

# Marcos Andr Goncalves

## List of Publications by Year in Descending Order

**Source:** <https://exaly.com/author-pdf/3389289/marcos-andre-goncalves-publications-by-year.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	Stroke Outcome Measurements From Electronic Medical Records: Cross-sectional Study on the Effectiveness of Neural and Nonneural Classifiers. <i>JMIR Medical Informatics</i> , <b>2021</b> , 9, e29120	3.6	0
158	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> , <b>2021</b> , 23, e27275	7.6	11
157	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> , <b>2021</b> , 58, 102481	6.3	7
156	A bias-variance analysis of state-of-the-art random forest text classifiers. <i>Advances in Data Analysis and Classification</i> , <b>2021</b> , 15, 379-405	1.8	1
155	Individualized extreme dominance (IndED): A new preference-based method for multi-objective recommender systems. <i>Information Sciences</i> , <b>2021</b> , 572, 558-573	7.7	2
154	Fixing the curse of the bad product descriptions: Search-boosted tag recommendation for E-commerce products. <i>Information Processing and Management</i> , <b>2020</b> , 57, 102289	6.3	7
153	Automatic Disambiguation of Author Names in Bibliographic Repositories. <i>Synthesis Lectures on Information Concepts, Retrieval, and Services</i> , <b>2020</b> , 12, 1-146	1.7	6
152	Extended pre-processing pipeline for text classification: On the role of meta-feature representations, sparsification and selective sampling. <i>Information Processing and Management</i> , <b>2020</b> , 57, 102263	6.3	9
151	FISATIO: A Fine-grained, Structured and Enriched Tourism Dataset for Indoor and Outdoor attractions. <i>Data in Brief</i> , <b>2020</b> , 28, 104906	1.2	1
150	Exploiting semantic relationships for unsupervised expansion of sentiment lexicons. <i>Information Systems</i> , <b>2020</b> , 94, 101606	2.7	4
149	Tagging and Tag Recommendation <b>2020</b> ,		1
148	A pragmatic approach to hierarchical categorization of research expertise in the presence of scarce information. <i>International Journal on Digital Libraries</i> , <b>2020</b> , 21, 61-73	1.4	5
147	Fine-grained tourism prediction: Impact of social and environmental features. <i>Information Processing and Management</i> , <b>2020</b> , 57, 102057	6.3	13
146	Exploiting syntactic and neighbourhood attributes to address cold start in tag recommendation. <i>Information Processing and Management</i> , <b>2019</b> , 56, 771-790	6.3	9
145	CluWords <b>2019</b> ,		11
144	Quality assessment of collaboratively-created web content with no manual intervention based on soft multi-view generation. <i>Expert Systems With Applications</i> , <b>2019</b> , 132, 226-238	7.8	0
143	Document Performance Prediction for Automatic Text Classification. <i>Lecture Notes in Computer Science</i> , <b>2019</b> , 132-139	0.9	

142	Hierarchical Clustering-Based Graphs for Large Scale Approximate Nearest Neighbor Search. <i>Pattern Recognition</i> , <b>2019</b> , 96, 106970	7.7	7
141	Similarity-Based Synthetic Document Representations for Meta-Feature Generation in Text Classification <b>2019</b> ,		3
140	Risk-Sensitive Learning to Rank with Evolutionary Multi-Objective Feature Selection. <i>ACM Transactions on Information Systems</i> , <b>2019</b> , 37, 1-34	4.8	5
139	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> , <b>2019</b> , 70, 242-255	2.7	5
138	Bag of textual graphs (BoTG): A general graph-based text representation model. <i>Journal of the Association for Information Science and Technology</i> , <b>2019</b> , 70, 817-829	2.7	5
137	Parallel rule-based selective sampling and on-demand learning to rank. <i>Concurrency Computation Practice and Experience</i> , <b>2019</b> , 31, e4464	1.4	0
136	A Thorough Evaluation of Distance-Based Meta-Features for Automated Text Classification. <i>IEEE Transactions on Knowledge and Data Engineering</i> , <b>2018</b> , 30, 2242-2256	4.2	11
135	BLOSS: Effective meta-blocking with almost no effort. <i>Information Systems</i> , <b>2018</b> , 75, 75-89	2.7	14
134	A Genetic Programming approach for feature selection in highly dimensional skewed data. <i>Neurocomputing</i> , <b>2018</b> , 273, 554-569	5.4	42
133	A Feature-Oriented Sentiment Rating for Mobile App Reviews <b>2018</b> ,		16
132	Exploiting efficient and effective lazy Semi-Bayesian strategies for text classification. <i>Neurocomputing</i> , <b>2018</b> , 307, 153-171	5.4	7
131	Semantically-Enhanced Topic Modeling <b>2018</b> ,		4
130	Improving random forests by neighborhood projection for effective text classification. <i>Information Systems</i> , <b>2018</b> , 77, 1-21	2.7	27
129	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> , <b>2017</b> , 68, 286-308	2.7	19
128	Diversity-based interactive learning meets multimodality. <i>Neurocomputing</i> , <b>2017</b> , 259, 159-175	5.4	5
127	A Two-Stage Machine learning approach for temporally-robust text classification. <i>Information Systems</i> , <b>2017</b> , 69, 40-58	2.7	8
126	Incremental author name disambiguation by exploiting domain-specific heuristics. <i>Journal of the Association for Information Science and Technology</i> , <b>2017</b> , 68, 931-945	2.7	20
125	A survey on tag recommendation methods. <i>Journal of the Association for Information Science and Technology</i> , <b>2017</b> , 68, 830-844	2.7	29

124	Stacking Bagged and Boosted Forests for Effective Automated Classification <b>2017</b> ,		8
123	Ranked batch-mode active learning. <i>Information Sciences</i> , <b>2017</b> , 379, 313-337	7.7	17
122	Automatic Hierarchical Categorization of Research Expertise Using Minimum Information. <i>Lecture Notes in Computer Science</i> , <b>2017</b> , 103-115	0.9	2
121	Compression-Based Selective Sampling for Learning to Rank <b>2016</b> ,		3
120	Incorporating Risk-Sensitiveness into Feature Selection for Learning to Rank <b>2016</b> ,		5
119	SentiBench - a benchmark comparison of state-of-the-practice sentiment analysis methods. <i>EPJ Data Science</i> , <b>2016</b> , 5,	3.4	182
118	Generalized BROOF-L2R <b>2016</b> ,		7
117	On interactive learning-to-rank for IR: Overview, recent advances, challenges, and directions. <i>Neurocomputing</i> , <b>2016</b> , 208, 3-24	5.4	11
116	A quantitative analysis of the temporal effects on automatic text classification. <i>Journal of the Association for Information Science and Technology</i> , <b>2016</b> , 67, 1639-1667	2.7	7
115	Early Prediction of Scholar Popularity <b>2016</b> ,		11
114	A multimodal query expansion based on genetic programming for visually-oriented e-commerce applications. <i>Information Processing and Management</i> , <b>2016</b> , 52, 783-800	6.3	9
113	Exploiting New Sentiment-Based Meta-level Features for Effective Sentiment Analysis <b>2016</b> ,		31
112	Beyond Relevance. <i>ACM Transactions on Intelligent Systems and Technology</i> , <b>2016</b> , 7, 1-34	8	17
111	On the combination of "off-the-shelf" sentiment analysis methods <b>2016</b> ,		7
110	Dissecting a Scholar Popularity Ranking into Different Knowledge Areas. <i>Lecture Notes in Computer Science</i> , <b>2016</b> , 253-265	0.9	0
109	The LExR Collection for Expertise Retrieval in Academia <b>2016</b> ,		8
108	TrendLearner: Early prediction of popularity trends of user generated content. <i>Information Sciences</i> , <b>2016</b> , 349-350, 172-187	7.7	29
107	On cold start for associative tag recommendation. <i>Journal of the Association for Information Science and Technology</i> , <b>2016</b> , 67, 83-105	2.7	14

106	Automatic Methods for Disambiguating Author Names in Bibliographic Data Repositories <b>2015</b> ,		1
105	On the combination of domain-specific heuristics for author name disambiguation: the nearest cluster method. <i>International Journal on Digital Libraries</i> , <b>2015</b> , 16, 229-246	1.4	23
104	Predicting the popularity of micro-reviews: A Foursquare case study. <i>Information Sciences</i> , <b>2015</b> , 325, 355-374	7.7	20
103	Combining Classifiers and User Feedback for Disambiguating Author Names <b>2015</b> ,		2
102	A Practical and Effective Sampling Selection Strategy for Large Scale Deduplication. <i>IEEE Transactions on Knowledge and Data Engineering</i> , <b>2015</b> , 27, 2305-2319	4.2	12
101	On the Impact of Academic Factors on Scholar Popularity: A Cross-Area Study. <i>Lecture Notes in Computer Science</i> , <b>2015</b> , 139-152	0.9	1
100	On Tag Recommendation for Expertise Profiling <b>2015</b> ,		9
99	A Soft Computing Approach for Learning to Aggregate Rankings <b>2015</b> ,		4
98	Parallel Lazy Semi-Naive Bayes Strategies for Effective and Efficient Document Classification <b>2015</b> ,		6
97	A Spatiotemporal Analysis of Brazilian Science from the Perspective of Researchers' Career Trajectories. <i>PLoS ONE</i> , <b>2015</b> , 10, e0141528	3.7	9
96	A Two-stage active learning method for learning to rank. <i>Journal of the Association for Information Science and Technology</i> , <b>2014</b> , 65, 109-128	2.7	7
95	Self-training author name disambiguation for information scarce scenarios. <i>Journal of the Association for Information Science and Technology</i> , <b>2014</b> , 65, 1257-1278	2.7	25
94	Personalized and object-centered tag recommendation methods for Web 2.0 applications. <i>Information Processing and Management</i> , <b>2014</b> , 50, 524-553	6.3	34
93	Characterizing scholar popularity: A case study in the Computer Science research community <b>2014</b> ,		8
92	He votes or she votes? Female and male discursive strategies in Twitter political hashtags. <i>PLoS ONE</i> , <b>2014</b> , 9, e87041	3.7	16
91	How you post is who you are <b>2014</b> ,		3
90	On the Dynamics of Social Media Popularity. <i>ACM Transactions on Internet Technology</i> , <b>2014</b> , 14, 1-23	3.8	50
89	Diversity-driven learning for multimodal image retrieval with relevance feedback <b>2014</b> ,		1

88	Quality assessment of collaborative content with minimal information <b>2014</b> ,		5
87	Combining domain-specific heuristics for author name disambiguation <b>2014</b> ,		3
86	On Efficient Meta-Level Features for Effective Text Classification <b>2014</b> ,		6
85	Multimodal retrieval with relevance feedback based on genetic programming. <i>Multimedia Tools and Applications</i> , <b>2014</b> , 69, 991-1019	2.5	15
84	Exploiting user feedback to learn to rank answers in q&a forums <b>2013</b> ,		48
83	An evolutionary approach to complex schema matching. <i>Information Systems</i> , <b>2013</b> , 38, 302-316	2.7	5
82	Using early view patterns to predict the popularity of youtube videos <b>2013</b> ,		199
81	A relevance feedback approach for the author name disambiguation problem <b>2013</b> ,		6
80	Temporal contexts: Effective text classification in evolving document collections. <i>Information Systems</i> , <b>2013</b> , 38, 388-409	2.7	6
79	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> , <b>2013</b> , 19, 147-160	1.9	3
78	Assessing the quality of textual features in social media. <i>Information Processing and Management</i> , <b>2013</b> , 49, 222-247	6.3	35
77	Tuning large scale deduplication with reduced effort <b>2013</b> ,		4
76	A tool for generating synthetic authorship records for evaluating author name disambiguation methods. <i>Information Sciences</i> , <b>2012</b> , 206, 42-62	7.7	15
75	Cost-effective on-demand associative author name disambiguation. <i>Information Processing and Management</i> , <b>2012</b> , 48, 680-697	6.3	36
74	Active associative sampling for author name disambiguation <b>2012</b> ,		12
73	On MultiView-Based Meta-learning for Automatic Quality Assessment of Wiki Articles. <i>Lecture Notes in Computer Science</i> , <b>2012</b> , 234-246	0.9	2
72	A Genetic Programming Approach to Record Deduplication. <i>IEEE Transactions on Knowledge and Data Engineering</i> , <b>2012</b> , 24, 399-412	4.2	49
71	Sentiment-based influence detection on Twitter. <i>Journal of the Brazilian Computer Society</i> , <b>2012</b> , 18, 169-183	1.9	39

70	Exploiting relevance, novelty and diversity in tag recommendation <b>2012</b> ,		2
69	A brief survey of automatic methods for author name disambiguation. <i>SIGMOD Record</i> , <b>2012</b> , 41, 15-26	1.1	170
68	Advertisement selection for online videos <b>2012</b> ,		2
67	A gender based study of tagging behavior in twitter <b>2012</b> ,		14
66	Improving On-Demand Learning to Rank through Parallelism. <i>Lecture Notes in Computer Science</i> , <b>2012</b> , 526-537	0.9	2
65	Automatic Vandalism Detection in Wikipedia with Active Associative Classification. <i>Lecture Notes in Computer Science</i> , <b>2012</b> , 138-143	0.9	2
64	An unsupervised heuristic-based approach for bibliographic metadata deduplication. <i>Information Processing and Management</i> , <b>2011</b> , 47, 706-718	6.3	12
63	Can the vector space model be used to identify biological entity activities?. <i>BMC Genomics</i> , <b>2011</b> , 12 Suppl 4, S1	4.5	2
62	A generic Web-based entity resolution framework. <i>Journal of the Association for Information Science and Technology</i> , <b>2011</b> , 62, 919-932		8
61	A relevance feedback method based on genetic programming for classification of remote sensing images. <i>Information Sciences</i> , <b>2011</b> , 181, 2671-2684	7.7	43
60	Calibrated lazy associative classification. <i>Information Sciences</i> , <b>2011</b> , 181, 2656-2670	7.7	21
59	A source independent framework for research paper recommendation <b>2011</b> ,		66
58	Relevance feedback based on genetic programming for image retrieval. <i>Pattern Recognition Letters</i> , <b>2011</b> , 32, 27-37	4.7	59
57	Word co-occurrence features for text classification. <i>Information Systems</i> , <b>2011</b> , 36, 843-858	2.7	60
56	GreenWiki <b>2011</b> ,		4
55	Building a research social network from an individual perspective <b>2011</b> ,		5
54	Rule-Based Active Sampling for Learning to Rank. <i>Lecture Notes in Computer Science</i> , <b>2011</b> , 240-255	0.9	10
53	Automatic Assessment of Document Quality in Web Collaborative Digital Libraries. <i>Journal of Data and Information Quality</i> , <b>2011</b> , 2, 1-30	2.5	23

52	Tuning Genetic Programming parameters with factorial designs <b>2010</b> ,		15
51	Exploiting co-occurrence and information quality metrics to recommend tags in web 2.0 applications <b>2010</b> ,		4
50	Effective self-training author name disambiguation in scholarly digital libraries <b>2010</b> ,		44
49	ONDUX <b>2010</b> ,		19
48	Demand-Driven Tag Recommendation. <i>Lecture Notes in Computer Science</i> , <b>2010</b> , 402-417	0.9	14
47	On Popularity in the Blogosphere. <i>IEEE Internet Computing</i> , <b>2010</b> , 14, 42-49	2.4	14
46	Learning to rank for content-based image retrieval <b>2010</b> ,		31
45	Geographical classification of documents using evidence from Wikipedia <b>2010</b> ,		13
44	PaMS: A component-based service for finding the missing full text of articles cataloged in a digital library. <i>Information Systems</i> , <b>2010</b> , 35, 544-556	2.7	1
43	Classifying documents with link-based bibliometric measures. <i>Information Retrieval</i> , <b>2010</b> , 13, 315-345	1.8	9
42	Assessing the quality of scientific conferences based on bibliographic citations. <i>Scientometrics</i> , <b>2010</b> , 83, 133-155	3	23
41	On the Quality of Information for Web 2.0 Services. <i>IEEE Internet Computing</i> , <b>2010</b> , 14, 47-55	2.4	15
40	An unsupervised heuristic-based hierarchical method for name disambiguation in bibliographic citations. <i>Journal of the Association for Information Science and Technology</i> , <b>2010</b> , 61, 1853-1870		66
39	Using structural information to improve search in Web collections. <i>Journal of the Association for Information Science and Technology</i> , <b>2010</b> , 61, 2503-2513		5
38	Using web information for author name disambiguation <b>2009</b> ,		55
37	Automatic quality assessment of content created collaboratively by web communities <b>2009</b> ,		52
36	Finding what is missing from a digital library: A case study in the Computer Science field. <i>Information Processing and Management</i> , <b>2009</b> , 45, 380-391	6.3	7
35	A flexible approach for extracting metadata from bibliographic citations. <i>Journal of the Association for Information Science and Technology</i> , <b>2009</b> , 60, 1144-1158		9



34	A genetic programming framework for content-based image retrieval. <i>Pattern Recognition</i> , <b>2009</b> , 42, 283-292	7.7	120
33	Automatic evaluation of digital libraries with 5SQual. <i>Journal of Informetrics</i> , <b>2009</b> , 3, 102-123	3.1	26
32	A Genre-Aware Approach to Focused Crawling. <i>World Wide Web</i> , <b>2009</b> , 12, 285-319	2.9	9
31	An evolutionary approach for combining different sources of evidence in search engines. <i>Information Systems</i> , <b>2009</b> , 34, 276-289	2.7	18
30	Learning to assess the quality of scientific conferences <b>2009</b> ,		12
29	SyGAR A Synthetic Data Generator for Evaluating Name Disambiguation Methods. <i>Lecture Notes in Computer Science</i> , <b>2009</b> , 437-441	0.9	3
28	On-Demand Associative Cross-Language Information Retrieval. <i>Lecture Notes in Computer Science</i> , <b>2009</b> , 165-173	0.9	
27	A digital library environment for integrating, disseminating and exploring ecological data. <i>Ecological Informatics</i> , <b>2008</b> , 3, 295-308	4.2	6
26	Learning to rank at query-time using association rules <b>2008</b> ,		33
25	The impact of term selection in genre-aware focused crawling <b>2008</b> ,		2
24	Exploiting temporal contexts in text classification <b>2008</b> ,		12
23	Replica identification using genetic programming <b>2008</b> ,		16
22	Keeping a digital library clean <b>2008</b> ,		19
21	From concepts to implementation and visualization <b>2008</b> ,		3
20	Towards a digital library theory: a formal digital library ontology. <i>International Journal on Digital Libraries</i> , <b>2008</b> , 8, 91-114	1.4	33
19	A Web services-based framework for building componentized digital libraries. <i>Journal of Systems and Software</i> , <b>2008</b> , 81, 809-822	3.3	5
18	What is a good digital library? A quality model for digital libraries. <i>Information Processing and Management</i> , <b>2007</b> , 43, 1416-1437	6.3	69
17	Transitioning from the ecological fieldwork to an online repository: a digital library solution and evaluation. <i>International Journal on Digital Libraries</i> , <b>2007</b> , 7, 109-112	1.4	3

16	FLUX-CIM <b>2007</b> ,		38
15	Computing block importance for searching on web sites <b>2007</b> ,		19
14	A combined component approach for finding collection-adapted ranking functions based on genetic programming <b>2007</b> ,		38
13	Exploiting Genre in Focused Crawling. <i>Lecture Notes in Computer Science</i> , <b>2007</b> , 62-73	0.9	12
12	Link-based similarity measures for the classification of Web documents. <i>Journal of the Association for Information Science and Technology</i> , <b>2006</b> , 57, 208-221		38
11	A comparative study of citations and links in document classification <b>2006</b> ,		28
10	Learning to deduplicate <b>2006</b> ,		17
9	Learning to advertise <b>2006</b> ,		92
8	Design, Implementation, and Evaluation of a Wizard Tool for Setting Up Component-Based Digital Libraries. <i>Lecture Notes in Computer Science</i> , <b>2006</b> , 135-146	0.9	4
7	Incremental, Semi-automatic, Mapping-Based Integration of Heterogeneous Collections into Archaeological Digital Libraries: Megiddo Case Study. <i>Lecture Notes in Computer Science</i> , <b>2005</b> , 139-150	0.9	4
6	Streams, structures, spaces, scenarios, societies (5s). <i>ACM Transactions on Information Systems</i> , <b>2004</b> , 22, 270-312	4.8	166
5	BDBComp <b>2004</b> ,		10
4	The effectiveness of automatically structured queries in digital libraries <b>2004</b> ,		13
3	Recommender Systems Research: A Connection-Centric Survey. <i>Journal of Intelligent Information Systems</i> , <b>2004</b> , 23, 107-143	2.1	91
2	Prototyping Digital Libraries Handling Heterogeneous Data Sources – The ETANA-DL Case Study. <i>Lecture Notes in Computer Science</i> , <b>2004</b> , 186-197	0.9	13
1	Visual Semantic Modeling of Digital Libraries. <i>Lecture Notes in Computer Science</i> , <b>2003</b> , 325-337	0.9	8