Daniel S Katz

List of Publications by Year
in descending order

115 papers	3,679 citations	394421 19 h-index	197818 49 g-index
140 all docs	140 docs citations	140 times ranked	4333 citing authors

#	Article	IF	Citations
1	Pegasus: A Framework for Mapping Complex Scientific Workflows onto Distributed Systems. Scientific Programming, 2005, 13, 219-237.	0.7	912
2	Swift: A language for distributed parallel scripting. Parallel Computing, 2011, 37, 633-652.	2.1	319
3	A Component Architecture for High-Performance Scientific Computing. International Journal of High Performance Computing Applications, 2006, 20, 163-202.	3.7	154
4	Software citation principles. PeerJ Computer Science, 0, 2, e86.	4.5	150
5	Parsl. , 2019, , .		138
6	A multi-disciplinary perspective on emergent and future innovations in peer review. F1000Research, 2017, 6, 1151.	1.6	134
7	Montage: a grid portal and software toolkit for science-grade astronomical image mosaicking. International Journal of Computational Science and Engineering, 2009, 4, 73.	0.5	110
8	Crops In Silico: Generating Virtual Crops Using an Integrative and Multi-scale Modeling Platform. Frontiers in Plant Science, 2017, 8, 786.	3.6	102
9	Ten Simple Rules for Taking Advantage of Git and GitHub. PLoS Computational Biology, 2016, 12, e1004947.	3.2	96
10	Swift/T: Large-Scale Application Composition via Distributed-Memory Dataflow Processing. , 2013, , .		92
11	Four simple recommendations to encourage best practices in research software. F1000Research, 2017, 6, 876.	1.6	88
12	A Roadmap for HEP Software and Computing R&D for the 2020s. Computing and Software for Big Science, 2019, 3, 1.	2.9	85
13	Montage: a grid-enabled engine for delivering custom science-grade mosaics on demand. , 2004, 5493, 221.		75
14	A multi-disciplinary perspective on emergent and future innovations in peer review. F1000Research, 2017, 6, 1151.	1.6	62
15	Workflow task clustering for best effort systems with Pegasus. , 2008, , .		60
16	Exploring Automatic, Online Failure Recovery for Scientific Applications at Extreme Scales. , 2014, , .		59
17	Enabling real-time multi-messenger astrophysics discoveries with deep learning. Nature Reviews Physics, 2019, 1, 600-608.	26.6	53
18	Journal of Open Source Software (JOSS): design and first-year review. PeerJ Computer Science, 2018, 4, e147.	4.5	42

#	Article	IF	Citations
19	The Pegasus portal. , 2005, , .		36
20	The global impact of science gateways, virtual research environments and virtual laboratories. Future Generation Computer Systems, 2019, 95, 240-248.	7.5	36
21	Enforcing public data archiving policies in academic publishing: A study of ecology journals. Big Data and Society, 2019, 6, 205395171983625.	4.5	32
22	Accelerated, scalable and reproducible AI-driven gravitational wave detection. Nature Astronomy, 2021, 5, 1062-1068.	10.1	31
23	EnLIGHTened Computing: An architecture for co-allocating network, compute, and other grid resources for high-end applications. , 2007, , .		27
24	The Four Pillars of Research Software Engineering. IEEE Software, 2021, 38, 97-105.	1.8	27
25	Design and evaluation of the gemtc framework for GPU-enabled many-task computing. , 2014, , .		26
26	Turbine: A Distributed-memory Dataflow Engine for High Performance Many-task Applications. Fundamenta Informaticae, 2013, 128, 337-366.	0.4	23
27	Recognizing the value of software: a software citation guide. F1000Research, 2020, 9, 1257.	1.6	23
28	A Workflow-Aware Storage System: An Opportunity Study. , 2012, , .		22
29	Community Organizations: Changing the Culture in Which Research Software Is Developed and Sustained. Computing in Science and Engineering, 2019, 21, 8-24.	1.2	22
30	Convergence of artificial intelligence and high performance computing on NSF-supported cyberinfrastructure. Journal of Big Data, 2020, 7, .	11.0	22
31	Scheduling many-task workloads on supercomputers: Dealing with trailing tasks. , 2010, , .		19
32	Turbine. , 2012, , .		19
33	Swift/T. , 2013, , .		18
34	Understanding Software in Research: Initial Results from Examining Nature and a Call for Collaboration. , 2017, , .		16
35	Distributed computing practice for largeâ€scale science and engineering applications. Concurrency Computation Practice and Experience, 2013, 25, 1559-1585.	2.2	15
36	Conceptualization of a US Research Software Sustainability Institute (URSSI). Computing in Science and Engineering, 2018, 20, 4-9.	1.2	15

#	Article	IF	Citations
37	Report on the Third Workshop on Sustainable Software for Science: Practice and Experiences (WSSSPE3). Journal of Open Research Software, 2016, 4, 37.	5.9	15
38	Design and analysis of data management in scalable parallel scripting. , 2012, , .		14
39	Parallelizing the execution of sequential scripts. , 2013, , .		14
40	Swift/T. ACM SIGPLAN Notices, 2013, 48, 309-310.	0.2	14
41	Research Software Development & Management in Universities: Case Studies from Manchester's RSDS Group, Illinois' NCSA, and Notre Dame's CRC. , 2019, , .		14
42	A multi-disciplinary perspective on emergent and future innovations in peer review. F1000Research, 0, 6, 1151.	1.6	14
43	A Community Roadmap for Scientific Workflows Research and Development. , 2021, , .		14
44	Scalable Parallel Programming in Python with Parsl. , 2019, , .		13
45	Transitive Credit and JSON-LD. Journal of Open Research Software, 2015, 3, .	5.9	12
46	A FAIR and AI-ready Higgs boson decay dataset. Scientific Data, 2022, 9, 31.	5.3	12
47	AME. , 2011, , .		11
48	Evaluating Online Global Recovery with Fenix Using Application-Aware In-Memory Checkpointing Techniques. , 2016, , .		11
49	The Challenge and Promise of Software Citation for Credit, Identification, Discovery, and Reuse. Journal of Data and Information Quality, 2016, 7, 1-5.	2.1	11
50	BioWorkbench: a high-performance framework for managing and analyzing bioinformatics experiments. PeerJ, 2018, 6, e5551.	2.0	11
51	Generating Complex Astronomy Workflows. , 2007, , 19-38.		10
52	A Social Content Delivery Network for Scientific Cooperation: Vision, Design, and Architecture. , 2012, , .		10
53	Cyberinfrastructure Usage Modalities on the TeraGrid. , 2011, , .		9
54	Supporting High-Performance and High-Throughput Computing for Experimental Science. Computing and Software for Big Science, 2019, 3, 1.	2.9	9

#	Article	IF	Citations
55	Software Citation in Theory and Practice. Lecture Notes in Computer Science, 2018, , 289-296.	1.3	9
56	MTC envelope. , 2013, , .		9
57	Fourth Workshop on Sustainable Software for Science: Practice and Experiences (WSSSPE4). Journal of Open Research Software, 2018, 6, 10.	5.9	9
58	A survey of the state of the practice for research software in the United States. PeerJ Computer Science, 0, 8, e963.	4.5	9
59	Application skeletons: Construction and use in eScience. Future Generation Computer Systems, 2016, 59, 114-124.	7.5	8
60	The importance of software citation. F1000Research, 2020, 9, 1257.	1.6	8
61	A Fresh Perspective on Developing and Executing DAG-Based Distributed Applications: A Case-Study of SAGA-Based Montage. , 2009, , .		7
62	Managing genomic variant calling workflows with Swift/T. PLoS ONE, 2019, 14, e0211608.	2.5	7
63	Report on the Second Workshop on Sustainable Software for Science: Practice and Experiences (WSSSPE2). Journal of Open Research Software, 2016, 4, 7.	5.9	7
64	Constructing a Social Content Delivery Network for eScience. , 2013, , .		6
65	Using Application Skeletons to Improve eScience Infrastructure. , 2014, , .		6
66	Integrating Abstractions to Enhance the Execution of Distributed Applications. , 2016, , .		6
67	Evaluating Distributed Execution of Workloads. , 2017, , .		6
68	BOSS-LDG: A Novel Computational Framework That Brings Together Blue Waters, Open Science Grid, Shifter and the LIGO Data Grid to Accelerate Gravitational Wave Discovery. , 2017, , .		6
69	The principles of tomorrow's university. F1000Research, 2018, 7, 1926.	1.6	6
70	<title>Architecture for access to a compute-intensive image mosaic service in the NVO</title>. , 2002, 4846, 91.		5
71	Critical perspectives on large-scale distributed applications and production Grids. , 2009, , .		5
72	Job and data clustering for aggregate use of multiple production cyberinfrastructures. , 2012, , .		5

#	Article	IF	Citations
73	Recent advances in e-Science. Future Generation Computer Systems, 2013, 29, 548.	7.5	5
74	Interlanguage parallel scripting for distributed-memory scientific computing. , 2015, , .		5
75	Publish your software: Introducing the Journal of Open Source Software (JOSS). Computing in Science and Engineering, 2018, 20, 84-88.	1.2	5
76	Software Must be Recognised as an Important Output of Scholarly Research. International Journal of Digital Curation, 2022, 16, 6.	0.2	5
77	Optimization of a parallel ocean general circulation model. , 1997, , .		4
78	MTC envelope. , 2013, , .		4
79	Evaluating storage systems for scientific data in the cloud. , 2014, , .		4
80	Introducing distributed dynamic dataâ€intensive (D3) science: Understanding applications and infrastructure. Concurrency Computation Practice and Experience, 2017, 29, e4032.	2.2	4
81	Quantifying the Impact of Memory Errors in Deep Learning. , 2019, , .		4
82	The State of Sustainable Research Software: Learning from the Workshop on Sustainable Software for Science: Practice and Experiences (WSSSPE5.1). Journal of Open Research Software, 2019, 7, 11.	5.9	4
83	The challenges of theory-software translation. F1000Research, 2020, 9, 1192.	1.6	4
84	Science applications of the Montage image mosaic engine. Proceedings of the International Astronomical Union, 2006, 2, 621-621.	0.0	3
85	Louisiana: a model for advancing regional e-Research through cyberinfrastructure. Philosophical Transactions Series A, Mathematical, Physical, and Engineering Sciences, 2009, 367, 2459-2469.	3.4	3
86	JETS: Language and System Support for Many-Parallel-Task Workflows. Journal of Grid Computing, 2013, 11, 341-360.	3.9	3
87	A social content delivery network for eâ€Science. Concurrency Computation Practice and Experience, 2017, 29, e3854.	2.2	3
88	Mapping the Research Software Sustainability Space. , 2018, , .		3
89	High Performance Computing Systems for Autonomous Spaceborne Missions. International Journal of High Performance Computing Applications, 2001, 15, 282-296.	3.7	2
90	An innovative application execution toolkit for multicluster grids. , 2009, , .		2

#	Article	IF	Citations
91	Novel submission modes for tightly coupled jobs across distributed resources for reduced time-to-solution. Philosophical Transactions Series A, Mathematical, Physical, and Engineering Sciences, 2009, 367, 2545-2556.	3.4	2
92	Special Section: Grid computing, high-performance and distributed applications. Future Generation Computer Systems, 2010, 26, 257-258.	7.5	2
93	Toward Interlanguage Parallel Scripting for Distributed-Memory Scientific Computing. , 2015, , .		2
94	Understanding the multifaceted geospatial software ecosystem: a survey approach. International Journal of Geographical Information Science, 2020, , 1-19.	4.8	2
95	Software Training in HEP. Computing and Software for Big Science, 2021, 5, 22.	2.9	2
96	Embedded/Real-Time Systems. International Journal of High Performance Computing Applications, 2001, 15, 186-190.	3.7	1
97	Data-intensive CyberShake computations on an opportunistic cyberinfrastructure. , 2011, , .		1
98	Pilot abstractions for compute, data, and network. , 2012, , .		1
99	On Replica Placement in a Social CDN for e-Science. , 2014, , .		1
100	DA-TC: a novel application execution model in multicluster systems. Cluster Computing, 2014, 17, 371-387.	5.0	1
101	Building a Sustainable Structure for Research Software Engineering Activities. , 2018, , .		1
102	Introduction to Accelerating Scientific Discovery With Reusable Software. Computing in Science and Engineering, 2019, 21, 5-7.	1.2	1
103	Sustaining Research Software via Research Software Engineers and Professional Associations. , 2021, , .		1
104	Addressing Research Software Sustainability via Institutes. , 2021, , .		1
105	Application Skeleton: Generating Synthetic Applications for Infrastructure Research. Journal of Open Source Software, 2016, 1, 17.	4.6	1
106	Real-time HEP analysis with funcX, a high-performance platform for function as a service. EPJ Web of Conferences, 2020, 245, 07046.	0.3	1
107	Porting Ordinary Applications to Blue Gene/Q Supercomputers. , 2015, , .		0
108	Software Citations and the ACAT Community. Journal of Physics: Conference Series, 2018, 1085, 022010.	0.4	0

#	Article	IF	Citations
109	Extended Abstract. ACM SIGAda Ada Letters, 2021, 40, 73-75.	0.1	0
110	Research Software Sustainability and Citation. , 2021, , .		0
111	Extreme Scale Survey Simulation with Python Workflows. , 2021, , .		0
112	Data-Oriented Distributed Computing for Science: Reality and Possibilities. Lecture Notes in Computer Science, 2006, , 1119-1124.	1.3	0
113	Topic 1: Support Tools and Environments. Lecture Notes in Computer Science, 2012, , 3-3.	1.3	0
114	Toward Interoperable Cyberinfrastructure: Common Descriptions for Computational Resources and Applications. , 2020, , .		0
115	Software citation. Information Services and Use, 2021, 41, 39-42.	0.2	0