

Walter Daelemans

List of Publications by Year in descending order

Source: <https://exaly.com/author-pdf/7100828/publications.pdf>

Version: 2024-02-01

121
papers

2,400
citations

331670

21
h-index

302126

39
g-index

132
all docs

132
docs citations

132
times ranked

1944
citing authors

#	ARTICLE	IF	CITATIONS
1	Automatic detection of cyberbullying in social media text. PLoS ONE, 2018, 13, e0203794.	2.5	174
2	Predicting age and gender in online social networks. , 2011, , .		142
3	Forgetting Exceptions is Harmful in Language Learning. Machine Learning, 1999, 34, 11-41.	5.4	128
4	Improving Accuracy in Word Class Tagging through the Combination of Machine Learning Systems. Computational Linguistics, 2001, 27, 199-229.	3.3	120
5	BioGraph: unsupervised biomedical knowledge discovery via automated hypothesis generation. Genome Biology, 2011, 12, R57.	9.6	109
6	IGTree: Using Trees for Compression and Classification in Lazy Learning Algorithms. Artificial Intelligence Review, 1997, 11, 407-423.	15.7	89
7	Authenticating the writings of Julius Caesar. Expert Systems With Applications, 2016, 63, 86-96.	7.6	72
8	Evaluating and understanding text-based stock price prediction models. Information Processing and Management, 2014, 50, 426-441.	8.6	67
9	Automatic monitoring of cyberbullying on social networking sites: From technological feasibility to desirability. Telematics and Informatics, 2015, 32, 89-97.	5.8	58
10	Data integration of structured and unstructured sources for assigning clinical codes to patient stays. Journal of the American Medical Informatics Association: JAMIA, 2016, 23, e11-e19.	4.4	56
11	Dutch plural inflection: The exception that proves the analogy†. Cognitive Psychology, 2007, 54, 283-318.	2.2	52
12	Memory-based morphological analysis. , 1999, , .		49
13	Explanation in Computational Stylometry. Lecture Notes in Computer Science, 2013, , 451-462.	1.3	45
14	Belgian economic policy uncertainty index: Improvement through text mining. International Journal of Forecasting, 2018, 34, 355-365.	6.5	45
15	CliCR: a Dataset of Clinical Case Reports for Machine Reading Comprehension. , 2018, , .		42
16	Measuring the complexity of writing systems*. Journal of Quantitative Linguistics, 1994, 1, 178-188.	1.2	40
17	Assessment of NER solutions against the first and second CALBC Silver Standard Corpus. Journal of Biomedical Semantics, 2011, 2, S11.	1.6	39
18	Cross-Genre Authorship Verification Using Unmasking. English Studies, 2012, 93, 340-356.	0.2	38

#	ARTICLE	IF	CITATIONS
19	Media coverage in times of political crisis: A text mining approach. Expert Systems With Applications, 2012, 39, 11616-11622.	7.6	37
20	Combined Optimization of Feature Selection and Algorithm Parameters in Machine Learning of Language. Lecture Notes in Computer Science, 2003, , 84-95.	1.3	33
21	Selecting relevant features from the electronic health record for clinical code prediction. Journal of Biomedical Informatics, 2017, 74, 92-103.	4.3	31
22	Patient representation learning and interpretable evaluation using clinical notes. Journal of Biomedical Informatics, 2018, 84, 103-113.	4.3	29
23	Unsupervised Context-Sensitive Spelling Correction of Clinical Free-Text with Word and Character N-Gram Embeddings. , 2017, , .		29
24	Overview of PAN 2019: Bots and Gender Profiling, Celebrity Profiling, Cross-Domain Authorship Attribution and Style Change Detection. Lecture Notes in Computer Science, 2019, , 402-416.	1.3	26
25	Logistic-based patient grouping for multi-disciplinary treatment. Artificial Intelligence in Medicine, 2002, 26, 87-107.	6.5	24
26	Introduction to the special issue on memory-based language processing. Journal of Experimental and Theoretical Artificial Intelligence, 1999, 11, 287-296.	2.8	22
27	Recent Advances in Example-Based Machine Translation. Computational Linguistics, 2004, 30, 516-520.	3.3	22
28	Fine-Grained Emotion Detection in Suicide Notes: A Thresholding Approach to Multi-Label Classification. Biomedical Informatics Insights, 2012, 5s1, BII.S8966.	4.6	22
29	Multimodular Text Normalization of Dutch User-Generated Content. ACM Transactions on Intelligent Systems and Technology, 2016, 7, 1-22.	4.5	22
30	Artificial intelligence tools for grammar and spelling instruction. Instructional Science, 1987, 16, 319-336.	2.0	21
31	Counting trees in Random Forests: Predicting symptom severity in psychiatric intake reports. Journal of Biomedical Informatics, 2017, 75, S112-S119.	4.3	19
32	Deep Transfer Learning for Art Classification Problems. Lecture Notes in Computer Science, 2019, , 631-646.	1.3	18
33	Current limitations in cyberbullying detection: On evaluation criteria, reproducibility, and data scarcity. Language Resources and Evaluation, 2021, 55, 597-633.	2.7	18
34	7. A comparison of Analogical Modeling to Memory-Based Language Processing. Human Cognitive Processing, 2002, , 157-179.	0.1	17
35	Implicit Schemata and Categories in Memory-based Language Processing. Language and Speech, 2013, 56, 309-328.	1.1	16
36	Using Distributed Representations to Disambiguate Biomedical and Clinical Concepts. , 2016, , .		16

#	ARTICLE	IF	CITATIONS
37	Sarcasm Detection Using an Ensemble Approach. , 2020, , .		15
38	A Short Review of Ethical Challenges in Clinical Natural Language Processing. , 2017, , .		13
39	Facilitatory Effects of Multi-Word Units in Lexical Processing and Word Learning: A Computational Investigation. <i>Frontiers in Psychology</i> , 2017, 8, 555.	2.1	12
40	Linguistic Accommodation in Teenagersâ€™ Social Media Writing: Convergence Patterns in Mixed-gender Conversations. <i>Journal of Quantitative Linguistics</i> , 2022, 29, 241-268.	1.2	11
41	Memory-based lexical acquisition and processing. <i>Lecture Notes in Computer Science</i> , 1995, , 85-98.	1.3	10
42	Memory-based learning. , 1997, , .		10
43	Kernel-Based Logical and Relational Learning with kLog for Hedge Cue Detection. <i>Lecture Notes in Computer Science</i> , 2012, , 347-357.	1.3	10
44	Simple Queries as Distant Labels for Predicting Gender on Twitter. , 2017, , .		10
45	Are we there yet? Exploring clinical domain knowledge of BERT models. , 2021, , .		9
46	Character-Level Transformer-Based Neural Machine Translation. , 2020, , .		9
47	Adolescentsâ€™ perceptions of social media writing: Has non-standard become the new standard?. <i>European Journal of Applied Linguistics</i> , 2019, 7, 189-224.	0.6	9
48	Using rule-induction techniques to model pronunciation variation in Dutch. <i>Computer Speech and Language</i> , 2004, 18, 1-23.	4.3	8
49	Memory-Based Learning. , 2010, , 154-179.		8
50	Lemmatization for variation-rich languages using deep learning. <i>Digital Scholarship in the Humanities</i> , 0, , fqw034.	0.7	8
51	EmoLabel: Semi-Automatic Methodology for Emotion Annotation of Social Media Text. <i>IEEE Transactions on Affective Computing</i> , 2022, 13, 579-591.	8.3	8
52	Children Probably Store Short Rather Than Frequent or Predictable Chunks: Quantitative Evidence From a Corpus Study. <i>Frontiers in Psychology</i> , 2019, 10, 80.	2.1	7
53	Lexical Patterns in Adolescentsâ€™ Online Writing: The Impact of Age, Gender, and Education. <i>Written Communication</i> , 2020, 37, 365-400.	1.3	7
54	Generative Art Inspired by Nature, Using NodeBox. <i>Lecture Notes in Computer Science</i> , 2011, , 264-272.	1.3	7

#	ARTICLE	IF	CITATIONS
55	Adolescents' social background and non-standard writing in online communication. Dutch Journal of Applied Linguistics, 2018, 7, 2-25.	0.3	7
56	Rule induction for global explanation of trained models. , 2018, , .		7
57	Transcription of out-of-vocabulary words in large vocabulary speech recognition based on phoneme-to-grapheme conversion. , 2002, , .		6
58	Assigning clinical codes with data-driven concept representation on Dutch clinical free text. Journal of Biomedical Informatics, 2017, 69, 118-127.	4.3	6
59	Unsupervised concept extraction from clinical text through semantic composition. Journal of Biomedical Informatics, 2019, 91, 103120.	4.3	6
60	Improving Hate Speech Type and Target Detection with Hateful Metaphor Features. , 2021, , .		6
61	Semantic and Syntactic Features for Dutch Coreference Resolution. , 2008, , 351-361.		6
62	On the Limits of Sentence Compression by Deletion. Lecture Notes in Computer Science, 2010, , 45-66.	1.3	6
63	A Neural Network for Hyphenation. , 1992, , 1647-1650.		6
64	Assessing the Stylistic Properties of Neurally Generated Text in Authorship Attribution. , 2017, , .		6
65	Default inheritance in an object-oriented representation of linguistic categories. International Journal of Human Computer Studies, 1994, 41, 149-177.	5.6	5
66	Evaluating Content-Independent Features for Personality Recognition. , 2014, , .		5
67	Lexical category acquisition is facilitated by uncertainty in distributional co-occurrences. PLoS ONE, 2018, 13, e0209449.	2.5	5
68	Improving Cross-Domain Hate Speech Detection by Reducing the False Positive Rate. , 2021, , .		5
69	Improving sequence segmentation learning by predicting trigrams. , 2005, , .		5
70	Selective impairment of adjective order constraints as overeager abstraction: An elaboration on Kemmerer et al. (2009). Journal of Neurolinguistics, 2013, 26, 46-72.	1.1	4
71	Evolution of the PAN Lab on Digital Text Forensics. The Kluwer International Series on Information Retrieval, 2019, , 461-485.	1.0	4
72	Abstraction and generalization. , 2005, , 104-147.		3

#	ARTICLE	IF	CITATIONS
73	Highlights of the BioTM 2010 workshop on advances in bio text mining. BMC Bioinformatics, 2010, 11, .	2.6	3
74	Robust Rhymes? The Stability of Authorial Style in Medieval Narratives*. Journal of Quantitative Linguistics, 2012, 19, 54-76.	1.2	3
75	The strategic impact of META-NET on the regional, national and international level. Language Resources and Evaluation, 2016, 50, 351-374.	2.7	3
76	Discourse lexicon induction for multiple languages and its use for gender profiling. Digital Scholarship in the Humanities, 2019, 34, 208-220.	0.7	3
77	Automatic classification of social media reports on violent incidents in South Africa using machine learning. South African Journal of Science, 2020, 116, .	0.7	3
78	Interlocutors's Age Impacts Teenagers's Online Writing Style: Accommodation in Intra- and Intergenerational Online Conversations. Frontiers in Artificial Intelligence, 2021, 4, 738278.	3.4	3
79	deLearyous: An Interactive Application for Interpersonal Communication Training. Communications in Computer and Information Science, 2012, , 87-90.	0.5	3
80	Using Wiktionary to Build an Italian Part-of-Speech Tagger. Lecture Notes in Computer Science, 2014, , 1-8.	1.3	3
81	Predicting Adolescents's Educational Track from Chat Messages on Dutch Social Media. , 2018, , .		3
82	Evaluating Hybrid Versus Data-Driven Coreference Resolution. , 2007, , 137-150.		3
83	An Ensemble Approach for Dutch Cross-Domain Hate Speech Detection. Lecture Notes in Computer Science, 2022, , 3-15.	1.3	3
84	Effects of online abstraction on adjective order preferences. Language, Cognition and Neuroscience, 2015, 30, 816-831.	1.2	2
85	Simulating speech processing with cochlear implants: How does channel interaction affect learning in neural networks?. PLoS ONE, 2019, 14, e0212134.	2.5	2
86	Comparing automated content analysis methods to distinguish issue communication by political parties on Twitter. Computational Communication Research, 2021, 3, 1-27.	2.0	2
87	A Chunk-Driven Bootstrapping Approach to Extracting Translation Patterns. Lecture Notes in Computer Science, 2010, , 394-405.	1.3	2
88	Chapter 10. Acquisition of phonological variables of a Flemish dialect by children raised in Standard Dutch. Studies in Language Variation, 2017, , 267-304.	0.2	2
89	Towards the Improvement of Automatic Emotion Pre-annotation with Polarity and Subjective Information. , 2017, , .		2
90	Transfer Learning for Digital Heritage Collections: Comparing Neural Machine Translation at the Subword-level and Character-level. , 2020, , .		2

#	ARTICLE	IF	CITATIONS
91	Disambiguation of the Neuter Pronoun and Its Effect on Pronominal Coreference Resolution. , 2007, , 48-55.		2
92	A Model of Dutch Morphophonology and its Applications. AI Communications, 1988, 1, 18-25.	1.2	1
93	Memory-Based Learning in Natural Language Processing. , 2005, , 3-14.		1
94	Computer-mediated communication (CMC) and social media corpora: Introduction. European Journal of Applied Linguistics, 2019, 7, 157-162.	0.6	1
95	Scalable Few-Shot Learning of Robust Biomedical Name Representations. , 2021, , .		1
96	Comparing Automated Content Analysis Methods To Distinguish Issue Communication by Political Parties on Twitter. SSRN Electronic Journal, 0, , .	0.4	1
97	Multi-modal Label Retrieval for the Visual Arts: The Case of Iconclass. , 2021, , .		1
98	Lazy and Eager Relational Learning Using Graph-Kernels. Lecture Notes in Computer Science, 2014, , 171-184.	1.3	1
99	Computational Language Analysis for Assessment of Schizophrenia. , 2017, , .		1
100	A Deep Generative Approach to Native Language Identification. , 2020, , .		1
101	Communicating across educational boundaries: accommodation patterns in adolescentsâ€™ online interactions. Applied Linguistics Review, 2024, 15, 1-29.	0.9	1
102	Predicting COVID-19 Symptoms From Free Text in Medical Records Using Artificial Intelligence: Feasibility Study. JMIR Medical Informatics, 2022, 10, e37771.	2.6	1
103	Detecting Vaccine Skepticism on Twitter Using Heterogeneous Information Networks. Lecture Notes in Computer Science, 2022, , 370-381.	1.3	1
104	Memory and Similarity. , 2005, , 26-56.		0
105	Application to morpho-phonology. , 2005, , 57-84.		0
106	Application to shallow parsing. , 2005, , 85-103.		0
107	Inspirations from linguistics and artificial intelligence. , 2005, , 15-25.		0
108	Guest Editorsâ€™ introduction: special issue of selected papers from ECML PKDD 2008. Machine Learning, 2008, 72, 155-156.	5.4	0

#	ARTICLE	IF	CITATIONS
109	Guest Editorsâ€™ Introduction: Special issue of Selected Papers from ECML PKDD 2008. Data Mining and Knowledge Discovery, 2008, 17, 1-2.	3.7	0
110	BioGraph: Knowledge Discovery and Exploration in the Biomedical Domain. , 2011, , .		0
111	Literary Detective Work on the Computer. Michael P. Oakes.. Digital Scholarship in the Humanities, 2016, , fqw055.	0.7	0
112	Exploring the Classification of Security Events using Sparse and Dense Representation of Text. , 2020, , .		0
113	Transfer Learning with Style Transfer between the Photorealistic and Artistic Domain. IS&T International Symposium on Electronic Imaging, 2021, 33, 41-1-41-9.	0.4	0
114	Contextual explanation rules for neural clinical classifiers. , 2021, , .		0
115	COREA: Coreference Resolution for Extracting Answers for Dutch. Theory and Applications of Natural Language Processing, 2013, , 115-128.	0.3	0
116	Outomatiese genreklassifikasie vir Afrikaans. South African Journal of Science and Technology, 2014, 33, , .	0.1	0
117	POS Tagging. , 2017, , 985-989.		0
118	Stylene: an Environment for Stylometry and Readability Research for Dutch. , 2017, , 195-209.		0
119	Revisiting neural relation classification in clinical notes with external information. , 2018, , .		0
120	From Strings to Other Things: Linking the Neighborhood and Transposition Effects in Word Reading. , 2018, , .		0
121	Metameric. Mental Lexicon, 2018, 13, 333-353.	0.5	0