



MANCHESTER
1824

UNIVERSITY OF
OXFORD

UNIVERSITY OF
Southampton



www.software.ac.uk

SSI and Communities of Practice

Shoaib Sufi - Software Sustainability Institute – shoaib.sufi@software.ac.uk
Maximising Impact and Reproducibility of Acoustic Modelling Software Workshop, 13 July 2020



The Software Sustainability Institute

- A national facility for cultivating world-class research through software
- **“Better Software, Better Research”**
- Software code/processes/community reaches boundaries in its development that prevent improvement, growth and adoption
- Providing the expertise and services needed to negotiate to the next stage
- Programmes, events, policy and tools to support the community developing and using research software
- **We advocate for all things Research Software**



bit.ly/BetterSoftwareTshirt

Teams

Software

Helping the community to develop software that meets the needs of reliable, reproducible, and reusable research

Policy

Collecting evidence on and promoting the place of software in research & sharing with stakeholders

Outreach

Exploiting our platform to enable engagement, delivery & uptake

Training

Delivering essential software skills to researchers, partnering with institutions, doctoral schools and the community

Community

Developing Communities of Practice by supporting the right people to understand and address topical issues

Activities

Software

75+ project consultancies
200+ evaluations
4 surgeries

Policy

1500+ RSEs engaged
Involved in UKRI long-term strategy
On 29 national and international committees

Outreach

170+ external contributors
20k unique visitors/month
7.5k followers (Twitter)

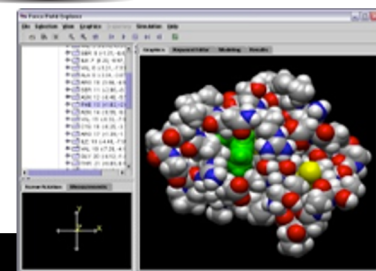
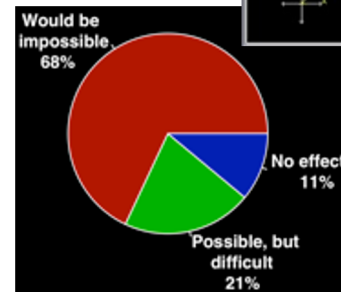
Training

300+ Carpentry workshops
7000+ learners, 250+ instructors
80+ guides

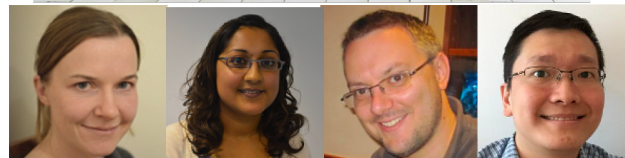
Community

140+ Fellows
30+ workshops organised

The
“7/10”



www.software.ac.uk



Methods of Engagement

Goal	Method
Engagement with domains	Fellowship Programme, (nascent) Support for CoP's
Helping research software grow	Evaluation, Consultancy, Open Call
Capability and Capacity building	Carpentries courses, instructor training, guides
Networking & cross-pollination	Collaborations Workshop, other events
Recognising Software Roles in research	Research Software Engineering Society, RSEConf
Partnering and outreach	Partnering to deliver policy on - software citation, careers, sustainability Website: blog, news, articles, social media.

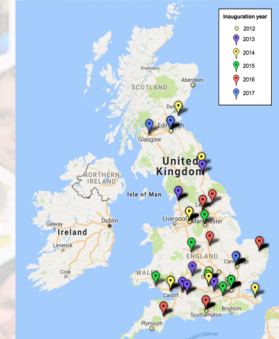


Fourth conference of
Research
Software
Engineers



Fellowship Programme

- <https://www.software.ac.uk/fellowship-programme>
- £3000 over 15 months
- Attend and/or organise research software events that are domain specific and/or cross cutting – esp, Communities of Practice related.
- All Career stages from PhD student to Professor
- 141 Fellows, running annually since 2012
- Competitive – Fellows 2021 applications open Dec 2020
 - five applications per place
- Evaluation of the Programme:
 - <https://f1000research.com/articles/7-1599/v1> (2012-2016 77 surveyed, 26 responses):
 - Over 90% said it benefitted them
 - Over 70% said it benefitted their careers
 - Personal: community, recognition, skills
 - Domains: tailored training, highlighting importance of sustainable practice
 - Institutions: Seeding local RSE groups/function, improved practice/interdisciplinarity



f1000Research F1000Research 2018, 7:1599 Last updated: 17 MAY 2019

[Check for updates](#)

RESEARCH ARTICLE
Raising the status of software in research: A survey-based evaluation of the Software Sustainability Institute Fellowship Programme [version 1; peer review: 3 approved with reservations]

Shoab Sufi, Caroline Jay
School of Computer Science, University of Manchester, Manchester, Greater Manchester, M13 9PL, UK

V1 First published: 03 Oct 2016, 7:1599 (<https://doi.org/10.21956/f1000research.10221.1>)
Latest published: 03 Oct 2016, 7:1599 (<https://doi.org/10.21956/f1000research.10221.1>)

Abstract
Background: This paper reports the results of an evaluation of the Software Sustainability Institute's Fellowship Programme, which focused on identifying and categorising the benefits that the fellowship has afforded its recipients, via a series of open questions.
Methods: The evaluation took the form of a survey open to people awarded Fellowships between 2012 and 2016, which asked people to report the effect that the programme had had on them, their institutions, their research domains and their careers.
Results: The results show that the Fellowship plays a wide-ranging role in supporting communities of best practice and skills transfer, and that a significant benefit is the way it has raised the profile of software in research, and those people who develop and advocate for it.
Conclusions: The evaluation of the programme has shown the need to support research software in situ and credit the engineers and researchers who are working in this important area that supports reproducibility, reuse and the integrity of research investments.

Keywords: research software, fellowship

Open Peer Review

Reviewer Status	1	2	3
Invited Reviewers	?	?	?
version 1	?	?	?
approved	?	?	?
reservations	?	?	?

1. Colin C. Venters University of Huddersfield, Huddersfield, UK
2. Dan Shuler University of California, Berkeley, Berkeley, USA
3. Lisa Carlshaw McInnes, Argonne National Laboratory, Argonne, USA

Any reports and responses or comments on the article can be found at the end of the article.

This article is included in the Science Policy Research gateway.

Page 1 of 18

Open Call for Projects

- 30 open call project since 2012
 - 3 applications per place
 - 4.5 months per project on average
- Helped projects with
 - Correctness
 - Usability
 - Maintainability
 - Architecture / scalability
 - Development Processes
 - Code Management
 - Open Development strategy

*“Help with setting up and testing the training material has been **fundamental to the role of MONC to the community ... I found SSI to be extremely efficient and accurate in their work**”*
- Adrian Hill, MONC

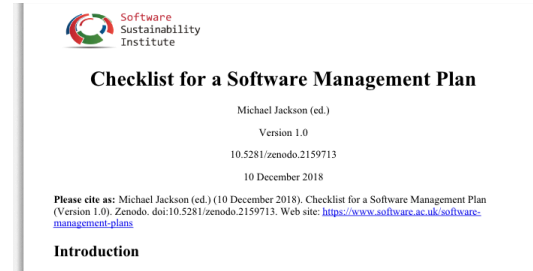
www.software.ac.uk/open-call-projects

Completed projects

- **LUX-ZEPLIN**: We provided consultancy to LUX-ZEPLIN's researchers at University College London on migrating data storage and analysis software from Microsoft Excel to a database-centred solution. Read the **case study**.
- **Synote**: We've worked with Synote to improve their current development processes and infrastructure, and helped develop general best practice guidelines for development, deployment and testing.
- **ionomicsHUB**: The Institute has assisted in the process of migrating the iHUB plant genomics service from Purdue University in the US to the University of Nottingham in the UK, to ensure the continued use and sustainability of the service for the community.
- **recipy**: We've worked with the developers of recipy - an innovative way to track provenance in Python - to develop an automated test suite for recipy as a precursor to expanding the development of recipy and promoting recipy more widely.
- **DataSHIELD**: We've helped improve the DataSHIELD software, provided advice on software development processes, reviewed its project governance, and helped investigate funding support models.
- **Mantid**: We've conducted development and community reviews of the Mantid neutron and muon data analysis software to feed into a high-level review of the project to shape its future development.
- **ReproPhylo**: The Institute helped to make the software more sustainable by conducting a technical review of the software and testing infrastructure, and assisted in the exploration of ways to implement extensibility to interface with user tools.
- **Clipper**: We have conducted a sustainability assessment of the Clipper video and audio clip sharing software, documentation and collaboration processes, to help enhance its longer-term maintenance.
- **MONC - prompting adoption and ongoing development**: we worked with EPCC and the Met Office to promote the uptake, and ongoing development, of the Met Office NERC cloud (MONC) model within the atmospheric sciences community. We assessed how easy it is to deploy MONC, helped set up a MONC virtual machine and advised on setting up resources for engaging with and supporting researchers. Read the **case study**.
- **APES - improving performance and interoperability of chemistry models**: together with the EPCC, the Institute helped to enable key computational chemistry software packages to work together and run on supercomputers such as HECToR (and its upcoming successor ARCHER) to improve modeling performance.
- **CGPACK**: We've helped to improve the CGPACK microstructure simulations library for new users and developers, automating its build process and helping to design a regression test suite to ensure results correctness. We also assessed the overall sustainability of the software.
- **RTI-VIPS**: We have dramatically simplified the deployment process for their Reflectance Transformation Imaging software, packaging these complex components into an installer designed for use by researchers in the humanities.

Software Management Plans

<https://www.software.ac.uk/software-management-plans>



What you Plan to do

Software Evaluation

Software evaluation guide:

<https://www.software.ac.uk/resources/guides-everything/software-evaluation-guide>

Online sustainability evaluation:

<https://www.software.ac.uk/resources/online-sustainability-evaluation>

- Fill in form (15 mins)
- Receive tailored advice
- Building
- Installing
- Testing
- Documentation
- Support
- Portability
- Contributor policy
- Copyright
- Licenses

What you are doing

Q1 - What your software does

Question 1.1: Does your website and documentation provide a clear, high-level overview of your software? *

- Yes
 No

Question 1.2: Does your website and documentation clearly describe the type of user who should use your software? *

- Yes
 No

Question 1.3: Do you publish case studies to show how your software has been used by yourself and others? *

- Yes
 No

Q2 - Your project's and software's identity

Question 2.1: Is the name of your project/software unique? *

- Yes
 No

Q3 - Availability of your software

Question 3.1: Is your software available as a package that can be deployed without building it? *

- Yes
 No

Question 3.2: Is your software available for free? *

- Yes
 No

Question 3.3: Is your source code publicly available to download, either as a downloadable bundle or via access to a source code repository? *

- Yes
 No

Question 3.4: Is your software hosted in an established, third-party repository like GitHub (<https://github.com>), BitBucket (<https://bitbucket.org>), LaunchPad (<https://launchpad.net>) or SourceForge (<https://sourceforge.net>)? *

- Yes
 No

Capability building

Collaboration with the Carpentries - an international collaboration and community around creating and delivering open training materials. Delivering essential software and data skills to researchers for conducting reproducible and reliable research through running **workshops** and publishing **training guides and tutorials**.

Capacity building

Through “instructor training” programme and Centres for Doctoral Training. Support for spreading the Carpentries training to a broad range of domains.

Community support

Support for existing and emerging training communities and make them sustainable (e.g. Library and High Performance Computing (HPC), Social Sciences Carpentries, RSEs, Fellows). Run training community events (CarpentryConnect). Seed financial support for institutional Carpentry membership and low level admin and workshop coordination support.

Training



THE
CARPENTRIES

**80+ training guides
viewed by 250K+ people**



Software
Sustainability
Institute

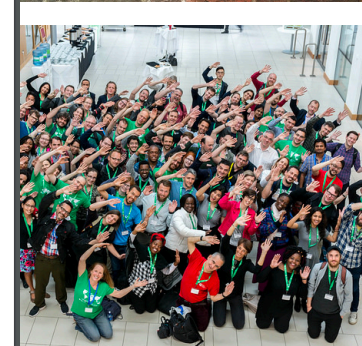
***Helped train 7000+
researchers and 250
instructors through
300+ workshops at
over 40 UK institutions***

Most popular guides

1. Software Evaluation Guide (over 65k unique visits)
2. Choosing a repository for your software project (over 50k unique visits)
3. How to cite and describe software (over 25k unique visits)
4. Developing maintainable software (over 25k unique visits)
5. In which journals should I publish my software (over 22k unique visits)

Events

- Collaborations Workshop: <http://www.software.ac.uk/cw>
 - Premier event / unconference since, since 2009
 - Researchers, Developers, funders, publishers, trainers, project leaders
 - Themed: Culture Change, Productivity and Sustainability, interoperability, training
- Exploring and discussing topical matters
 - Software and licensing, Software and Credit, Research Data Visualisation, Docker – www.software.ac.uk/workshops
- Bringing experts together
 - Measuring the Impact of Workshops meeting
 - <https://www.software.ac.uk/miw>
 - Workshop on Sustainable Software (WOSS19)
 - <https://www.software.ac.uk/wosss19>
- Enhancing Community
 - Carpentry Connect, Manchester 2019
 - <https://www.software.ac.uk/ccmcr19>



Research Software Engineer



Software Sustainability Institute

Those who write software in a research context who identify more with the software than the research (more or less)

Started in 2012 as an idea at CW12

1. Recognition
2. A career path, albeit still informal
3. A growth in the number of software experts
4. Greater access to skills for researchers

210,000

UK researchers

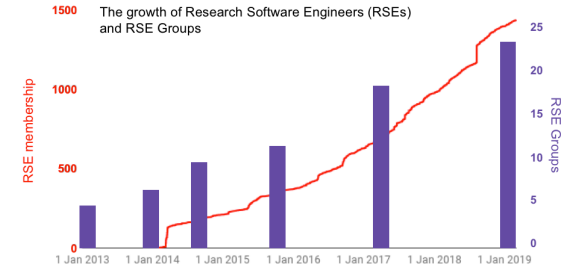
~150,000

researchers need help

115

researchers per RSE

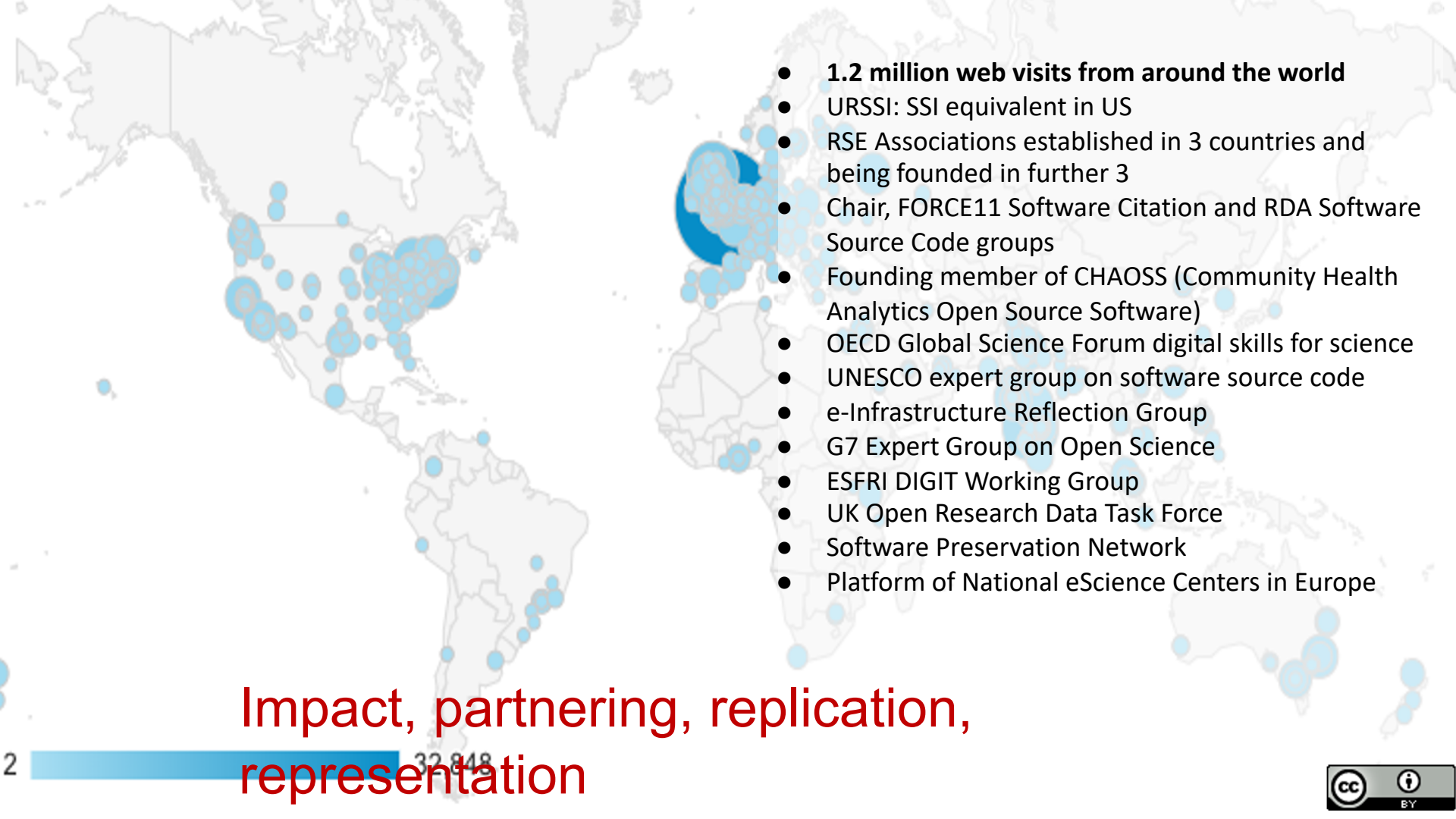
More RSE capacity needed



RSE Slack channel has 1400+ people on



Research Software Engineer Fellowships II



- **1.2 million web visits from around the world**
- URSSI: SSI equivalent in US
- RSE Associations established in 3 countries and being founded in further 3
- Chair, FORCE11 Software Citation and RDA Software Source Code groups
- Founding member of CHAOSS (Community Health Analytics Open Source Software)
- OECD Global Science Forum digital skills for science
- UNESCO expert group on software source code
- e-Infrastructure Reflection Group
- G7 Expert Group on Open Science
- ESFRI DIGIT Working Group
- UK Open Research Data Task Force
- Software Preservation Network
- Platform of National eScience Centers in Europe

Impact, partnering, replication,
representation

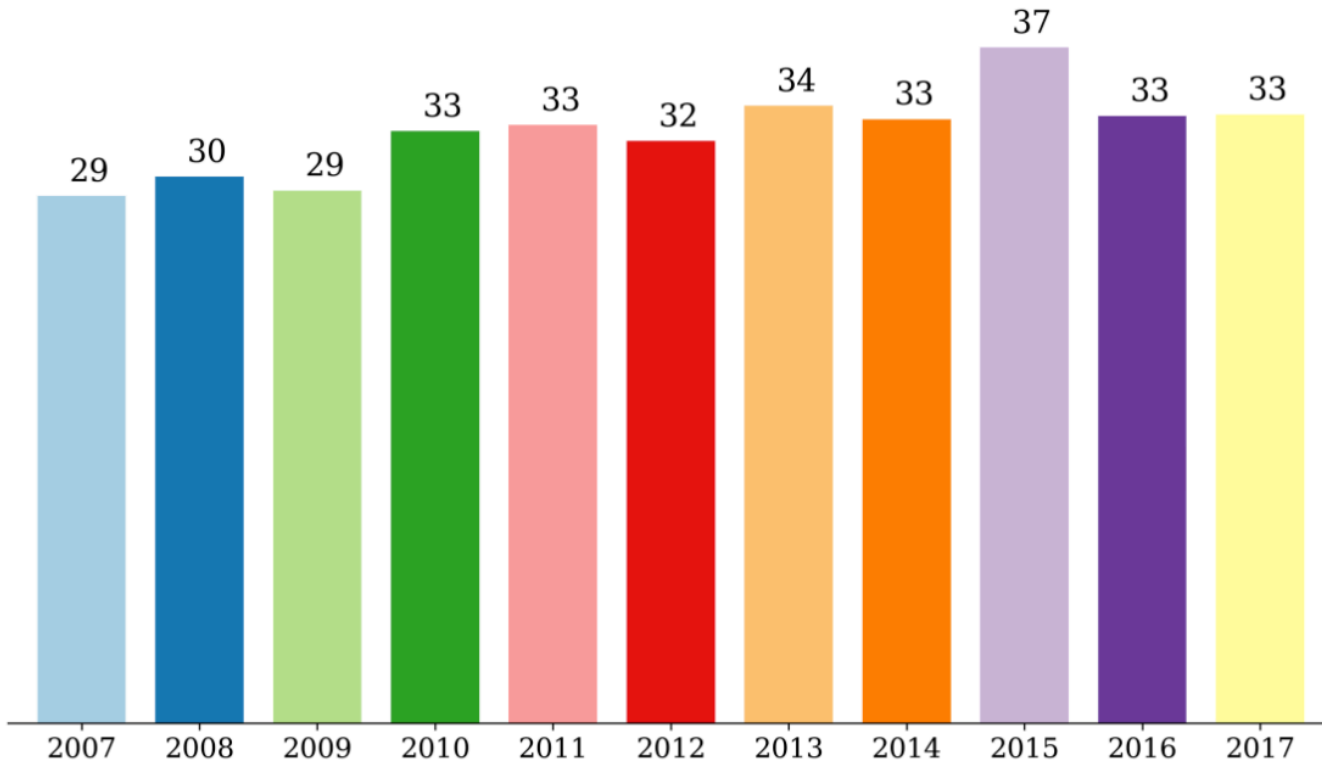


Promoting Software as a Research Output

Percentage of cross-council funding in software-reliant research



www.software.ac.uk



Clearly software is important to research but it's not always recognised

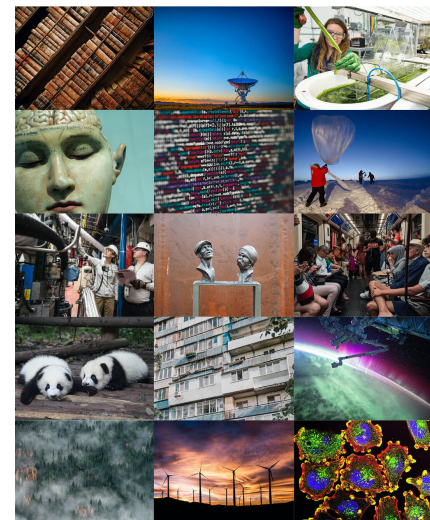
Celebrating all research outputs -

hidden
REF



- REF overlooks people that are vital to the success of research
- 97% of the submissions are publications – other outputs are accepted – artefact of risk averse strategy
- Hidden REF – recognize people & outputs that are often missed in research evaluation
- Submissions are 300 words, need to be at a research institution - collect and announce categories → submissions open → winners
- You can help! – via social media, nominating yourself for the panel, suggest a new category

hidden-ref.org



Recognising those vital to research, advances research.

Why are better practices important?



Growth in a Time of Debt

	B	C	I	J	K	L	M
2			Real GDP growth				
3			Debt/GDP				
4	Country	Coverage	30 or less	30 to 60	60 to 90	90 or above	30 or less
26			3.7	3.0	3.5	1.7	5.5
27	Minimum		1.6	0.3	1.3	-1.8	0.8
28	Maximum		5.4	4.9	10.2	3.6	13.3
29							
30	US	1946-2009	n.a.	3.4	3.3	-2.0	n.a.
31	UK	1946-2009	n.a.	2.4	2.5	2.4	n.a.
32	Sweden	1946-2009	3.6	2.9	2.7	n.a.	6.3
33	Spain	1946-2009	1.5	3.4	4.2	n.a.	9.9
34	Portugal	1952-2009	4.8	2.5	0.3	n.a.	7.9
35	New Zealand	1948-2009	2.5	2.9	3.9	-7.9	2.6
36	Netherlands	1956-2009	4.1	2.7	1.1	n.a.	6.4
37	Norway	1947-2009	3.4	5.1	n.a.	n.a.	5.4
38	Japan	1946-2009	7.0	4.0	1.0	0.7	7.0
39	Italy	1951-2009	5.4	2.1	1.8	1.0	5.6
40	Ireland	1948-2009	4.4	4.5	4.0	2.4	2.9
41	Greece	1970-2009	4.0	0.3	2.7	2.9	13.3
42	Germany	1946-2009	3.9	0.9	n.a.	n.a.	3.2
43	France	1949-2009	4.9	2.7	3.0	n.a.	5.2
44	Finland	1946-2009	3.8	2.4	5.5	n.a.	7.0
45	Denmark	1950-2009	3.5	1.7	2.4	n.a.	5.6
46	Canada	1951-2009	1.9	3.6	4.1	n.a.	2.2
47	Belgium	1947-2009	n.a.	4.2	3.1	2.6	n.a.
48	Austria	1948-2009	5.2	3.3	-3.8	n.a.	5.7
49	Australia	1951-2009	3.2	4.9	4.0	n.a.	5.9
50							
51			4.1	2.8	2.8	=AVERAGE(L30:L44)	

"All I can hope is that future historians note that one of the core empirical points providing the intellectual foundation for the global move to austerity in the early 2010s was based on someone accidentally not updating a row formula in Excel"

- Mike Konczal, Roosevelt Institute



Communities of Practice



“Groups of people who share a concern or a passion for something they do and learn how to do it better as they interact regularly.” Jean Lave and Etienne Wenger in 1991 and further elaborated in 1998

”Three components are required in order to be a CoP: (1) the **domain**, (2) the **community**, and (3) the **practice**.”

Reference - <https://www.learning-theories.com/communities-of-practice-lave-and-wenger.html>

The areas of tailored support for CoPs



www.software.ac.uk

- State of Software Skills
- Training needs
- Collaborations needs

(remember – it's early days and things will develop)



Are these things always happening

- For analysis, new models, new tools:
 - Version Control
 - Reproducibility
 - Testing
 - Documentation (e.g. for handover)
 - Clear Licensing
- This is about culture not just methods/programming

- Assess Software Skills

- The survey that you have been sent is based on work we have previously done
- Understand where the focus needs to be
- E.g. an earlier survey and report by the Computational Acoustics SIG reported Feb 2020 by Amelia Gully:

Main challenges

- **Skills:** need technical skills combined with domain-specific knowledge.
- **Software:** steep learning curves, specific requirements, lack of new approaches, “quirks”.
- **Technology:** computational expense, calibration and validation, complex simulation problems.

Potential solutions

- **Knowledge base:** for validation & benchmarking, material properties, database of experts.
- **Training:** resources for specific methods and software, data and code management, work with ECRs...
- **Standards:** improved interoperability.

And
then ...

Training Needs



www.software.ac.uk

- What do ECR's need
 - Surveying the specifics of where they are at and what they need
 - What curricula need developing
- Are there open formats for sharing curricula
- How can we develop from learners to experts – is there a path
- What are the first steps in making this happen
 - Are Software and Data Carpentry appropriate formats – is there a need for data handling training and automation skills enhancement
 - Are toolboxes needed (not code but problem/solution how'tos) - i.e. what was previously mentioned as a knowledge base – how do we go about making one?



- Collaborations needs
 - survey and/or explore what they are
 - General problems
 - Are the right people talking together
 - problem holder / solution providers
 - what about mutual credit vs 'free effort'
 - how about interoperability
 - Collecting lessons learned

Conclusion



www.software.ac.uk

- Have a long history of working with and developing CoP's
- We are moving in the direction of helping advise CoP's wrt to Research Software perspectives

Acknowledgements

The SSI team/*alumni*:

- Aleksandra Nenadic
- *Aleksandra Pawlik*
- *Alexander Hay*
- *Arno Proeme*
- Carole Goble
- Claire Wyatt
- Clem Hadfield
- Dave De Roure
- *Devasena Prasad*
- Giacomo Peru
- Graeme Smith
- *Iain Emsley*
- Jacalyn Laird
- James Graham
- John Robinson
- Les Carr
- *Malcolm Atkinson*
- *Malcolm Illingworth*

- Mario Antonioletti
- Mark Parsons
- Mike Jackson
- Olivier Philippe
- *Priyanka Singh*
- Rachael Ainsworth
- *Raniere Silva*
- *Rob Baxter*
- *Robin Wilson*
- Sam Mangham
- Selina Aragon
- Shoab Sufi
- Simon Hettrick
- Stephen Crouch
- *Tim Parkinson*
- *Toni Collis*
- *Plus the SSI Fellows and RSE community*

Scientific software:

- Dan Katz
- Heather Piowowar
- James Howison
- Jeff Carver
- Jennifer Schopf
- Kaitlin Thaney
- Martin Fenner
- Victoria Stodden
- Patrick Aerts
- WSSSPE community
- WOSSS Community

Software/Data Carpentry

- Greg Wilson
- Jonah Duckles
- Tracy Teal
- Instructor Community

Supported by the UK Research Councils through grants EP/H043160/1, EP/N006410/1 and EP/S021779/1 .
Additional project funding received from Jisc.





Questions