

Guoliang Li

List of Publications by Year in descending order

Source: <https://exaly.com/author-pdf/3064759/publications.pdf>

Version: 2024-02-01

183
papers

6,603
citations

159585

30
h-index

161849

54
g-index

186
all docs

186
docs citations

186
times ranked

2626
citing authors

#	ARTICLE	IF	CITATIONS
1	EASE. , 2008, , .		290
2	Truth inference in crowdsourcing. Proceedings of the VLDB Endowment, 2017, 10, 541-552.	3.8	278
3	Crowdsourced Data Management: A Survey. IEEE Transactions on Knowledge and Data Engineering, 2016, 28, 2296-2319.	5.7	217
4	Can we beat the prefix filtering?. , 2012, , .		157
5	QASCA. , 2015, , .		155
6	Leveraging transitive relations for crowdsourced joins. , 2013, , .		154
7	An End-to-End Automatic Cloud Database Tuning System Using Deep Reinforcement Learning. , 2019, , .		149
8	iCrowd. , 2015, , .		144
9	Effective keyword search for valuable lcas over xml documents. , 2007, , .		143
10	A Survey of Traffic Prediction: from Spatio-Temporal Data to Intelligent Transportation. Data Science and Engineering, 2021, 6, 63-85.	6.4	138
11	Efficient interactive fuzzy keyword search. , 2009, , .		133
12	QTune. Proceedings of the VLDB Endowment, 2019, 12, 2118-2130.	3.8	129
13	Pass-join. Proceedings of the VLDB Endowment, 2011, 5, 253-264.	3.8	127
14	G-Tree: An Efficient and Scalable Index for Spatial Search on Road Networks. IEEE Transactions on Knowledge and Data Engineering, 2015, 27, 2175-2189.	5.7	119
15	Fast-join: An efficient method for fuzzy token matching based string similarity join. , 2011, , .		101
16	String similarity search and join: a survey. Frontiers of Computer Science, 2016, 10, 399-417.	2.4	100
17	DeepEye: Towards Automatic Data Visualization. , 2018, , .		100
18	Making data visualization more efficient and effective: a survey. VLDB Journal, 2020, 29, 93-117.	4.1	99

#	ARTICLE	IF	CITATIONS
19	An end-to-end learning-based cost estimator. Proceedings of the VLDB Endowment, 2019, 13, 307-319.	3.8	97
20	String similarity joins. Proceedings of the VLDB Endowment, 2014, 7, 625-636.	3.8	91
21	DESKS: Direction-Aware Spatial Keyword Search. , 2012, , .		89
22	Entity matching. Proceedings of the VLDB Endowment, 2011, 4, 622-633.	3.8	85
23	Approximate Query Processing: What is New and Where to Go?. Data Science and Engineering, 2018, 3, 379-397.	6.4	83
24	Trie-join. Proceedings of the VLDB Endowment, 2010, 3, 1219-1230.	3.8	81
25	DOCS. Proceedings of the VLDB Endowment, 2016, 10, 361-372.	3.8	79
26	Crowdsourced POI labelling: Location-aware result inference and Task Assignment. , 2016, , .		73
27	MassJoin: A mapreduce-based method for scalable string similarity joins. , 2014, , .		70
28	Crowdsourced Data Management. , 2017, , .		70
29	Cost-Effective Crowdsourced Entity Resolution. , 2016, , .		67
30	DITA. , 2018, , .		66
31	PriRadar: A Privacy-Preserving Framework for Spatial Crowdsourcing. IEEE Transactions on Information Forensics and Security, 2020, 15, 299-314.	6.9	66
32	G-tree. , 2013, , .		65
33	An Efficient Ride-Sharing Framework for Maximizing Shared Route. IEEE Transactions on Knowledge and Data Engineering, 2018, 30, 219-233.	5.7	63
34	Efficient type-ahead search on relational data. , 2009, , .		62
35	CDB. , 2017, , .		58
36	Database Meets Artificial Intelligence: A Survey. IEEE Transactions on Knowledge and Data Engineering, 2022, 34, 1096-1116.	5.7	55

#	ARTICLE	IF	CITATIONS
37	Seal. Proceedings of the VLDB Endowment, 2012, 5, 824-835.	3.8	52
38	Trajectory-driven Influential Billboard Placement. , 2018, , .		51
39	Effective Travel Time Estimation: When Historical Trajectories over Road Networks Matter. , 2020, , .		51
40	Effective keyword-based selection of relational databases. , 2007, , .		50
41	Signature-Based Trajectory Similarity Join. IEEE Transactions on Knowledge and Data Engineering, 2017, 29, 870-883.	5.7	50
42	A pivotal prefix based filtering algorithm for string similarity search. , 2014, , .		49
43	Efficient fuzzy full-text type-ahead search. VLDB Journal, 2011, 20, 617-640.	4.1	48
44	Trie-join: a trie-based method for efficient string similarity joins. VLDB Journal, 2012, 21, 437-461.	4.1	48
45	Distributed In-memory Trajectory Similarity Search and Join on Road Network. , 2019, , .		48
46	Interactive and Deterministic Data Cleaning. , 2016, , .		47
47	Supporting Database Applications as a Service. Proceedings - International Conference on Data Engineering, 2009, , .	0.0	46
48	Top-k string similarity search with edit-distance constraints. , 2013, , .		46
49	VizNet. , 2019, , .		46
50	AI Meets Database: AI4DB and DB4AI. , 2021, , .		43
51	Query performance prediction for concurrent queries using graph embedding. Proceedings of the VLDB Endowment, 2020, 13, 1416-1428.	3.8	43
52	DeepEye. , 2018, , .		42
53	Human-in-the-loop data integration. Proceedings of the VLDB Endowment, 2017, 10, 2006-2017.	3.8	41
54	Querying shortest paths on time dependent road networks. Proceedings of the VLDB Endowment, 2019, 12, 1249-1261.	3.8	40

#	ARTICLE	IF	CITATIONS
55	Effective location identification from microblogs. , 2014, , .		39
56	Natural Language to Visualization by Neural Machine Translation. IEEE Transactions on Visualization and Computer Graphics, 2022, 28, 217-226.	4.4	39
57	Hike. , 2017, , .		37
58	Crowdsourced top-k algorithms. Proceedings of the VLDB Endowment, 2016, 9, 612-623.	3.8	36
59	Interactive SQL query suggestion: Making databases user-friendly. , 2011, , .		35
60	Two birds with one stone: An efficient hierarchical framework for top-k and threshold-based string similarity search. , 2015, , .		35
61	Crowdsourcing-based real-time urban traffic speed estimation: From trends to speeds. , 2016, , .		35
62	Faerie. , 2011, , .		34
63	Extending string similarity join to tolerant fuzzy token matching. ACM Transactions on Database Systems, 2014, 39, 1-45.	2.8	34
64	Human-in-the-loop Outlier Detection. , 2020, , .		34
65	Privacy-Preserving Top-k Spatial Keyword Queries in Untrusted Cloud Environments. IEEE Transactions on Services Computing, 2015, , 1-1.	4.6	33
66	V-Tree: Efficient kNN Search on Moving Objects with Road-Network Constraints. , 2017, , .		32
67	Automatic View Generation with Deep Learning and Reinforcement Learning. , 2020, , .		32
68	Progressive Keyword Search in Relational Databases. Proceedings - International Conference on Data Engineering, 2009, , .	0.0	31
69	A partition-based method for string similarity joins with edit-distance constraints. ACM Transactions on Database Systems, 2013, 38, 1-33.	2.8	29
70	Scalable column concept determination for web tables using large knowledge bases. Proceedings of the VLDB Endowment, 2013, 6, 1606-1617.	3.8	28
71	Efficient Filtering Algorithms for Location-Aware Publish/Subscribe. IEEE Transactions on Knowledge and Data Engineering, 2015, 27, 950-963.	5.7	27
72	A partial-order-based framework for cost-effective crowdsourced entity resolution. VLDB Journal, 2018, 27, 745-770.	4.1	27

#	ARTICLE	IF	CITATIONS
73	SAIL: Structure-aware indexing for effective and progressive top-k keyword search over XML documents. Information Sciences, 2009, 179, 3745-3762.	6.9	26
74	Finding Top-k Answers in Keyword Search over Relational Databases Using Tuple Units. IEEE Transactions on Knowledge and Data Engineering, 2011, 23, 1781-1794.	5.7	24
75	Location-aware instant search. , 2012, , .		24
76	Top-k Spatio-Textual Similarity Join. IEEE Transactions on Knowledge and Data Engineering, 2016, 28, 551-565.	5.7	24
77	Synthesizing Natural Language to Visualization (NL2VIS) Benchmarks from NL2SQL Benchmarks. , 2021, , .		24
78	Incremental Quality Inference in Crowdsourcing. Lecture Notes in Computer Science, 2014, , 453-467.	1.3	24
79	A unified framework for string similarity search with edit-distance constraint. VLDB Journal, 2017, 26, 249-274.	4.1	23
80	Crowdsourcing Database Systems: Overview and Challenges. , 2019, , .		23
81	A learned query rewrite system using Monte Carlo tree search. Proceedings of the VLDB Endowment, 2021, 15, 46-58.	3.8	23
82	Providing built-in keyword search capabilities in RDBMS. VLDB Journal, 2011, 20, 1-19.	4.1	22
83	Supporting efficient top-k queries in type-ahead search. , 2012, , .		22
84	Steerable Self-Driving Data Visualization. IEEE Transactions on Knowledge and Data Engineering, 2022, 34, 475-490.	5.7	22
85	K-Join: Knowledge-Aware Similarity Join. IEEE Transactions on Knowledge and Data Engineering, 2016, 28, 3293-3308.	5.7	21
86	Finding and ranking compact connected trees for effective keyword proximity search in XML documents. Information Systems, 2010, 35, 186-203.	3.6	20
87	Efficient Fuzzy Type-Ahead Search in XML Data. IEEE Transactions on Knowledge and Data Engineering, 2012, 24, 882-895.	5.7	20
88	Efficient Similarity Join and Search on Multi-Attribute Data. , 2015, , .		20
89	Cleaning Relations Using Knowledge Bases. , 2017, , .		20
90	CDB. Proceedings of the VLDB Endowment, 2018, 11, 1926-1929.	3.8	20

#	ARTICLE	IF	CITATIONS
91	A unified framework for approximate dictionary-based entity extraction. VLDB Journal, 2015, 24, 143-167.	4.1	19
92	Relational data synthesis using generative adversarial networks. Proceedings of the VLDB Endowment, 2020, 13, 1962-1975.	3.8	19
93	Supporting Search-As-You-Type Using SQL in Databases. IEEE Transactions on Knowledge and Data Engineering, 2013, 25, 461-475.	5.7	18
94	Incentive-Based Entity Collection Using Crowdsourcing. , 2018, , .		18
95	Interactive Cleaning for Progressive Visualization through Composite Questions. , 2020, , .		18
96	Suggesting Topic-Based Query Terms as You Type. , 2010, , .		17
97	DeepTrack: Monitoring and exploring spatio-temporal data. Proceedings of the VLDB Endowment, 2020, 13, 2841-2844.	3.8	17
98	Selective data acquisition in the wild for model charging. Proceedings of the VLDB Endowment, 2022, 15, 1466-1478.	3.8	17
99	An Efficient Trie-based Method for Approximate Entity Extraction with Edit-Distance Constraints. , 2012, , .		16
100	Efficient parallel partition-based algorithms for similarity search and join with edit distance constraints. , 2013, , .		16
101	Sybil Defense in Crowdsourcing Platforms. , 2017, , .		16
102	Learned Cardinality Estimation for Similarity Queries. , 2021, , .		16
103	An effective 3-in-1 keyword search method over heterogeneous data sources. Information Systems, 2011, 36, 248-266.	3.6	15
104	A Prefix-Filter based Method for Spatio-Textual Similarity Join. IEEE Transactions on Knowledge and Data Engineering, 2014, 26, 2354-2367.	5.7	15
105	Machine learning for databases. Proceedings of the VLDB Endowment, 2021, 14, 3190-3193.	3.8	15
106	Sailer. , 2008, , .		14
107	An Autonomous Materialized View Management System with Deep Reinforcement Learning. , 2021, , .		14
108	$\text{\$}\hbox {CDBTune}^{\{+\}}\text{\$}$: An efficient deep reinforcement learning-based automatic cloud database tuning system. VLDB Journal, 2021, 30, 959-987.	4.1	13

#	ARTICLE	IF	CITATIONS
109	Parallel Structural Join Algorithm on Shared-Memory Multi-Core Systems. , 2008, , .		12
110	GINIX: Generalized inverted index for keyword search. Tsinghua Science and Technology, 2013, 18, 77-87.	6.1	12
111	META. Proceedings of the VLDB Endowment, 2016, 9, 828-839.	3.8	12
112	Effective Result Inference for Context-Sensitive Tasks in Crowdsourcing. Lecture Notes in Computer Science, 2016, , 33-48.	1.3	12
113	Context-aware result inference in crowdsourcing. Information Sciences, 2018, 460-461, 346-363.	6.9	12
114	$\langle \text{mathsf{B}}^{\{p\}} \rangle$ $\langle \text{mathsf{Tree}} \rangle$: A Predictive $\langle \text{mathsf{B}}^{\{+ \}} \rangle$ $\langle \text{mathsf{Tree}} \rangle$ for Reducing Writes on Phase Change Memory. IEEE Transactions on Knowledge and Data Engineering, 2014, 26, 2368-2381.	5.7	11
115	CrowdOTA: An Online Task Assignment System in Crowdsourcing. , 2018, , .		11
116	VisClean. Proceedings of the VLDB Endowment, 2020, 13, 2821-2824.	3.8	11
117	Data Management for Machine Learning: A Survey. IEEE Transactions on Knowledge and Data Engineering, 2022, , 1-1.	5.7	11
118	ELAN: An Efficient Location-Aware Analytics System. Big Data Research, 2016, 5, 16-21.	4.2	10
119	Bounded Approximate Query Processing. IEEE Transactions on Knowledge and Data Engineering, 2019, 31, 2262-2276.	5.7	10
120	Evaluating Public Anxiety for Topic-Based Communities in Social Networks. IEEE Transactions on Knowledge and Data Engineering, 2022, 34, 1191-1205.	5.7	10
121	Interactive search in XML data. , 2009, , .		10
122	T-Crowd: Effective Crowdsourcing for Tabular Data. , 2018, , .		9
123	Outdated Fact Detection in Knowledge Bases. , 2020, , .		9
124	CrowdRL: An End-to-End Reinforcement Learning Framework for Data Labelling. , 2021, , .		9
125	Incremental Mining of Frequent Query Patterns from XML Queries for Caching. IEEE International Conference on Data Mining, 2006, , .	0.0	8
126	Adaptive Database Schema Design for Multi-Tenant Data Management. IEEE Transactions on Knowledge and Data Engineering, 2014, 26, 2079-2093.	5.7	8

#	ARTICLE	IF	CITATIONS
127	Breaking out of the MisMatch trap. , 2014, , .		8
128	Distilling relations using knowledge bases. VLDB Journal, 2018, 27, 497-519.	4.1	8
129	Domain Adaptation for Deep Entity Resolution. , 2022, , .		8
130	A User-Friendly Patent Search Paradigm. IEEE Transactions on Knowledge and Data Engineering, 2013, 25, 1439-1443.	5.7	7
131	A general framework to resolve the MisMatch problem in XML keyword search. VLDB Journal, 2015, 24, 493-518.	4.1	7
132	An Efficient Ride-Sharing Framework for Maximizing Shared Routes. , 2018, , .		7
133	Towards Automatic Mathematical Exercise Solving. Data Science and Engineering, 2019, 4, 179-192.	6.4	7
134	A Crowdsourcing Framework for Collecting Tabular Data. IEEE Transactions on Knowledge and Data Engineering, 2020, 32, 2060-2074.	5.7	7
135	CrowdChart: Crowdsourced Data Extraction From Visualization Charts. IEEE Transactions on Knowledge and Data Engineering, 2021, 33, 3537-3549.	5.7	7
136	An effective and versatile keyword search engine on heterogenous data sources. Proceedings of the VLDB Endowment, 2008, 1, 1452-1455.	3.8	7
137	Data Management at Huawei: Recent Accomplishments and Future Challenges. , 2019, , .		6
138	A game-based framework for crowdsourced data labeling. VLDB Journal, 2020, 29, 1311-1336.	4.1	6
139	A Human-Machine Method for Web Table Understanding. Lecture Notes in Computer Science, 2013, , 179-189.	1.3	6
140	LearnedSQLGen: Constraint-aware SQL Generation using Reinforcement Learning. , 2022, , .		6
141	Efficient Algorithms for Top-k Keyword Queries on Spatial Databases. , 2011, , .		5
142	Crowdsourced Selection on Multi-Attribute Data. , 2017, , .		5
143	Representation Learning from Limited Educational Data with Crowdsourced Labels. IEEE Transactions on Knowledge and Data Engineering, 2020, , 1-1.	5.7	5
144	Crowdsourcing-based Data Extraction from Visualization Charts. , 2020, , .		5

#	ARTICLE	IF	CITATIONS
145	HTAP Databases: What is New and What is Next. , 2022, , .		5
146	Exploiting Even Partition to Accelerate Structure Join. , 2006, , .		4
147	Exploit sequencing views in semantic cache to accelerate xpath query evaluation. , 2007, , .		4
148	An Effective Semantic Cache for Exploiting XPath Query/View Answerability. Journal of Computer Science and Technology, 2010, 25, 347-361.	1.5	4
149	Fast Similar Subgraph Search with Maximum Common Connected Subgraph Constraints. , 2013, , .		4
150	Removing the mismatch headache in XML keyword search. , 2013, , .		4
151	Top-k spatio-textual similarity join. , 2016, , .		4
152	A Review on the Architecture of Software Defined Network. Chinese Journal of Electronics, 2018, 27, 1111-1117.	1.5	4
153	GPSC: A Grid-Based Privacy-Reserving Framework for Online Spatial Crowdsourcing. IEEE Transactions on Knowledge and Data Engineering, 2022, 34, 5378-5390.	5.7	4
154	Sevi: Speech-to-Visualization through Neural Machine Translation. , 2022, , .		4
155	A Semantic Cache Framework for Secure XML Queries. Journal of Computer Science and Technology, 2008, 23, 988-997.	1.5	3
156	Incremental sequence-based frequent query pattern mining from XML queries. Data Mining and Knowledge Discovery, 2009, 18, 472-516.	3.7	3
157	Keyword searches in data-centric XML documents using tree partitioning. Tsinghua Science and Technology, 2009, 14, 7-18.	6.1	3
158	KEMB: A Keyword-Based XML Message Broker. IEEE Transactions on Knowledge and Data Engineering, 2011, 23, 1035-1049.	5.7	3
159	K-Join: Knowledge-Aware Similarity Join. , 2017, , .		3
160	Thinking and Modeling for Big Data from the Perspective of the I Ching. International Journal of Information Technology and Decision Making, 2017, 16, 1451-1463.	3.9	3
161	A Novel Cost-Based Model for Data Repairing. , 2017, , .		3
162	Discovering Mis-Categorized Entities. , 2018, , .		3

#	ARTICLE	IF	CITATIONS
163	MathGraph: A Knowledge Graph for Automatically Solving Mathematical Exercises. Lecture Notes in Computer Science, 2019, , 760-776.	1.3	3
164	Efficient XML Keyword Search: From Graph Model to Tree Model. Lecture Notes in Computer Science, 2013, , 25-39.	1.3	3
165	Interactively discovering and ranking desired tuples by data exploration. VLDB Journal, 2022, 31, 753-777.	4.1	3
166	Effective Indices for Efficient Approximate String Search and Similarity Join. , 2008, , .		2
167	Cleaning Your Wrong Google Scholar Entries. , 2018, , .		2
168	A framework for image dark data assessment. World Wide Web, 2020, 23, 2079-2105.	4.0	2
169	SCEND: An Efficient Semantic Cache to Adequately Explore Answerability of Views. Lecture Notes in Computer Science, 2006, , 460-473.	1.3	2
170	Influential Factors on Collective Anxiety of Online Topic-Based Communities. Frontiers in Psychology, 2021, 12, 740065.	2.1	2
171	LCA-Based Keyword Search for Effectively Retrieving "Information Unit" from Web Pages. , 2008, , .		1
172	Crowdsourced Data Management. , 2018, , .		1
173	Location and Trajectory Identification from Microblogs. Journal of Computer Science and Technology, 2019, 34, 727-746.	1.5	1
174	Mis-categorized entities detection. VLDB Journal, 2021, 30, 515-536.	4.1	1
175	Crowdsourced Entity Alignment: A Decision Theory Based Approach. Lecture Notes in Computer Science, 2017, , 19-36.	1.3	1
176	Automatic URL completion and prediction using fuzzy type-ahead search. , 2009, , .		1
177	RNE: computing shortest paths using road network embedding. VLDB Journal, 2022, 31, 507-528.	4.1	1
178	Special Issue of DASFAA 2019. Data Science and Engineering, 2019, 4, 177-178.	6.4	0
179	Knowledge Base Error Detection with Relation Sensitive Embedding. Lecture Notes in Computer Science, 2019, , 725-741.	1.3	0
180	Adaptive algorithms for crowd-aided categorization. VLDB Journal, 0, , 1.	4.1	0

#	ARTICLE	IF	CITATIONS
181	Crowdsourced Operators. , 2018, , 97-154.		0
182	Cost Control. , 2018, , 45-62.		0
183	AutoView: An Autonomous Materialized View Management System with Encoder-Reducer. IEEE Transactions on Knowledge and Data Engineering, 2022, , 1-1.	5.7	0