Paul Edward Rayson

List of Publications by Year
in descending order

77 papers	2,453 citations	394421 19 h-index	302126 39 g-index
83 all docs	83 docs citations	83 times ranked	1476 citing authors

#	Article	IF	Citations
1	From key words to key semantic domains. International Journal of Corpus Linguistics, 2008, 13, 519-549.	1.4	472
2	Comparing corpora using frequency profiling. , 2000, , .		190
3	Social Differentiation in the Use of English Vocabulary. International Journal of Corpus Linguistics, 1997, 2, 133-152.	1.4	159
4	The online use of Violence and Journey metaphors by patients with cancer, as compared with health professionals: a mixed methods study. BMJ Supportive and Palliative Care, 2017, 7, 60-66.	1.6	159
5	An Exploratory Study of Information Retrieval Techniques in Domain Analysis. , 2008, , .		90
6	A computer-assisted study of the use of Violence metaphors for cancer and end of life by patients, family carers and health professionals. International Journal of Corpus Linguistics, 0, , 205-231.	1.4	85
7	Semantics-based composition for aspect-oriented requirements engineering. , 2007, , .		84
8	In search of meaning: Lessons, resources and next steps for computational analysis of financial discourse. Journal of Business Finance and Accounting, 2019, 46, 265-306.	2.7	74
9	Shallow knowledge as an aid to deep understanding in early phase requirements engineering. IEEE Transactions on Software Engineering, 2005, 31, 969-981.	5.6	70
10	EA-Miner. , 2005, , .		63
11	Geoparsing, GIS, and Textual Analysis: Current Developments in Spatial Humanities Research. International Journal of Humanities and Arts Computing, 2015, 9, 1-14.	0.4	57
12	Comparing and combining a semantic tagger and a statistical tool for MWE extraction. Computer Speech and Language, 2005, 19, 378-397.	4.3	52
13	Retrieving, classifying and analysing narrative commentary in unstructured (glossy) annual reports published as PDF files. Accounting and Business Research, 2020, 50, 6-34.	1.8	52
14	Analyzing the semantic content and persuasive composition of extremist media: A case study of texts produced during the Gaza conflict. Information Systems Frontiers, 2011, 13, 61-73.	6.4	45
15	A flexible framework to experiment with ontology learning techniques. Knowledge-Based Systems, 2008, 21, 192-199.	7.1	44
16	The language of Islamic extremism. International Journal of Corpus Linguistics, 2012, 17, 259-286.	1.4	42
17	Who Am I? Analyzing Digital Personas in Cybercrime Investigations. Computer, 2013, 46, 54-61.	1.1	42
18	Guidelines for normalising Early Modern English corpora: Decisions and justifications. ICAME Journal, 2015, 39, 5-24.	0.6	42

#	ARTICLE	IF	CITATIONS
19	REVERE: Support for Requirements Synthesis from Documents. Information Systems Frontiers, 2002, 4, 343-353.	6.4	39
20	Multiword expressions: hard going or plain sailing?. Language Resources and Evaluation, 2010, 44, 1-5.	2.7	32
21	A Systematic Survey of Online Data Mining Technology Intended for Law Enforcement. ACM Computing Surveys, 2015, 48, 1-54.	23.0	31
22	Can you detect early dementia from an email? A proof of principle study of daily computer use to detect cognitive and functional decline. International Journal of Geriatric Psychiatry, 2018, 33, 867-874.	2.7	29
23	EA-Miner: Towards Automation in Aspect-Oriented Requirements Engineering. , 2007, , 4-39.		27
24	COUNTER: corpus of Urdu news text reuse. Language Resources and Evaluation, 2017, 51, 777-803.	2.7	24
25	Artefacts as designed, artefacts as used: resources for uncovering activity dynamics. Cognition, Technology and Work, 2005, 7, 76-87.	3.0	20
26	Automatically Analyzing Large Texts in a <scp>GIS</scp> Environment: The Registrar General's Reports and Cholera in the 19th Century. Transactions in GIS, 2015, 19, 296-320.	2.3	20
27	Measuring Short Text Reuse for the Urdu Language. IEEE Access, 2018, 6, 7412-7421.	4.2	19
28	Extracting multiword expressions with a semantic tagger. , 2003, , .		19
29	A time-sensitive historical thesaurus-based semantic tagger for deep semantic annotation. Computer Speech and Language, 2017, 46, 113-135.	4.3	18
30	A framework for P2P application development. Computer Communications, 2008, 31, 387-401.	5.1	16
31	Classification of Short Text Comments by Sentiment and Actionability for VoiceYourView. , 2010, , .		16
32	Customising geoparsing and georeferencing for historical texts. , 2013, , .		16
33	Safeguarding Cyborg Childhoods: Incorporating the On/Offline Behaviour of Children into Everyday Social Work Practices. British Journal of Social Work, 2014, 44, 596-614.	1.4	16
34	Supporting Law Enforcement in Digital Communities through Natural Language Analysis. Lecture Notes in Computer Science, 2008, , 122-134.	1.3	16
35	A tool suite for aspect-oriented requirements engineering. , 2006, , .		15
36	The ParlaMint corpora of parliamentary proceedings. Language Resources and Evaluation, 2023, 57, 415-448.	2.7	15

#	Article	IF	Citations
37	MasakhaNER: Named Entity Recognition for African Languages. Transactions of the Association for Computational Linguistics, 2021, 9, 1116-1131.	4.8	14
38	Corpus Tools and Methods, Today and Tomorrow: Incorporating Linguists' Manual Annotations. Literary and Linguistic Computing, 2007, 23, 163-180.	0.6	13
39	Natural Language Processing Methods and Bipolar Disorder: Scoping Review. JMIR Mental Health, 2022, 9, e35928.	3.3	12
40	A deeply annotated testbed for geographical text analysis. , 2017, , .		11
41	The REVERE Project: Experiments with the Application of Probabilistic NLP to Systems Engineering. Lecture Notes in Computer Science, 2001, , 288-300.	1.3	11
42	Sentiment analysis tools should take account of the number of exclamation marks!!!. , 2015, , .		10
43	Social Media Monitoring of the COVID-19 Pandemic and Influenza Epidemic With Adaptation for Informal Language in Arabic Twitter Data: Qualitative Study. JMIR Medical Informatics, 2021, 9, e27670.	2.6	10
44	Comparing corpora using frequency profiling. , 2000, , .		10
45	Discovering affect-laden requirements to achieve system acceptance. , 2014, , .		9
46	Metaphor, Popular Science, and Semantic Tagging: Distant reading with the<i>Historical Thesaurus of English</i>. Digital Scholarship in the Humanities, 0, , fqv045.	0.7	9
47	CLEU â€•A Crossâ€‑Language Englishâ€‑Urdu Corpus and Benchmark for Text Reuse Experiments. Journal of the Association for Information Science and Technology, 2018, 70, 729.	2.9	9
48	Understanding who uses Reddit: Profiling individuals with a self-reported bipolar disorder diagnosis. , 2021, , .		9
49	â€œi didnâ€™t spel that wrong did i. Oopsâ€‑ Lingvisticae Investigationes, 2012, 35, 367-388.	0.3	8
50	Automatic error tagging of spelling mistakes in learner corpora. Studies in Corpus Linguistics, 2011, , 109-126.	0.2	8
51	Differentiating Act from Ideology: Evidence from Messages For and Against Violent Extremism. Negotiation and Conflict Management Research, 2012, 5, 289-306.	1.0	7
52	Dealing with heterogeneous big data when geoparsing historical corpora. , 2014, , .		7
53	Exploring fine-grained sentiment values in online product reviews. , 2015, , .		7
54	Known and unknown requirements in healthcare. Requirements Engineering, 2020, 25, 1-20.	3.1	7

#	Article	IF	Citations
55	Experiments in 17th century English: manual versus automatic conceptual history. Literary and Linguistic Computing, 2012, 27, 395-408.	0.6	6
56	Improving the precision of corpus methods. , 2010, , 279-290.		6
57	ASSIST. , 2006, , .		6
58	UNLT: Urdu Natural Language Toolkit. Natural Language Engineering, 2023, 29, 942-977.	2.5	6
59	Combining Mouse and Keyboard Events with Higher Level Desktop Actions to Detect Mild Cognitive Impairment. , 2016, , .		5
60	Measuring MWE compositionality using semantic annotation. , 2006, , .		5
61	A Flexible Framework To Experiment With Ontology Learning Techniques. , 2007, , 153-166.		4
62	A Service-Indepenent Model for Linking Online User Profile Information. , 2014, , .		3
63	A word sense disambiguation corpus for Urdu. Language Resources and Evaluation, 2019, 53, 397-418.	2.7	3
64	Problematising characteristicness. International Journal of Corpus Linguistics, 2021, 26, 305-335.	1.4	3
65	Language Independent Evaluation of Translation Style and Consistency: Comparing Human and Machine Translations of Camusâ€™ Novel â€œThe Strangerâ€. Lecture Notes in Computer Science, 2014, , 116-124.	1.3	3
66	Analysing Keyword Lists. , 2020, , 119-139.		3
67	Annotated web as corpus. , 2006, , .		2
68	18. Automatic extraction of translation equivalents of phrasal and light verbs in English and Russian. , 2008, , 293-309.		2
69	Uncovering Environmental Change in the English Lake District: Using Computational Techniques to Trace the Presence and Documentation of Historical Flora. Digital Scholarship in the Humanities, 2021, 36, 736-756.	0.7	2
70	Assessment of non-directed computer-use behaviours in the home can indicate early cognitive impairment: A proof of principle longitudinal study. Aging and Mental Health, 2023, 27, 193-202.	2.8	2
71	Children Online. International Journal of Corpus Linguistics, 2012, 17, 443-481.	1.4	1
72	Heterogeneous Narrative Content in Annual Reports Published as PDF Files: Extraction, Classification and Incremental Predictive Ability. SSRN Electronic Journal, 0, , .	0.4	1

#	Article	IF	Citations
73	â€œi didnâ€™t spel that wrong did i. Oopsâ€• Contemporary Discourses of Hate and Radicalism Across Space and Genres, 2014, , 217-237.	0.0	1
74	What is middleware made of?. , 2012, , .		0
75	lexiDB: A scalable corpus database management system. , 2016, , .		0
76	Developing Multilingual Automatic Semantic Annotation Systems. , 2019, , 94-109.		0
77	Textual variations affect human judgements of sentiment values. Electronic Commerce Research and Applications, 2022, 53, 101149.	5.0	0