

Marta Mattoso

List of Publications by Year in descending order

Source: <https://exaly.com/author-pdf/2811662/publications.pdf>

Version: 2024-02-01

154
papers

2,149
citations

361296

20
h-index

377752

34
g-index

162
all docs

162
docs citations

162
times ranked

1330
citing authors

#	ARTICLE	IF	CITATIONS
1	A Survey of Data-Intensive Scientific Workflow Management. Journal of Grid Computing, 2015, 13, 457-493.	2.5	202
2	SciCumulus: A Lightweight Cloud Middleware to Explore Many Task Computing Paradigm in Scientific Workflows. , 2010, , .		89
3	A Provenance-based Adaptive Scheduling Heuristic for Parallel Scientific Workflows in Clouds. Journal of Grid Computing, 2012, 10, 521-552.	2.5	79
4	Towards supporting the life cycle of large scale scientific experiments. International Journal of Business Process Integration and Management, 2010, 5, 79.	0.2	75
5	Adaptive Normalization: A novel data normalization approach for non-stationary time series. , 2010, , .		74
6	An algebraic approach for data-centric scientific workflows. Proceedings of the VLDB Endowment, 2011, 4, 1328-1339.	2.1	56
7	Towards a Taxonomy of Provenance in Scientific Workflow Management Systems. , 2009, , .		50
8	Building Reliable Web Services Compositions. Lecture Notes in Computer Science, 2003, , 59-72.	1.0	48
9	Dynamic steering of HPC scientific workflows: A survey. Future Generation Computer Systems, 2015, 46, 100-113.	4.9	46
10	Multi-objective scheduling of Scientific Workflows in multisite clouds. Future Generation Computer Systems, 2016, 63, 76-95.	4.9	46
11	Grid Data Management: Open Problems and New Issues. Journal of Grid Computing, 2007, 5, 273-281.	2.5	43
12	Chiron: a parallel engine for algebraic scientific workflows. Concurrency Computation Practice and Experience, 2013, 25, 2327-2341.	1.4	43
13	Capturing and querying workflow runtime provenance with PROV. , 2013, , .		43
14	Parallel OLAP query processing in database clusters with data replication. Distributed and Parallel Databases, 2009, 25, 97-123.	1.0	41
15	Towards a Taxonomy for Cloud Computing from an e-Science Perspective. Computer Communications and Networks, 2010, , 47-62.	0.8	34
16	A new genomic taxonomy system for the <i>Synechococcus</i> collective. Environmental Microbiology, 2020, 22, 4557-4570.	1.8	32
17	An adaptive parallel execution strategy for cloud-based scientific workflows. Concurrency Computation Practice and Experience, 2012, 24, 1531-1550.	1.4	31
18	BaMBa: towards the integrated management of Brazilian marine environmental data. Database: the Journal of Biological Databases and Curation, 2015, 2015, bav088.	1.4	30

#	ARTICLE	IF	CITATIONS
19	The use of mediation and ontology technologies for software component information retrieval. , 2001, , .		26
20	SciPhy: A Cloud-Based Workflow for Phylogenetic Analysis of Drug Targets in Protozoan Genomes. Lecture Notes in Computer Science, 2011, , 66-70.	1.0	26
21	Exploring many task computing in scientific workflows. , 2009, , .		24
22	ProvManager: a provenance management system for scientific workflows. Concurrency Computation Practice and Experience, 2012, 24, 1513-1530.	1.4	24
23	Performance evaluation of parallel strategies in public clouds: A study with phylogenomic workflows. Future Generation Computer Systems, 2013, 29, 1816-1825.	4.9	24
24	In situ visualization and data analysis for turbidity currents simulation. Computers and Geosciences, 2018, 110, 23-31.	2.0	24
25	Using ontologies for domain information retrieval. , 0, , .		22
26	MTCProv: a practical provenance query framework for many-task scientific computing. Distributed and Parallel Databases, 2012, 30, 351-370.	1.0	21
27	Data-centric iteration in dynamic workflows. Future Generation Computer Systems, 2015, 46, 114-126.	4.9	20
28	Raw data queries during data-intensive parallel workflow execution. Future Generation Computer Systems, 2017, 75, 402-422.	4.9	20
29	Managing structural genomic workflows using Web services. Data and Knowledge Engineering, 2005, 53, 45-74.	2.1	19
30	OLAP Query Processing in a Database Cluster. Lecture Notes in Computer Science, 2004, , 355-362.	1.0	18
31	Odyssey-Search: A multi-agent system for component information search and retrieval. Journal of Systems and Software, 2006, 79, 204-215.	3.3	18
32	Automatic composition of Web services with contingency plans. , 2004, , .		17
33	Planning spatial workflows to optimize grid performance. , 2006, , .		17
34	ProtozoaDB: dynamic visualization and exploration of protozoan genomes. Nucleic Acids Research, 2007, 36, D547-D552.	6.5	17
35	Provenance management in Swift. Future Generation Computer Systems, 2011, 27, 775-780.	4.9	17
36	A Distribution Design Methodology for Object DBMS. Distributed and Parallel Databases, 2004, 16, 45-90.	1.0	15

#	ARTICLE	IF	CITATIONS
37	Kairos: An Architecture for Securing Authorship and Temporal Information of Provenance Data in Grid-Enabled Workflow Management Systems. , 2008, , .		15
38	Provenance Services for Distributed Workflows. , 2008, , .		15
39	Data parallelism in bioinformatics workflows using Hydra. , 2010, , .		15
40	Optimizing Phylogenetic Analysis Using SciHm Cloud-based Scientific Workflow. , 2011, , .		15
41	Supporting dynamic parameter sweep in adaptive and user-steered workflow. , 2011, , .		15
42	Efficient Scheduling of Scientific Workflows Using Hot Metadata in a Multisite Cloud. IEEE Transactions on Knowledge and Data Engineering, 2019, 31, 1940-1953.	4.0	15
43	Exploring Molecular Evolution Reconstruction Using a Parallel Cloud Based Scientific Workflow. Lecture Notes in Computer Science, 2012, , 179-191.	1.0	15
44	Dfanalyzer. Proceedings of the VLDB Endowment, 2018, 11, 2082-2085.	2.1	15
45	User-steering of HPC workflows. , 2013, , .		14
46	Provenance Data in the Machine Learning Lifecycle in Computational Science and Engineering. , 2019, , .		14
47	Algebraic dataflows for big data analysis. , 2013, , .		13
48	Managing structural genomic workflows using Web services. Data and Knowledge Engineering, 2005, 53, 45-74.	2.1	13
49	Towards a Cost Model for Scheduling Scientific Workflows Activities in Cloud Environments. , 2011, , .		12
50	UNCERTAINTY QUANTIFICATION IN COMPUTATIONAL PREDICTIVE MODELS FOR FLUID DYNAMICS USING A WORKFLOW MANAGEMENT ENGINE. , 2012, 2, 53-71.		12
51	Designing a parallel cloud based comparative genomics workflow to improve phylogenetic analyses. Future Generation Computer Systems, 2013, 29, 2205-2219.	4.9	12
52	Workflow provenance in the lifecycle of scientific machine learning. Concurrency Computation Practice and Experience, 2022, 34, e6544.	1.4	12
53	Efficiently Processing XML Queries over Fragmented Repositories with PartiX. Lecture Notes in Computer Science, 2006, , 150-163.	1.0	12
54	Parallel query processing for OLAP in grids. Concurrency Computation Practice and Experience, 2008, 20, 2039-2048.	1.4	11

#	ARTICLE	IF	CITATIONS
55	Dimensioning the virtual cluster for parallel scientific workflows in clouds. , 2013, , .		11
56	Exploring Large Scale Receptor-Ligand Pairs in Molecular Docking Workflows in HPC Clouds. , 2014, , .		11
57	Capturing and Analyzing Provenance from Spark-based Scientific Workflows with SAMbA-RaP. Future Generation Computer Systems, 2020, 112, 658-669.	4.9	11
58	BioWorkbench: a high-performance framework for managing and analyzing bioinformatics experiments. PeerJ, 2018, 6, e5551.	0.9	11
59	Comparison and versioning of scientific workflows. , 2009, , .		10
60	A Performance Evaluation of X-Ray Crystallography Scientific Workflow Using SciCumulus. , 2011, , .		10
61	Analyzing related raw data files through dataflows. Concurrency Computation Practice and Experience, 2016, 28, 2528-2545.	1.4	10
62	Keeping track of user steering actions in dynamic workflows. Future Generation Computer Systems, 2019, 99, 624-643.	4.9	10
63	Data reduction in scientific workflows using provenance monitoring and user steering. Future Generation Computer Systems, 2020, 110, 481-501.	4.9	10
64	OdysseyShare: an environment for collaborative component-based development. , 0, , .		9
65	Parallelism in Bioinformatics Workflows. Lecture Notes in Computer Science, 2005, , 583-597.	1.0	9
66	XCraft. , 2008, , .		9
67	Evaluating parameter sweep workflows in high performance computing. , 2012, , .		9
68	DfAnalyzer: Runtime dataflow analysis tool for Computational Science and Engineering applications. SoftwareX, 2020, 12, 100592.	1.2	9
69	Mediating heterogeneous Web services. , 0, , .		8
70	A Strategy for Provenance Gathering in Distributed Scientific Workflows. , 2009, , .		8
71	Neural networks cartridges for data mining on time series. , 2009, , .		8
72	Many task computing for orthologous genes identification in protozoan genomes using Hydra. Concurrency Computation Practice and Experience, 2011, 23, 2326-2337.	1.4	8

#	ARTICLE	IF	CITATIONS
73	An Orthology-Based Analysis of Pathogenic Protozoa Impacting Global Health: An Improved Comparative Genomics Approach with Prokaryotes and Model Eukaryote Orthologs. OMICS A Journal of Integrative Biology, 2014, 18, 524-538.	1.0	8
74	A Practical Roadmap for Provenance Capture and Data Analysis in Spark-Based Scientific Workflows. , 2018, , .		8
75	A workflow for seismic imaging with quantified uncertainty. Computers and Geosciences, 2020, 145, 104615.	2.0	8
76	High-Performance Query Processing of a Real-World OLAP Database with ParGRES. Lecture Notes in Computer Science, 2008, , 188-200.	1.0	8
77	Experiment Line: Software Reuse in Scientific Workflows. Lecture Notes in Computer Science, 2009, , 264-272.	1.0	8
78	Using Domain-Specific Data to Enhance Scientific Workflow Steering Queries. Lecture Notes in Computer Science, 2012, , 152-167.	1.0	8
79	A Survey on XML Fragmentation. SIGMOD Record, 2014, 43, 24-35.	0.7	8
80	A Lightweight Middleware Monitor for Distributed Scientific Workflows. , 2008, , .		7
81	Efficient Runtime Capture of Multiworkflow Data Using Provenance. , 2019, , .		7
82	Using Explicit Control Processes in Distributed Workflows to Gather Provenance. Lecture Notes in Computer Science, 2008, , 186-199.	1.0	7
83	Enabling Re-executions of Parallel Scientific Workflows Using Runtime Provenance Data. Lecture Notes in Computer Science, 2012, , 229-232.	1.0	7
84	Structural genomic workflows supported by Web services. , 0, , .		6
85	Applying Theory Revision to the Design of Distributed Databases. Lecture Notes in Computer Science, 2003, , 57-74.	1.0	6
86	OrthoSearch. , 2008, , .		6
87	Handling Failures in Parallel Scientific Workflows Using Clouds. , 2012, , .		6
88	Towards an Adaptive and Distributed Architecture for Managing Workflow Provenance Data. , 2014, , .		6
89	Data Analytics in Bioinformatics: Data Science in Practice for Genomics Analysis Workflows. , 2015, , .		6
90	Managing hot metadata for scientific workflows on multisite clouds. , 2016, , .		6

#	ARTICLE	IF	CITATIONS
91	Uncertainty quantification in numerical simulation of particle-laden flows. Computational Geosciences, 2016, 20, 265-281.	1.2	6
92	An Opportunistic Algorithm for Scheduling Workflows on Grids. , 2006, , 1-12.		6
93	Webcomposer: a tool for the composition and execution of web service-based workflows. , 0, , .		5
94	Applying reinforcement learning to scheduling strategies in an actual grid environment. International Journal of High Performance Systems Architecture, 2009, 2, 116.	0.2	5
95	Detecting distant homologies on protozoans metabolic pathways using scientific workflows. International Journal of Data Mining and Bioinformatics, 2010, 4, 256.	0.1	5
96	Exploring provenance in high performance scientific computing. , 2011, , .		5
97	Discovering drug targets for neglected diseases using a pharmacophylogenomic cloud workflow. , 2012, , .		5
98	Runtime Dynamic Structural Changes of Scientific Workflows in Clouds. , 2013, , .		5
99	Exploratory Analysis of Raw Data Files through Dataflows. , 2014, , .		5
100	STINGRAY: system for integrated genomic resources and analysis. BMC Research Notes, 2014, 7, 132.	0.6	5
101	Scientific Workflow Scheduling with Provenance Data in a Multisite Cloud. Lecture Notes in Computer Science, 2017, , 80-112.	1.0	5
102	Mining a large database with a parallel database server. Intelligent Data Analysis, 1999, 3, 437-451.	0.4	4
103	ARAXA: Storing and managing Active XML documents. Web Semantics, 2010, 8, 209-224.	2.2	4
104	SciLightning: A Cloud Provenance-Based Event Notification for Parallel Workflows. Lecture Notes in Computer Science, 2014, , 352-365.	1.0	4
105	Improving workflow design by mining reusable tasks. Journal of the Brazilian Computer Society, 2015, 21, .	0.8	4
106	Capturing Provenance for Runtime Data Analysis in Computational Science and Engineering Applications. Lecture Notes in Computer Science, 2018, , 183-187.	1.0	4
107	Provenance of Dynamic Adaptations in User-Steered Dataflows. Lecture Notes in Computer Science, 2018, , 16-29.	1.0	4
108	GExpLine: A Tool for Supporting Experiment Composition. Lecture Notes in Computer Science, 2010, , 251-259.	1.0	4

#	ARTICLE	IF	CITATIONS
109	Athena: Text Mining Based Discovery of Scientific Workflows in Disperse Repositories. Lecture Notes in Computer Science, 2012, , 104-121.	1.0	4
110	An architecture for managing distributed scientific resources. , 0, , .		3
111	RL-Based Scheduling Strategies in Actual Grid Environments. , 2008, , .		3
112	Adaptive hybrid partitioning for OLAP query processing in a database cluster. International Journal of High Performance Computing and Networking, 2008, 5, 251.	0.4	3
113	Provenance traces from Chiron parallel workflow engine. , 2013, , .		3
114	Provenance traces of the swift parallel scripting system. , 2013, , .		3
115	Deriving scientific workflows from algebraic experiment lines: A practical approach. Future Generation Computer Systems, 2017, 68, 111-127.	4.9	3
116	Provenance Supporting Hyperparameter Analysis in Deep Neural Networks. Lecture Notes in Computer Science, 2021, , 20-38.	1.0	3
117	An Environment to Define and Execute In-Silico Workflows Using Web Services. Lecture Notes in Computer Science, 2005, , 288-291.	1.0	3
118	Apuama: Combining Intra-query and Inter-query Parallelism in a Database Cluster. Lecture Notes in Computer Science, 2006, , 649-661.	1.0	3
119	Estimating Costs of Path Expression Evaluation in Distributed Object Databases. Lecture Notes in Computer Science, 2002, , 351-360.	1.0	3
120	Towards Supporting Provenance Gathering and Querying in Different Database Approaches. Lecture Notes in Computer Science, 2015, , 254-257.	1.0	3
121	Clouds and Reproducibility: A Way to Go to Scientific Experiments?. Computer Communications and Networks, 2017, , 127-151.	0.8	3
122	Scientific Workflow Scheduling with Provenance Support in Multisite Cloud. Lecture Notes in Computer Science, 2017, , 206-219.	1.0	3
123	Ariane: An Awareness Mechanism for Shared Databases. Lecture Notes in Computer Science, 2004, , 92-104.	1.0	3
124	Software components retrieval through mediators and web search. Journal of the Brazilian Computer Society, 2002, 8, 55-63.	0.8	3
125	Adding domain data to code profiling tools to debug workflow parallel execution. Future Generation Computer Systems, 2020, 110, 422-439.	4.9	2
126	Distributed in-memory data management for workflow executions. PeerJ Computer Science, 2021, 7, e527.	2.7	2

#	ARTICLE	IF	CITATIONS
127	Análise de Hiperparâmetros em Aplicações de Aprendizado Profundo por meio de Dados de Proveniência. , 0, , .		2
128	A parallel spatial join framework using PMR-quadtrees. , 0, , .		1
129	Digging Database Statistics and Costs Parameters for Distributed Query Processing. Lecture Notes in Computer Science, 2003, , 301-318.	1.0	1
130	Preface to the Special Issue on Grid Data Management. Journal of Grid Computing, 2007, 5, 271-272.	2.5	1
131	Using Ontologies to Support Deep Water Oil Exploration Scientific Workflows. , 2009, , .		1
132	Improving Many-Task computing in scientific workflows using P2P techniques. , 2010, , .		1
133	Applying Provenance to Protect Attribution in Distributed Computational Scientific Experiments. Lecture Notes in Computer Science, 2015, , 139-151.	1.0	1
134	Towards a Theory Revision Approach for the Vertical Fragmentation of Object Oriented Databases. Lecture Notes in Computer Science, 2002, , 216-226.	1.0	1
135	On the Usage of Structural Information in Constrained Semi-Supervised Clustering of XML Documents. , 2008, , 67-86.		1
136	Virtual Partitioning. , 2009, , 3340-3341.		1
137	A P2P Approach to Many Tasks Computing for Scientific Workflows. Lecture Notes in Computer Science, 2011, , 327-339.	1.0	1
138	Exploiting the Parallel Execution of Homology Workflow Alternatives in HPC Compute Clouds. Lecture Notes in Computer Science, 2015, , 336-350.	1.0	1
139	Provenance in Workflows. , 2018, , 2912-2916.		1
140	BioProv - A provenance library for bioinformatics workflows. Journal of Open Source Software, 2021, 6, 3622.	2.0	1
141	Mechanisms for specifying communication behavior in object oriented database systems. , 2000, , .		0
142	On the performance of the position() XPath function. , 2013, , .		0
143	Database Clusters. , 2009, , 700-704.		0
144	ARAXA: Storing and Managing Active XML Documents. SSRN Electronic Journal, 0, , .	0.4	0

#	ARTICLE	IF	CITATIONS
145	Towards a Threat Model for Provenance in e-Science. Lecture Notes in Computer Science, 2010, , 277-279.	1.0	0
146	A Provenance-Based Approach to Resource Discovery in Distributed Molecular Dynamics Workflows. Lecture Notes in Computer Science, 2010, , 66-80.	1.0	0
147	Provenance in Workflows. , 2017, , 1-5.		0
148	Database Clusters. , 2017, , 1-5.		0
149	Virtual Partitioning. , 2017, , 1-2.		0
150	Enhancing Energy Production with Exascale HPC Methods. Communications in Computer and Information Science, 2017, , 233-246.	0.4	0
151	Virtual Partitioning. , 2018, , 4462-4463.		0
152	Database Clusters. , 2018, , 940-944.		0
153	Experiencing DfAnalyzer for Runtime Analysis of Phylogenomic Dataflows. Lecture Notes in Computer Science, 2020, , 105-116.	1.0	0
154	A horizontal partitioning-based method for frequent pattern mining in transport timetable. Expert Systems, 0, , e12881.	2.9	0