

# Kalina Bontcheva

## List of Publications by Year in descending order

Source: <https://exaly.com/author-pdf/551792/publications.pdf>

Version: 2024-02-01

112  
papers

3,798  
citations

279487

23  
h-index

214527

47  
g-index

121  
all docs

121  
docs citations

121  
times ranked

2738  
citing authors

#	ARTICLE	IF	CITATIONS
1	Detection and Resolution of Rumours in Social Media. ACM Computing Surveys, 2019, 51, 1-36.	16.1	478
2	Getting More Out of Biomedical Documents with GATE's Full Lifecycle Open Source Text Analytics. PLoS Computational Biology, 2013, 9, e1002854.	1.5	261
3	Analysis of named entity recognition and linking for tweets. Information Processing and Management, 2015, 51, 32-49.	5.4	232
4	Stance Detection with Bidirectional Conditional Encoding. , 2016, , .		192
5	SemEval-2017 Task 8: RumourEval: Determining rumour veracity and support for rumours. , 2017, , .		162
6	GATE. , 2001, , .		133
7	Evolving GATE to meet new challenges in language engineering. Natural Language Engineering, 2004, 10, 349-373.	2.1	131
8	SemEval-2019 Task 7: RumourEval, Determining Rumour Veracity and Support for Rumours. , 2019, , .		103
9	Where's @wally?. , 2013, , .		97
10	Discourse-aware rumour stance classification in social media using sequential classifiers. Information Processing and Management, 2018, 54, 273-290.	5.4	89
11	Hawkes Processes for Continuous Time Sequence Classification: an Application to Rumour Stance Classification in Twitter. , 2016, , .		75
12	GATE Teamware: a web-based, collaborative text annotation framework. Language Resources and Evaluation, 2013, 47, 1007-1029.	1.8	73
13	Making sense of social media streams through semantics: A survey. Semantic Web, 2014, 5, 373-403.	1.1	69
14	A Natural Language Query Interface to Structured Information. , 2008, , 361-375.		60
15	Adapting SVM for data sparseness and imbalance: a case study in information extraction. Natural Language Engineering, 2009, 15, 241-271.	2.1	55
16	Architectural elements of language engineering robustness. Natural Language Engineering, 2002, 8, 257-274.	2.1	54
17	CLOnE: Controlled Language for Ontology Editing. Lecture Notes in Computer Science, 2007, , 142-155.	1.0	53
18	Microblog-genre noise and impact on semantic annotation accuracy. , 2013, , .		52

#	ARTICLE	IF	CITATIONS
19	Automatic Report Generation from Ontologies: The MIAKT Approach. Lecture Notes in Computer Science, 2004, , 324-335.	1.0	50
20	Crowdsourcing research opportunities. , 2012, , .		50
21	Generalisation in named entity recognition: A quantitative analysis. Computer Speech and Language, 2017, 44, 61-83.	2.9	47
22	A framework for real-time semantic social media analysis. Web Semantics, 2017, 44, 75-88.	2.2	44
23	Sub-story detection in Twitter with hierarchical Dirichlet processes. Information Processing and Management, 2017, 53, 989-1003.	5.4	42
24	GATECloud.net: a platform for large-scale, open-source text processing on the cloud. Philosophical Transactions Series A, Mathematical, Physical, and Engineering Sciences, 2013, 371, 20120071.	1.6	37
25	Classifying Tweet Level Judgements of Rumours in Social Media. , 2015, , .		36
26	MÄmir: An open-source semantic search framework for interactive information seeking and discovery. Web Semantics, 2015, 30, 52-68.	2.2	35
27	Multimedia indexing through multi-source and multi-language information extraction: the MUMIS project. Data and Knowledge Engineering, 2004, 48, 247-264.	2.1	34
28	Crowdsourcing the Annotation of Rumourous Conversations in Social Media. , 2015, , .		33
29	Hierarchical, perceptron-like learning for ontology-based information extraction. , 2007, , .		31
30	Using uneven margins SVM and perceptron for information extraction. , 2005, , .		29
31	Simple Open Stance Classification for Rumour Analysis. , 2017, , .		28
32	Gaussian Processes for Rumour Stance Classification in Social Media. ACM Transactions on Information Systems, 2019, 37, 1-24.	3.8	27
33	Classification aware neural topic model for COVID-19 disinformation categorisation. PLoS ONE, 2021, 16, e0247086.	1.1	27
34	Mining Information for Instance Unification. Lecture Notes in Computer Science, 2006, , 329-342.	1.0	27
35	USFD at SemEval-2016 Task 6: Any-Target Stance Detection on Twitter with Autoencoders. , 2016, , .		27
36	Corpus Linguistics and South Asian Languages: Corpus Creation and Tool Development. Literary and Linguistic Computing, 2004, 19, 509-524.	0.6	26

#	ARTICLE	IF	CITATIONS
37	Overview of the Special Issue on Trust and Veracity of Information in Social Media. ACM Transactions on Information Systems, 2016, 34, 1-5.	3.8	26
38	Experiments with geographic knowledge for information extraction. , 2003, , .		25
39	The evolution of argumentation mining: From models to social media and emerging tools. Information Processing and Management, 2019, 56, 102055.	5.4	25
40	RoundTrip Ontology Authoring. Lecture Notes in Computer Science, 2008, , 50-65.	1.0	25
41	Generating Tailored Textual Summaries from Ontologies. Lecture Notes in Computer Science, 2005, , 531-545.	1.0	24
42	Improving habitability of natural language interfaces for querying ontologies with feedback and clarification dialogues. Web Semantics, 2013, 19, 1-21.	2.2	24
43	Which politicians receive abuse? Four factors illuminated in the UK general election 2019. EPJ Data Science, 2020, 9, .	1.5	24
44	Point Process Modelling of Rumour Dynamics in Social Media. , 2015, , .		21
45	Opinion analysis for business intelligence applications. , 2008, , .		19
46	Natural Language Processing for the Semantic Web. Synthesis Lectures on the Semantic Web: Theory and Technology, 2016, 6, 1-194.	5.0	19
47	Mental health-related conversations on social media and crisis episodes: a time-series regression analysis. Scientific Reports, 2020, 10, 1342.	1.6	19
48	Next generation knowledge access. Journal of Knowledge Management, 2005, 9, 64-84.	3.2	18
49	Classifying <scp>T</scp>witter favorites: Like, bookmark, or Thanks?. Journal of the Association for Information Science and Technology, 2016, 67, 17-25.	1.5	18
50	Natural Language Technology for Information Integration in Business Intelligence. , 2007, , 366-380.		16
51	Crowdsourcing Named Entity Recognition and Entity Linking Corpora. , 2017, , 875-892.		16
52	Team Bertha von Suttner at SemEval-2019 Task 4: Hyperpartisan News Detection using ELMo Sentence Representation Convolutional Network. , 2019, , .		16
53	Tailoring Automatically Generated Hypertext. User Modeling and User-Adapted Interaction, 2005, 15, 135-168.	2.9	15
54	Using GATE as an environment for teaching NLP. , 2002, , .		15

#	ARTICLE	IF	CITATIONS
55	Large-scale, parallel automatic patent annotation. , 2008, , .		13
56	Semantic Annotations and Retrieval: Manual, Semiautomatic, and Automatic Generation. , 2011, , 77-116.		12
57	Hyperlocal Home Location Identification of Twitter Profiles. , 2017, , .		12
58	ResToRinG CaPitaLiZaTion in #TweeTs. , 2015, , .		11
59	Rumour verification through recurring information and an inner-attention mechanism. Online Social Networks and Media, 2019, 13, 100045.	2.3	11
60	Journalist-in-the-Loop: Continuous Learning as a Service for Rumour Analysis. , 2019, , .		11
61	Knowledge management and human language: crossing the chasm. Journal of Knowledge Management, 2005, 9, 108-131.	3.2	10
62	Using @Twitter Conventions to Improve #LOD-Based Named Entity Disambiguation. Lecture Notes in Computer Science, 2015, , 171-186.	1.0	10
63	Semantic Annotation and Human Language Technology. , 2006, , 29-50.		9
64	Games with a Purpose or Mechanised Labour?. , 2013, , .		9
65	Semantic Enrichment and Search: A Case Study on Environmental Science Literature. D-Lib Magazine, 2015, 21, .	0.5	9
66	Understanding climate change tweets: an open source toolkit for social media analysis. , 0, , .		9
67	Semantic Information Access. , 2006, , 139-169.		8
68	Using Human Language Technology for Automatic Annotation and Indexing of Digital Library Content. Lecture Notes in Computer Science, 2002, , 613-625.	1.0	8
69	Semantic Search over Documents and Ontologies. Lecture Notes in Computer Science, 2014, , 31-53.	1.0	8
70	Modeling Tweet Arrival Times using Log-Gaussian Cox Processes. , 2015, , .		8
71	Using a text engineering framework to build an extendable and portable IE-based summarisation system. , 2002, , .		8
72	Longitudinal Modeling of Social Media with Hawkes Process Based on Users and Networks. , 2017, , .		7

#	ARTICLE	IF	CITATIONS
73	Semantic Web and Human Computation: The status of an emerging field. <i>Semantic Web</i> , 2018, 9, 291-302.	1.1	7
74	USFD: Twitter NER with Drift Compensation and Linked Data. , 2015, , .		7
75	Using HLT for acquiring, retrieving and publishing knowledge in AKT. , 2001, , .		7
76	Rapid customization of an information extraction system for a surprise language. <i>ACM Transactions on Asian Language Information Processing</i> , 2003, 2, 295-300.	0.8	6
77	Adapting Support Vector Machines for F-term-based Classification of Patents. <i>ACM Transactions on Asian Language Information Processing</i> , 2008, 7, 1-19.	0.8	6
78	SoBigData. , 2018, , .		6
79	Automatic Summarization of Online Debates. , 2017, , .		6
80	Passive-Aggressive Sequence Labeling with Discriminative Post-Editing for Recognising Person Entities in Tweets. , 2014, , .		6
81	Vindication, virtue, and vitriol. <i>Journal of Computational Social Science</i> , 2020, 3, 401-443.	1.4	5
82	Weverify: Wider and Enhanced Verification for You Project Overview and Tools. , 2020, , .		5
83	Adapting a Robust Multi-genre NE System for Automatic Content Extraction. <i>Lecture Notes in Computer Science</i> , 2002, , 264-273.	1.0	5
84	Human Language Technologies. , 2009, , 37-49.		5
85	Stance Classification in Out-of-Domain Rumours: A Case Study Around Mental Health Disorders. <i>Lecture Notes in Computer Science</i> , 2017, , 53-64.	1.0	5
86	OLLIE. , 2003, , .		4
87	Transition of legacy systems to semantically enabled applications: TAO method and tools. <i>Semantic Web</i> , 2012, 3, 157-168.	1.1	4
88	Real-time Social Media Analytics through Semantic Annotation and Linked Open Data. , 2015, , .		4
89	Extracting Information from Social Media with GATE. , 2016, , 133-158.		4
90	User profiling with geo-located posts and demographic data. , 2016, , .		4

#	ARTICLE	IF	CITATIONS
91	Towards Enhanced Usability of Natural Language Interfaces to Knowledge Bases. <i>Annals of Information Systems</i> , 2010, , 105-133.	0.5	4
92	Collaborative Web-Based Tools for Multi-layer Text Annotation. , 2017, , 229-256.		4
93	Gold Standard Online Debates Summaries and First Experiments Towards Automatic Summarization of Online Debate Data. <i>Lecture Notes in Computer Science</i> , 2018, , 495-505.	1.0	4
94	eTranslation's Submissions to the WMT 2019 News Translation Task. , 2019, , .		4
95	CA manager framework. , 2009, , .		3
96	A Framework for Real-Time Semantic Social Media Analysis. <i>SSRN Electronic Journal</i> , 2017, , .	0.4	3
97	Indexing and querying linguistic metadata and document content. <i>Current Issues in Linguistic Theory</i> , 2007, , 35-44.	0.1	3
98	A Human-annotated Dataset for Evaluating Tweet Ranking Algorithms. , 2015, , .		2
99	Quantifying Media Influence and Partisan Attention on Twitter During the UK EU Referendum. <i>Lecture Notes in Computer Science</i> , 2018, , 274-290.	1.0	2
100	Generation of multilingual explanations from conceptual graphs. <i>Current Issues in Linguistic Theory</i> , 1997, , 365.	0.1	2
101	Improving Habitability of Natural Language Interfaces for Querying Ontologies With Feedback and Clarification Dialogues. <i>SSRN Electronic Journal</i> , 0, , .	0.4	2
102	Understanding Human Preferences for Summary Designs in Online Debates Domain. <i>Polibits</i> , 0, 54, 79-85.	0.0	2
103	MMmir: An Open-Source Semantic Search Framework for Interactive Information Seeking and Discovery. <i>SSRN Electronic Journal</i> , 0, , .	0.4	1
104	Predicting News Source Credibility. , 2019, , .		1
105	Linguistic Analysis Model for Monitoring User Reaction on Satirical News for Brazilian Portuguese. <i>Lecture Notes in Computer Science</i> , 2020, , 313-320.	1.0	1
106	Emerging language technologies and the rediscovery of the past: a research agenda. <i>International Journal on Digital Libraries</i> , 2005, 5, 309-316.	1.1	0
107	Multistage BiCross encoder for multilingual access to COVID-19 health information. <i>PLoS ONE</i> , 2021, 16, e0256874.	1.1	0
108	Reuse and challenges in evaluating language generation systems. , 2003, , .		0

#	ARTICLE	IF	CITATIONS
109	Session details: RDSM 2015. , 2015, , .		0
110	GATE: An Open-source NLP Toolkit for Mining Social Media. , 2016, , 499-511.		0
111	WeVerify: Wider and Enhanced Verification for You - Project Overview and Tool Demonstration. , 2019, , .		0
112	An Adoption of a Contradiction Detection Task to Assist the Summarization of Online Debates. , 2020, , .		0