

# 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	Workflow provenance in the lifecycle of scientific machine learning. <i>Concurrency Computation Practice and Experience</i> , 2022, 34, e6544.	1.4	12
2	Distributed in-memory data management for workflow executions. <i>PeerJ Computer Science</i> , 2021, 7, e527.	2.7	2
3	Provenance Supporting Hyperparameter Analysis in Deep Neural Networks. <i>Lecture Notes in Computer Science</i> , 2021, , 20-38.	1.0	3
4	BioProv - A provenance library for bioinformatics workflows. <i>Journal of Open Source Software</i> , 2021, 6, 3622.	2.0	1
5	Data reduction in scientific workflows using provenance monitoring and user steering. <i>Future Generation Computer Systems</i> , 2020, 110, 481-501.	4.9	10
6	Adding domain data to code profiling tools to debug workflow parallel execution. <i>Future Generation Computer Systems</i> , 2020, 110, 422-439.	4.9	2
7	A workflow for seismic imaging with quantified uncertainty. <i>Computers and Geosciences</i> , 2020, 145, 104615.	2.0	8
8	A new genomic taxonomy system for the <i>Synechococcus</i> collective. <i>Environmental Microbiology</i> , 2020, 22, 4557-4570.	1.8	32
9	DfAnalyzer: Runtime dataflow analysis tool for Computational Science and Engineering applications. <i>SoftwareX</i> , 2020, 12, 100592.	1.2	9
10	Capturing and Analyzing Provenance from Spark-based Scientific Workflows with SAMbA-RaP. <i>Future Generation Computer Systems</i> , 2020, 112, 658-669.	4.9	11
11	Experiencing DfAnalyzer for Runtime Analysis of Phylogenomic Dataflows. <i>Lecture Notes in Computer Science</i> , 2020, , 105-116.	1.0	0
12	Keeping track of user steering actions in dynamic workflows. <i>Future Generation Computer Systems</i> , 2019, 99, 624-643.	4.9	10
13	Provenance Data in the Machine Learning Lifecycle in Computational Science and Engineering. , 2019, , .		14
14	Efficient Runtime Capture of Multiworkflow Data Using Provenance. , 2019, , .		7
15	Efficient Scheduling of Scientific Workflows Using Hot Metadata in a Multisite Cloud. <i>IEEE Transactions on Knowledge and Data Engineering</i> , 2019, 31, 1940-1953.	4.0	15
16	In situ visualization and data analysis for turbidity currents simulation. <i>Computers and Geosciences</i> , 2018, 110, 23-31.	2.0	24
17	A Practical Roadmap for Provenance Capture and Data Analysis in Spark-Based Scientific Workflows. , 2018, , .		8
18	Capturing Provenance for Runtime Data Analysis in Computational Science and Engineering Applications. <i>Lecture Notes in Computer Science</i> , 2018, , 183-187.	1.0	4

#	ARTICLE	IF	CITATIONS
19	Provenance of Dynamic Adaptations in User-Steered Dataflows. Lecture Notes in Computer Science, 2018, , 16-29.	1.0	4
20	Dfanalyzer. Proceedings of the VLDB Endowment, 2018, 11, 2082-2085.	2.1	15
21	BioWorkbench: a high-performance framework for managing and analyzing bioinformatics experiments. PeerJ, 2018, 6, e5551.	0.9	11
22	Provenance in Workflows. , 2018, , 2912-2916.		1
23	Virtual Partitioning. , 2018, , 4462-4463.		0
24	Database Clusters. , 2018, , 940-944.		0
25	Raw data queries during data-intensive parallel workflow execution. Future Generation Computer Systems, 2017, 75, 402-422.	4.9	20
26	Deriving scientific workflows from algebraic experiment lines: A practical approach. Future Generation Computer Systems, 2017, 68, 111-127.	4.9	3
27	Clouds and Reproducibility: A Way to Go to Scientific Experiments?. Computer Communications and Networks, 2017, , 127-151.	0.8	3
28	Scientific Workflow Scheduling with Provenance Support in Multisite Cloud. Lecture Notes in Computer Science, 2017, , 206-219.	1.0	3
29	Scientific Workflow Scheduling with Provenance Data in a Multisite Cloud. Lecture Notes in Computer Science, 2017, , 80-112.	1.0	5
30	Provenance in Workflows. , 2017, , 1-5.		0
31	Database Clusters. , 2017, , 1-5.		0
32	Virtual Partitioning. , 2017, , 1-2.		0
33	Enhancing Energy Production with Exascale HPC Methods. Communications in Computer and Information Science, 2017, , 233-246.	0.4	0
34	Analyzing related raw data files through dataflows. Concurrency Computation Practice and Experience, 2016, 28, 2528-2545.	1.4	10
35	Managing hot metadata for scientific workflows on multisite clouds. , 2016, , .		6
36	Uncertainty quantification in numerical simulation of particle-laden flows. Computational Geosciences, 2016, 20, 265-281.	1.2	6

#	ARTICLE	IF	CITATIONS
37	Multi-objective scheduling of Scientific Workflows in multisite clouds. Future Generation Computer Systems, 2016, 63, 76-95.	4.9	46
38	BaMBa: towards the integrated management of Brazilian marine environmental data. Database: the Journal of Biological Databases and Curation, 2015, 2015, bav088.	1.4	30
39	Improving workflow design by mining reusable tasks. Journal of the Brazilian Computer Society, 2015, 21, .	0.8	4
40	Data-centric iteration in dynamic workflows. Future Generation Computer Systems, 2015, 46, 114-126.	4.9	20
41	Dynamic steering of HPC scientific workflows: A survey. Future Generation Computer Systems, 2015, 46, 100-113.	4.9	46
42	Data Analytics in Bioinformatics: Data Science in Practice for Genomics Analysis Workflows. , 2015, , .		6
43	A Survey of Data-Intensive Scientific Workflow Management. Journal of Grid Computing, 2015, 13, 457-493.	2.5	202
44	Applying Provenance to Protect Attribution in Distributed Computational Scientific Experiments. Lecture Notes in Computer Science, 2015, , 139-151.	1.0	1
45	Towards Supporting Provenance Gathering and Querying in Different Database Approaches. Lecture Notes in Computer Science, 2015, , 254-257.	1.0	3
46	Exploiting the Parallel Execution of Homology Workflow Alternatives in HPC Compute Clouds. Lecture Notes in Computer Science, 2015, , 336-350.	1.0	1
47	Exploratory Analysis of Raw Data Files through Dataflows. , 2014, , .		5
48	Towards an Adaptive and Distributed Architecture for Managing Workflow Provenance Data. , 2014, , .		6
49	Exploring Large Scale Receptor-Ligand Pairs in Molecular Docking Workflows in HPC Clouds. , 2014, , .		11
50	SciLightning: A Cloud Provenance-Based Event Notification for Parallel Workflows. Lecture Notes in Computer Science, 2014, , 352-365.	1.0	4
51	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
52	STINGRAY: system for integrated genomic resources and analysis. BMC Research Notes, 2014, 7, 132.	0.6	5
53	A Survey on XML Fragmentation. SIGMOD Record, 2014, 43, 24-35.	0.7	8
54	Dimensioning the virtual cluster for parallel scientific workflows in clouds. , 2013, , .		11

#	ARTICLE	IF	CITATIONS
55	Performance evaluation of parallel strategies in public clouds: A study with phylogenomic workflows. <i>Future Generation Computer Systems</i> , 2013, 29, 1816-1825.	4.9	24
56	Designing a parallel cloud based comparative genomics workflow to improve phylogenetic analyses. <i>Future Generation Computer Systems</i> , 2013, 29, 2205-2219.	4.9	12
57	Chiron: a parallel engine for algebraic scientific workflows. <i>Concurrency Computation Practice and Experience</i> , 2013, 25, 2327-2341.	1.4	43
58	Runtime Dynamic Structural Changes of Scientific Workflows in Clouds. , 2013, , .		5
59	User-steering of HPC workflows. , 2013, , .		14
60	Capturing and querying workflow runtime provenance with PROV. , 2013, , .		43
61	Provenance traces from Chiron parallel workflow engine. , 2013, , .		3
62	Provenance traces of the swift parallel scripting system. , 2013, , .		3
63	On the performance of the position() XPath function. , 2013, , .		0
64	Algebraic dataflows for big data analysis. , 2013, , .		13
65	Handling Failures in Parallel Scientific Workflows Using Clouds. , 2012, , .		6
66	Evaluating parameter sweep workflows in high performance computing. , 2012, , .		9
67	MTCProv: a practical provenance query framework for many-task scientific computing. <i>Distributed and Parallel Databases</i> , 2012, 30, 351-370.	1.0	21
68	A Provenance-based Adaptive Scheduling Heuristic for Parallel Scientific Workflows in Clouds. <i>Journal of Grid Computing</i> , 2012, 10, 521-552.	2.5	79
69	Discovering drug targets for neglected diseases using a pharmacophylogenomic cloud workflow. , 2012, , .		5
70	UNCERTAINTY QUANTIFICATION IN COMPUTATIONAL PREDICTIVE MODELS FOR FLUID DYNAMICS USING A WORKFLOW MANAGEMENT ENGINE. , 2012, 2, 53-71.		12
71	ProvManager: a provenance management system for scientific workflows. <i>Concurrency Computation Practice and Experience</i> , 2012, 24, 1513-1530.	1.4	24
72	An adaptive parallel execution strategy for cloud-based scientific workflows. <i>Concurrency Computation Practice and Experience</i> , 2012, 24, 1531-1550.	1.4	31

#	ARTICLE	IF	CITATIONS
73	Athena: Text Mining Based Discovery of Scientific Workflows in Disperse Repositories. Lecture Notes in Computer Science, 2012, , 104-121.	1.0	4
74	Exploring Molecular Evolution Reconstruction Using a Parallel Cloud Based Scientific Workflow. Lecture Notes in Computer Science, 2012, , 179-191.	1.0	15
75	Using Domain-Specific Data to Enhance Scientific Workflow Steering Queries. Lecture Notes in Computer Science, 2012, , 152-167.	1.0	8
76	Enabling Re-executions of Parallel Scientific Workflows Using Runtime Provenance Data. Lecture Notes in Computer Science, 2012, , 229-232.	1.0	7
77	Optimizing Phylogenetic Analysis Using SciHm Cloud-based Scientific Workflow. , 2011, , .		15
78	A Performance Evaluation of X-Ray Crystallography Scientific Workflow Using SciCumulus. , 2011, , .		10
79	Towards a Cost Model for Scheduling Scientific Workflows Activities in Cloud Environments. , 2011, , .		12
80	Supporting dynamic parameter sweep in adaptive and user-steered workflow. , 2011, , .		15
81	Provenance management in Swift. Future Generation Computer Systems, 2011, 27, 775-780.	4.9	17
82	Many task computing for orthologous genes identification in protozoan genomes using Hydra. Concurrency Computation Practice and Experience, 2011, 23, 2326-2337.	1.4	8
83	Exploring provenance in high performance scientific computing. , 2011, , .		5
84	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
85	An algebraic approach for data-centric scientific workflows. Proceedings of the VLDB Endowment, 2011, 4, 1328-1339.	2.1	56
86	A P2P Approach to Many Tasks Computing for Scientific Workflows. Lecture Notes in Computer Science, 2011, , 327-339.	1.0	1
87	Detecting distant homologies on protozoans metabolic pathways using scientific workflows. International Journal of Data Mining and Bioinformatics, 2010, 4, 256.	0.1	5
88	ARAXA: Storing and managing Active XML documents. Web Semantics, 2010, 8, 209-224.	2.2	4
89	Adaptive Normalization: A novel data normalization approach for non-stationary time series. , 2010, , .		74
90	Data parallelism in bioinformatics workflows using Hydra. , 2010, , .		15

#	ARTICLE	IF	CITATIONS
91	Towards supporting the life cycle of large scale scientific experiments. International Journal of Business Process Integration and Management, 2010, 5, 79.	0.2	75
92	SciCumulus: A Lightweight Cloud Middleware to Explore Many Task Computing Paradigm in Scientific Workflows. , 2010, , .		89
93	Improving Many-Task computing in scientific workflows using P2P techniques. , 2010, , .		1
94	Towards a Taxonomy for Cloud Computing from an e-Science Perspective. Computer Communications and Networks, 2010, , 47-62.	0.8	34
95	GExpLine: A Tool for Supporting Experiment Composition. Lecture Notes in Computer Science, 2010, , 251-259.	1.0	4
96	Towards a Threat Model for Provenance in e-Science. Lecture Notes in Computer Science, 2010, , 277-279.	1.0	0
97	A Provenance-Based Approach to Resource Discovery in Distributed Molecular Dynamics Workflows. Lecture Notes in Computer Science, 2010, , 66-80.	1.0	0
98	Using Ontologies to Support Deep Water Oil Exploration Scientific Workflows. , 2009, , .		1
99	Parallel OLAP query processing in database clusters with data replication. Distributed and Parallel Databases, 2009, 25, 97-123.	1.0	41
100	Exploring many task computing in scientific workflows. , 2009, , .		24
101	A Strategy for Provenance Gathering in Distributed Scientific Workflows. , 2009, , .		8
102	Towards a Taxonomy of Provenance in Scientific Workflow Management Systems. , 2009, , .		50
103	Neural networks cartridges for data mining on time series. , 2009, , .		8
104	Comparison and versioning of scientific workflows. , 2009, , .		10
105	Applying reinforcement learning to scheduling strategies in an actual grid environment. International Journal of High Performance Systems Architecture, 2009, 2, 116.	0.2	5
106	Experiment Line: Software Reuse in Scientific Workflows. Lecture Notes in Computer Science, 2009, , 264-272.	1.0	8
107	Database Clusters. , 2009, , 700-704.		0
108	Virtual Partitioning. , 2009, , 3340-3341.		1

#	ARTICLE	IF	CITATIONS
109	Parallel query processing for OLAP in grids. <i>Concurrency Computation Practice and Experience</i> , 2008, 20, 2039-2048.	1.4	11
110	RL-Based Scheduling Strategies in Actual Grid Environments. , 2008, , .		3
111	Kairos: An Architecture for Securing Authorship and Temporal Information of Provenance Data in Grid-Enabled Workflow Management Systems. , 2008, , .		15
112	XCraft. , 2008, , .		9
113	OrthoSearch. , 2008, , .		6
114	Provenance Services for Distributed Workflows. , 2008, , .		15
115	A Lightweight Middleware Monitor for Distributed Scientific Workflows. , 2008, , .		7
116	Adaptive hybrid partitioning for OLAP query processing in a database cluster. <i>International Journal of High Performance Computing and Networking</i> , 2008, 5, 251.	0.4	3
117	Using Explicit Control Processes in Distributed Workflows to Gather Provenance. <i>Lecture Notes in Computer Science</i> , 2008, , 186-199.	1.0	7
118	High-Performance Query Processing of a Real-World OLAP Database with ParGRES. <i>Lecture Notes in Computer Science</i> , 2008, , 188-200.	1.0	8
119	On the Usage of Structural Information in Constrained Semi-Supervised Clustering of XML Documents. , 2008, , 67-86.		1
120	ProtozoaDB: dynamic visualization and exploration of protozoan genomes. <i>Nucleic Acids Research</i> , 2007, 36, D547-D552.	6.5	17
121	Preface to the Special Issue on Grid Data Management. <i>Journal of Grid Computing</i> , 2007, 5, 271-272.	2.5	1
122	Grid Data Management: Open Problems and New Issues. <i>Journal of Grid Computing</i> , 2007, 5, 273-281.	2.5	43
123	Odyssey-Search: A multi-agent system for component information search and retrieval. <i>Journal of Systems and Software</i> , 2006, 79, 204-215.	3.3	18
124	Planning spatial workflows to optimize grid performance. , 2006, , .		17
125	Efficiently Processing XML Queries over Fragmented Repositories with PartiX. <i>Lecture Notes in Computer Science</i> , 2006, , 150-163.	1.0	12
126	Apuama: Combining Intra-query and Inter-query Parallelism in a Database Cluster. <i>Lecture Notes in Computer Science</i> , 2006, , 649-661.	1.0	3



#	ARTICLE	IF	CITATIONS
127	An Opportunistic Algorithm for Scheduling Workflows on Grids. , 2006, , 1-12.		6
128	Managing structural genomic workflows using Web services. Data and Knowledge Engineering, 2005, 53, 45-74.	2.1	19
129	Parallelism in Bioinformatics Workflows. Lecture Notes in Computer Science, 2005, , 583-597.	1.0	9
130	An Environment to Define and Execute In-Silico Workflows Using Web Services. Lecture Notes in Computer Science, 2005, , 288-291.	1.0	3
131	Managing structural genomic workflows using Web services. Data and Knowledge Engineering, 2005, 53, 45-74.	2.1	13
132	A Distribution Design Methodology for Object DBMS. Distributed and Parallel Databases, 2004, 16, 45-90.	1.0	15
133	OLAP Query Processing in a Database Cluster. Lecture Notes in Computer Science, 2004, , 355-362.	1.0	18
134	Automatic composition of Web services with contingency plans. , 2004, , .		17
135	Ariane: An Awareness Mechanism for Shared Databases. Lecture Notes in Computer Science, 2004, , 92-104.	1.0	3
136	Digging Database Statistics and Costs Parameters for Distributed Query Processing. Lecture Notes in Computer Science, 2003, , 301-318.	1.0	1
137	Applying Theory Revision to the Design of Distributed Databases. Lecture Notes in Computer Science, 2003, , 57-74.	1.0	6
138	Building Reliable Web Services Compositions. Lecture Notes in Computer Science, 2003, , 59-72.	1.0	48
139	Estimating Costs of Path Expression Evaluation in Distributed Object Databases. Lecture Notes in Computer Science, 2002, , 351-360.	1.0	3
140	Software components retrieval through mediators and web search. Journal of the Brazilian Computer Society, 2002, 8, 55-63.	0.8	3
141	Towards a Theory Revision Approach for the Vertical Fragmentation of Object Oriented Databases. Lecture Notes in Computer Science, 2002, , 216-226.	1.0	1
142	The use of mediation and ontology technologies for software component information retrieval. , 2001, , .		26
143	Mechanisms for specifying communication behavior in object oriented database systems. , 2000, , .		0
144	Mining a large database with a parallel database server. Intelligent Data Analysis, 1999, 3, 437-451.	0.4	4

#	ARTICLE	IF	CITATIONS
145	A parallel spatial join framework using PMR-quadtrees. , 0, , .		1
146	Using ontologies for domain information retrieval. , 0, , .		22
147	An architecture for managing distributed scientific resources. , 0, , .		3
148	Mediating heterogeneous Web services. , 0, , .		8
149	OdysseyShare: an environment for collaborative component-based development. , 0, , .		9
150	Structural genomic workflows supported by Web services. , 0, , .		6
151	Webcomposer: a tool for the composition and execution of web service-based workflows. , 0, , .		5
152	ARAXA: Storing and Managing Active XML Documents. SSRN Electronic Journal, 0, , .	0.4	0
153	Análise de Hiperparâmetros em Aplicações de Aprendizado Profundo por meio de Dados de Proveniência. , 0, , .		2
154	A horizontal partitioning-based method for frequent pattern mining in transport timetable. Expert Systems, 0, , e12881.	2.9	0