

# Sherif Sakr

## List of Publications by Year in descending order

Source: <https://exaly.com/author-pdf/3594332/publications.pdf>

Version: 2024-02-01

165  
papers

3,420  
citations

218677

26  
h-index

182427

51  
g-index

181  
all docs

181  
docs citations

181  
times ranked

3348  
citing authors

#	ARTICLE	IF	CITATIONS
1	DIA: User-defined interval analytics on distributed streams. Information Systems, 2022, 104, 101679.	3.6	3
2	Exploiting time series of Sentinel-1 and Sentinel-2 to detect grassland mowing events using deep learning with reject region. Scientific Reports, 2022, 12, 983.	3.3	12
3	Interpretability in healthcare: A comparative study of local machine learning interpretability techniques. Computational Intelligence, 2021, 37, 1633-1650.	3.2	58
4	SDDM: an interpretable statistical concept drift detection method for data streams. Journal of Intelligent Information Systems, 2021, 56, 459-484.	3.9	12
5	DLBench: a comprehensive experimental evaluation of deep learning frameworks. Cluster Computing, 2021, 24, 2017-2038.	5.0	32
6	Graph Generators. ACM Computing Surveys, 2021, 53, 1-30.	23.0	23
7	cSmartML: A Meta Learning-Based Framework for Automated Selection and Hyperparameter Tuning for Clustering. , 2021, , .		3
8	Benchmarking big data systems: A survey. Computer Communications, 2020, 149, 241-251.	5.1	16
9	A First Step Towards a Streaming Linked Data Life-Cycle. Lecture Notes in Computer Science, 2020, , 634-650.	1.3	5
10	Large-Scale Graph Processing Systems. , 2020, , 59-93.		0
11	Towards making sense of Spark-SQL performance for processing vast distributed RDF datasets. , 2020, , .		4
12	Large-Scale Machine/Deep Learning Frameworks. , 2020, , 117-126.		0
13	Large-Scale Stream Processing Systems. , 2020, , 95-115.		0
14	General-Purpose Big Data Processing Systems. , 2020, , 17-43.		0
15	Large-Scale Processing Systems of Structured Data. , 2020, , 45-58.		0
16	RDF Data Storage and Query Processing Schemes. ACM Computing Surveys, 2019, 51, 1-36.	23.0	69
17	On the interpretability of machine learning-based model for predicting hypertension. BMC Medical Informatics and Decision Making, 2019, 19, 146.	3.0	141
18	CARDIORESPIRATORY FITNESS AND INCIDENT STROKE TYPES: THE FIT (HENRY FORD EXERCISE TESTING) PROJECT. Journal of the American College of Cardiology, 2019, 73, 1697.	2.8	0

#	ARTICLE	IF	CITATIONS
19	Artificial intelligence for plaque characterization: A scientific exercise looking for a clinical application. <i>Atherosclerosis</i> , 2019, 288, 158-159.	0.8	4
20	An Outlook to Declarative Languages for Big Steaming Data. , 2019, , .		4
21	Predictors of in-hospital length of stay among cardiac patients: A machine learning approach. <i>International Journal of Cardiology</i> , 2019, 288, 140-147.	1.7	110
22	The Interplay of the Global Atherosclerotic Cardiovascular Disease Risk Scoring and Cardiorespiratory Fitness for the Prediction of All-Cause Mortality and Myocardial Infarction: The Henry Ford Exercise Testing Project (The FIT Project). <i>American Journal of Cardiology</i> , 2019, 124, 511-517.	1.6	4
23	Calculation of Average Road Speed Based on Car-to-Car Messaging. , 2019, , .		0
24	D \$\$^2\$\$\$ 2 IA: Stream Analytics on User-Defined Event Intervals. <i>Lecture Notes in Computer Science</i> , 2019, , 346-361.	1.3	1
25	Native Distributed RDF Systems. , 2019, , 1171-1178.		0
26	Big SQL systems: an experimental evaluation. <i>Cluster Computing</i> , 2019, 22, 1347-1377.	5.0	12
27	ILIME: Local and Global Interpretable Model-Agnostic Explainer of Black-Box Decision. <i>Lecture Notes in Computer Science</i> , 2019, , 53-68.	1.3	11
28	Framework-Based Scale-Out RDF Systems. , 2019, , 771-777.		0
29	NoSQL Database Systems. , 2019, , 1193-1198.		0
30	Dagstuhl Seminar on Big Stream Processing. <i>SIGMOD Record</i> , 2019, 47, 36-39.	1.2	0
31	Distributed RDF Query Processing. , 2018, , 51-83.		0
32	Prognostic value of exercise capacity among patients with treated depression: The Henry Ford Exercise Testing (FIT) Project. <i>Clinical Cardiology</i> , 2018, 41, 532-538.	1.8	3
33	Cardiorespiratory Fitness and Cardiovascular Disease Prevention: an Update. <i>Current Atherosclerosis Reports</i> , 2018, 20, 1.	4.8	134
34	A Differentiated Caching Mechanism to Enable Primary Storage Deduplication in Clouds. <i>IEEE Transactions on Parallel and Distributed Systems</i> , 2018, 29, 1202-1216.	5.6	21
35	HDM: A Composable Framework for Big Data Processing. <i>IEEE Transactions on Big Data</i> , 2018, 4, 150-163.	6.1	12
36	Business Process Analytics and Big Data Systems: A Roadmap to Bridge the Gap. <i>IEEE Access</i> , 2018, 6, 77308-77320.	4.2	25

#	ARTICLE	IF	CITATIONS
37	Editorial for Special Issue of Journal of Big Data Research on "Big Medical/Healthcare Data Analytics"; Big Data Research, 2018, 13, 1-2.	4.2	3
38	NoSQL Database Systems. , 2018, , 1-6.		0
39	Big Data Systems Meet Machine Learning Challenges: Towards Big Data Science as a Service. Big Data Research, 2018, 14, 1-11.	4.2	96
40	Using machine learning on cardiorespiratory fitness data for predicting hypertension: The Henry Ford Exercise Testing (FIT) Project. PLoS ONE, 2018, 13, e0195344.	2.5	76
41	HDM-MC in-Action: A Framework for Big Data Analytics across Multiple Clusters. , 2018, , .		0
42	Predictive Model for the Incidence of Hyperkalemia for Congestive Heart Failure Patients on Spironolactone. , 2018, , .		1
43	Linked Data. , 2018, , .		43
44	Centralized RDF Query Processing. , 2018, , 33-49.		2
45	Stream Processing Languages in the Big Data Era. SIGMOD Record, 2018, 47, 29-40.	1.2	21
46	Native Distributed RDF Systems. , 2018, , 1-8.		0
47	NoSQL Database Systems. , 2018, , 1-6.		0
48	Framework-Based Scale-Out RDF Systems. , 2018, , 1-7.		0
49	Big Data Processing Stacks. IT Professional, 2017, 19, 34-41.	1.5	12
50	Handbook of Big Data Technologies. , 2017, , .		49
51	Big Data Storage and Data Models. , 2017, , 3-29.		9
52	Non-native RDF Storage Engines. , 2017, , 339-364.		0
53	Big Data Programming Models. , 2017, , 31-63.		10
54	USING MACHINE LEARNING TO DEFINE THE ASSOCIATION BETWEEN CARDIORESPIRATORY FITNESS AND ALL-CAUSE MORTALITY: THE FIT (HENRY FORD EXERCISE TESTING) PROJECT. Journal of the American College of Cardiology, 2017, 69, 1612.	2.8	2

#	ARTICLE	IF	CITATIONS
55	On business process monitoring using cross-flow coordination. <i>Service Oriented Computing and Applications</i> , 2017, 11, 203-215.	1.6	5
56	Cardiorespiratory fitness and incident heart failure: The Henry Ford Exercise Testing (FIT) Project. <i>American Heart Journal</i> , 2017, 185, 35-42.	2.7	47
57	Structural XML Query Processing. <i>ACM Computing Surveys</i> , 2017, 50, 1-41.	23.0	8
58	Towards Big Data Analytics across Multiple Clusters. , 2017, , .		4
59	Higher Fitness Is Strongly Protective in Patients with Family History of Heart Disease: The FIT Project. <i>American Journal of Medicine</i> , 2017, 130, 367-371.	1.5	8
60	The impact of digoxin on mortality in patients with chronic systolic heart failure: A propensity-matched cohort study. <i>International Journal of Cardiology</i> , 2017, 228, 214-218.	1.7	13
61	Using Machine Learning to Define the Association between Cardiorespiratory Fitness and All-Cause Mortality (from the Henry Ford Exercise Testing Project). <i>American Journal of Cardiology</i> , 2017, 120, 2078-2084.	1.6	22
62	Predicting diabetes mellitus using SMOTE and ensemble machine learning approach: The Henry Ford Exercise Testing (FIT) project. <i>PLoS ONE</i> , 2017, 12, e0179805.	2.5	194
63	Comparison of machine learning techniques to predict all-cause mortality using fitness data: the Henry ford exercise testing (FIT) project. <i>BMC Medical Informatics and Decision Making</i> , 2017, 17, 174.	3.0	59
64	Invited Talks. , 2017, , .		0
65	On the Spectrum of Web Scale Data Management. , 2017, , 487-509.		1
66	Big Data 2.0 Processing Systems: Taxonomy and Open Challenges. <i>Journal of Grid Computing</i> , 2016, 14, 379-405.	3.9	46
67	Large-Scale Graph Processing Using Apache Giraph. , 2016, , .		15
68	Process Analytics. , 2016, , .		30
69	DREAM in Action. , 2016, , .		18
70	Tools, Use Cases, and Discussions. , 2016, , 135-150.		0
71	A distributed query execution engine of big attributed graphs. <i>SpringerPlus</i> , 2016, 5, 665.	1.2	1
72	On Analyzing the Impact of Authors and Their Collaboration Patterns in the Major Computer Algorithms Research Conferences. <i>Collnet Journal of Scientometrics and Information Management</i> , 2016, 10, 155-173.	0.8	1

#	ARTICLE	IF	CITATIONS
73	International conferences on computer system: Analysis of EuroSys, SOSP, and OSDI during 2006-2014. Collnet Journal of Scientometrics and Information Management, 2016, 10, 175-195.	0.8	1
74	Towards a Comprehensive Data Analytics Framework for Smart Healthcare Services. Big Data Research, 2016, 4, 44-58.	4.2	126
75	Network-based social coordination of business processes. Information Systems, 2016, 58, 56-74.	3.6	29
76	Building Pipelines for Heterogeneous Execution Environments for Big Data Processing. IEEE Software, 2016, 33, 60-67.	1.8	34
77	An Anti-Pattern-based Runtime Business Process Compliance Monitoring Framework. International Journal of Advanced Computer Science and Applications, 2016, 7, .	0.7	6
78	A Cloud-Based Platform for Democratizing and Socializing the Benchmarking Process. International Journal of Advanced Computer Science and Applications, 2016, 7, .	0.7	0
79	DREAM. Proceedings of the VLDB Endowment, 2015, 8, 654-665.	3.8	81
80	New generation of big data processing systems: Technologies, challenges and opportunities. , 2015, , .		0
81	Compliance Monitoring as a Service: Requirements, Architecture and Implementation. , 2015, , .		1
82	EnterpriseÂ2.0: Research Challenges and Opportunities. Lecture Notes in Business Information Processing, 2015, , 16-30.	1.0	3
83	Composable and efficient functional big data processing framework. , 2015, , .		3
84	Big Data Processing Systems: State-of-the-Art and Open Challenges. , 2015, , .		3
85	Runtime self-monitoring approach of business process compliance in cloud environments. Cluster Computing, 2015, 18, 1503-1526.	5.0	11
86	CDPort: A Portability Framework for NoSQL Datastores. Arabian Journal for Science and Engineering, 2015, 40, 2531-2553.	1.1	14
87	Large scale graph processing systems: survey and an experimental evaluation. Cluster Computing, 2015, 18, 1189-1213.	5.0	81
88	Runtime detection of business process compliance violations. , 2015, , .		20
89	Big Graph Processing Systems: State-of-the-Art and Open Challenges. , 2015, , .		6
90	A Framework for Consumer-Centric SLA Management of Cloud-Hosted Databases. IEEE Transactions on Services Computing, 2015, 8, 534-549.	4.6	41

#	ARTICLE	IF	CITATIONS
91	Towards an Extensible Middleware for Database Benchmarking. Lecture Notes in Computer Science, 2015, , 82-96.	1.3	12
92	A Framework of Enriching Business Processes Life-Cycle with Tagging Information. Lecture Notes in Computer Science, 2015, , 309-313.	1.3	7
93	SUPER: Social-Based Business Process Management Framework. Lecture Notes in Computer Science, 2015, , 413-417.	1.3	6
94	How to Make Business Processes "Socialize". EAI Endorsed Transactions on Industrial Networks and Intelligent Systems, 2015, 2, 150284.	1.9	3
95	CDPort. , 2014, , .		10
96	Cloud Data Management. , 2014, , .		24
97	Hybrid query execution engine for large attributed graphs. Information Systems, 2014, 41, 45-73.	3.6	8
98	Cloud-hosted databases: technologies, challenges and opportunities. Cluster Computing, 2014, 17, 487-502.	5.0	54
99	5th International Workshop on Graph Data Management: Techniques and applications (GDM 2014). , 2014, , .		0
100	SLA-Driven Database Replication on Virtualized Database Servers. , 2014, , 97-118.		1
101	Towards Comprehensive Measurement of Consistency Guarantees for Cloud-Hosted Data Storage Services. Lecture Notes in Computer Science, 2014, , 32-47.	1.3	16
102	MapReduce Family of Large-Scale Data-Processing Systems. , 2014, , 39-106.		5
103	The Family of Map-Reduce. , 2014, , 1-39.		1
104	Cloud-Hosted Data Storage Systems. , 2014, , 21-45.		0
105	Big Data Processing Systems. , 2014, , 135-176.		2
106	Is Your Cloud-Hosted Database Truly Elastic?. , 2013, , .		4
107	Incorporating Uncertainty into In-Cloud Application Deployment Decisions for Availability. , 2013, , .		2
108	Querying Process Models Repositories by Aggregated Graph Search. Lecture Notes in Business Information Processing, 2013, , 573-585.	1.0	10

#	ARTICLE	IF	CITATIONS
109	Improving Availability of Cloud-Based Applications through Deployment Choices. , 2013, , .		2
110	The family of mapreduce and large-scale data processing systems. ACM Computing Surveys, 2013, 46, 1-44.	23.0	127
111	Modeling performance of a parallel streaming engine. , 2013, , .		17
112	Consumer-centric SLA manager for cloud-hosted databases. , 2013, , .		5
113	Availability analysis for deployment of in-cloud applications. , 2013, , .		7
114	4th international workshop on graph data management: Techniques and application (GDM 2013) [front matter]. , 2013, , .		0
115	G-SPARQL. , 2012, , .		40
116	An architecture framework for application-managed scaling of cloud-hosted relational databases. , 2012, , .		1
117	GDM 2012 Workshop Introduction. , 2012, , .		0
118	AdaptRDF: adaptive storage management for RDF databases. International Journal of Web Information Systems, 2012, 8, 234-250.	2.4	9
119	Application-Managed Replication Controller for Cloud-Hosted Databases. , 2012, , .		6
120	On efficient processing of BPMN-Q queries. Computers in Industry, 2012, 63, 867-881.	9.9	18
121	Application-Managed Database Replication on Virtualized Cloud Environments. , 2012, , .		9
122	SLA-Based and Consumer-centric Dynamic Provisioning for Cloud Databases. , 2012, , .		42
123	On understanding the economics and elasticity challenges of deploying business applications on public cloud infrastructure. Journal of Internet Services and Applications, 2012, 3, 173-193.	2.1	49
124	A decade of database conferences: a look inside the program committees. Scientometrics, 2012, 91, 173-184.	3.0	7
125	Trade-Off Analysis of Elasticity Approaches for Cloud-Based Business Applications. Lecture Notes in Computer Science, 2012, , 468-482.	1.3	7
126	CloudDB AutoAdmin: Towards a Truly Elastic Cloud-Based Data Store. , 2011, , .		20



#	ARTICLE	IF	CITATIONS
127	MyDeepWeb: An Integration Service for Your OWN Deep Web Data. , 2011, , .		0
128	One Size Does Not Fit All: A Group-Based Service Selection for Web-Based Business Processes. , 2011, , .		1
129	A Survey of Large Scale Data Management Approaches in Cloud Environments. IEEE Communications Surveys and Tutorials, 2011, 13, 311-336.	39.4	323
130	Partial process models to manage business process variants. International Journal of Business Process Integration and Management, 2011, 5, 240.	0.0	9
131	A decade of database research publications: a look inside. Scientometrics, 2011, 88, 521-533.	3.0	5
132	Liquid benchmarks. , 2011, , .		1
133	Liquid Benchmarks: Towards an Online Platform for Collaborative Assessment of Computer Science Research Results. Lecture Notes in Computer Science, 2011, , 10-24.	1.3	4
134	On Maintaining Consistency of Process Model Variants. Lecture Notes in Business Information Processing, 2011, , 289-300.	1.0	12
135	A Query Language for Analyzing Business Processes Execution. Lecture Notes in Computer Science, 2011, , 281-297.	1.3	51
136	On Efficient Evaluation of XML Queries. , 2011, , 239-293.		0
137	Relational processing of RDF queries. SIGMOD Record, 2010, 38, 23-28.	1.2	100
138	Graph indexing and querying: a review. International Journal of Web Information Systems, 2010, 6, 101-120.	2.4	24
139	Towards a comprehensive assessment for selectivity estimation approaches of XML queries. International Journal of Web Engineering and Technology, 2010, 6, 58.	0.2	7
140	Efficient Relational Techniques for Processing Graph Queries. Journal of Computer Science and Technology, 2010, 25, 1237-1255.	1.5	11
141	An efficient features-based processing technique for supergraph queries. , 2010, , .		0
142	A framework for querying graph-based business process models. , 2010, , .		40
143	An Experimental Evaluation of Relational RDF Storage and Querying Techniques. Lecture Notes in Computer Science, 2010, , 215-226.	1.3	14
144	Querying Graph-Based Repositories of Business Process Models. Lecture Notes in Computer Science, 2010, , 33-44.	1.3	20

#	ARTICLE	IF	CITATIONS
145	Efficient and Adaptable Query Workload-Aware Management for RDF Data. Lecture Notes in Computer Science, 2010, , 390-399.	1.3	1
146	GDM2010 Workshop Organizersâ€™ Message. Lecture Notes in Computer Science, 2010, , 1-1.	1.3	0
147	XML compression techniques: A survey and comparison. Journal of Computer and System Sciences, 2009, 75, 303-322.	1.2	71
148	Storing and Querying Graph Data Using Efficient Relational Processing Techniques. Lecture Notes in Business Information Processing, 2009, , 379-392.	1.0	3
149	GraphREL: A Decomposition-Based and Selectivity-Aware Relational Framework for Processing Sub-graph Queries. Lecture Notes in Computer Science, 2009, , 123-137.	1.3	25
150	FeedRank: A Semantic-Based Management System of Web Feeds. Lecture Notes in Computer Science, 2009, , 126-133.	1.3	2
151	Cardinality-Aware Purely Relational XQuery Processor. Journal of Database Management, 2009, 20, 76-125.	1.5	2
152	An Empirical Evaluation of XML Compression Tools. Lecture Notes in Computer Science, 2009, , 49-63.	1.3	2
153	XML Tree Structure Compression. , 2008, , .		12
154	Algebraâ€based XQuery cardinality estimation. International Journal of Web Information Systems, 2008, 4, 6-47.	2.4	6
155	Dependable cardinality forecasts for XQuery. Proceedings of the VLDB Endowment, 2008, 1, 463-477.	3.8	20
156	Improving the Relational Evaluation of XML Queries by Means of Path Summaries. Lecture Notes in Computer Science, 2008, , 378-386.	1.3	1
157	XSelMark: A Micro-benchmark for Selectivity Estimation Approaches of XML Queries. Lecture Notes in Computer Science, 2008, , 735-744.	1.3	1
158	A SQL. , 2007, , .		19
159	Towards a Framework for Mapping Between UML/OCL and XML/XQuery. Lecture Notes in Computer Science, 2004, , 241-259.	1.3	4
160	XQuery on SQL Hosts. , 2004, , 252-263.		48
161	Large-Scale Data Management Techniques in Cloud Computing Platforms. , 0, , 85-123.		0
162	An Overview of Graph Indexing and Querying Techniques. , 0, , 222-239.		1

#	ARTICLE	IF	CITATIONS
163	Relational Techniques for Storing and Querying RDF Data. Advances in Data Mining and Database Management Book Series, 0, , 269-285.	0.5	0
164	Querying Graph Databases. Advances in Data Mining and Database Management Book Series, 0, , 304-322.	0.5	0
165	An Overview of Graph Indexing and Querying Techniques. Advances in Data Mining and Database Management Book Series, 0, , 71-88.	0.5	3