

Marcos Andr Goncalves

List of Publications by Citations

Source: <https://exaly.com/author-pdf/3389289/marcos-andre-goncalves-publications-by-citations.pdf>

Version: 2024-04-27

This document has been generated based on the publications and citations recorded by exaly.com. For the latest version of this publication list, visit the link given above.

The third column is the impact factor (IF) of the journal, and the fourth column is the number of citations of the article.

159
papers

3,286
citations

31
h-index

50
g-index

175
ext. papers

3,920
ext. citations

3.5
avg, IF

5.47
L-index

#	Paper	IF	Citations
159	Using early view patterns to predict the popularity of youtube videos 2013 ,		199
158	SentiBench - a benchmark comparison of state-of-the-practice sentiment analysis methods. <i>EPJ Data Science</i> , 2016 , 5,	3.4	182
157	A brief survey of automatic methods for author name disambiguation. <i>SIGMOD Record</i> , 2012 , 41, 15-26	1.1	170
156	Streams, structures, spaces, scenarios, societies (5s). <i>ACM Transactions on Information Systems</i> , 2004 , 22, 270-312	4.8	166
155	A genetic programming framework for content-based image retrieval. <i>Pattern Recognition</i> , 2009 , 42, 283-292	7.7	120
154	Learning to advertise 2006 ,		92
153	Recommender Systems Research: A Connection-Centric Survey. <i>Journal of Intelligent Information Systems</i> , 2004 , 23, 107-143	2.1	91
152	What is a good digital library? A quality model for digital libraries. <i>Information Processing and Management</i> , 2007 , 43, 1416-1437	6.3	69
151	A source independent framework for research paper recommendation 2011 ,		66
150	An unsupervised heuristic-based hierarchical method for name disambiguation in bibliographic citations. <i>Journal of the Association for Information Science and Technology</i> , 2010 , 61, 1853-1870		66
149	Word co-occurrence features for text classification. <i>Information Systems</i> , 2011 , 36, 843-858	2.7	60
148	Relevance feedback based on genetic programming for image retrieval. <i>Pattern Recognition Letters</i> , 2011 , 32, 27-37	4.7	59
147	Using web information for author name disambiguation 2009 ,		55
146	Automatic quality assessment of content created collaboratively by web communities 2009 ,		52
145	On the Dynamics of Social Media Popularity. <i>ACM Transactions on Internet Technology</i> , 2014 , 14, 1-23	3.8	50
144	A Genetic Programming Approach to Record Deduplication. <i>IEEE Transactions on Knowledge and Data Engineering</i> , 2012 , 24, 399-412	4.2	49
143	Exploiting user feedback to learn to rank answers in q&a forums 2013 ,		48

142	Effective self-training author name disambiguation in scholarly digital libraries 2010 ,		44
141	A relevance feedback method based on genetic programming for classification of remote sensing images. <i>Information Sciences</i> , 2011 , 181, 2671-2684	7.7	43
140	A Genetic Programming approach for feature selection in highly dimensional skewed data. <i>Neurocomputing</i> , 2018 , 273, 554-569	5.4	42
139	Sentiment-based influence detection on Twitter. <i>Journal of the Brazilian Computer Society</i> , 2012 , 18, 169-183	1.9	39
138	Link-based similarity measures for the classification of Web documents. <i>Journal of the Association for Information Science and Technology</i> , 2006 , 57, 208-221		38
137	FLUX-CIM 2007 ,		38
136	A combined component approach for finding collection-adapted ranking functions based on genetic programming 2007 ,		38
135	Cost-effective on-demand associative author name disambiguation. <i>Information Processing and Management</i> , 2012 , 48, 680-697	6.3	36
134	Assessing the quality of textual features in social media. <i>Information Processing and Management</i> , 2013 , 49, 222-247	6.3	35
133	Personalized and object-centered tag recommendation methods for Web 2.0 applications. <i>Information Processing and Management</i> , 2014 , 50, 524-553	6.3	34
132	Learning to rank at query-time using association rules 2008 ,		33
131	Towards a digital library theory: a formal digital library ontology. <i>International Journal on Digital Libraries</i> , 2008 , 8, 91-114	1.4	33
130	Exploiting New Sentiment-Based Meta-level Features for Effective Sentiment Analysis 2016 ,		31
129	Learning to rank for content-based image retrieval 2010 ,		31
128	A survey on tag recommendation methods. <i>Journal of the Association for Information Science and Technology</i> , 2017 , 68, 830-844	2.7	29
127	TrendLearner: Early prediction of popularity trends of user generated content. <i>Information Sciences</i> , 2016 , 349-350, 172-187	7.7	29
126	A comparative study of citations and links in document classification 2006 ,		28
125	Improving random forests by neighborhood projection for effective text classification. <i>Information Systems</i> , 2018 , 77, 1-21	2.7	27

124	Automatic evaluation of digital libraries with 5SQual. <i>Journal of Informetrics</i> , 2009 , 3, 102-123	3.1	26
123	Self-training author name disambiguation for information scarce scenarios. <i>Journal of the Association for Information Science and Technology</i> , 2014 , 65, 1257-1278	2.7	25
122	On the combination of domain-specific heuristics for author name disambiguation: the nearest cluster method. <i>International Journal on Digital Libraries</i> , 2015 , 16, 229-246	1.4	23
121	Automatic Assessment of Document Quality in Web Collaborative Digital Libraries. <i>Journal of Data and Information Quality</i> , 2011 , 2, 1-30	2.5	23
120	Assessing the quality of scientific conferences based on bibliographic citations. <i>Scientometrics</i> , 2010 , 83, 133-155	3	23
119	Calibrated lazy associative classification. <i>Information Sciences</i> , 2011 , 181, 2656-2670	7.7	21
118	Incremental author name disambiguation by exploiting domain-specific heuristics. <i>Journal of the Association for Information Science and Technology</i> , 2017 , 68, 931-945	2.7	20
117	Predicting the popularity of micro-reviews: A Foursquare case study. <i>Information Sciences</i> , 2015 , 325, 355-374	7.7	20
116	A general multiview framework for assessing the quality of collaboratively created content on web 2.0. <i>Journal of the Association for Information Science and Technology</i> , 2017 , 68, 286-308	2.7	19
115	ONDUX 2010,		19
114	Keeping a digital library clean 2008 ,		19
113	Computing block importance for searching on web sites 2007 ,		19
112	An evolutionary approach for combining different sources of evidence in search engines. <i>Information Systems</i> , 2009 , 34, 276-289	2.7	18
111	Beyond Relevance. <i>ACM Transactions on Intelligent Systems and Technology</i> , 2016 , 7, 1-34	8	17
110	Ranked batch-mode active learning. <i>Information Sciences</i> , 2017 , 379, 313-337	7.7	17
109	Learning to deduplicate 2006 ,		17
108	A Feature-Oriented Sentiment Rating for Mobile App Reviews 2018 ,		16
107	He votes or she votes? Female and male discursive strategies in Twitter political hashtags. <i>PLoS ONE</i> , 2014 , 9, e87041	3.7	16

106	Replica identification using genetic programming 2008 ,		16
105	A tool for generating synthetic authorship records for evaluating author name disambiguation methods. <i>Information Sciences</i> , 2012 , 206, 42-62	7.7	15
104	Multimodal retrieval with relevance feedback based on genetic programming. <i>Multimedia Tools and Applications</i> , 2014 , 69, 991-1019	2.5	15
103	Tuning Genetic Programming parameters with factorial designs 2010 ,		15
102	On the Quality of Information for Web 2.0 Services. <i>IEEE Internet Computing</i> , 2010 , 14, 47-55	2.4	15
101	BLOSS: Effective meta-blocking with almost no effort. <i>Information Systems</i> , 2018 , 75, 75-89	2.7	14
100	Demand-Driven Tag Recommendation. <i>Lecture Notes in Computer Science</i> , 2010 , 402-417	0.9	14
99	On Popularity in the Blogosphere. <i>IEEE Internet Computing</i> , 2010 , 14, 42-49	2.4	14
98	A gender based study of tagging behavior in twitter 2012 ,		14
97	On cold start for associative tag recommendation. <i>Journal of the Association for Information Science and Technology</i> , 2016 , 67, 83-105	2.7	14
96	Geographical classification of documents using evidence from Wikipedia 2010 ,		13
95	The effectiveness of automatically structured queries in digital libraries 2004 ,		13
94	Prototyping Digital Libraries Handling Heterogeneous Data Sources □The ETANA-DL Case Study. <i>Lecture Notes in Computer Science</i> , 2004 , 186-197	0.9	13
93	Fine-grained tourism prediction: Impact of social and environmental features. <i>Information Processing and Management</i> , 2020 , 57, 102057	6.3	13
92	A Practical and Effective Sampling Selection Strategy for Large Scale Deduplication. <i>IEEE Transactions on Knowledge and Data Engineering</i> , 2015 , 27, 2305-2319	4.2	12
91	Active associative sampling for author name disambiguation 2012 ,		12
90	An unsupervised heuristic-based approach for bibliographic metadata deduplication. <i>Information Processing and Management</i> , 2011 , 47, 706-718	6.3	12
89	Learning to assess the quality of scientific conferences 2009 ,		12

88	Exploiting temporal contexts in text classification 2008 ,		12
87	Exploiting Genre in Focused Crawling. <i>Lecture Notes in Computer Science</i> , 2007 , 62-73	0.9	12
86	CluWords 2019 ,		11
85	A Thorough Evaluation of Distance-Based Meta-Features for Automated Text Classification. <i>IEEE Transactions on Knowledge and Data Engineering</i> , 2018 , 30, 2242-2256	4.2	11
84	On interactive learning-to-rank for IR: Overview, recent advances, challenges, and directions. <i>Neurocomputing</i> , 2016 , 208, 3-24	5.4	11
83	Early Prediction of Scholar Popularity 2016 ,		11
82	Impact of Big Data Analytics on People's Health: Overview of Systematic Reviews and Recommendations for Future Studies. <i>Journal of Medical Internet Research</i> , 2021 , 23, e27275	7.6	11
81	Rule-Based Active Sampling for Learning to Rank. <i>Lecture Notes in Computer Science</i> , 2011 , 240-255	0.9	10
80	BDBComp 2004 ,		10
79	Exploiting syntactic and neighbourhood attributes to address cold start in tag recommendation. <i>Information Processing and Management</i> , 2019 , 56, 771-790	6.3	9
78	On Tag Recommendation for Expertise Profiling 2015 ,		9
77	Extended pre-processing pipeline for text classification: On the role of meta-feature representations, sparsification and selective sampling. <i>Information Processing and Management</i> , 2020 , 57, 102263	6.3	9
76	A multimodal query expansion based on genetic programming for visually-oriented e-commerce applications. <i>Information Processing and Management</i> , 2016 , 52, 783-800	6.3	9
75	A flexible approach for extracting metadata from bibliographic citations. <i>Journal of the Association for Information Science and Technology</i> , 2009 , 60, 1144-1158		9
74	A Genre-Aware Approach to Focused Crawling. <i>World Wide Web</i> , 2009 , 12, 285-319	2.9	9
73	Classifying documents with link-based bibliometric measures. <i>Information Retrieval</i> , 2010 , 13, 315-345	1.8	9
72	A Spatiotemporal Analysis of Brazilian Science from the Perspective of Researchers' Career Trajectories. <i>PLoS ONE</i> , 2015 , 10, e0141528	3.7	9
71	A Two-Stage Machine learning approach for temporally-robust text classification. <i>Information Systems</i> , 2017 , 69, 40-58	2.7	8

70	Characterizing scholar popularity: A case study in the Computer Science research community 2014 ,		8
69	Stacking Bagged and Boosted Forests for Effective Automated Classification 2017 ,		8
68	A generic Web-based entity resolution framework. <i>Journal of the Association for Information Science and Technology</i> , 2011 , 62, 919-932		8
67	Visual Semantic Modeling of Digital Libraries. <i>Lecture Notes in Computer Science</i> , 2003 , 325-337	0.9	8
66	The LExR Collection for Expertise Retrieval in Academia 2016 ,		8
65	Fixing the curse of the bad product descriptions: Search-boosted tag recommendation for E-commerce products. <i>Information Processing and Management</i> , 2020 , 57, 102289	6.3	7
64	Generalized BROOF-L2R 2016 ,		7
63	A quantitative analysis of the temporal effects on automatic text classification. <i>Journal of the Association for Information Science and Technology</i> , 2016 , 67, 1639-1667	2.7	7
62	Exploiting efficient and effective lazy Semi-Bayesian strategies for text classification. <i>Neurocomputing</i> , 2018 , 307, 153-171	5.4	7
61	Hierarchical Clustering-Based Graphs for Large Scale Approximate Nearest Neighbor Search. <i>Pattern Recognition</i> , 2019 , 96, 106970	7.7	7
60	A Two-stage active learning method for learning to rank. <i>Journal of the Association for Information Science and Technology</i> , 2014 , 65, 109-128	2.7	7
59	Finding what is missing from a digital library: A case study in the Computer Science field. <i>Information Processing and Management</i> , 2009 , 45, 380-391	6.3	7
58	On the combination of "off-the-shelf" sentiment analysis methods 2016 ,		7
57	On the cost-effectiveness of neural and non-neural approaches and representations for text classification: A comprehensive comparative study. <i>Information Processing and Management</i> , 2021 , 58, 102481	6.3	7
56	Automatic Disambiguation of Author Names in Bibliographic Repositories. <i>Synthesis Lectures on Information Concepts, Retrieval, and Services</i> , 2020 , 12, 1-146	1.7	6
55	A relevance feedback approach for the author name disambiguation problem 2013 ,		6
54	Temporal contexts: Effective text classification in evolving document collections. <i>Information Systems</i> , 2013 , 38, 388-409	2.7	6
53	On Efficient Meta-Level Features for Effective Text Classification 2014 ,		6

52	A digital library environment for integrating, disseminating and exploring ecological data. <i>Ecological Informatics</i> , 2008 , 3, 295-308	4.2	6
51	Parallel Lazy Semi-Naive Bayes Strategies for Effective and Efficient Document Classification 2015 ,		6
50	Diversity-based interactive learning meets multimodality. <i>Neurocomputing</i> , 2017 , 259, 159-175	5.4	5
49	Incorporating Risk-Sensitiveness into Feature Selection for Learning to Rank 2016 ,		5
48	An evolutionary approach to complex schema matching. <i>Information Systems</i> , 2013 , 38, 302-316	2.7	5
47	Quality assessment of collaborative content with minimal information 2014 ,		5
46	Building a research social network from an individual perspective 2011 ,		5
45	Using structural information to improve search in Web collections. <i>Journal of the Association for Information Science and Technology</i> , 2010 , 61, 2503-2513		5
44	A Web services-based framework for building componentized digital libraries. <i>Journal of Systems and Software</i> , 2008 , 81, 809-822	3.3	5
43	Risk-Sensitive Learning to Rank with Evolutionary Multi-Objective Feature Selection. <i>ACM Transactions on Information Systems</i> , 2019 , 37, 1-34	4.8	5
42	10SENT: A stable sentiment analysis method based on the combination of off-the-shelf approaches. <i>Journal of the Association for Information Science and Technology</i> , 2019 , 70, 242-255	2.7	5
41	Bag of textual graphs (BoTG): A general graph-based text representation model. <i>Journal of the Association for Information Science and Technology</i> , 2019 , 70, 817-829	2.7	5
40	A pragmatic approach to hierarchical categorization of research expertise in the presence of scarce information. <i>International Journal on Digital Libraries</i> , 2020 , 21, 61-73	1.4	5
39	Tuning large scale deduplication with reduced effort 2013 ,		4
38	Exploiting co-occurrence and information quality metrics to recommend tags in web 2.0 applications 2010 ,		4
37	GreenWiki 2011 ,		4
36	A Soft Computing Approach for Learning to Aggregate Rankings 2015 ,		4
35	Exploiting semantic relationships for unsupervised expansion of sentiment lexicons. <i>Information Systems</i> , 2020 , 94, 101606	2.7	4

34	Semantically-Enhanced Topic Modeling 2018 ,		4
33	Incremental, Semi-automatic, Mapping-Based Integration of Heterogeneous Collections into Archaeological Digital Libraries: Megiddo Case Study. <i>Lecture Notes in Computer Science</i> , 2005 , 139-150	0.9	4
32	Design, Implementation, and Evaluation of a Wizard Tool for Setting Up Component-Based Digital Libraries. <i>Lecture Notes in Computer Science</i> , 2006 , 135-146	0.9	4
31	Compression-Based Selective Sampling for Learning to Rank 2016 ,		3
30	Similarity-Based Synthetic Document Representations for Meta-Feature Generation in Text Classification 2019 ,		3
29	How you post is who you are 2014 ,		3
28	Combining domain-specific heuristics for author name disambiguation 2014 ,		3
27	Evaluation of parameters for combining multiple textual sources of evidence for Web image retrieval using genetic programming. <i>Journal of the Brazilian Computer Society</i> , 2013 , 19, 147-160	1.9	3
26	From concepts to implementation and visualization 2008 ,		3
25	Transitioning from the ecological fieldwork to an online repository: a digital library solution and evaluation. <i>International Journal on Digital Libraries</i> , 2007 , 7, 109-112	1.4	3
24	SyGAR Δ Synthetic Data Generator for Evaluating Name Disambiguation Methods. <i>Lecture Notes in Computer Science</i> , 2009 , 437-441	0.9	3
23	Combining Classifiers and User Feedback for Disambiguating Author Names 2015 ,		2
22	On MultiView-Based Meta-learning for Automatic Quality Assessment of Wiki Articles. <i>Lecture Notes in Computer Science</i> , 2012 , 234-246	0.9	2
21	Can the vector space model be used to identify biological entity activities?. <i>BMC Genomics</i> , 2011 , 12 Suppl 4, S1	4.5	2
20	Exploiting relevance, novelty and diversity in tag recommendation 2012 ,		2
19	Advertisement selection for online videos 2012 ,		2
18	The impact of term selection in genre-aware focused crawling 2008 ,		2
17	Automatic Hierarchical Categorization of Research Expertise Using Minimum Information. <i>Lecture Notes in Computer Science</i> , 2017 , 103-115	0.9	2

16	Improving On-Demand Learning to Rank through Parallelism. <i>Lecture Notes in Computer Science</i> , 2012 , 526-537	0.9	2
15	Automatic Vandalism Detection in Wikipedia with Active Associative Classification. <i>Lecture Notes in Computer Science</i> , 2012 , 138-143	0.9	2
14	Individualized extreme dominance (IndED): A new preference-based method for multi-objective recommender systems. <i>Information Sciences</i> , 2021 , 572, 558-573	7.7	2
13	Automatic Methods for Disambiguating Author Names in Bibliographic Data Repositories 2015 ,		1
12	On the Impact of Academic Factors on Scholar Popularity: A Cross-Area Study. <i>Lecture Notes in Computer Science</i> , 2015 , 139-152	0.9	1
11	Diversity-driven learning for multimodal image retrieval with relevance feedback 2014 ,		1
10	PaMS: A component-based service for finding the missing full text of articles cataloged in a digital library. <i>Information Systems</i> , 2010 , 35, 544-556	2.7	1
9	FISETIO: A Fine-grained, Structured and Enriched Tourism Dataset for Indoor and Outdoor attractions. <i>Data in Brief</i> , 2020 , 28, 104906	1.2	1
8	Tagging and Tag Recommendation 2020 ,		1
7	A bias-variance analysis of state-of-the-art random forest text classifiers. <i>Advances in Data Analysis and Classification</i> , 2021 , 15, 379-405	1.8	1
6	Quality assessment of collaboratively-created web content with no manual intervention based on soft multi-view generation. <i>Expert Systems With Applications</i> , 2019 , 132, 226-238	7.8	0
5	Stroke Outcome Measurements From Electronic Medical Records: Cross-sectional Study on the Effectiveness of Neural and Nonneural Classifiers. <i>JMIR Medical Informatics</i> , 2021 , 9, e29120	3.6	0
4	Dissecting a Scholar Popularity Ranking into Different Knowledge Areas. <i>Lecture Notes in Computer Science</i> , 2016 , 253-265	0.9	0
3	Parallel rule-based selective sampling and on-demand learning to rank. <i>Concurrency Computation Practice and Experience</i> , 2019 , 31, e4464	1.4	0
2	Document Performance Prediction for Automatic Text Classification. <i>Lecture Notes in Computer Science</i> , 2019 , 132-139	0.9	
1	On-Demand Associative Cross-Language Information Retrieval. <i>Lecture Notes in Computer Science</i> , 2009 , 165-173	0.9	