

# Erhard Rahm

## List of Publications by Year in descending order

Source: <https://exaly.com/author-pdf/3671767/publications.pdf>

Version: 2024-02-01

161  
papers

7,988  
citations

172457  
29  
h-index

76900  
74  
g-index

177  
all docs

177  
docs citations

177  
times ranked

4065  
citing authors

#	ARTICLE	IF	CITATIONS
1	Distributed temporal graph analytics with GRADOOP. VLDB Journal, 2022, 31, 375-401.	4.1	13
2	LEAPME: Learning-based Property Matching with Embeddings. Data and Knowledge Engineering, 2022, 137, 101943.	3.4	4
3	Multi-source dataset of e-commerce products with attributes for property matching. Data in Brief, 2022, 41, 107884.	1.0	0
4	Enhancing Cross-lingual Semantic Annotations using Deep Network Sentence Embeddings. , 2021, , .		1
5	Optimization of the Mainzelliste software for fast privacy-preserving record linkage. Journal of Translational Medicine, 2021, 19, 33.	4.4	4
6	Towards the smart use of embedding and instance features for property matching. , 2021, , .		0
7	Incremental clustering techniques for multi-party Privacy-Preserving Record Linkage. Data and Knowledge Engineering, 2020, 128, 101809.	3.4	13
8	Evolution Analysis of Large Graphs with Gradoop. Communications in Computer and Information Science, 2020, , 402-408.	0.5	2
9	Incremental Multi-source Entity Resolution for Knowledge Graph Completion. Lecture Notes in Computer Science, 2020, , 393-408.	1.3	11
10	Evaluating Cross-lingual Semantic Annotation for Medical Forms. , 2020, , .		4
11	Informativeness-Based Active Learning for Entity Resolution. Communications in Computer and Information Science, 2020, , 125-141.	0.5	1
12	ERGAN: Generative Adversarial Networks for Entity Resolution. , 2020, , .		1
13	Analyzing Temporal Graphs with Gradoop. Datenbank-Spektrum, 2019, 19, 199-208.	1.3	5
14	Big Data Competence Center ScaDS Dresden/Leipzig: Overview and selected research activities. Datenbank-Spektrum, 2019, 19, 5-16.	1.3	5
15	ScaDS Research on Scalable Privacy-preserving Record Linkage. Datenbank-Spektrum, 2019, 19, 31-40.	1.3	6
16	BIGGR: Bringing Gradoop to Applications. Datenbank-Spektrum, 2019, 19, 51-60.	1.3	2
17	Evaluation of metadata representations in RDF stores. Semantic Web, 2019, 10, 205-229.	1.9	21
18	PRIMAT. Proceedings of the VLDB Endowment, 2019, 12, 1826-1829.	3.8	5

#	ARTICLE	IF	CITATIONS
19	Combining Semantic and Lexical Measures to Evaluate Medical Terms Similarity. Lecture Notes in Computer Science, 2019, , 17-32.	1.3	2
20	A Learning-Based Approach to Combine Medical Annotation Results. Lecture Notes in Computer Science, 2019, , 135-143.	1.3	1
21	Holistic Schema Matching. , 2019, , 960-965.		3
22	Large-Scale Entity Resolution. , 2019, , 1100-1105.		2
23	Large-Scale Schema Matching. , 2019, , 1105-1110.		0
24	Incremental Clustering on Linked Data. , 2018, , .		13
25	ScaDS Dresden/Leipzig â€“ A competence center for collaborative big data research. IT - Information Technology, 2018, 60, 327-333.	0.9	0
26	Evolving semantic annotations through multiple versions of controlled medical terminologies. Health and Technology, 2018, 8, 361-376.	3.6	2
27	Distributed Privacy-Preserving Record Linkage Using Pivot-Based Filter Techniques. , 2018, , .		4
28	Smart Medical Information Technology for Healthcare (SMITH). Methods of Information in Medicine, 2018, 57, e92-e105.	1.2	89
29	Post-processing Methods for High Quality Privacy-Preserving Record Linkage. Lecture Notes in Computer Science, 2018, , 263-278.	1.3	6
30	Large-Scale Schema Matching. , 2018, , 1-6.		2
31	Using Link Features for Entity Clustering in Knowledge Graphs. Lecture Notes in Computer Science, 2018, , 576-592.	1.3	37
32	Declarative and distributed graph analytics with GRADOOP. Proceedings of the VLDB Endowment, 2018, 11, 2006-2009.	3.8	19
33	Parallel Privacy-preserving Record Linkage using LSH-based Blocking. , 2018, , .		14
34	Scalable Matching and Clustering of Entities with FAMER. Complex Systems Informatics and Modeling Quarterly, 2018, , 61-83.	0.9	18
35	Holistic Schema Matching. , 2018, , 1-5.		1
36	Privacy-Preserving Record Linkage for Big Data: Current Approaches and Research Challenges. , 2017, , 851-895.		73

#	ARTICLE	IF	CITATIONS
37	Management and Analysis of Big Graph Data: Current Systems and Open Challenges. , 2017, , 457-505.		36
38	Cypher-based Graph Pattern Matching in Gradoop. , 2017, , .		14
39	Comparative Evaluation of Distributed Clustering Schemes for Multi-source Entity Resolution. Lecture Notes in Computer Science, 2017, , 278-293.	1.3	27
40	Distributed Holistic Clustering on Linked Data. Lecture Notes in Computer Science, 2017, , 371-382.	1.3	7
41	Mining and ranking of generalized multi-dimensional frequent subgraphs. , 2017, , .		6
42	DIMSpan. , 2017, , .		9
43	Evaluating and Improving Annotation Tools for Medical Forms. Lecture Notes in Computer Science, 2017, , 1-16.	1.3	4
44	Towards a Multi-level Approach for the Maintenance of Semantic Annotations. , 2017, , .		0
45	Graph Mining for Complex Data Analytics. , 2016, , .		6
46	Analyzing extended property graphs with Apache Flink. , 2016, , .		20
47	Scalable Privacy-Preserving Linking of Multiple Databases Using Counting Bloom Filters. , 2016, , .		14
48	Holistic Entity Clustering for Linked Data. , 2016, , .		18
49	A survey of current Link Discovery frameworks. Semantic Web, 2016, 8, 419-436.	1.9	131
50	Big Data Analytics. IT - Information Technology, 2016, 58, 155-156.	0.9	5
51	The Case for Holistic Data Integration. Lecture Notes in Computer Science, 2016, , 11-27.	1.3	32
52	Evolution of biomedical ontologies and mappings: Overview of recent approaches. Computational and Structural Biotechnology Journal, 2016, 14, 333-340.	4.1	33
53	A Reuse-Based Annotation Approach for Medical Documents. Lecture Notes in Computer Science, 2016, , 135-150.	1.3	8
54	Speeding up Privacy Preserving Record Linkage for Metric Space Similarity Measures. Datenbank-Spektrum, 2016, 16, 227-236.	1.3	9

#	ARTICLE	IF	CITATIONS
55	Semi-Automatic Identification of Counterfeit Offers in Online Shopping Platforms. Journal of Internet Commerce, 2016, 15, 59-75.	5.5	2
56	Leveraging the Impact of Ontology Evolution on Semantic Annotations. Lecture Notes in Computer Science, 2016, , 68-82.	1.3	12
57	Automatic Extraction of Semantic Relations from Wikipedia. International Journal on Artificial Intelligence Tools, 2015, 24, 1540010.	1.0	9
58	Architekturen für verteiltes und paralleles Datenmanagement. EXamen Press, 2015, , 43-78.	0.0	0
59	FoodBroker - Generating Synthetic Datasets for Graph-Based Business Analytics. Lecture Notes in Computer Science, 2015, , 145-155.	1.3	8
60	Annotating Medical Forms Using UMLS. Lecture Notes in Computer Science, 2015, , 55-69.	1.3	6
61	Effective Composition of Mappings for Matching Biomedical Ontologies. Lecture Notes in Computer Science, 2015, , 176-190.	1.3	5
62	A Clustering-Based Framework to Control Block Sizes for Entity Resolution. , 2015, , .		37
63	Extracting Semantic Concept Relations from Wikipedia. , 2014, , .		16
64	A multi-part matching strategy for mapping LOINC with laboratory terminologies. Journal of the American Medical Informatics Association: JAMIA, 2014, 21, 792-800.	4.4	17
65	Graph-based data integration and business intelligence with BIIG. Proceedings of the VLDB Endowment, 2014, 7, 1577-1580.	3.8	17
66	Discovering product counterfeits in online shops. Journal of Data and Information Quality, 2014, 5, 1-3.	2.1	4
67	Enriching ontology mappings with semantic relations. Data and Knowledge Engineering, 2014, 93, 1-18.	3.4	58
68	BIIG: Enabling business intelligence with integrated instance graphs. , 2014, , .		10
69	Iterative Computation of Connected Graph Components with MapReduce. Datenbank-Spektrum, 2014, 14, 107-117.	1.3	11
70	Target-driven merging of taxonomies with Atom. Information Systems, 2014, 42, 1-14.	3.6	29
71	LinkLion: A Link Repository for the Web of Data. Lecture Notes in Computer Science, 2014, , 439-443.	1.3	15
72	Parallel Entity Resolution with Dedoop. Datenbank-Spektrum, 2013, 13, 23-32.	1.3	31

#	ARTICLE	IF	CITATIONS
73	COntoDiff: generation of complex evolution mappings for life science ontologies. Journal of Biomedical Informatics, 2013, 46, 15-32.	4.3	60
74	Don't match twice. , 2013, , .		14
75	When to Reach for the Cloud: Using Parallel Hardware for Link Discovery. Lecture Notes in Computer Science, 2013, , 275-289.	1.3	11
76	Semi-automatic Adaptation of Mappings between Life Science Ontologies. Lecture Notes in Computer Science, 2013, , 90-104.	1.3	14
77	Semantic Enrichment of Ontology Mappings: A Linguistic-Based Approach. Lecture Notes in Computer Science, 2013, , 42-55.	1.3	9
78	The Scholarly Impact of CLEF (2000-2009). Lecture Notes in Computer Science, 2013, , 1-12.	1.3	23
79	Impact of ontology evolution on functional analyses. Bioinformatics, 2012, 28, 2671-2677.	4.1	33
80	Tailoring entity resolution for matching product offers. , 2012, , .		49
81	Dedoop. Proceedings of the VLDB Endowment, 2012, 5, 1878-1881.	3.8	107
82	A Self-Configuring Schema Matching System. , 2012, , .		25
83	CODEX: exploration of semantic changes between ontology versions. Bioinformatics, 2012, 28, 895-896.	4.1	14
84	Entity Search Strategies for Mashup Applications. , 2012, , .		7
85	Load Balancing for MapReduce-based Entity Resolution. , 2012, , .		101
86	Multi-pass sorted neighborhood blocking with MapReduce. Computer Science - Research and Development, 2012, 27, 45-63.	2.7	58
87	Towards a Benchmark for Ontology Merging. Lecture Notes in Computer Science, 2012, , 124-133.	1.3	27
88	WETSUIT. Proceedings of the VLDB Endowment, 2012, 5, 1970-1973.	3.8	9
89	Recent Advances in Schema and Ontology Evolution. , 2011, , 149-190.		54
90	GOMMA: a component-based infrastructure for managing and analyzing life science ontologies and their evolution. Journal of Biomedical Semantics, 2011, 2, 6.	1.6	49

#	ARTICLE	IF	CITATIONS
91	Block-based load balancing for entity resolution with MapReduce. , 2011, , .		18
92	Rule-based construction of matching processes. , 2011, , .		3
93	PDFMeat. , 2011, , .		4
94	Learning-based entity resolution with MapReduce. , 2011, , .		11
95	ATOM: Automatic target-driven ontology merging. , 2011, , .		35
96	Towards Large-Scale Schema and Ontology Matching. , 2011, , 3-27.		100
97	AMC - A framework for modelling and comparing matching systems as matching processes. , 2011, , .		37
98	Affiliation analysis of database publications. SIGMOD Record, 2011, 40, 26-31.	1.2	14
99	Restricting the overlap of Top-N sets in schema matching. , 2011, , .		1
100	Generic schema matching, ten years later. Proceedings of the VLDB Endowment, 2011, 4, 695-701.	3.8	139
101	Frameworks for entity matching: A comparison. Data and Knowledge Engineering, 2010, 69, 197-210.	3.4	294
102	Management of evolving semantic grid metadata within a collaborative platform. Information Sciences, 2010, 180, 1837-1849.	6.9	4
103	Evaluation of entity resolution approaches on real-world match problems. Proceedings of the VLDB Endowment, 2010, 3, 484-493.	3.8	275
104	Learning-Based Approaches for Matching Web Data Entities. IEEE Internet Computing, 2010, 14, 23-31.	3.3	26
105	Discovering Evolving Regions in Life Science Ontologies. Lecture Notes in Computer Science, 2010, , 19-34.	1.3	5
106	On Matching Large Life Science Ontologies in Parallel. Lecture Notes in Computer Science, 2010, , 35-49.	1.3	36
107	Rewrite techniques for performance optimization of schema matching processes. , 2010, , .		28
108	Comparative evaluation of entity resolution approaches with FEVER. Proceedings of the VLDB Endowment, 2009, 2, 1574-1577.	3.8	25

#	ARTICLE	IF	CITATIONS
109	Efficient Management of Biomedical Ontology Versions. Lecture Notes in Computer Science, 2009, , 574-583.	1.3	9
110	OnEX: Exploring changes in life science ontologies. BMC Bioinformatics, 2009, 10, 250.	2.6	23
111	Convergent validity of bibliometric Google Scholar data in the field of chemistry – Citation counts for papers that were accepted by Angewandte Chemie International Edition or rejected but published elsewhere, using Google Scholar, Science Citation Index, Scopus, and Chemical Abstracts. Journal of Informetrics, 2009, 3, 27-35.	2.9	76
112	MediGRID: Towards a user friendly secured grid infrastructure. Future Generation Computer Systems, 2009, 25, 326-336.	7.5	63
113	Estimating the Quality of Ontology-Based Annotations by Considering Evolutionary Changes. Lecture Notes in Computer Science, 2009, , 71-87.	1.3	9
114	Comparing the scientific impact of conference and journal publications in computer science. Information Services and Use, 2008, 28, 127-128.	0.2	14
115	Analyzing the Evolution of Life Science Ontologies and Mappings. Lecture Notes in Computer Science, 2008, , 11-27.	1.3	23
116	A Platform for Collaborative Management of Semantic Grid Metadata. Studies in Computational Intelligence, 2008, , 115-125.	0.9	2
117	Quickmig. , 2007, , .		50
118	FUNC: a package for detecting significant associations between gene sets and ontological annotations. BMC Bioinformatics, 2007, 8, 41.	2.6	180
119	Matching large schemas: Approaches and evaluation. Information Systems, 2007, 32, 857-885.	3.6	183
120	Instance-Based Matching of Large Life Science Ontologies. , 2007, , 172-187.		20
121	Parameterized XPath Views. Lecture Notes in Computer Science, 2007, , 125-137.	1.3	0
122	Dynamic Fusion of Web Data. Lecture Notes in Computer Science, 2007, , 14-16.	1.3	6
123	An online bibliography on schema evolution. SIGMOD Record, 2006, 35, 30-31.	1.2	65
124	BioFuice: Mapping-Based Data Integration in Bioinformatics. Lecture Notes in Computer Science, 2006, , 124-135.	1.3	11
125	An Integrated Platform for Analyzing Molecular-Biological Data Within Clinical Studies. Lecture Notes in Computer Science, 2006, , 399-410.	1.3	4
126	Adaptive website recommendations with AWESOME. VLDB Journal, 2005, 14, 357-372.	4.1	3



#	ARTICLE	IF	CITATIONS
127	Automatic Optimization of Web Recommendations Using Feedback and Ontology Graphs. Lecture Notes in Computer Science, 2005, , 375-386.	1.3	4
128	Supporting executable mappings in model management. , 2005, , .		56
129	Citation analysis of database publications. SIGMOD Record, 2005, 34, 48-53.	1.2	64
130	Schema and ontology matching with COMA++. , 2005, , .		395
131	Hybrid Integration of Molecular-Biological Annotation Data. Lecture Notes in Computer Science, 2005, , 208-223.	1.3	3
132	Matching large XML schemas. SIGMOD Record, 2004, 33, 26-31.	1.2	76
133	AgentWork: a workflow system supporting rule-based workflow adaptation. Data and Knowledge Engineering, 2004, 51, 223-256.	3.4	241
134	Flexible Integration of Molecular-Biological Annotation Data: The GenMapper Approach. Lecture Notes in Computer Science, 2004, , 811-822.	1.3	18
135	Dynamic query scheduling in parallel data warehouses. Concurrency Computation Practice and Experience, 2003, 15, 1169-1190.	2.2	7
136	Developing metadata-intensive applications with Rondo. Web Semantics, 2003, 1, 47-74.	2.9	21
137	Comparison of Schema Matching Evaluations. Lecture Notes in Computer Science, 2003, , 221-237.	1.3	150
138	Rondo. , 2003, , .		175
139	COMA – A system for flexible combination of schema matching approaches. , 2002, , 610-621.		580
140	A survey of approaches to automatic schema matching. VLDB Journal, 2001, 10, 334-350.	4.1	2,288
141	XMach-1: A Benchmark for XML Data Management. Informatik Aktuell, 2001, , 264-273.	0.6	53
142	Data Warehouse Scenarios for Model Management. Lecture Notes in Computer Science, 2000, , 1-15.	1.3	41
143	Evaluierung von Data Warehouse-Werkzeugen. , 2000, , 43-57.		2
144	Datenbanksysteme. , 1999, , .		3

#	ARTICLE	IF	CITATIONS
145	Dynamic load balancing in parallel database systems. Lecture Notes in Computer Science, 1996, , 37-52.	1.3	20
146	Analysis of parallel scan processing in Shared Disk database systems. Lecture Notes in Computer Science, 1995, , 485-500.	1.3	5
147	Cache management for shared sequential data access. Information Systems, 1993, 18, 197-213.	3.6	0
148	Empirical performance evaluation of concurrency and coherency control protocols for database sharing systems. ACM Transactions on Database Systems, 1993, 18, 333-377.	2.8	51
149	Parallel query processing in shared disk database systems. SIGMOD Record, 1993, 22, 32-37.	1.2	22
150	High performance cache management for sequential data access. , 1992, , .		1
151	Performance evaluation of extended storage architectures for transaction processing. , 1992, , .		14
152	Performance evaluation of extended storage architectures for transaction processing. SIGMOD Record, 1992, 21, 308-317.	1.2	2
153	High performance cache management for sequential data access. Performance Evaluation Review, 1992, 20, 243-244.	0.6	0
154	A Framework for workload allocation in distributed transaction processing systems. Journal of Systems and Software, 1992, 18, 171-190.	4.5	28
155	Performance evaluation of parallel transaction processing in shared nothing database systems. Lecture Notes in Computer Science, 1992, , 295-310.	1.3	10
156	Architekturansätze zur Unterstützung heterogener Datenbanken. Informatik Aktuell, 1992, , 106-118.	0.6	0
157	Mehrrechner-Datenbanksysteme. Informatik-Fachberichte, 1988, , 6-19.	0.2	0
158	A Reliable and Efficient Synchronization Protocol for Database Sharing Systems. Informatik-Fachberichte, 1987, , 336-347.	0.2	0
159	Primary copy synchronization for DB-Sharing. Information Systems, 1986, 11, 275-286.	3.6	24
160	Concurrency Control in DB-Sharing Systems. Informatik-Fachberichte, 1986, , 617-632.	0.2	2
161	Developing Metadata-Intensive Applications with Rondo. SSRN Electronic Journal, 0, , .	0.4	1