www.software.ac.uk

We have to keep an open mind

Slides: https://doi.org/10.6084/m9.figshare.9885398

20th September 2019, FOSS4GUK, Edinburgh

Neil Chue Hong (@npch), Software Sustainability Institute

ORCID: 0000-0002-8876-7606 | N.ChueHong@software.ac.uk

Supported by:

BBSRC

EPSRC Engineering and Physical Sciences Research Council

Arts & Humanities Research Council

E·S·R·C Economic & Social Research Council

MRC Medical Research Council

NERC Science of the Environment

Science & Technology Facilities Council

Slides licensed under CC-BY where indicated

Why is open good?

Why is open difficult?

Slides: https://doi.org/10.6084/m9.figshare.9885398

I've spent 20 years working on projects that rely on open

diarhea
diereah
dyereeah
diarrheah
dioreah
diarrhoea

DOI:10.1016/j.ijmedinf.2010.04.002

Open == fun! #UKSnow EH8 0/10

http://www.jstor.org/stable/25704679

Photo: NASA/GSFC

2009: Norwich

Warren, Price, Goswami (2010). http://adsabs.harvard.edu/abs/2010AGUFMPA33A1603W

LIVE BBC NEWS CHANNEL

Page last updated at 14:13 GMT, Friday, 20 November 2009

E-mail this to a friend Printable version

Hackers target leading climate rese

The e-mail system of one of the world's leading climate research units has been breached by hackers.

Experts warn that universities' e-mail systems are vulnerable to attacks

E-mails reportedly from the University of East Anglia's Climatic Research Unit (CRU), including personal exchanges, appeared on the internet on Thursday.

A university spokesman confirmed the email system had been hacked and that information was taken and published without permission.

An investigation was underway and the police had been informed, he added.

"We are aware that information from a server used for research information in one area of the university has been made available on public websites," the spokesman stated.

"Because of the volume of this information we cannot currently confirm that all of this material is genuine.

"This information has been obtained and published without our permission and we took immediate action to remove the server in question from operation.

"ClimateGate"

"Scientists need to take steps to make available all the data that support their work and full methodological workings, including their computer codes."
- Phil Willis MP, Inquiry Chairman

"Easily-identified bugs, no visible test mechanism, [code that] was not apparently under version control and was poorly documented".
- John Graham-Cumming

By not making code and data open, it couldn’t be scrutinised and improved.

So why is open not the norm?

Barriers to sharing

Code		Data
77%	Time to document and clean up	54%
52%	Dealing with questions from users	34%
44%	Not receiving attribution	42%
40%	Possibility of patents	-
34%	Legal Barriers (ie. copyright)	41%
-	Time to verify release with admin	38%
30%	Potential loss of future publications	35%
30%	Competitors may get an advantage	33%
20%	Web/disk space limitations	29%

Victoria Stodden, http://www.stodden.net/AMP2011/

"I'm sorry, but our computer code was not written with an eye toward distributing for other people to use. The codes are not documented and we don't have the time or resources to document them."

"When you approach a PI for the source codes and raw data, you better explain who you are, whom you work for, why you need the data and what you are going to do with it."

"I have to say that this is a very unusual request without any explanation! Please ask your supervisor to send me an email with a detailed, and I mean detailed, explanation."

Stodden, Seiler, Ma (2018). https://doi.org/10.1073/pnas.1708290115

Science AAAS

“This particular project was something I wrote a couple years ago to help me out with a workflow… **I’d put it up on Github**, so that others could potentially use it or use the code.

So I went to **see what people were saying** about this project. It seemed liked I’d done something **fundamentally wrong**, so stupid that it flabbergasts someone…
So of course **I start sobbing**.

Then I see these people’s follower count, and I sob harder. I can’t help but think of potential future employers that are no longer potential.”

http://www.software.ac.uk/blog/2013-01-25-haters-gonna-hate-why-you-shouldnt-be-ashamed-releasing-your-code

Software Sustainability Institute (est 2010)

Why is some software successful, and some isn‘t?

https://www.software.ac.uk/blog/2019-09-12-making-open-source-research-software-visible-path-better-sustainability

Arguing for sustainability

The argument for contributing to open source is that it's not a public good but a common-pool resource because it takes some effort to maintain the resource so everyone can benefit

Allison Randal
Ex-President, OSI
allisonrandal.com

Elinor Ostrom
Donella Meadows
Nadia Eghbal

	Excludable	Non-excludable
Rivalrous	**Private goods** food, clothing, cars, parking spaces	**Common-pool resources** fish stocks, timber, coal
Non-rivalrous	**Club goods** cinemas, private parks, satellite television	**Public goods** free-to-air television, air, national defense

https://www.rhyslindmark.com/club-goods-digital-infrastructure-and-blockchains

Photo by Ales Krivec on Unsplash

Software is a secret garden

Photo by Paul Hanaoka on Unsplash

Photo by Denny Müller on Unsplash

SSI Fellowship Programme 2020

Become an ambassador of
good software practice

https://www.software.ac.uk/programmes-and-events/fellowship-programme

Geospatial Data Curriculum

This workshop is co-developed with the National Ecological Observatory Network managing and understanding spatial data formats, understanding coordinate refere R for analysis and visualization.

Join the geospatial curriculum email list to get updates and be involved in conversa

Interested in teaching these materials? We have an onboarding video and accompa these lessons. After watching this video, please contact team@carpentries.org so t tor. Instructors who have completed onboarding will be given priority status for tea workshops.

DATA CARPENTRY

BUILDING COMMUNITIES TEACHING UNIVERSAL DATA LITERACY

THE CARPENTRIES

Lessons

Lesson	Site	Repository	Reference	Instructor Notes	Maintainer(s)
Geospatial Workshop Overview					Arthur Endsley, Anne Fouilloux, Jeff Hollister, Stace Maples, Chris Prener, Joseph Stachelek, Michael Sumner, Michele Tobias, Leah Wasser
Introduction to Geospatial Concepts					Chris Prener, Tyson Swetnam
Introduction to R for Geospatial Data					Lachlan Deer, Juan Fung
Introduction to Geospatial Raster and Vector Data with R					Lauren O'Brien, Joseph Stachelek, Jane Wyngaard

https://datacarpentry.org/lessons/#geospatial-curriculum

https://github.com/carpentries-incubator/geospatial-python

Open is good
Open is difficult

We have to talk about it more
(and not just with each other)

To keep an open mind:

1. Acknowledge your biases
2. Be kind to others
   and to yourself
3. Celebrate openness!

Slides: https://doi.org/10.6084/m9.figshare.9885398

Acknowledgements

The SSI team/*alumni*:
- Aleksandra Nenadic
- *Aleksandra Pawlik*
- *Alexander Hay*
- *Arno Proeme*
- Carole Goble
- Caroline Jay
- Claire Wyatt
- Clem Hadfield
- Dave De Roure
- *Devasena Prasad*
- Giacomo Peru
- Graeme Smith
- *Iain Emsley*
- 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*
- Shoaib 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
- WSSSPE community

The Carpentries:
- Greg Wilson
- Jonah Duckles
- Tracy Teal
- Instructor Community

Additional Input:
- Allison Randal
- Dr Greg
- Tom Russell

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

CC BY

About the Institute

Software Sustainability Institute

www.software.ac.uk

A national facility for cultivating better, more sustainable, research software to enable world-class research

- Software reaches boundaries in its development cycle that prevent improvement, growth and adoption
- Providing the expertise and services needed to negotiate to the next stage
- Developing the policy and tools to support the community developing and using research software

Supported by the UK Research Councils through grants EP/H043160/1, EP/N006410/1 and EP/S021779/1

CC BY

Software

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

Training

Delivering essential software skills to researchers via CDTs, institutions & doctoral schools

Outreach

Exploiting our platform to enable engagement, delivery & uptake

Collecting evidence on the community's software use & sharing with stakeholders

Policy

Bringing together the right people to understand and address topical issues

Community

Software

Consultancy

50+ projects

Advice

130+ evaluations
4 surgeries

Training

Courses

35+ UK SWC workshops
1000+ learners

Guides

80+ guides
50,000 readers

Outreach

Website & blog

150+ contributed articles
20,000 unique visitors per month
3,000 Twitter followers

Research

740 researchers
50,000 grants analysed

300+ RSEs engaged

Campaigns

BETTER SOFTWARE BETTER RESEARCH

2100 signatures

13 issues highlighted

Policy

Journal of open research software

Workshops

20+ workshops organised

Fellowship

61 domain ambassadors

Community

Ideas to change the culture

1. Persuade your organisation to associate its brand with open source software (inc scientific software)
2. Promote working on *existing* OSS projects as part of UG and Masters programs, so it's seen as valuable career development for students and staff

CC BY

Ideas to change the culture

3. Agree a set %age of Research Software Engineering group time to contribute to OSS important to your organisation
4. Ask your information services (IT services, library) department to implement subscriptions to the most significant research software your organisation relies on

CC BY

Take home messages

www.software.ac.uk

- Sustainability of research software should focus on *research* as the primary stakeholder, but…
- Many decision-makers for research software so…
- We need to shift the way that open source research software development is *perceived*
  - Celebrate OSS contributions
  - Recognise contributing to OSS as career development
  - Acceptance of new models to contribute to maintenance and support of software

Slides: https://doi.org/10.6084/m9.figshare.8986499

CC BY

Additional References

www.software.ac.uk

- Allison Randal - Open source influences on technology innovation: https://www.youtube.com/watch?v=f4nCMVfNqqM
- Rhys Lindmark – Club Goods, Digital Infrastructure, and Blockchains: https://www.rhyslindmark.com/club-goods-digital-infrastructure-and-blockchains/
- Juan Nunez-Iglesias – Why citations are not enough for open source software: https://ilovesymposia.com/2019/05/28/why-citations-are-not-enough-for-open-source-software/
- Tidelift – 2018 Professional Open Source Survey https://tidelift.com/about/2018-Tidelift-professional-open-source-survey-results
- Robert A Muenchen – The Popularity of Data Science Software: http://r4stats.com/articles/popularity/

CC BY