The Research Software Engineering Revolution

Gabriele Bozzola, PhD


TIMEO

Combien il est en sus de terre

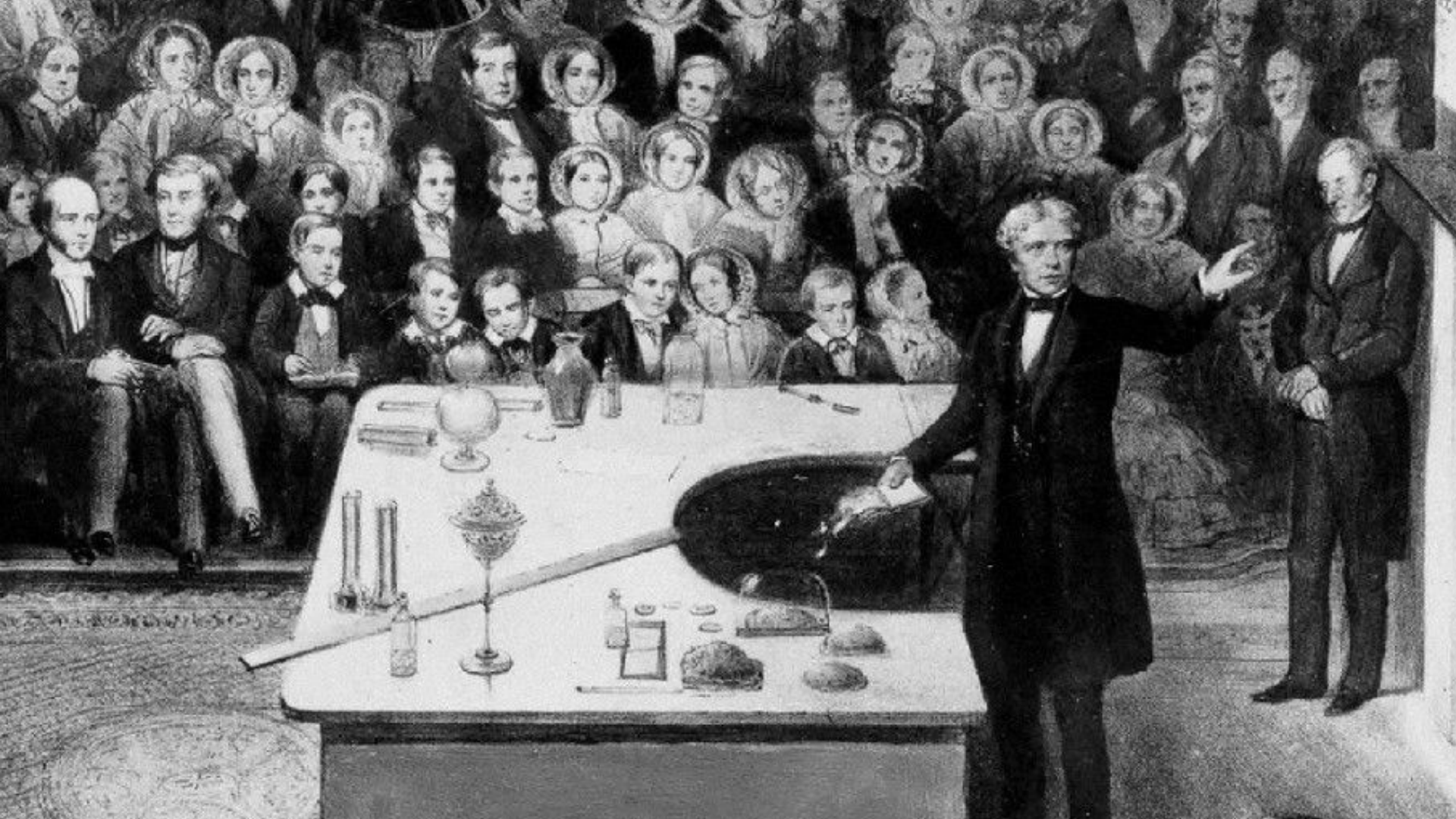

NULLIUS
IN
VERBA

PHILOSOPHIÆ NATURALIS PRINCIPIA MATHEMATICA

Autore *Is. Newton* Trin. Coll. Cantab. Soc. Matheseos Professore Lucasiano, & Societatis Regalis Sodali.

IMPRIMATUR.
S. PEPYS, *Reg. Soc.* PRÆSES.
Julii 5. 1686.

LONDINI,
Jussu *Societatis Regiæ* ac Typis *Josephi Streater*. Prostat apud plures Bibliopolas. *Anno* MDCLXXXVII.



AMERICAN
Journal of Mathematics
PURE AND APPLIED.

EDITOR IN CHIEF,
J. J. SYLVESTER, LL. D., F. R. S., *Corr. Mem. Inst. of France.*

ASSOCIATE EDITOR IN CHARGE,
WILLIAM E. STORY, PH. D., (*Leipsic.*)

WITH THE CO-OPERATION OF
BENJAMIN PEIRCE, LL. D., F. R. S., PROFESSOR OF MATHEMATICS IN HARVARD UNIVERSITY, In Mechanics,
SIMON NEWCOMB, LL. D., F. R. S., CORR. MEM. INST. OF FRANCE, SUPERINTENDENT OF THE AMERICAN EPHEMERIS, In Astronomy,
AND
H. A. ROWLAND, C. E., In Physics.

PUBLISHED UNDER THE AUSPICES OF THE
JOHNS HOPKINS UNIVERSITY.

Πάντα γα μὰν τὰ γιγνωσκόμενα ἀριθμὸν ἔχοντι.—*Philolaos.*

Volume I.

BALTIMORE:
PRINTED FOR THE EDITORS BY JOHN MURPHY & CO.

B. WESTERMANN & CO., D. VAN NOSTRAND, TRÜBNER & CO., *New York.*
A. WILLIAMS & CO., *Boston.* FERREE & CO., *Philadelphia.* A. ASHER & CO., *Berlin.*
GAUTHIER-VILLARS, *Paris.*

1878.

Reprinted with the permission of The Johns Hopkins Press

Johnson Reprint Corporation Kraus Reprint Corporation

Springer

ELSEVIER

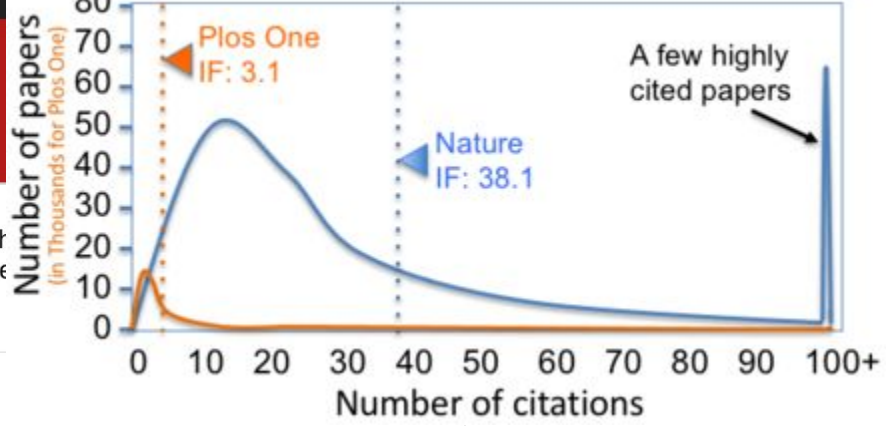
SCIENCE = PAPERS

arXiv

arXiv is a free distribution service and an open-access archive for nearly 2.4 million sch
mathematics, computer science, quantitative biology, quantitative finance, statistics, ele
science, and economics. Materials on this site are not peer-reviewed by arXiv.

Subject search and browse:

Physics | Search | Form Interface | Catchup



Physics

Albert Einstein

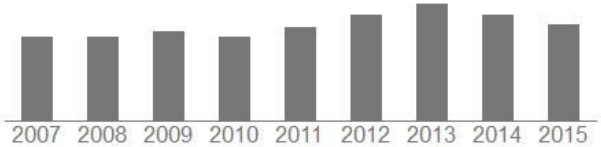
Follow

Institute of Advanced Studies, Princeton
Physics
No verified email

Title 1–20	Cited by	Year
Can quantum-mechanical description of physical reality be considered complete?	13983	1935

Google Scholar

Citation indices	All	Since 2010
Citations	95625	31937
h-index	108	67
i10-index	365	219

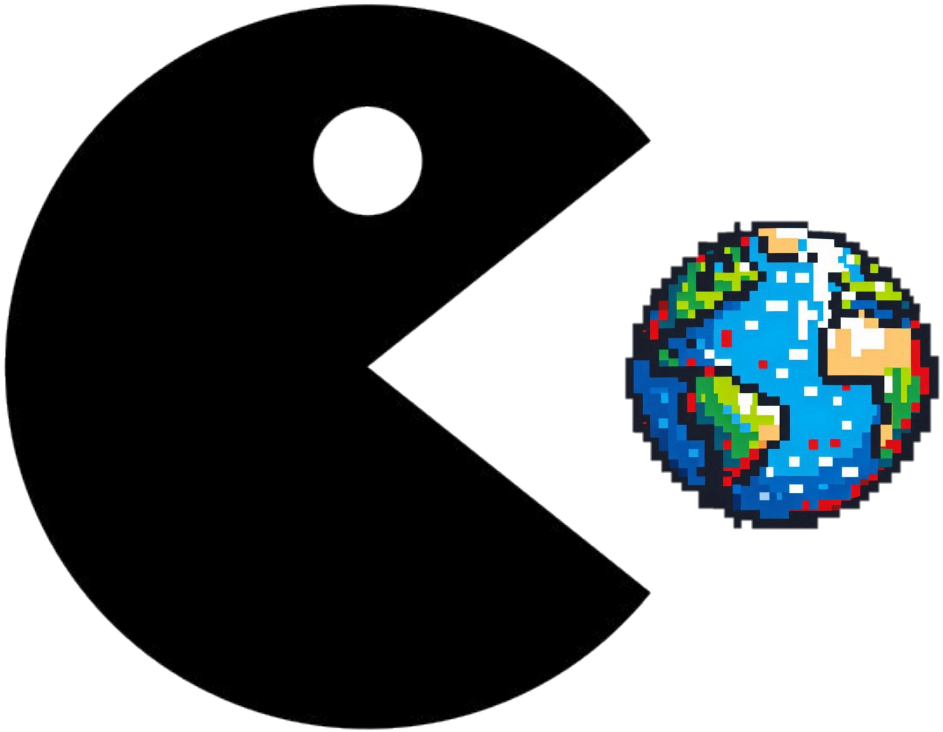


SCIENCE = PAPERS

THIS IS A PROBLEM!

TOPIC FOR ANOTHER TALK…

SOFTWARE IS EATING THE WORLD



YouTube

facebook

skype™

Of 2,000 scientists Jo Hannay and colleagues surveyed online,[2] 91 percent said using scientific software is important for their own research, 84 percent said developing scientific software is important for their own research, 53.5 percent claimed to spend more time developing scientific software than they did 10 years ago, and 38 percent spend at least one fifth of their time developing software. Scientists aren't just using software; they are its prime producers.

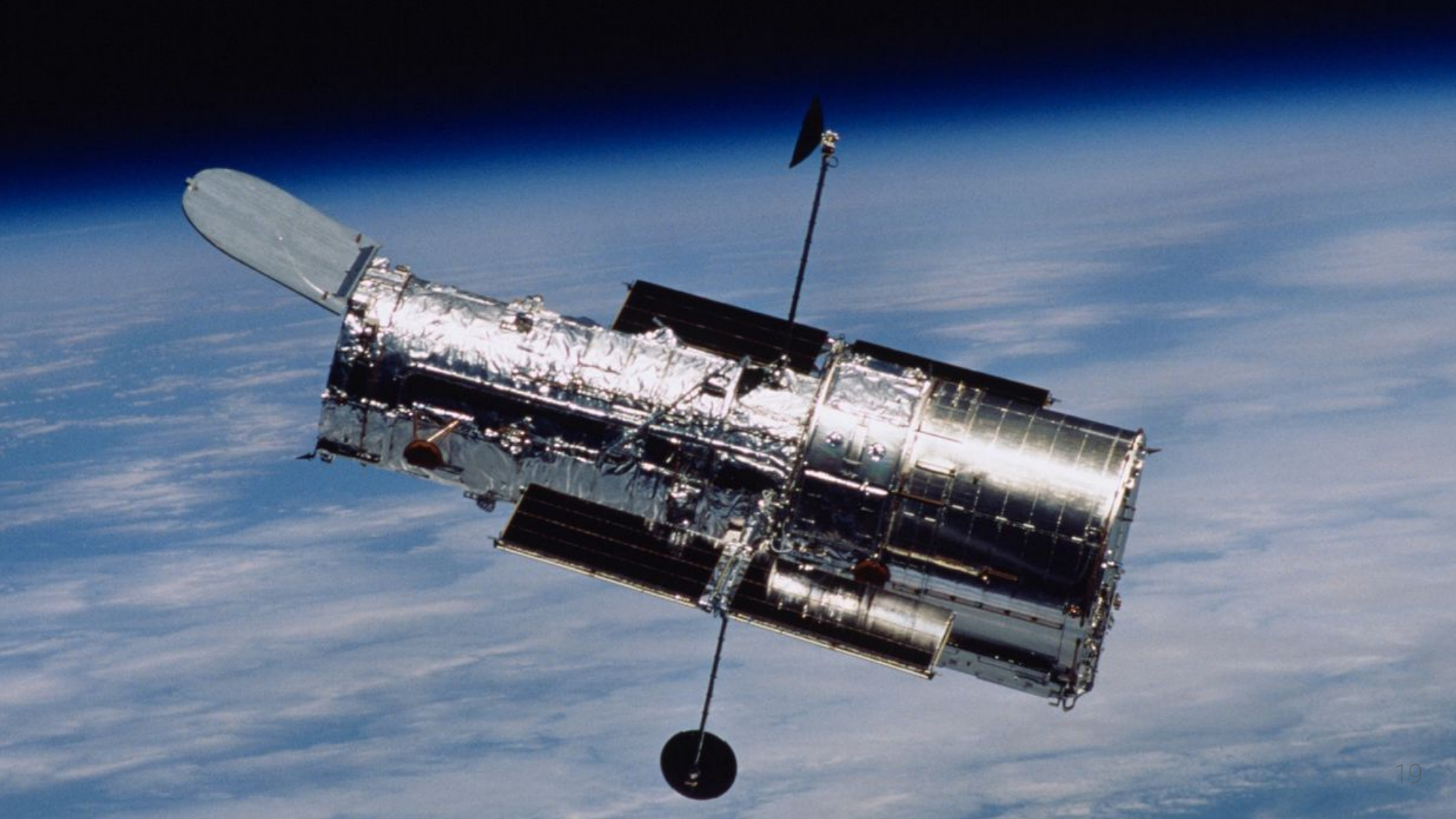
Research software is critical to supporting science. Between 1998-2016, the NSF made more than 18k awards totaling $9.6 billion related to research software.

MODERN SCIENCE RUNS ON SOFTWARE

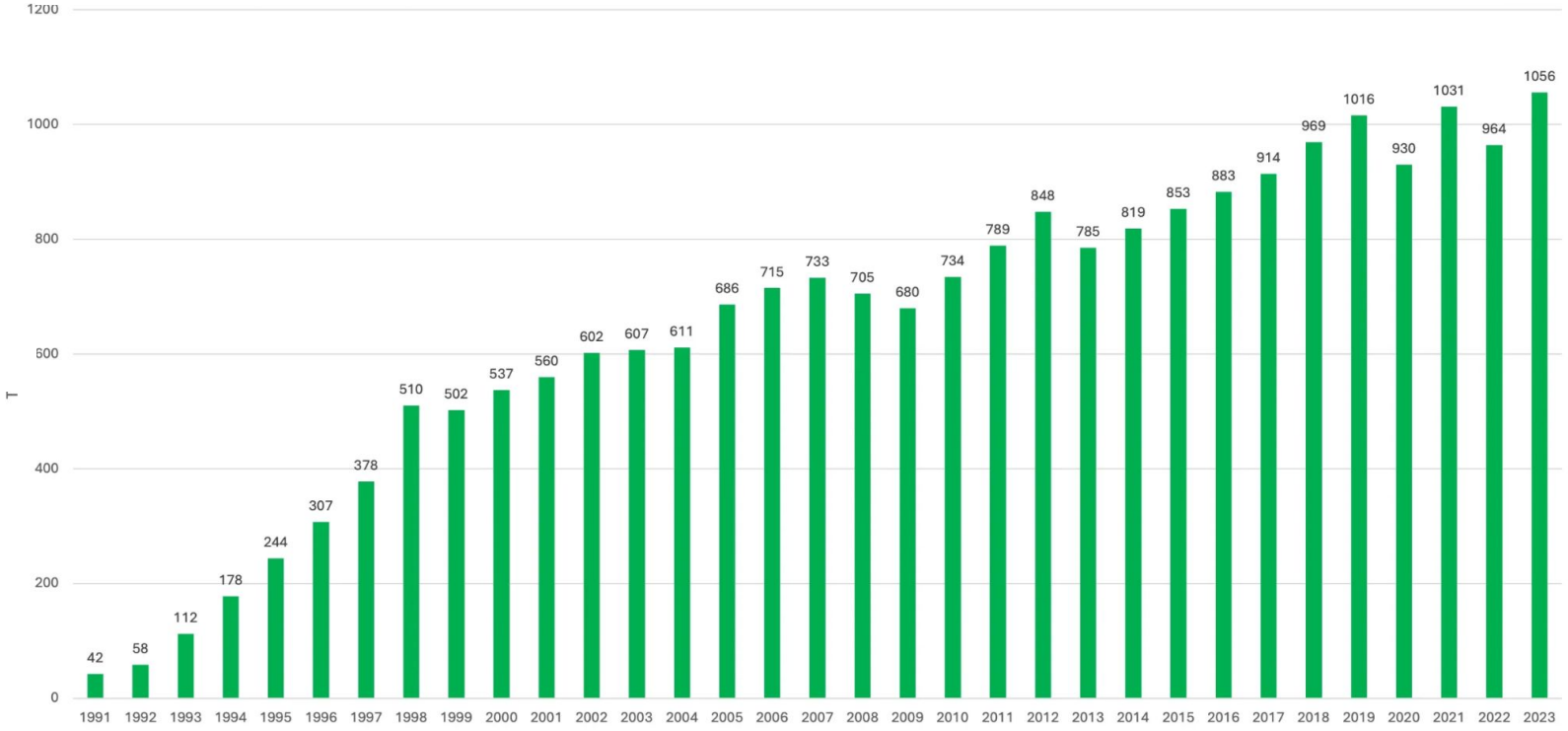
Research software plays such a critical role in day to day research that a comprehensive survey reports 90–95% of researchers in the US and the UK rely upon it and more than 60% were unable to continue working if such software stopped functioning (Hettrick, 2014).

38%

of total spend invested into software-reliant research in 2017

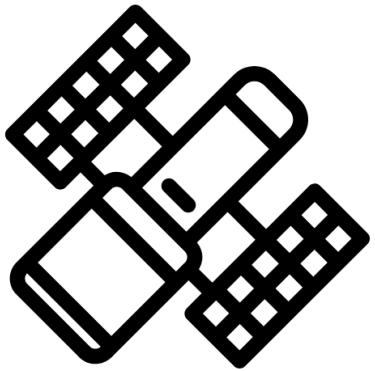


Hubble Space Telescope Peer-Reviewed Scientific Papers Published



Data from NASA

21,000 papers in 30+ years



≈

22,000 citations in 4 years

NumPy

Image credits for HST icon: Fourup Corporate

10,000 citations in 10 years

≈

AstroPy

9,000 citations in 28 years

≈

SExtractor

12,000 citations in 12 years

emcee

≈

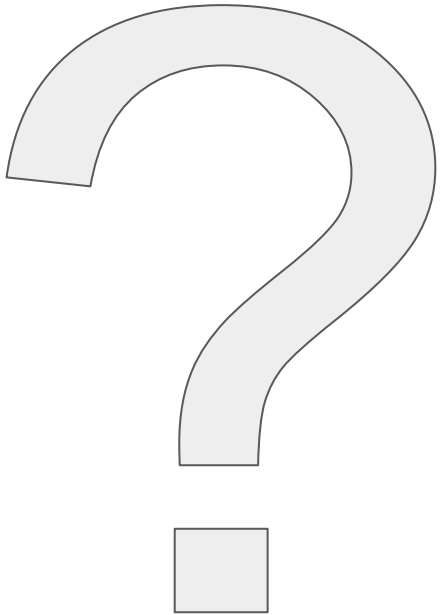
6,000 citations in 17 years

GADGET-2

≈

UNRECOGNIZED CONTRIBUTIONS

SOFTWARE THAT WAS NOT BUILT



Do not spend too much time on software!

Software-oriented people leave academia

Impactful software is not built/maintained

Quality of research software is lowered

Frequency of using *best* practices As a follow-up question, we asked the respondents how frequently they used a number of standard software engineering practices. The response options were *Never*, *Sometimes*, *Half of the time*, *Most of the time*, *Always*. The following list reports those who responded *Most of the time* or *Always* for the most commonly used practices (in decreasing order):

- Continuous Integration – 54% (54/100)
- Use of coding standards – 54% (54/100)
- Architecture or Design – 51% (52/101)
- Requirements – 43% (43/101)
- Peer code review – 34% (34/99)

RQ7. *How reproducible are notebooks?*
Answer: We were able to successfully run 24.11% of the unambiguous execution order Python notebooks. This number is close to the results of a previous reproducibility study [32] about general computer systems research (24.9%). However, the rate is way smaller (4.03%) when we count only notebooks that produce the same results. The most common causes of failures were related to missing dependencies, the presence of hidden states and out-of-order executions, and data accessibility.

THE STATE OF RESEARCH SOFTWARE IS NOT GREAT

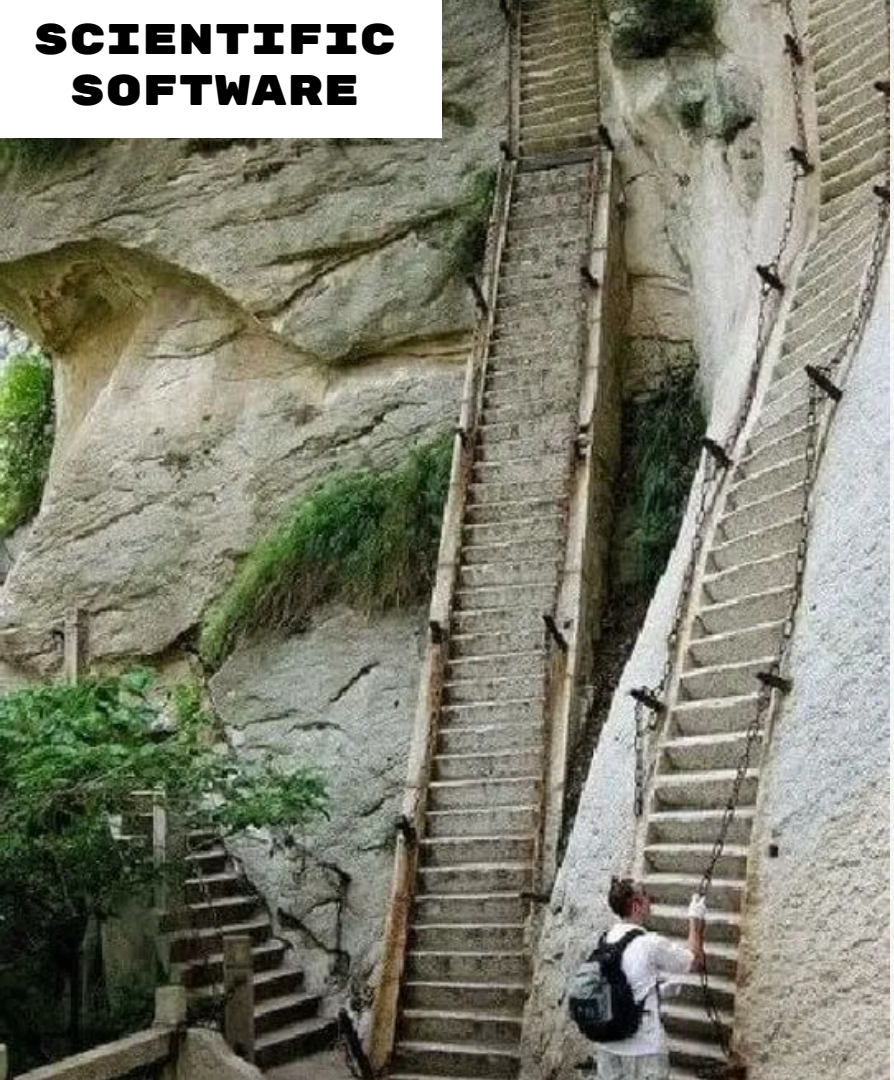
We find that 74% of R files failed to complete without error in the initial execution, while 56% failed when code cleaning was applied, showing that many errors can be prevented with good coding practices.

Wikipedia
https://en.wikipedia.org › wiki › Replication_crisis

Replication crisis

The **replication crisis** is an ongoing methodological crisis in which the results of many scientific studies are difficult or impossible to reproduce.

SCIENTIFIC SOFTWARE



THE ENTIRE ACADEMIC SYSTEM

UNREWARDED SOFTWARE CONTRIBUTIONS

Better incentives

Better software

Better research

2010:
WIND OF CHANGE

SOFTWARE SUSTAINABILITY INSTITUTE

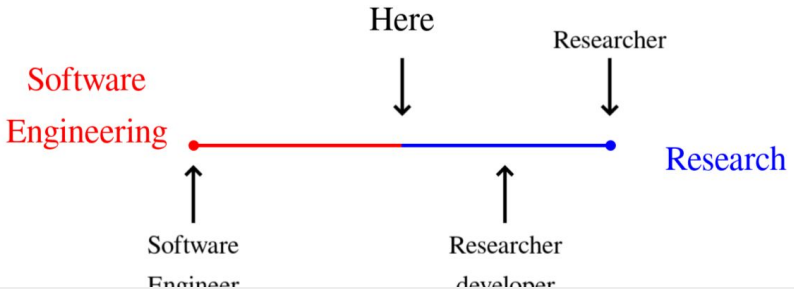
SSI was "***founded on the premise that helping individuals and institutions understand the vital role that software plays in research would accelerate progress in every field of scientific and academic endeavour***"

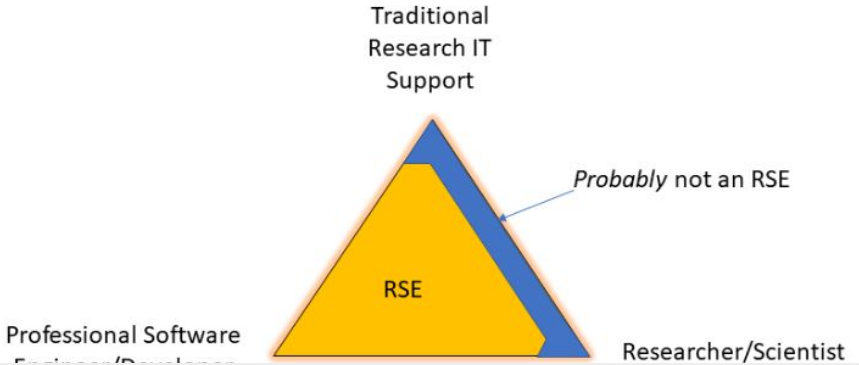
2012, Oxford, collaboration meeting:

Why is there no career for software developers in academia?

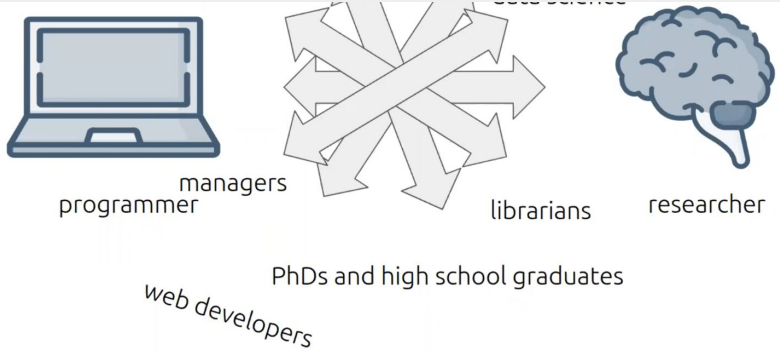
Research Software Engineers

Analyst Developer Analyst Programmer Analyst Programmer - SITS (x 3) Analyst/Programmer Applications Developer Applied Scientist Architectural Robotics Developer Assistant Data Programmer Assistant Project Manager Atmospheric Correction and Radiative Transfer Model Scientist Audio Software Developer - KTP Associate Bioinformatician Bioinformatician In Potato Genomics and Genetics Bioinformatician/Computational Bioscientist in Microbiology Bioinformaticians Bioinformatics Analyst Bioinformatics Postdoctoral Researcher Bioinformatics scientist Biometric Software Systems Developer Biorespository Software Developer C++ / 3D Graphics Software Engineer Casebooks Project Editor (Research Assistant/Associate) Climate Researcher (Research Associate) Clinical Study Programmer CoMPLEX Research Associate Computational Biologist / Bioinformatician Computational Scientist Computational Scientist in Computational Fluid Dynamics & Industrial Applications Computational Scientist in Structural Mechanics and Industrial Applications Computer Scientist Computer Vision Researcher Content Developer/Programmer Control Engineer-IMG - 3 posts CREATe Data Specialist Data Analyst Data Integration Coordinator Data Manager x3 Database and Software Engineer Database Manager/Researcher Database Programmer Digital Media Technician E-Learning Portal Manager (KTP Associate) e-Learning Systems Development Analyst e-Learning Systems Development Analyst (Moodle, SQL) E-Learning Web Developer E-Portfolio Learning Technologist Embedded Systems Engineer Engineering Technician Environmental Scientist EPSRC Studentship on Algorithmic Construction of Finsler-Lyapunov Functions Experimental Officer in Bioinformatics Experimental Psychologist Finance Assistant Gaia Alerts Software Developer Gaia Software Developer (Gaia UK Team) GIS Applications Specialist Graduate Programmer / Software Developer Graphics Programmer Health Data Manager / Scientist High Throughput Bioinformatician High Throughput Sequencing Bioinformatician (Two posts) HIVE Manager/ HIVE Co-ordinator HIVE Senior Researcher and Technical Lead Hydro-informatics Scientific Software Developer Image Analysis Manager for Cancer Imaging Information Systems Developer Instrumentation Engineer Investigator Statistician IT Developer IT Services Manager IT Services Specialist (e-Learning Systems) IT Support Technician (Unix / Windows Systems) Knowledge Transfer Partnership (KTP) Associate: Innovent Technologies LTD Knowledge Transfer Partnerships (KTP) Associate - Software Developer KTP Associate - Robot Vision Scientist (Research Fellow) KTP Associate (Fixed Term Contract for 24 months) KTP Associate (Precision Agriculture Data Analyst) KTP Associate â€" Graduate Web Developer KTP Associate: Electronics / Robotics Engineer Learning Technologist Leicester Respiratory BRU IT Developer Linguist / Psycholinguist Maker Space Technician Marie Curie Early Stage Researcher Marie Curie Early Stage Researcher in Radar Rainfall for Integrated Water Quality Modelling Marine Earth Observation Scientists Medical Statistician Medical Statistician/Senior Medical Statistician Metrology E[illegible]nistrator (Fixed-term) NIHR Research Methods Fellow PDRA on EU Project on Automated Multisensor Surveillan[illegible]st Doctoral Research Worker Post Doctoral Researcher in the application of Digital Technology Post-Doctoral Res[illegible]octoral Research Associate (Pathogen Genomics) Post-Doctoral Research Fellow Postdoctoral Fellow - populati[illegible]ow in Cancer Therapeutics Postdoctoral Research Assistant Postdoctoral Research Associate Postdoctoral Research Fellow Postdoctoral Research Scientist Postdoctoral Researcher in Declarative (Logic and Functional) Programming Postdoctoral Researcher Postdoctoral Scientist Postdoctoral statistician Postdoctoral Training Fellow - Statistical and Computational Genetics of Autism Principal / Senior Bioinformatician Principal Bioinformatician Product Development Engineer (Rail) Publishing Portal Web Developer Radio Frequency Engineer Reader in Computer Science Reporting Analyst Research (Software) Engineer Research Assistant Research Associate Research Fellow Research Image Data Manager, Biomedical Engineering Research Officer Research Officer â€" Social Protection Research postgraduate Research Programmer Research Scientist Research Scientist / Senior Research Scientist Research Scientist in Machine Learning and Computer Vision Research Software Developer Research Software Developer for the Herchel Smith Professor of Organic Chemistry Research Software Engineer Research Studentship Research Worker Researcher SAP Trainee Technical Analyst Scientific Officer with Michela Garofalo Scientist SEAHA Studentship: Extracting epidemiological data from collections SEEG Data Archive Manager Senior / Research Associate in Clinical Integration and Image Analysis for Fetal Surgery Senior Analyst Programmer (Business Analysis) Senior Analyst/Programmer Senior Bioinformatician Senior Bioinformatician / Bioinformatician Senior Computational Statistician - Spatial Models Senior Data Acquisition Scientist / Data Acquisition Scientist Senior Data Manager Senior Database Administrator Senior IT Developer Analyst Senior Mathematical Modeller Senior Media Developer Senior Postdoctoral Researcher - Evolutionary and Computational Analysis of Infectious Disease (Phylodynamics) Senior Research Assistant Senior Research Associate Senior Research Associate â€" Molecular Modelling & Simmulation Senior Research Associate in Quantitative Clinical MRI Senior Research Fellow Senior Research Fellow/Research Fellow in Vibration Diagnostics and Prognostics/Digital Signal Processing Senior Research Laboratory Technician Senior Research Technician Senior Software Developer in Bioinformatics Senior Software Engineer / Software Engineer Senior Statistical Epidemiologist Senior Systems Administrator Senior Technician / Demonstrator (UCMK) Senior Web Developer SharePoint Developer Software Developer Software Developer (Bioinformatics) Software Developer (KTP Associate) Software Developer x 2 Software Developer/Programmer Software Developers in e-Learning Software Engineer Software/ Database Developer (KTP Associate) Sports Programme Manager Statistical Geneticist Statistical Programmer/Data Scientist Statistician Statistician/Epidemiologist Student and Enrolment Services Manager SWCAR Information Assistant System Administrator Systems Developer Systems, Data and Applications Team Leader Teaching Fellow in Computational Methods UTRCI Research Scientist, Control Systems Web Application Programmer Web Developer





RSE = Those who regularly use expertise in programming to advance research



According to Simon Hattrick, Ian Cosden, and Vanessa Sochat respectively

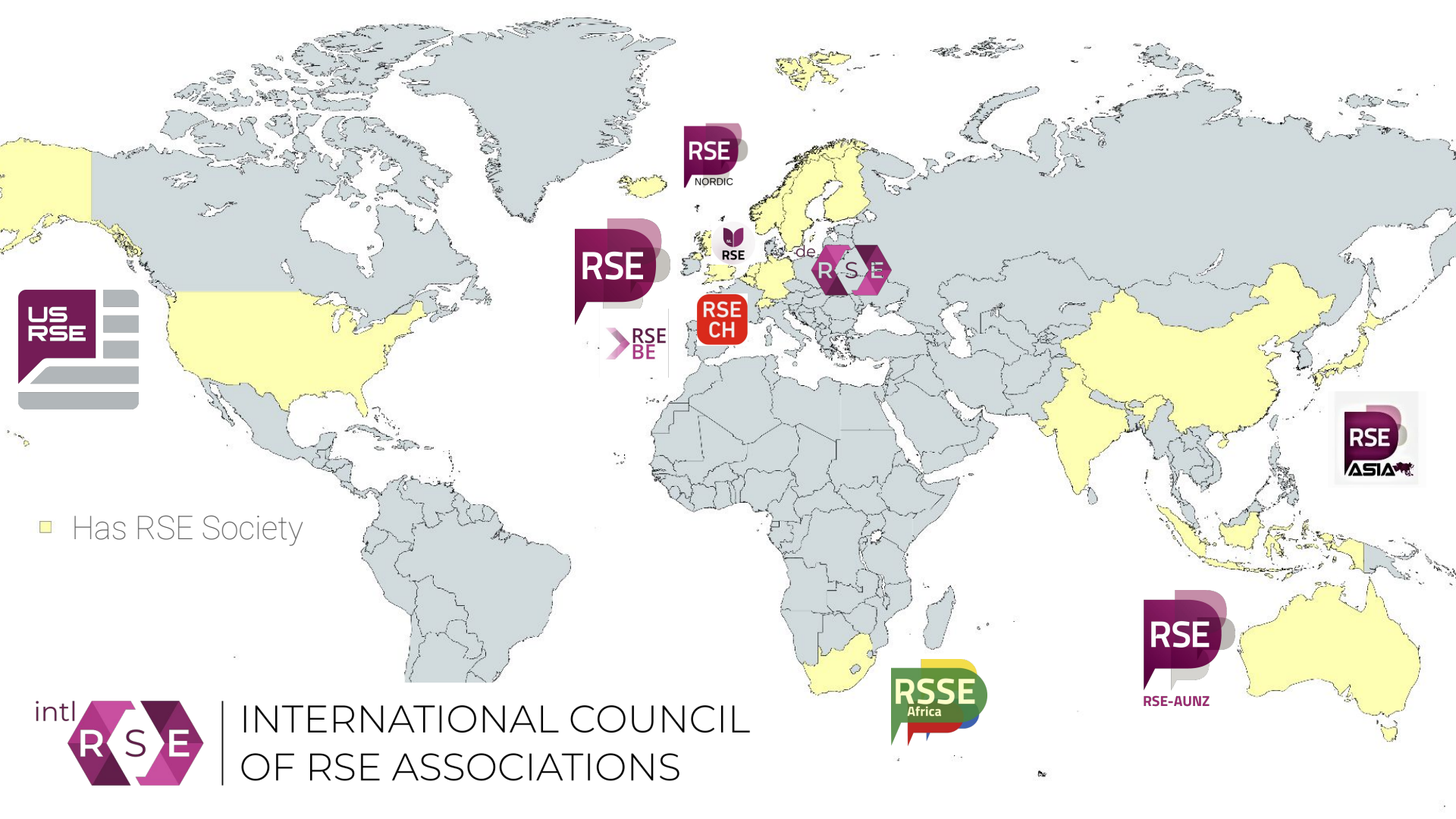
SOFTWARE SUSTAINABILITY INSTITUTE

RSE

US RSE

2010 2012 2014 2016 2018 2021

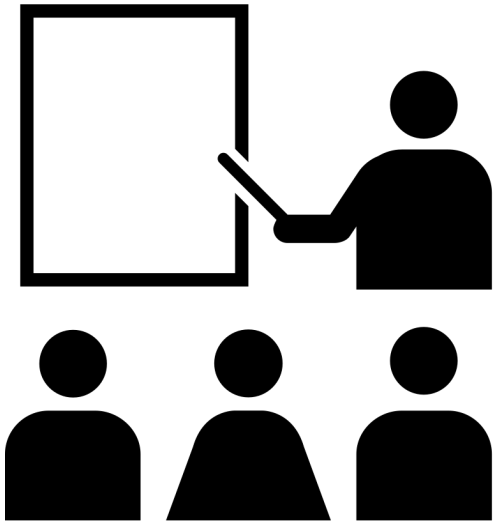
intl RSE



- Has RSE Society

intl RSE | INTERNATIONAL COUNCIL OF RSE ASSOCIATIONS

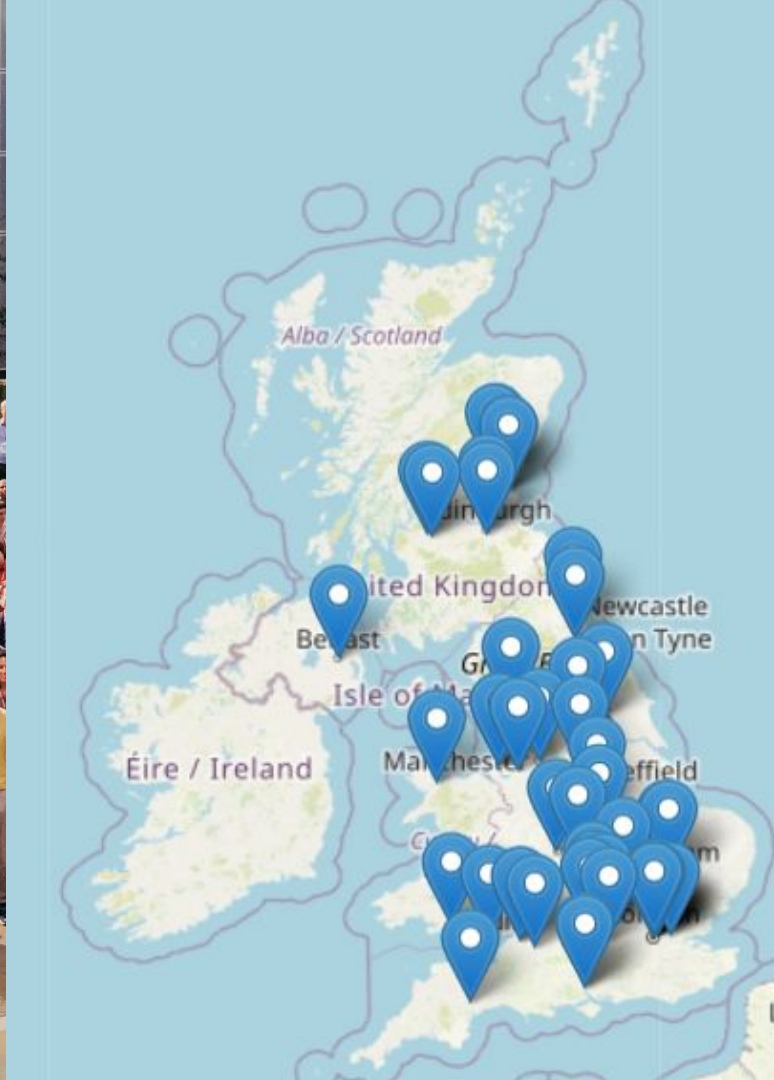

BETTER
SOFTWARE
BETTER
RESEARCH
www.software.ac.uk



The Society of Research Software Engineering was founded on the belief that a world which relies on software must recognise the people who develop it.

Our mission is to establish a research environment that recognises the vital role of software in research.

We work to increase software skills across everyone in research, to promote collaboration between researchers and software experts, and to support the creation of an academic career path for Research Software Engineers.


Alba / Scotland
Éire / Ireland

R1 Universities with RSE Groups

As of 2024-09-24, **12 out of 146** R1 universities appear to have RSE groups based on search results. If you find errors, please, open an issue. See code here.

Search universities...

University	RSE Group	Link
Arizona State University	✅	Link
Auburn University	❌	
Baylor University	❌	
Binghamton University	❌	

sbozzolo.github.io/has_rse

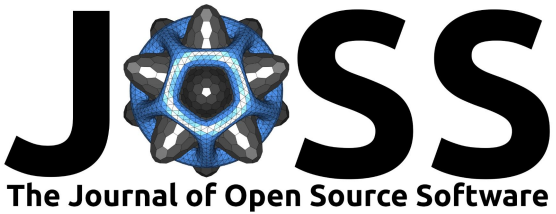
ALFRED P. SLOAN
FOUNDATION

CHAN
ZUCKERBERG
INITIATIVE

SIMONS
FOUNDATION

USRSE'25

Philadelphia, PA
October 6-8, 2025

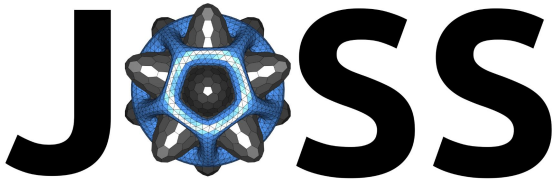

JOSS
The Journal of Open Source Software


ROpenSci


FORRT


software
carpentry


pyOpenSci



The Journal of Open Source Software is a **developer friendly**, open access journal for research software packages.

Committed to publishing quality research software with zero article processing charges or subscription fees.

Submit a paper to JOSS

[REVIEW]: kuibit: Analyzing Einstein Toolkit simulations with Python #3099

New issue

Closed

whedon opened on Mar 12, 2021 · edited by whedon

Edits

Submitting author: @Sbozzolo (Gabriele Bozzola)
Repository: https://github.com/Sbozzolo/kuibit/
Version: 1.0.0
Editor: @eloisabentivegna
Reviewer: @Yurlungur, @eloisabentivegna
Archive: 10.5281/zenodo.4681119

Assignees

Yurlungur

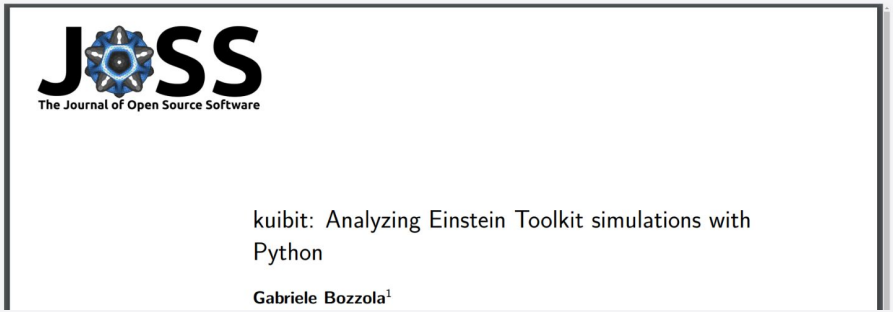
eloisabentivegna

Labels

Python accepted published recommend-accept review

kuibit: Analyzing Einstein Toolkit simulations with Python

Submitted 30 January 2021 • Published 13 April 2021



Software repository

Paper review

Download paper

Software archive

Review
Editor: @eloisabentivegna (all papers)
Reviewers: @yurlungur (all reviews), @eloisabentivegna (all reviews)

Authors
Gabriele Bozzola (0000-0003-3696-6408)

JOSS Papers

All Papers 3255 | Published Papers 2882 | Active Papers 373

Search by title, tag, author, or language

PUBLISHED Published 1 day ago

AixWeather: A Weather Data Generation Tool for Building Energy System Simulations. Pull, Transform, Export.

Python

@martinraetz

DOI 10.21105/joss.07344

PUBLISHED Published 2 days ago

ExaDEM: a HPC application based on exaNBody targeting scalable DEM simulations with complex particle shapes

Python C++ Gnuplot

@rprat-pro

DOI 10.21105/joss.07484

PUBLISHED Published 5 days ago

DisCoTec: Distributed higher-dimensional HPC simulations with the sparse grid combination technique

C++ Python

@freifrauvonbleifrei

DOI 10.21105/joss.07018

The Journal of Open Source Education

The Journal of Open Source Education is an **educator friendly** journal for publishing open-source educational materials and software.

Submit a paper to JOSE

🎉 Volunteer to review

Explore Papers

Documentation

Learn More

Recently Published Papers 65

PUBLISHED Published about 2 months ago

The University of Toronto Climate Downscaling Workflow: Tools and Resources for Climate Change Impact Analysis

Jupyter Notebook JavaScript

@mikemorris12

DOI 10.21105/jose.00243

Peer review of Python software to support open science

Python Package Peer Review

Community Partnerships

Simplifying Packaging

pyOpenSci

ROpenSci

Framework for
Open and
Reproducible
Research
Training

FORRT

The Carpentries teaches foundational coding and data science skills to researchers worldwide.

THE CARPENTRIES

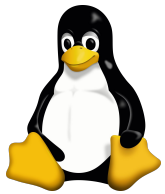
oftware carpentry

The Carpentries has built a community of

4942	183	4529	70
INSTRUCTORS	TRAINERS	WORKSHOPS	COUNTRIES

RSE

Openness
Collaboration
Reproduciability
Accessibility



RSE Opportunities

Current RSE Openings

1. Staff or Senior Research Engineer or Scientist: University of Illinois Urbana-Champaign, Urbana, IL *Posted: Sep 26, 2024*
2. Research Software Engineer II (RSEII), HPC & AI Software Engineer: Dartmouth College ITC, Hanover, NH (Hybrid) *Posted: Sep 25, 2024*
3. Research Software Developer: National Center for Ecological Analysis & Synthesis, UC Santa Barbara, Santa Barbara, CA *Posted: Sep 23, 2024*
4. Quantum Architect, Simulation & Software: PsiQuantum, Palo Alto, CA or remote *Posted: Sep 21, 2024*
5. Engineering Scientist: Texas Advanced Computing Center, Austin, Texas *Posted: Sep 20, 2024*
6. Research Software Engineer II: Center for Digital Humanities at Princeton University, Princeton, NJ *Posted: Sep 18, 2024*
7. Computer Systems Engineer: Lawrence Berkeley National Laboratory, Berkeley, CA *Posted: Sep 11, 2024*
8. Senior Full-Stack Geospatial Developer: Eagle Rock Analytics, Sacramento, CA / Remote *Posted: Sep 10, 2024*
9. Research Scientist II/Senior: Georgia Institute of Technology, Remote Eligible *Posted: Sep 10, 2024*

Join us in our mission to push for research world where

- Code is treated as a first-class research output
- Software development is recognized and rewarded
- Research code is FAIR (Findable, Accessible, Interoperable, and Reusable)
- Researchers are equipped with the necessary skills and tools to develop high-quality, sustainable software
- There are meaningful career paths for those who wish to focus on software

How to start helping:

Be the first one to **recognize the value of research software**

gbozzola@caltech.edu
sbozzolo.github.io
linkedin.com/gabrielebozzola

Connect and find your community

Check these out for grants, fellowships, trainings, events, and additional resources

- US RSE (us-rse.org)
- Research Software Alliance (researchsoft.org)
- Software Sustainability Institute (software.ac.uk)
- Better Scientific Software (bssw.io)