Anna Korhonen

List of Publications by Year in descending order

Source: https://exaly.com/author-pdf/11771984/publications.pdf

Version: 2024-02-01

55	2,560	361045	276539
papers	citations	20 h-index	41 g-index

57	57	57	1761
all docs	docs citations	times ranked	citing authors

#	Article	IF	Citations
1	SimLex-999: Evaluating Semantic Models With (Genuine) Similarity Estimation. Computational Linguistics, 2015, 41, 665-695.	2.5	664
2	A large-scale classification of English verbs. Computers and the Humanities, 2008, 42, 21-40.	1.4	187
3	Semantic Specialization of Distributional Word Vector Spaces using Monolingual and Cross-Lingual Constraints. Transactions of the Association for Computational Linguistics, 2017, 5, 309-324.	3.2	117
4	Link prediction in drug-target interactions network using similarity indices. BMC Bioinformatics, 2017, 18, 39.	1.2	92
5	SimVerb-3500: A Large-Scale Evaluation Set of Verb Similarity. , 2016, , .		89
6	Statistical Metaphor Processing. Computational Linguistics, 2013, 39, 301-353.	2.5	86
7	Cancer Hallmarks Analytics Tool (CHAT): a text mining approach to organize and evaluate scientific literature on cancer. Bioinformatics, 2017, 33, 3973-3981.	1.8	85
8	A systematic literature review of automatic Alzheimerâ€™s disease detection from speech and language. Journal of the American Medical Informatics Association: JAMIA, 2020, 27, 1784-1797.	2.2	82
9	Learning to Understand Phrases by Embedding the Dictionary. Transactions of the Association for Computational Linguistics, 2016, 4, 17-30.	3.2	81
10	Zone analysis in biology articles as a basis for information extraction. International Journal of Medical Informatics, 2006, 75, 468-487.	1.6	68
11	Modeling Language Variation and Universals: A Survey on Typological Linguistics for Natural Language Processing. Computational Linguistics, 2019, 45, 559-601.	2.5	60
12	Text Mining for Literature Review and Knowledge Discovery in Cancer Risk Assessment and Research. PLoS ONE, 2012, 7, e33427.	1.1	60
13	Automatic semantic classification of scientific literature according to the hallmarks of cancer. Bioinformatics, 2016, 32, 432-440.	1.8	56
14	HyperLex: A Large-Scale Evaluation of Graded Lexical Entailment. Computational Linguistics, 2017, 43, 781-835.	2.5	49
15	LION LBD: a literature-based discovery system for cancer biology. Bioinformatics, 2019, 35, 1553-1561.	1.8	47
16	Improving Multi-Modal Representations Using Image Dispersion: Why Less is Sometimes More. , 2014, , .		44
17	Unsupervised and constrained Dirichlet process mixture models for verb clustering. , 2009, , .		38
18	On the Relation between Linguistic Typology and (Limitations of) Multilingual Language Modeling. , 2018, , .		34

#	Article	IF	Citations
19	Improving verb clustering with automatically acquired selectional preferences. , 2009, , .		34
20	Dependency parsing of learner English. International Journal of Corpus Linguistics, 2018, 23, 28-54.	0.6	31
21	Gender differences in cancer susceptibility: role of oxidative stress. Carcinogenesis, 2016, 37, 985-992.	1.3	28
22	Semantically motivated subcategorization acquisition. , 2002, , .		27
23	Multi-Modal Models for Concrete and Abstract Concept Meaning. Transactions of the Association for Computational Linguistics, 2014, 2, 285-296.	3.2	26
24	Morph-fitting: Fine-Tuning Word Vector Spaces with Simple Language-Specific Rules. , 2017, , .		26
25	A Quantitative Empirical Analysis of the Abstract/Concrete Distinction. Cognitive Science, 2014, 38, 162-177.	0.8	24
26	Language Modeling for Morphologically Rich Languages: Character-Aware Modeling for Word-Level Prediction. Transactions of the Association for Computational Linguistics, 2018, 6, 451-465.	3.2	24
27	Adversarial Propagation and Zero-Shot Cross-Lingual Transfer of Word Vector Specialization. , 2018, , .		23
28	Exploring big educational learner corpora for SLA research. International Journal of Learner Corpus Research, 2015, 1, 96-129.	0.4	22
29	Bio-SimVerb and Bio-SimLex: wide-coverage evaluation sets of word similarity in biomedicine. BMC Bioinformatics, 2018, 19, 33.	1.2	21
30	Isomorphic Transfer of Syntactic Structures in Cross-Lingual NLP. , 2018, , .		21
31	The first step in the development of text mining technology for cancer risk assessment: identifying and organizing scientific evidence in risk assessment literature. BMC Bioinformatics, 2009, 10, 303.	1.2	20
32	Towards Unrestricted, Large-Scale Acquisition of Feature-Based Conceptual Representations from Corpus Data. Research on Language and Computation, 2009, 7, 137-170.	0.4	18
33	Conceptual metaphor theory meets the data: a corpus-based human annotation study. Language Resources and Evaluation, 2013, 47, 1261-1284.	1.8	18
34	Verb Class Discovery from Rich Syntactic Data. , 2008, , 16-27.		18
35	Exocrine Pancreatic Carcinogenesis and Autotaxin Expression. PLoS ONE, 2012, 7, e43209.	1.1	17
36	A comparison and user-based evaluation of models of textual information structure in the context of cancer risk assessment. BMC Bioinformatics, 2011, 12, 69.	1.2	15

#	Article	IF	Citations
37	Active learning-based information structure analysis of full scientific articles and two applications for biomedical literature review. Bioinformatics, 2013, 29, 1440-1447.	1.8	15
38	Text mining for improved exposure assessment. PLoS ONE, 2017, 12, e0173132.	1.1	13
39	Multi-SimLex: A Large-Scale Evaluation of Multilingual and Crosslingual Lexical Semantic Similarity. Computational Linguistics, 2021, 46, 847-897.	2.5	13
40	Probabilistic Distributional Semantics with Latent Variable Models. Computational Linguistics, 2014, 40, 587-631.	2.5	12
41	Grouping chemicals for health risk assessment: A text mining-based case study of polychlorinated biphenyls (PCBs). Toxicology Letters, 2016, 241, 32-37.	0.4	12
42	Weakly supervised learning of information structure of scientific abstractsâ€”is it accurate enough to benefit real-world tasks in biomedicine?. Bioinformatics, 2011, 27, 3179-3185.	1.8	11
43	Anchoring and Agreement in Syntactic Annotations. , 2016, , .		11
44	Automatic Extraction of Property Norm-Like Data From Large Text Corpora. Cognitive Science, 2014, 38, 638-682.	0.8	10
45	Evaluation of carcinogenic modes of action for pesticides in fruit on the Swedish market using a text-mining tool. Frontiers in Pharmacology, 2014, 5, 145.	1.6	8
46	Is "Universal Syntax" Universally Useful for Learning Distributed Word Representations?. , 2016, , .		8
47	Parameter Space Factorization for Zero-Shot Learning across Tasks and Languages. Transactions of the Association for Computational Linguistics, 2021, 9, 410-428.	3.2	7
48	Automatic lexical classification: bridging research and practice. Philosophical Transactions Series A, Mathematical, Physical, and Engineering Sciences, 2010, 368, 3621-3632.	1.6	6
49	Semantic Data Set Construction from Human Clustering and Spatial Arrangement. Computational Linguistics, 2021, 47, 69-116.	2.5	6
50	Application of Text Mining in Risk Assessment of Chemical Mixtures: A Case Study of Polycyclic Aromatic Hydrocarbons (PAHs). Environmental Health Perspectives, 2021, 129, 67008.	2.8	6
51	Investigating the cross-lingual translatability of VerbNet-style classification. Language Resources and Evaluation, 2018, 52, 771-799.	1.8	5
52	Unsupervised discovery of information structure in biomedical documents. Bioinformatics, 2015, 31, 1084-1092.	1.8	4
53	Decoding Sentiment from Distributed Representations of Sentences. , 2017, , .		4
54	Native Language Identification on EFCAMDAT. , 0, , 159-184.		0

#	Article	IF	Citations
55	Subcategorization frame identification for learner English. International Journal of Corpus Linguistics, 2021, 26, 187-218.	0.6	0