954. J. Zakos, B. Verma, Xue Li, S. Kulkarni. Intelligent encoding of concepts in Web document retrieval 72-77. [Crossref]

1955. Abhishek Das, Ankit Jain. Indexing the World Wide Web 1-28. [Crossref]

1956. W.-H. Lin, R. Jin, A. Hauptmann. Web image retrieval re-ranking with relevance model 242-248. [Crossref]

1957. T. Imai, R. Schwartz, F. Kubala, L. Nguyen. Improved topic discrimination of broadcast news using a model of multiple simultaneous topics 727-730. [Crossref]

1958. Marco A. Alvarez, Xiaojun Qi, Changhui Yan. GO-Based Term Semantic Similarity 174-185. [Crossref]

1959. Sheng-Tang Wu, Yuefeng Li, Yue Xu, Binh Pham, Phoebe Chen. Automatic Pattern-Taxonomy Extraction for Web Mining 242-248. [Crossref]

1960. Daniela López De Luise, Débora Hisgen. MLW and Bilingualism 1-32. [Crossref]

1961. Anne R. Diekema. Design and Development Considerations for a Multilingual Digital Library 1222-1233. [Crossref]

1962. Dingyi Han, Ting Shen, Shicong Meng, Yong Yu. Cuckoo Ring: BalancingWorkload for Locality Sensitive Hash 49-56. [Crossref]

1963. Zhi-Hong Deng, Shi-Wei Tang, Dong-Qing Yang, Ming Zhang, Xiao-Bin Wu, Meng Yang. SRFW: a simple, fast and effective text classification algorithm 1267-1271. [Crossref]

1964. Yilu Zhou, Jialun Qin, Hsinchun Chen, J.F. Nunamaker. Multilingual Web Retrieval: An Experiment on a Multilingual Business Intelligence Portal 43a-43a. [Crossref]

1965. M.P. Oakes, M.J. Tayor. Re-use of instruction manual in an automated intermediary system 1226-1229. [Crossref]

1966. M. Ceccato, M. Marin, K. Mens, L. Moonen, P. Tonella, T. Tourwe. A Qualitative Comparison of Three Aspect Mining Techniques 13-22. [Crossref]

1967. A. Kuhn, S. Ducasse, T. Girba. Enriching Reverse Engineering with Semantic Clustering 133-142. [Crossref]

1968. V. Shanks, H.E. Williams. Fast categorisation of large document collections 194-204. [Crossref]

1969. Min Xu, Liang-Tien Chia, Haoran Yi, D. Rajan. Affective Content Detection in Sitcom Using Subtitle and Audio 129-134. [Crossref]

1970. R.F. Mihalcea, S.I. Mihalcea. Word semantics for information retrieval: moving one step closer to the Semantic Web 280-287. [Crossref]

1971. Wei Wang, C.V. Ravishankar. Adaptive data broadcasting in asymmetric communication environments 27-36. [Crossref]

1972. G.-N. Wang, M. Evens, D.B. Hier. On the evaluation of LITREF: a PC-based information retrieval system to support stroke diagnosis 548-555. [Crossref]

1973. H. Al-Mubaid. Context-Based Technique for Biomedical Term Classification 1577-1584. [Crossref]

1974. Alex Kohn, François Bry, Alexander Manta. Semantic Search on Unstructured Data 194-213. [Crossref]

1975. I.J. Aalbersberg, F. Sijstermans. High-quality and high-performance full-text document retrieval: the Parallel InfoGuide System 142-150. [Crossref]

1976. M.F. Wyle. A wide area network information filter 10-15. [Crossref]

1977. K. Kaikhah. Automatic text summarization with neural networks 40-44. [Crossref]

1978. K. Saito, R. Nakano. Extracting characteristic words of text using neural networks 1397-1402. [Crossref]

1979. Sebastian Furth, Joachim Baumeister. Semantification of Large Corpora of Technical Documentation 171-200. [Crossref]

1980. Yueqin Hang, Jie Shen, Yin Lin, Zhao Minya. Context information extraction of the query based on single document 431-434. [Crossref]

1981. O. Ibrahimov, I. Sethi, N. Dimitrova. The performance analysis of a Chi-square similarity measure for topic related clustering of noisy transcripts 285-288. [Crossref]

1982. D. Fragoudis, S.D. Likothanassis. Retriever: a self-training agent for intelligent information discovery 594-599. [Crossref]

1983. A.P. Nikora. Classifying Requirements: Towards a More Rigorous Analysis of Natural-Language Specifications 291-300. [Crossref]

1984. A. Honrado, R. Leon, R. O'Donnel, D. Sinclair. A word stemming algorithm for the Spanish language 139-145. [Crossref]

1985. R.A. Calvo. Factor semantics for document retrieval 198-203. [Crossref]

1986. M.M.S. Beg, R. Ali. A subjective cum objective measure of web search quality 365-370. [Crossref]

1987. Gaurav Jain, A. Ginwala, Y.A. Asiandogan. An approach to text classification using dimensionality reduction and combination of classifiers 564-569. [Crossref]

1988. C. Bayrak, H. Joshi. Learning Contextual Behavior of Text Data 299-304. [Crossref]

1989. V.M. Orengo, C. Huyck. A stemming algorithm for the portuguese language 186-193. [Crossref]

1990. Eric Poitras, Negar Fazeli. Mining the Edublogosphere to Enhance Teacher Professional Development 42-65. [Crossref]

1991. Elias Oliveira, Patrick Marques Ciarelli, Evandro Ottoni Teatini Salles. An Evolving System in the Text Classification Problem 467-486. [Crossref]

1992. K.M. Hammouda, M.S. Kamel. Phrase-based document similarity based on an index graph model 203-210. [Crossref]

1993. G.J.F. Jones, J.T. Foote, K. Sparck Jones, S.J. Young. Video mail retrieval: the effect of word spotting accuracy on precision 309-312. [Crossref]

1994. A. Kiani-B, M.-R. Akbarzadeh-T., M.H. Moeinzadeh. Intelligent Extractive Text Summarization Using Fuzzy Inference Systems 1-4. [Crossref]

1995. Yan Yang, M. Kamel, Fan Jin. A model of document clustering using ant colony algorithm and validity index 2730-2735. [Crossref]

1996. C. Borgelt, A. Nurnberger, R. Kruse. Fuzzy Learning Vector Quantization with Size and Shape Parameters 195-200. [Crossref]

1997. T.D. Breaux, J.W. Reed. Using Ontology in Hierarchical Information Clustering 111b-111b. [Crossref]

1998. Nian-Shing Chen, P. Kinshuk, Chun-Wang Wei, Hong-Jhe Chen. Mining e-Learning Domain Concept Map from Academic Articles 694-698. [Crossref]

1999. Lin Li, Huifan Xiao, Guandong Xu. Recommending Related Microblogs 202-210. [Crossref]

2000. Xiao-Shen Zheng, Pi-Lian He, Mei Tian, Fu-Yong Yuan. Algorithm of documents clustering based on minimum spanning tree 199-203. [Crossref]

2001. B. Grilheres, C. Beauce, S. Canu, S. Brunessaux. A Platform for Semantic Annotations and Ontology Population Using Conditional Random Fields 790-793. [Crossref]

2002. D. Bonino, F. Corno, L. Farinetti. DOSE: a distributed open semantic elaboration platform 580-588. [Crossref]

2003. Víctor Peinado, Álvaro Rodrigo, Fernando López-Ostenero. Multilingual Information Access 203-228. [Crossref]

2004. Mark Kilfoil, Ali Ghorbani. SWAMI 171-195. [Crossref]

2005. S. Al-Hawamdeh, C.J.C. Loke. Compound document processing system 640-644. [Crossref]

2006. P. Wolf, B. Raj. The MERL SpokenQuery information retrieval system a system for retrieving pertinent documents from a spoken query 317-320. [Crossref]

2007. F. Crestani. Effects of word recognition errors in spoken query processing 39-47. [Crossref]

2008. A.F. Smeaton, A.L. Spitz. Using character shape coding for information retrieval 974-978. [Crossref]

2009. Luying Liu, Jianchu Kang, Jing Yu, Zhongliang Wang. A Comparative Study on Unsupervised Feature Selection Methods for Text Clustering 597-601. [Crossref]

2010. Christian F. Hempelmann. NLP for Search 53-74. [Crossref]

2011. G.J.F. Jones, J.T. Foote, K. Spark Jones, S.J. Young. Robust talker-independent audio document retrieval 311-314. [Crossref]

2012. F.V. Paulovich, R. Minghim. Text Map Explorer: a Tool to Create and Explore Document Maps 245-251. [Crossref]

2013. L. Kerschberg, Wooju Kim, A. Scime. A semantic taxonomy-based personalizable meta-search agent 41-50. [Crossref]

2014. P. Kazienko, M. Kiewra. Integration of relational databases and web site content for product and page recommendation 111-116. [Crossref]

2015. S. Davies, S. Badem, M.D. Williams, R. King. "Google by reformulation": modeling search as successive refinement 435-440. [Crossref]

2016. A.E. Hassan, Jingwei Wu, R.C. Holt. Visualizing Historical Data Using Spectrographs 31-31. [Crossref]

2017. D.A. James. A system for unrestricted topic retrieval from radio news broadcasts 279-282. [Crossref]

2018. Andreas Schmidt, Martin Atzmueller, Martin Hollender. Data Preparation for Big Data Analytics: 157-170. [Crossref]

2019. K. Saito, R. Nakano. Weight sharing on naive bayes document model 576-581. [Crossref]

2020. S.M.F.DS. Mustapha, N. Idris, R. Abdullah. Embedding Information Retrieval and Nearest-Neighbour Algorithm into Automated Essay Grading System 169-172. [Crossref]

2021. Daniela López De Luise, Débora Hisgen. MLW and Bilingualism 555-586. [Crossref]

2022. U. Kruschwitz. Exploiting structure for intelligent Web search 9. [Crossref]

2023. E. Sanchez, J.-A. Santini. Fuzzy logic and the Internet: a fuzzy Web site evaluator 344-347. [Crossref]

2024. O. Hoeber, Xue Dong Yang. The Visual Exploration of Web Search Results Using HotMap 157-165. [Crossref]

2025. Roger Blake, Steven Gordon, G. Shankaranarayanan. The Role of Case-Based Research in Information Technology and Systems 200-220. [Crossref]

2026. B. Shanmukha Rao, S.V. Rao, G. Sajith. A user-profile assisted meta search engine 713-717. [Crossref]

2027. Biju Issac. Spam Detection Approaches with Case Study Implementation on Spam Corpora 194-212. [Crossref]

2028. Kaveh Hassani, Won-Sook Lee. A Universal Architecture for Migrating Cognitive Agents: 238-265. [Crossref]

2029. Y. Ohsawa, N.E. Benson, M. Yachida. KeyGraph: automatic indexing by co-occurrence graph based on building construction metaphor 12-18. [Crossref]

2030. E. Savia, S. Kaski, V. Tuulos, P. Myllymaki. On text-based estimation of document relevance 3275-3280. [Crossref]

2031. D. Tikk, Z.T. Kardkovacs, Z. Bansaghi. Topic Mapping: a Tool for Finding the Meaning of Internet Search Queries 227-232. [Crossref]

2032. Fei-Yue Wang, G. Lai, S. Tang. An application specific knowledge engine for researches in intelligent transportation systems 841-846. [Crossref]

2033. D.C.C. Poo, Teck Kang Toh, C.S.G. Khoo. Search interface for Z39.50 compliant online catalogs over the Internet 7. [Crossref]

2034. Man Lan, Sam-Yuan Sung, Hwee-Boon Low, Chew-Lim Tan. A comparati e study on term eighting schemes for te t categori ation 546-551. [Crossref]

2035. A. Myka, U. Guntzer. On automatic similarity linking in digital libraries 278-283. [Crossref]

2036. S.L.Y. Lam, Dik Lun Lee. Feature reduction for neural network based text categorization 195-202. [Crossref]

2037. J. Perkio, V. Tuulos, W. Buntine, H. Tirri. Multi-Faceted Information Retrieval System for Large Scale Email Archives 557-564. [Crossref]

2038. S. Kosinov. Evaluation of N-grams conflation approach in text-based information retrieval 136-142. [Crossref]

2039. T.W. Yan, H. Garcia-Molina. Index structures for information filtering under the vector space model 337-347. [Crossref]

2040. Giuliano Armano, Alessandro Giuliani, Eloisa Vargiu. Intelligent Techniques in Recommender Systems and Contextual Advertising 105-128. [Crossref]

2041. M. Shepherd, C. Watters, J. Duffy, R. Kaushik. Browsing and keyword-based profiles: a cautionary tale 9. [Crossref]

2042. Asmae Dami, Mohamed Fakir, Belaid Bouikhalene. Information Retrieval (IR) and Extracting Associative Rules 713-732. [Crossref]

2043. Qingyang Xu, Wanli Zuo. Extracting Precise Link Context Using NLP Parsing Technique 64-69. [Crossref]

2044. O. Hoeber, Xue-Dong Yang, Yiyu Yao. Visualization Support for Interactive Query Refinement 657-665. [Crossref]

2045. Y. Tsuboshita, H. Okamoto. Extracting information in a graded manner from a neural-network system with continuous attractors 3095-3100. [Crossref]

2046. D. Fragoudis, S.D. Likothanassis. Learning to identify interesting links in intelligent information discovery 410-413. [Crossref]

2047. E. Daly, A. Gray, M. Haahr. On Using Peer Profiles to Create Self-Organizing P2P Networks 588-592. [Crossref]

2048. C. Bighini, A. Carbonaro, G. Casadei. InLinx for document classification, sharing and recommendation 91-95. [Crossref]

2049. W.-u.-Rehman Chaudhry, F. Meziane. Information extraction from heterogeneous sources using domain ontologies 511-516. [Crossref]

2050. Xiaomin Ying, Ming Liu, Wenhua Dou. User-interested-keywords set discovery using rough sets for intelligent information agents 61-66. [Crossref]

2051. Debasis Ganguly, Johannes Leveling, Gareth J.F. Jones. Bengali (Bangla) Information Retrieval 273-301. [Crossref]

2052. W.K. Lai, K.M. Hoe, T.S.Y. Tai, M.C.Y. Seah. Classifying English Web pages with "smart" ant-like agents 411-416. [Crossref]

2053. Ying Liu, B.J. Ciliax, K. Borges, V. Dasigi, A. Ram, S.B. Navathe, R. Dingledine. Comparison of two schemes for automatic keyword extraction from MELINE for functional gene clustering 382-392. [Crossref]

2054. M.S. Vallim, J.M.A. Coello. An agent for Web information dissemination based on a genetic algorithm 3834-3839. [Crossref]

2055. A. Bagga, Jianying Hu, Jialin Zhong, G. Ramesh. Multi-source combined-media video tracking for summarization 818-821. [Crossref]

2056. T. Rachidi, O. Iraqi, M. Bouzoubaa, A.B. El Khattab, M. El Kourdi, A. Zahi, A. Bensaid. Barq: Distributed multilingual internet search engine with focus on Arabic language 428-435. [Crossref]

2057. Jalel Akaichi. Sentiment Classification: 1-26. [Crossref]

2058. Miguel Pragosa, Vitor Basto-Fernandes, Luísa Oliveira. Semantic Integration Framework for Resources Identification 205-219. [Crossref]

2059. S.M. Yang, Xiao-Bin Wu, Zhi-Hong Deng, Ming Zhang, Dong-Qing Yang. Relative term-frequency based feature selection for text categorization 1432-1436. [Crossref]

2060. M.E.S. Mendes, L. Sacks. Evaluating fuzzy clustering for relevance-based information access 648-653. [Crossref]

2061. P. Francq, A. Delchambre. Using Documents Assessments to Build Communities of Interests 327-333. [Crossref]

2062. T.W. Fox. Document vector compression and its application in document clustering 2029-2032. [Crossref]

2063. Kang Peng, S. Vucetic, Bo Han, Hongbo Xie, Z. Obradovic. Exploiting unlabeled data for improving accuracy of predictive data mining 267-274. [Crossref]

2064. S. Vadrevu, S. Nagarajan, F. Gelgi, H. Davulcu. Automated Metadata and Instance Extraction from News Web Sites 38-41. [Crossref]

2065. G. Mamakis, A.G. Malamos, A. Axaridou, Y. Kaliakatsos, A. Ware. An Algorithm for Automatic Content Summarization in Modern Greek Language 579-591. [Crossref]

2066. J. Caverlee, L. Liu, D. Buttler. Probe, cluster, and discover: focused extraction of QA-Pagelets from the deep Web 103-114. [Crossref]

2067. Simon Pickert, Philipp Sandner. Whom to Trust for Financial Advice? 175-207. [Crossref]

2068. Hua-Jun Zeng, Zheng Chen, Wei-Ying Ma. A unified framework for clustering heterogeneous Web objects 161-170. [Crossref]

2069. Wen-Lin Hsu, Sheau-Dong Lang. Feature reduction and database maintenance in NETNEWS classification 137-144. [Crossref]

2070. Anne R. Diekema. Design and Development Considerations for a Multilingual Digital Library 1-12. [Crossref]

2071. Yasuhiro Yamada, Kanji Kato, Sachio Hirokawa. Text Mining for Analysis of Interviews and Questionnaires 1390-1406. [Crossref]

2072. Yasuhiro Yamada, Kanji Kato, Sachio Hirokawa. Text Mining for Analysis of Interviews and Questionnaires 52-68. [Crossref]

2073. Yogesh Kumar Meena, Dinesh Gopalani. Statistical Features for Extractive Automatic Text Summarization 126-144. [Crossref]

2074. R. Amarasiri, D. Alahakoon, K. Smith, M. Premaratne. HDGSOMr: A High Dimensional Growing Self-Organizing Map Using Randomness for Efficient Web and Text Mining 215-221. [Crossref]

2075. A. Bilke, F. Naumann. Schema Matching Using Duplicates 69-80. [Crossref]

2076. J.-L. Gauvnin, L. Lamel, Y. De Kercadio, G. Adda. Transcription and indexation of broadcast data 1663-1666. [Crossref]

2077. S.J. Young, M.G. Brown, J.T. Foote, G.J.F. Jones, K. Sparck Jones. Acoustic indexing for multimedia retrieval and browsing 199-202. [Crossref]

2078. R. Silipo, F. Crestani. Prosodic stress and topic detection in spoken sentences 243-252. [Crossref]

2079. L. Gandy, S. Rahimi, B. Gupta. A Modified Competitive Agglomeration for Relational Data Algorithm 210-215. [Crossref]

2080. Robert J. Hendley, Barry Wilkins, Russell Beale. Organix 34-46. [Crossref]