Ulf Leser

List of Publications by Year
in descending order

93 papers	3,386 citations	147801 31 h-index	168389 53 g-index
97 all docs	97 docs citations	97 times ranked	4192 citing authors

#	Article	IF	Citations
1	Deep learning with word embeddings improves biomedical named entity recognition. Bioinformatics, 2017, 33, i37-i48.	4.1	400
2	The Stratosphere platform for big data analytics. VLDB Journal, 2014, 23, 939-964.	4.1	348
3	Ras-Mediated Deregulation of the Circadian Clock in Cancer. PLoS Genetics, 2014, 10, e1004338.	3.5	140
4	Fast and Accurate Time Series Classification with WEASEL. , 2017, , .		137
5	What makes a gene name? Named entity recognition in the biomedical literature. Briefings in Bioinformatics, 2005, 6, 357-369.	6.5	134
6	A Comprehensive Benchmark of Kernel Methods to Extract Protein€"Protein Interactions from Literature. PLoS Computational Biology, 2010, 6, e1000837.	3.2	124
7	ALIBABA: PubMed as a graph. Bioinformatics, 2006, 22, 2444-2445.	4.1	114
8	Completeness of integrated information sources. Information Systems, 2004, 29, 583-615.	3.6	100
9	DNA copy number changes define spatial patterns of heterogeneity in colorectal cancer. Nature Communications, 2017, 8, 14093.	12.8	85
10	A query language for biological networks. Bioinformatics, 2005, 21, ii33-ii39.	4.1	69
11	Comparative assessment of differential network analysis methods. Briefings in Bioinformatics, 2017, 18, bbw061.	6.5	65
12	GeneView: a comprehensive semantic search engine for PubMed. Nucleic Acids Research, 2012, 40, W585-W591.	14.5	62
13	DynamicCloudSim: Simulating heterogeneity in computational clouds. Future Generation Computer Systems, 2015, 46, 85-99.	7.5	54
14	HunFlair: an easy-to-use tool for state-of-the-art biomedical named entity recognition. Bioinformatics, 2021, 37, 2792-2794.	4.1	53
15	A hierarchical approach to model web query interfaces for web source integration. Proceedings of the VLDB Endowment, 2009, 2, 325-336.	3.8	52
16	Implications for Molecular Mechanisms of Glycoprotein Hormone Receptors Using a New Sequence-Structure-Function Analysis Resource. Molecular Endocrinology, 2007, 21, 574-580.	3.7	50
17	The transformer earthquake alerting model: a new versatile approach to earthquake early warning. Geophysical Journal International, 2021, 225, 646-656.	2.4	49
18	Finding Kinetic Parameters Using Text Mining. OMICS A Journal of Integrative Biology, 2004, 8, 131-152.	2.0	47

#	Article	IF	Citations
19	Gene mention normalization and interaction extraction with context models and sentence motifs. Genome Biology, 2008, 9, S14.	9.6	47
20	Cross talk between Wnt/β-catenin and Irf8 in leukemia progression and drug resistance. Journal of Experimental Medicine, 2013, 210, 2239-2256.	8.5	47
21	A survey on annotation tools for the biomedical literature. Briefings in Bioinformatics, 2014, 15, 327-340.	6.5	47
22	Regular Path Queries on Large Graphs. Lecture Notes in Computer Science, 2012, , 177-194.	1.3	44
23	Assembly of a Comprehensive Regulatory Network for the Mammalian Circadian Clock: A Bioinformatics Approach. PLoS ONE, 2015, 10, e0126283.	2.5	43
24	Predictive performance modeling for distributed batch processing using black box monitoring and machine learning. Information Systems, 2019, 82, 33-52.	3.6	43
25	Search, adapt, and reuse. SIGMOD Record, 2011, 40, 6-16.	1.2	39
26	HUNER: improving biomedical NER with pretraining. Bioinformatics, 2020, 36, 295-302.	4.1	38
27	Trends in Genome Compression. Current Bioinformatics, 2014, 9, 315-326.	1.5	38
28	Earthquake magnitude and location estimation from real time seismic waveforms with a transformer network. Geophysical Journal International, 2021, 226, 1086-1104.	2.4	37
29	RCSI. Proceedings of the VLDB Endowment, 2013, 6, 1534-1545.	3.8	36
30	State-of-the-art in string similarity search and join. SIGMOD Record, 2014, 43, 64-76.	1.2	33
31	Similarity search for scientific workflows. Proceedings of the VLDB Endowment, 2014, 7, 1143-1154.	3.8	33
32	Set similarity joins on mapreduce. Proceedings of the VLDB Endowment, 2018, 11, 1110-1122.	3.8	33
33	TEASER: early and accurate time series classification. Data Mining and Knowledge Discovery, 2020, 34, 1336-1362.	3.7	32
34	Question answering for Biology. Methods, 2015, 74, 36-46.	3.8	31
35	Efficiently Detecting Inclusion Dependencies. , 2007, , .		30
36	SETH detects and normalizes genetic variants in text. Bioinformatics, 2016, 32, 2883-2885.	4.1	30

#	Article	IF	Citations
37	Finding k-shortest paths with limited overlap. VLDB Journal, 2020, 29, 1023-1047.	4.1	29
38	Tools for managing and analyzing microarray data. Briefings in Bioinformatics, 2012, 13, 46-60.	6.5	28
39	Deregulation of the endogenous C/EBPβ LIP isoform predisposes to tumorigenesis. Journal of Molecular Medicine, 2015, 93, 39-49.	3.9	28
40	CellFinder: a cell data repository. Nucleic Acids Research, 2014, 42, D950-D958.	14.5	26
41	Data Management Challenges in Next Generation Sequencing. Datenbank-Spektrum, 2012, 12, 161-171.	1.3	25
42	Alternative routing. , 2015, , .		25
43	SOFA: An extensible logical optimizer for UDF-heavy data flows. Information Systems, 2015, 52, 96-125.	3.6	25
44	Basal subtype is predictive for response to cetuximab treatment in patient-derived xenografts of squamous cell head and neck cancer. International Journal of Cancer, 2017, 141, 1215-1221.	5.1	24
45	Optimization of Complex Dataflows with User-Defined Functions. ACM Computing Surveys, 2018, 50, 1-39.	23.0	23
46	Deep web integration with VisQI. Proceedings of the VLDB Endowment, 2010, 3, 1613-1616.	3.8	21
47	Classical Hodgkin's lymphoma shows epigenetic features of abortive plasma cell differentiation. Haematologica, 2011, 96, 863-870.	3.5	20
48	High-performance information extraction with AliBaba. , 2009, , .		18
49	SAASFEE. Proceedings of the VLDB Endowment, 2015, 8, 1892-1895.	3.8	17
50	Histone acetylation and DNA demethylation of T cells result in an anaplastic large cell lymphoma-like phenotype. Haematologica, 2013, 98, 247-254.	3.5	15
51	Estimating genome-wide regulatory activity from multi-omics data sets using mathematical optimization. BMC Systems Biology, 2017, 11, 41.	3.0	15
52	H3K9me3-mediated epigenetic regulation of senescence in mice predicts outcome of lymphoma patients. Nature Communications, 2020, 11, 3651.	12.8	15
53	Effective and efficient similarity search in scientific workflow repositories. Future Generation Computer Systems, 2016, 56, 584-594.	7.5	14
54	Low uncertainty multifeature magnitude estimation with 3-D corrections and boosting tree regression: application to North Chile. Geophysical Journal International, 2020, 220, 142-159.	2.4	14

#	Article	IF	Citations
55	Annotation and initial evaluation of a large annotated German oncological corpus. JAMIA Open, 2021, 4, ooab025.	2.0	14
56	Simple tricks for improving pattern-based information extraction from the biomedical literature. Journal of Biomedical Semantics, 2010, 1, 9.	1.6	12
57	Recognizing chemicals in patents: a comparative analysis. Journal of Cheminformatics, 2016, 8, 59.	6.1	12
58	Robust <i>in-silico</i> identification of cancer cell lines based on next generation sequencing. Oncotarget, 2017, 8, 34310-34320.	1.8	12
59	DNA methylation reveals distinct cells of origin for pancreatic neuroendocrine carcinomas and pancreatic neuroendocrine tumors. Genome Medicine, 2022, 14, 24.	8.2	12
60	MRCSI. Proceedings of the VLDB Endowment, 2015, 8, 461-472.	3.8	11
61	Reflection of successful anticancer drug development processes in the literature. Drug Discovery Today, 2016, 21, 1740-1744.	6.4	11
62	Computer-assisted curation of a human regulatory core network from the biological literature. Bioinformatics, 2015, 31, 1258-1266.	4.1	10
63	SoFIA: a data integration framework for annotating high-throughput datasets. Bioinformatics, 2016, 32, 2590-2597.	4.1	10
64	Variant information systems for precision oncology. BMC Medical Informatics and Decision Making, 2018, 18, 107.	3.0	10
65	BB-Tree: A Main-Memory Index Structure for Multidimensional Range Queries. , 2019, , .		10
66	Query Planning with Information Quality Bounds. , 2001, , 85-94.		10
67	On-Demand Indexing for Referential Compression of DNA Sequences. PLoS ONE, 2015, 10, e0132460.	2.5	8
68	How to improve information extraction from German medical records. IT - Information Technology, 2017, 59, 171-179.	0.9	8
69	VIST - a Variant-Information Search Tool for precision oncology. BMC Bioinformatics, 2019, 20, 429.	2.6	8
70	PEDL: extracting protein–protein associations using deep language models and distant supervision. Bioinformatics, 2020, 36, i490-i498.	4.1	8
71	Feedback-Based Resource Allocation for Batch Scheduling of Scientific Workflows. , 2019, , .		7
72	Inference of Surface Membrane Factors of HIV-1 Infection through Functional Interaction Networks. PLoS ONE, 2010, 5, e13139.	2.5	7

#	Article	IF	Citations
73	IXDB, an X chromosome integrated database (update). Nucleic Acids Research, 1999, 27, 123-127.	14.5	6
74	Large-scale entity representation learning for biomedical relationship extraction. Bioinformatics, 2021, 37, 236-242.	4.1	6
75	Sequence Factorization with Multiple References. PLoS ONE, 2015, 10, e0139000.	2.5	6
76	How well are protein structures annotated in secondary databases?. Proteins: Structure, Function and Bioinformatics, 2005, 60, 571-576.	2.6	5
77	Potential and Pitfalls of Domain-Specific Information Extraction at Web Scale. , 2016, , .		5
78	SOA-Based Integration of Text Mining Services. , 2009, , .		4
79	Modeling Data Flow Execution in a Parallel Environment. Lecture Notes in Computer Science, 2017, , 183-196.	1.3	4
80	Estimation of Transcription Factor Activity in Knockdown Studies. Scientific Reports, 2019, 9, 9593.	3.3	4
81	Algorithms for differential splicing detection using exon arrays: a comparative assessment. BMC Genomics, 2015, 16, 136.	2.8	3
82	Exploiting Automatic Vectorization to Employ SPMD on SIMD Registers. , 2018, , .		3
83	The Collaborative Research Center FONDA. Datenbank-Spektrum, 2021, 21, 255-260.	1.3	3
84	A structural keystone for drug design. Journal of Integrative Bioinformatics, 2006, 3, 21-31.	1.5	2
85	Elevated Flt3L Predicts Long-Term Survival in Patients with High-Grade Gastroenteropancreatic Neuroendocrine Neoplasms. Cancers, 2021, 13, 4463.	3.7	2
86	Robust in-silico identification of Cancer Cell Lines based on RNA and targeted DNA sequencing data. Scientific Reports, 2019, 9, 367.	3.3	1
87	DeepTable: a permutation invariant neural network for table orientation classification. Data Mining and Knowledge Discovery, 2020, 34, 1963-1983.	3.7	1
88	Graph Management in the Life Sciences. , 2009, , 1266-1271.		1
89	Public data sources for regulatory genomic features. Medizinische Genetik, 2021, 33, 167-177.	0.2	1
90	POS: Online learning for memory-aware scheduling of scientific workflows. , 2018, , .		0

#	Article	IF	Citations
91	Semantic Data Integration for Life Science Entities. , 2009, , 2555-2559.		0
92	Semantic Data Integration for Life Science Entities. , 2018, , 3387-3391.		0
93	Senescence Markers from a DLBCL-Reminiscent Mouse Lymphoma Model Predict Patient Outcome. Blood, 2018, 132, 2850-2850.	1.4	0