Computing & Software work recognition & citations in publications

Eduardo Rodrigues
University of Cincinnati

LHCb Tuesday Meeting, CERN, 15th Jan. 2019

What would you say are the core "products" of academic research? Most people, when asked this question, talk about research papers, trained scientists, books and perhaps even data. But this list misses a critical component of much of the research being done today: software.

https://physicsworld.com/a/why-we-should-give-credit-to-code-creators/

Is C&SW work recognition a hot topic ?

- ❑ VERY MUCH SO !!!
- ❑ Discussions in workshops, articles, blogs, HSF, etc.


IOP Physics World - the member magazine of the Institute of Physics

physicsworld.com

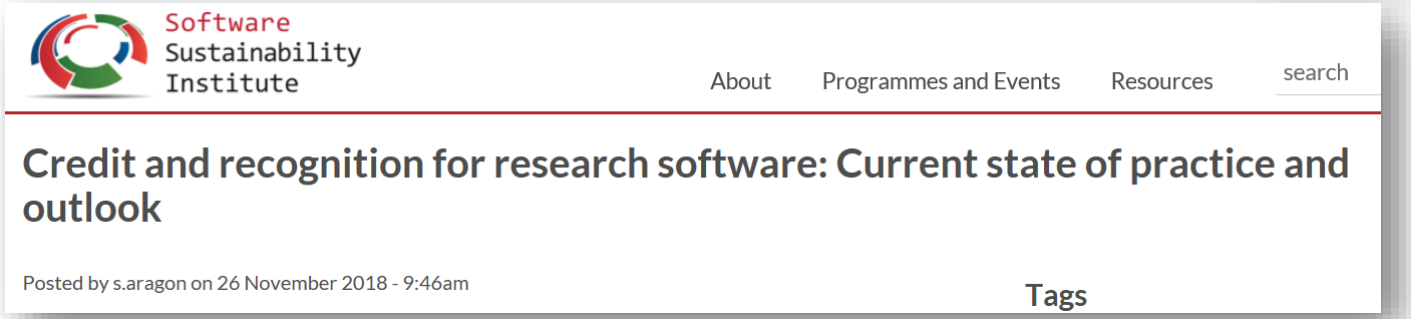
Home | News | Blog | Multimedia | In depth | Events

In depth browse
- Astronomy, astrophysics & cosmology
- Atomic, molecular & optical physics
- Biological & medical

Why we should give credit to code creators

Mar 9, 2017 2 comments

Taken from the March 2017 issue of *Physics World*


Software Sustainability Institute

About Programmes and Events Resources search

Credit and recognition for research software: Current state of practice and outlook

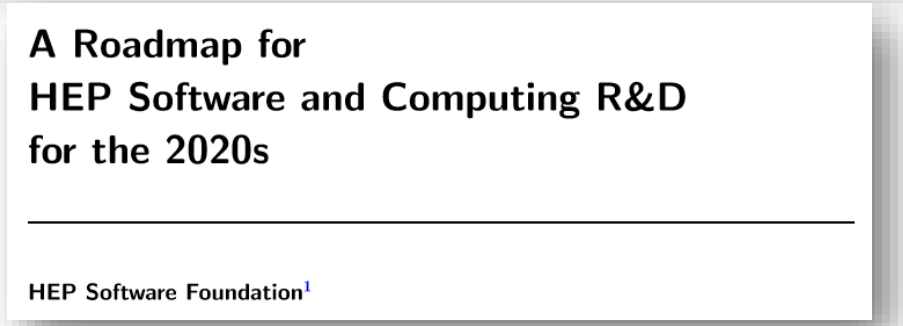
Posted by s.aragon on 26 November 2018 - 9:46am

Tags


September 25, 2017

Software Heritage and repository metadata: a software citation solution

This is a third (and last) in a series of short (ok, the first and second were actually short, even if this one isn't) blogs related to talks and discussions at the 10th RDA Plenary.


A Roadmap for HEP Software and Computing R&D for the 2020s

HEP Software Foundation[1]


Home

Software Sustainability Institute

About Programm

Making Software a First-Class Citizen in Research

Posted by s.aragon on 28 November 2018 - 9:29am

+ whole dedicated paper on careers, etc., see https://arxiv.org/abs/1807.02875

Why is it so ? Quick quiz …

Hardware and analysis work are recognised and properly cited (to a large extent)

Why is it so ? Quick quiz …

Hardware and analysis work are recognised and properly cited (to a large extent)

RIGHT

Why is it so ? Quick quiz …

Hardware and analysis work are recognised and properly cited (to a large extent)

RIGHT

Computing & software work are recognised and properly cited (to a large extent)

Why is it so ? Quick quiz …

Hardware and analysis work are recognised and properly cited (to a large extent)

RIGHT

Computing & software work are recognised and properly cited (to a large extent)

WRONG

Why is it so ? Quick quiz …

Hardware and analysis work are recognised and properly cited (to a large extent)

RIGHT

Computing & software work are recognised and properly cited (to a large extent)

WRONG

Computing & software work should be recognised and properly cited, as any other work

Why is it so ? Quick quiz …

Hardware and analysis work are recognised and properly cited (to a large extent)

RIGHT

Computing & software work are recognised and properly cited (to a large extent)

WRONG

Computing & software work should be recognised and properly cited, as any other work

RIGHT

Why is it so ? Quick quiz …

Hardware and analysis work are recognised and properly cited (to a large extent)

RIGHT

Computing & software work are recognised and properly cited (to a large extent)

WRONG

Computing & software work should be recognised and properly cited, as any other work

RIGHT

Do you disagree with any of the above?

GO BACK AND THINK OVER !

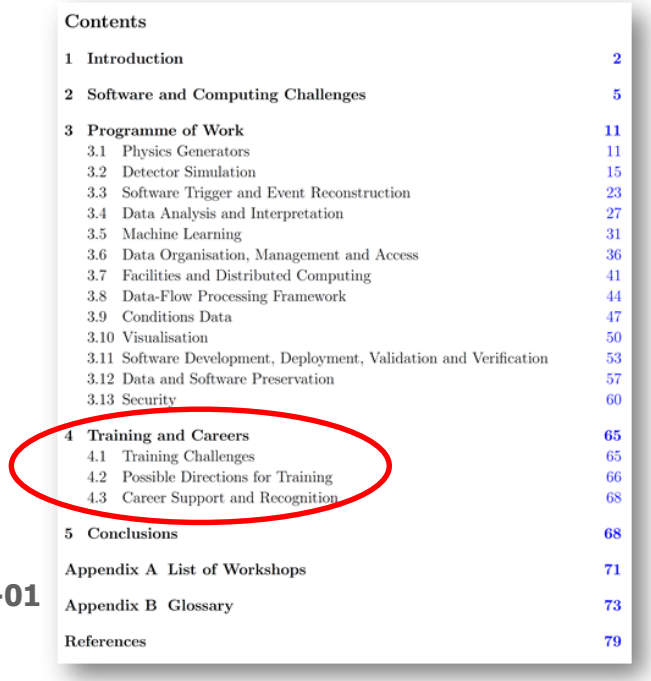
C&SW – career, recognition (and training)

- Career, recognition and training are all hot and important topics in the community
- Software needs to be recognised as another outcome/product of physics research
- System of academic crediting needs to be brought into the XXI century, an era when software is a fundamental part of doing science

> With software development becoming ever more important in physics research, Arfon Smith argues that we need to adopt better ways of recognizing those who contribute to this largely unrewarded activity.

https://physicsworld.com/a/why-we-should-give-credit-to-code-creators/

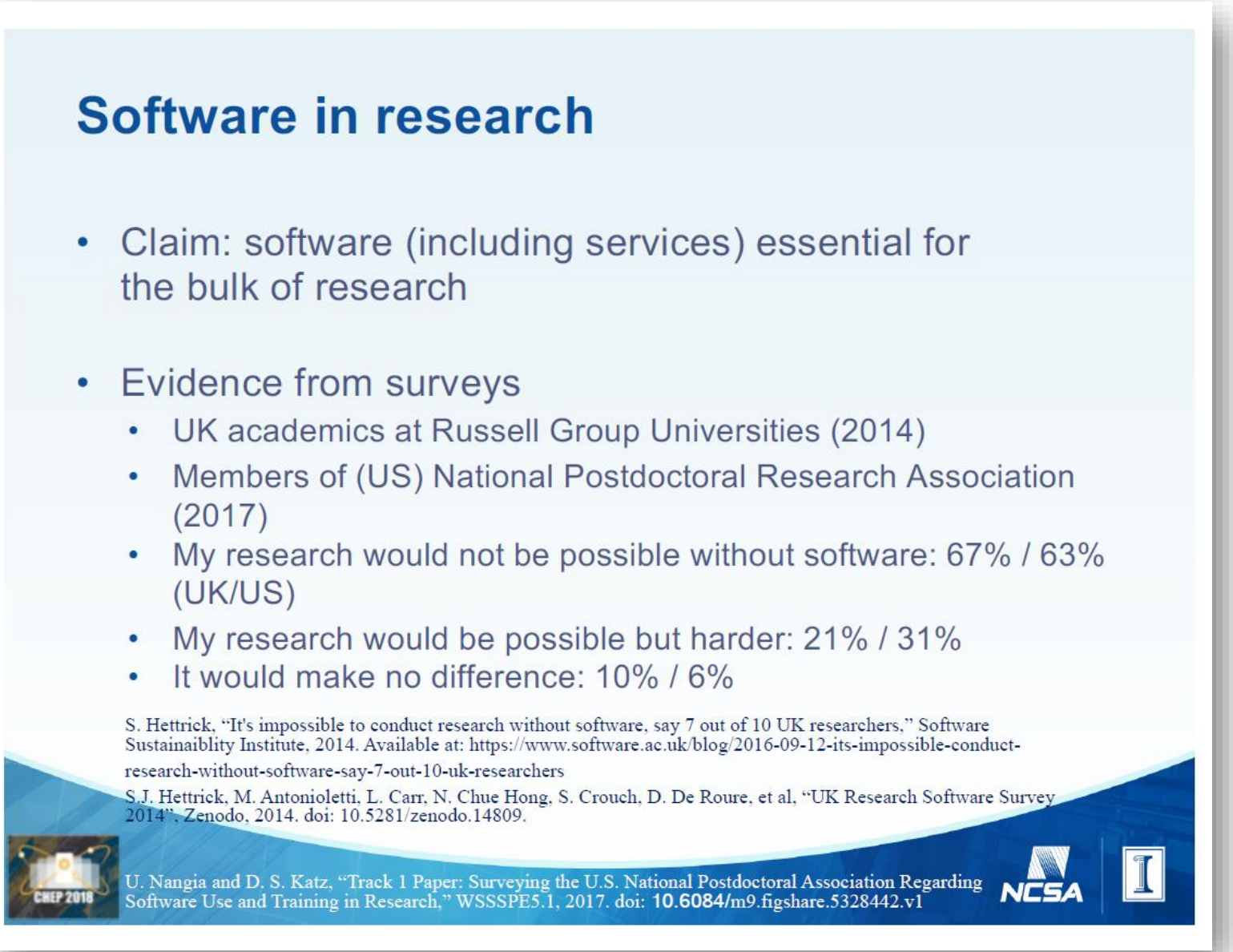
- The HSF Community White Paper
  "A Roadmap for HEP Software and Computing R&D for the 2020s"
  contains a chapter on these matters, which is backed up by a WG paper
  - CWP accepted for publication in CSBS journal
- See HSF page for details and links


Contents

HSF-CWP-2017-01
arXiv: 1712.06982

I want numbers to prove it … !

- Here it goes:
- Slide taken from Dan Katz's presentation at CHEP 2018
- Slides contain other interesting material too
- Worth a read …



If you still aren't convinced there is an issue …

- Take the TMVA package:
- For sure one of the most commonly used software packages in HEP in general and by LHC analysts in particular
- Yet one of its main authors had to leave the field (was in LHCb) because he did not manage to secure a job in HEP
- This is wrong!

- BTW, TMVA is often not cited in (our) papers whereas other software packages such as FastJet are
- This is wrong!

Why should I care ?

- Having better working conditions and clear recognition means is a benefit to all, regardless

- Our “Big Data experiments” depend on high-quality software
- We need to train and retain skillful physicists in C&SW

- (Taken from https://physicsworld.com/a/why-we-should-give-credit-to-code-creators/)
  Individuals most likely to be suffering a career penalty from spending time working on (open-source) software are some of the most employable people outside of academia
- Second, the work these individuals contribute to open source is highly visible, and discoverable, because of the significance of these tools in industry
- Third, with jobs in industry often paying two or three times more than postdoctoral-level salaries, many of the best and brightest young academics are leaving academia for industry

- The issue(s) may hit you personally at some point

Mis-conceptions and wrong/right attitudes – LHCb extracts

How to make C&SW work (more) citable?

- Seems obvious: for work to be citable it needs to be documented in an appropriate format in the 1st place !

- Work can mean many things, including also code documentation, testing, project management, etc.
- Description of research software (packages/applications/products in general) can be via technical notes, papers in relevant and peer-reviewed journals, conference presentations and proceedings
  - This will help research and analysis preservation, BTW

- Software may have dedicated papers (often on the methodology/idea) that are trivially cited via the standard paper citation mechanism, hence providing credit to authors
- But "academic citations not attached to publications are still a new concept and a cultural change is required to make citing software in research papers effective" [ref]
- What if such papers are not available …?
- New types of "journals" have seen the light, which rather focus on the description of research software

- Note: if there's a software paper, one can/should cite that too, but not in place of citing the software itself

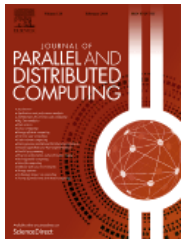
How to make C&SW work (more) citable? Examples of journals

Journals for work reports, ideas, proposals, etc.

- CSBS
- SoftwareX

Presents new concepts for large-scale, collaborative computing and software development for particle, astroparticle, and nuclear physics domains, as well as observational astronomy and cosmology, or high-brilliance light sources.

- Journal of Parallel and Distributed Computing
- Big Data Research



SoftwareX aims to acknowledge the impact of software on today's research practice, and on new scientific discoveries in almost all research domains. *SoftwareX* also aims to stress the importance of the software developers who are, in part, responsible for this impact.

- Springer **also has various journals for computer science & co. …**
- **Want a long list?**

  See In which journals should I publish my software?

  (Much of the material is well beyond physics)

How to make C&SW work (more) citable? Examples of journals

For code

- ❑ The Journal of Open Source Software (JOSS)
- ❑ Journal of Open Research Software (JORS)
- ❑ Zenodo
- ❑ **BTW, in Zenodo one can even set up "communities", which is a simple way to organise things**
  - Examples: the HSF and PyHEP2018 communities

The Journal of Open Source Software

A **developer friendly** journal for research software packages.

Journal of open research software

The *Journal of Open Research Software* (JORS) features peer reviewed Software Metapapers describing research software with high reuse potential. We are working with a number of specialist and institutional repositories to ensure that the associated software is professionally archived, preserved, and is openly available. Equally importantly, the software and the papers will be citable, and reuse will be tracked.

For data

- ❑ Zenodo **(also for proceedings, any immutable digital object)**
- ❑ Nature Scientific Data

SCIENTIFIC DATA

Scientific Data primarily publishes Data Descriptors, a new type of publication that provides detailed descriptions of research datasets, including the methods used to collect the data and technical analyses supporting the quality of the measurements.

zenodo

What can I upload?
All research outputs from all fields of science are welcome. In the upload form you can choose between types of files: publications (book, book section, conference paper, journal article, patent, preprint, report, thesis, technical note, working paper, etc.), posters, presentations, datasets, images (figures, plots, drawings, diagrams, photos), software, videos/audio and interactive materials such as lessons.

How and what is LHCb doing?

What more/else can LHCb do ?

Wrap up

- Do not ignore that recognition of work in C&SW is an issue far from being sorted out
- Recognise somebody else's work the same way you would like yours to be recognised by peers.
  It *does not* matter whether the work is analysis, software or hardware. At least it shouldn't.

- When you write your paper, please be sure to cite the software-related publications
  and software packages you use
- If you produce software that other people use, please make it citable and document how to cite it
- If you review papers, please complain about software that isn't cited