Padmini Srinivasan

List of Publications by Year
in descending order

91 papers	2,922 citations	304743 22 h-index	233421 45 g-index
94 all docs	94 docs citations	94 times ranked	2111 citing authors

#	ARTICLE	IF	CITATIONS
1	Text mining: Generating hypotheses from MEDLINE. Journal of the Association for Information Science and Technology, 2004, 55, 396-413.	2.6	219
2	Topical web crawlers. ACM Transactions on Internet Technology, 2004, 4, 378-419.	4.4	201
3	Hierarchical Text Categorization Using Neural Networks. Information Retrieval, 2002, 5, 87-118.	2.0	197
4	Quality through flow and immersion. , 2012, , .		132
5	Evaluating topic-driven web crawlers. , 2001, , .		118
6	Learning to crawl. ACM Transactions on Information Systems, 2005, 23, 430-462.	4.9	115
7	Crawling the Web. , 2004, , 153-177.		111
8	Mining MEDLINE for implicit links between dietary substances and diseases. Bioinformatics, 2004, 20, i290-i296.	4.1	108
9	The gene normalization task in BioCreative III. BMC Bioinformatics, 2011, 12, S2.	2.6	101
10	Automatic text categorization and its application to text retrieval. IEEE Transactions on Knowledge and Data Engineering, 1999, 11, 865-879.	5.7	98
11	Text Preprocessing for Text Mining in Organizational Research: Review and Recommendations. Organizational Research Methods, 2022, 25, 114-146.	9.1	84
12	Vocabulary mining for information retrieval: rough sets and fuzzy sets. Information Processing and Management, 2001, 37, 15-38.	8.6	83
13	A General Evaluation Framework for Topical Crawlers. Information Retrieval, 2005, 8, 417-447.	2.0	82
14	Link contexts in classifier-guided topical crawlers. IEEE Transactions on Knowledge and Data Engineering, 2006, 18, 107-122.	5.7	82
15	On the predictive ability of narrative disclosures in annual reports. European Journal of Operational Research, 2010, 202, 789-801.	5.7	81
16	Query expansion and MEDLINE. Information Processing and Management, 1996, 32, 431-443.	8.6	77
17	BioCreative III interactive task: an overview. BMC Bioinformatics, 2011, 12, S4.	2.6	65
18	Hierarchical neural networks for text categorization (poster abstract). , 1999, , .		60

#	ARTICLE	IF	CITATIONS
19	Engagement with Health Agencies on Twitter. PLoS ONE, 2014, 9, e112235.	2.5	58
20	Social media engagement analysis of U.S. Federal health agencies on Facebook. BMC Medical Informatics and Decision Making, 2017, 17, 49.	3.0	55
21	Life Satisfaction and the Pursuit of Happiness on Twitter. PLoS ONE, 2016, 11, e0150881.	2.5	52
22	GOP primary season on twitter. , 2013, , .		49
23	Detecting Wikipedia vandalism with active learning and statistical language models. , 2010, , .		45
24	Retrieval Feedback in MEDLINE. Journal of the American Medical Informatics Association: JAMIA, 1996, 3, 157-167.	4.4	40
25	The importance of rough approximations for information retrieval. International Journal of Man-Machine Studies, 1991, 34, 657-671.	0.7	37
26	MeSH: a window into full text for document summarization. Bioinformatics, 2011, 27, i120-i128.	4.1	35
27	Optimal document-indexing vocabulary for MEDLINE. Information Processing and Management, 1996, 32, 503-514.	8.6	33
28	Cross-language information retrieval with the UMLS metathesaurus. , 1998, , .		31
29	Online Asymmetric Active Learning with Imbalanced Data. , 2016, , .		24
30	Supervised learning models to predict firm performance with annual reports: An empirical study. Journal of the Association for Information Science and Technology, 2014, 65, 400-413.	2.9	22
31	Predicting Web Page Status. Information Systems Research, 2010, 21, 345-364.	3.7	20
32	Stock chatter: Using stock sentiment to predict price direction. Algorithmic Finance, 2013, 2, 169-196.	0.3	20
33	Retrieval with gene queries. BMC Bioinformatics, 2006, 7, 220.	2.6	18
34	Perceptions of presidential candidates' personalities in twitter. Journal of the Association for Information Science and Technology, 2016, 67, 249-267.	2.9	18
35	Integrated digital libraries: changing the essence of research, analysis and knowledge management. Journal of Knowledge Management, 1998, 2, 47-52.	5.1	17
36	Using Personal Traits For Brand Preference Prediction. , 2015, , .		17

#	Article	IF	Citations
37	Belief surveillance with Twitter. , 2012, , .		16
38	Revisiting The American Voter on Twitter. , 2017, , .		16
39	An investigation of content representation using text grammars. ACM Transactions on Information Systems, 1993, 11, 51-75.	4.9	15
40	Topic models and a revisit of text-related applications. , 2008, , .		13
41	Spam detection in online classified advertisements. , 2011, , .		13
42	Political speech in social media streams. , 2012, , .		13
43	Intelligent information retrieval using rough set approximations. Information Processing and Management, 1989, 25, 347-361.	8.6	12
44	Exploring the Forecasting Potential of Company Annual Reports. Proceedings of the American Society for Information Science and Technology, 2007, 43, 1-15.	0.2	12
45	Comparison of IPC and USPC classification systems in patent prior art searches. , 2010, , .		11
46	Status Locality on the Web: Implications for Building Focused Collections. Information Systems Research, 2013, 24, 802-821.	3.7	11
47	Bumps and Bruises. , 2017, , .		11
48	GO for gene documents. BMC Bioinformatics, 2007, 8, S3.	2.6	10
49	Ranking target objects of navigational queries. , 2006, , .		9
50	A Girl Has No Name: Automated Authorship Obfuscation using Mutant-X. Proceedings on Privacy Enhancing Technologies, 2019, 2019, 54-71.	2.8	9
51	Fuzzy versus probabilistic models for user relevance judgments. Journal of the Association for Information Science and Technology, 1990, 41, 264-271.	1.0	8
52	A quality-threshold data summarization algorithm. , 2008, , .		8
53	Using Annotations from Controlled Vocabularies to Find Meaningful Associations. , 2007, , 247-263.		8
54	A relevance-based topic model for news event tracking. , 2009, , .		8

#	ARTICLE	IF	CITATIONS
55	Distilling conceptual connections from MeSH co-occurrences. Studies in Health Technology and Informatics, 2004, 107, 808-12.	0.3	8
56	A comparison of two-poisson, inverse document frequency and discrimination value models of document representation. Information Processing and Management, 1990, 26, 269-278.	8.6	7
57	Accurate Detection of Automatically Spun Content via Stylometric Analysis. , 2017, , .		7
58	Using Classification Code Hierarchies for Patent Prior Art Searches. The Kluwer International Series on Information Retrieval, 2011, , 287-304.	1.0	7
59	A postmortem of suspended Twitter accounts in the 2016 U.S. presidential election. , 2019, , .		7
60	On generalizing the Two-Poisson Model. Journal of the Association for Information Science and Technology, 1990, 41, 61-66.	1.0	6
61	Adaptive Filtering of Newswire Stories using Two-Level Clustering. Information Retrieval, 2002, 5, 209-237.	2.0	6
62	Translating surveys to surveillance on social media. , 2014, , .		6
63	Comparing Crowd-Based, Game-Based, and Machine-Based Approaches in Initial Query and Query Refinement Tasks. Lecture Notes in Computer Science, 2013, , 495-506.	1.3	6
64	What's trending?. , 2011, , .		5
65	With a Little Help from the Crowd: Receiving Unauthorized Academic Assistance through Online Labor Markets. , 2012, , .		5
66	On the role of classification in patent invalidity searches. , 2009, , .		4
67	Exploiting Ontology Structure and Patterns of Annotation to Mine Significant Associations between Pairs of Controlled Vocabulary Terms. Lecture Notes in Computer Science, 2008, , 44-60.	1.3	4
68	News Event Modeling and Tracking in the Social Web with Ontological Guidance. , 2010, , .		3
69	Applying human computation mechanisms to information retrieval. Proceedings of the American Society for Information Science and Technology, 2012, 49, 1-10.	0.2	3
70	Adaptive classifiers, topic drifts and GO annotations. AMIA ... Annual Symposium proceedings, 2007, , 681-5.	0.2	3
71	Hybrid hierarchical classifiers for categorization of medical documents. Proceedings of the American Society for Information Science and Technology, 2005, 40, 65-70.	0.2	2
72	Manjal. , 2005, , .		2

#	Article	IF	Citations
73	Bayesian Embedding of Co-occurrence Data for Query-Based Visualization. , 2011, , .		2
74	A semantic approach to involve Twitter in LBD efforts. , 2012, , .		2
75	Ferret: a sentence-based literature scanning system. BMC Bioinformatics, 2015, 16, 198.	2.6	2
76	Disease Comorbidity Linkages between MEDLINE and Patient Data. , 2017, , .		2
77	Rough and Fuzzy Sets for Data Mining of a Controlled Vocabulary for Textual Retrieval. Studies in Fuzziness and Soft Computing, 2000, , 358-372.	0.8	2
78	Human Computation for Information Retrieval. , 2013, , 205-214.		2
79	GO for gene documents. , 2006, , .		1
80	Predicting Web Page Status. SSRN Electronic Journal, 0, , .	0.4	1
81	Data-Driven Methods for SMS-Based FAQ Retrieval. Lecture Notes in Computer Science, 2013, , 104-118.	1.3	1
82	A Latent Dirichlet Framework for Relevance Modeling. Lecture Notes in Computer Science, 2009, , 13-25.	1.3	1
83	My Word! Machine versus Human Computation Methods for Identifying and Resolving Acronyms. Computacion Y Sistemas, 2019, 23, .	0.3	1
84	Literature-based discovery: New trends and techniques. Sponsored by SIG HCI. Proceedings of the American Society for Information Science and Technology, 2005, 40, 497-497.	0.2	0
85	Ontological research and its applications to the biomedical domain. Proceedings of the American Society for Information Science and Technology, 2006, 42, n/a-n/a.	0.2	0
86	Identifying De Facto Prescription Norms from Clinical Data. , 2019, , .		0
87	Identifying De Facto Prescription Norms in a Hospital Setting: A study with antibiotics. , 2019, , .		0
88	Text Mining of Biological Resources. , 2009, , 3065-3068.		0
89	Hybrid Crowd-Machine Methods as Alternatives to Pooling and Expert Judgments. Lecture Notes in Computer Science, 2014, , 60-72.	1.3	0
90	Text Mining of Biological Resources. , 2017, , 1-5.		0

#	Article	IF	Citations
91	Text Mining of Biological Resources. , 2018, , 4092-4096.		0