

## Workshop Report

# Collaborations Workshop 2018 (CW18) Report - Culture Change, Productivity and Sustainability

Raniere Gaia Costa da Silva<sup>‡</sup>, Shoaib Ahmed Sufi<sup>‡</sup>, Selina Aragon Camarasa<sup>§</sup>

<sup>‡</sup> The Software Sustainability Institute, School of Computer Science, University of Manchester, Manchester, United Kingdom  
<sup>§</sup> Software Sustainability Institute, EPCC, University of Edinburgh, Edinburgh, United Kingdom

Corresponding author:

Reviewable v1

Received: 02 Oct 2018 | Published: 21 Jan 2019

Citation: Costa da Silva R, Sufi S, Aragon Camarasa S (2019) Collaborations Workshop 2018 (CW18) Report – Culture Change, Productivity and Sustainability. Research Ideas and Outcomes 5: e30250.

<https://doi.org/10.3897/rio.5.e30250>

## Abstract

The Collaborations Workshop 2018 (CW18) took place at The School of Mathematics, Cardiff University from the 26th to 28th March 2018. 90 people attended the event to discuss the themes of the workshop: culture change, productivity and sustainability. With a mix of lightning talks, keynotes, Q&A panels, discussions with speed blogging, collaborative ideas sessions, mini-workshops & demos, a social programme and a Hack Day in the agenda, CW18 was a feature-packed and immersive event.

## Keywords

unconference, hackday, research software

## Keynotes

This year, we had two keynote talks on the workshop. The first one, about culture change, was delivered by [Kirstie Whitaker](#), and [John Hammersley](#) presented the second one about productivity.

## Culture Change

Kirstie Whitaker is a well-known open science advocate who works as a Research Fellow at the Alan Turing Institute in London. Her inspirational talk (Whitaker 2018) touched on some of her personal experiences when trying to convince researchers to do things in the open or change her colleagues' minds during the [University and College Union \(UCU\) 2018 industrial action](#).

Kirstie mentioned three important activities to promote culture change: shape the agenda, mobilise citizens and invest in leaders. All three are important when doing the hard task of “turning around an 800 year-old ship” and you need to engage everyone onboard of the ship.

An important warning included in the talk was that “the idea of letting perfect stand in the way of progress or stand in the way of good is very very very damaging to the open community”. People seeking perfection are important for culture change because they are the ones pushing the ship forward, but everyone must remember to be kind, constructive and respectful of the effort other people are also making towards the change.

## Productivity

John Hammersley is the CEO and co-founder of [Overleaf](#), an online collaborative writing and reviewing tool for LaTeX documents that helps two million users worldwide to be more productive, and one of the sponsors of Collaborations Workshop. John explained that sharing LaTeX source files by email is not the most efficient way to collaborate due to technical barriers, e.g. users having different versions of LaTeX packages and synchronisation constraints – one author needs to wait for the other to send the document back before working on it again. Having an online central place to collaborate on LaTeX documents makes users more productive.

Researchers' productivity is not only measured by the number of articles they write. The number of times one article is cited also plays an important role in the current system of measuring productivity in research. John presented the audience with a chart (Hammersley 2018), see Fig. 1, from The Royal Society that shows the direct correlation between the number of citations and the number of different countries the co-authors are based in. Given that more international co-authors increase the likelihood of citations, we should invest in lowering the barrier to international collaboration. This is not only limited to access to online platforms but also funding to support, build and maintain international collaborators.

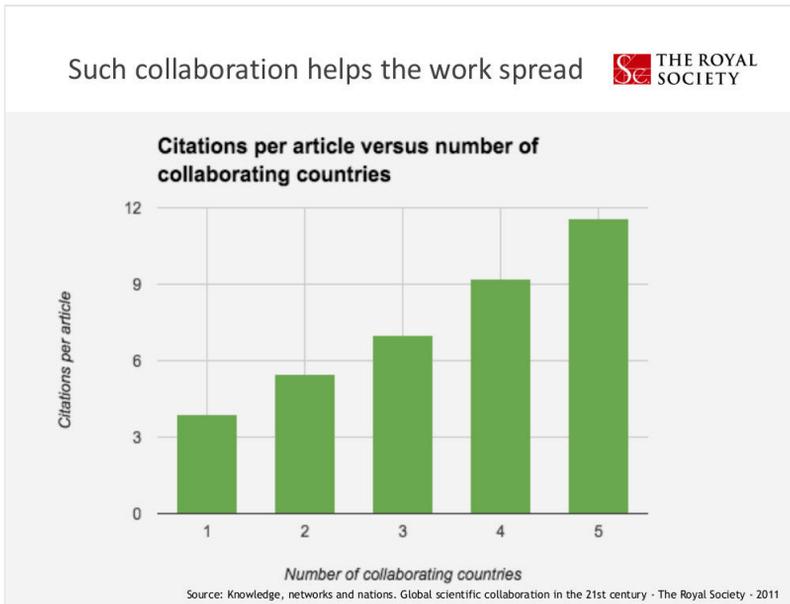


Figure 1. [doi](#)

"Such collaboration helps the work spread" slide from John Hammerley's presentation during Collaborations Workshop 2018.

## Talks on Sustainability Perspectives

There were three short talks covering different facets related to sustainable software in the context of culture change and productivity. We had [Daniel S. Katz](#) (National Center for Supercomputing Applications (NCSA), University of Illinois Urbana-Champaign), [Naomi Penfold](#) (eLife) and [Matthew Upson](#) (former UK Government Digital Service) speaking and representing the views of researchers, publishers and citizens, respectively.

### Researchers' Perspective

Daniel provided updates on the [Workshop on Sustainable Software for Science: Practice and Experiences](#) (WSSSPE) series and the [US Research Software Sustainability Institute](#) (URSSI). WSSSPE is an international community-driven organisation that promotes sustainable research software by addressing challenges related to the full lifecycle of research software through shared learning and community action. Daniel presented a diagram (Katz 2018), see Fig. 2, showing stakeholders of research software and their relationship and how the diagram ended up being covered by the WSSSPE4 working groups and [WSSSPE5.1 speed blogs](#). [WSSSPE6.1](#) will be a one-day workshop during the [IEEE eScience 2018 conference](#), on 29th October 2018 in Amsterdam, and everyone is welcome to attend the event.

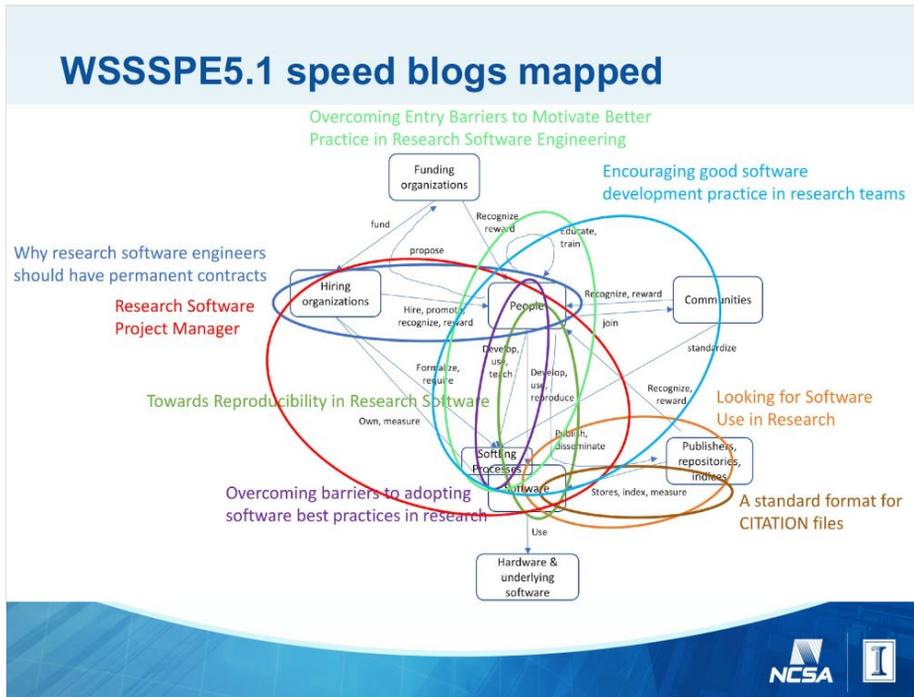


Figure 2. doi

WSSSPE5.1 speed blog mapped slide from Daniel S. Katz's presentation during Collaborations Workshop 2018.

URSSI is a new project being funded by the [National Science Foundation](#) (NSF) focussing on the entire research software ecosystem with plans to encourage a supportive and inclusive community.

### Publishers' Perspective

Naomi's talk (Penfold 2018) focussed on the way research is shared, the barriers that need to be removed, and how sharing contributes to sustainability. She highlighted that "shared data is often not machine-readable" because authors formatted it for human reading, see Fig. 3.

Naomi also talked about literate programming — the inclusion of code and data within the article — as one way to write sustainable research narratives and the challenge to long-term preservation of research code.

### Citizens' Perspective

Matthew started his talk (Upson 2018) by stating that the UK Government is a great open source user and contributor but is not open across all offices. When you promote open source and you have a lot of data, the natural next step would be to make your data openly available, something that the UK Government already does at <https://www.gov.uk/government/statistics>. Matthews showed in one of his slides, see Fig. 4, a simplified version of the pipeline to have the data published in gov.uk where all the manual and hard-to-replicate steps are “hidden”.



Figure 3. [doi](https://doi.org/10.1371/journal.pone.0197098.g004)

Data presentation in spreadsheets slide from Naomi Penfold's presentation during Collaborations Workshop 2018.

In 2016-2017, Matthew worked on his (crazy) "Reproducible Analytical Pipelines" idea (Upson 2017) that is now gaining traction in government and focusses on making the data published at gov.uk reproducible. Matthew's take-home message was that sustainability is a social problem, and, if you want to change the landscape, you need to talk a lot about your "crazy" sustainability ideas to people.

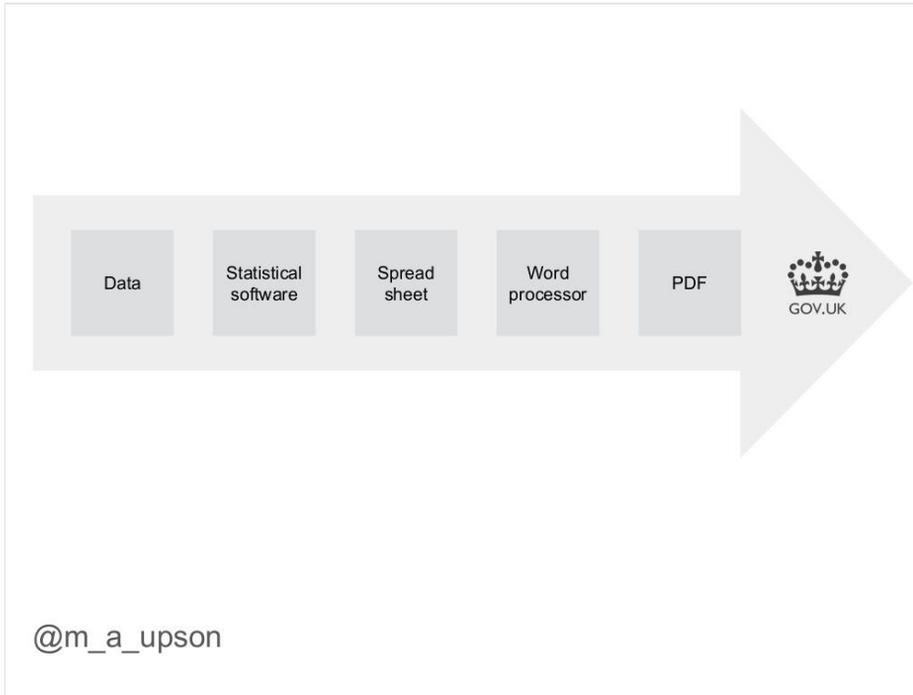


Figure 4. [doi](#)

gov.uk report pipeline slide from Matthew Upson's presentation during Collaborations Workshop 2018.

## Discussion Panels

### Q&A on Culture Change and Productivity

Institute Director, [Neil Chue Hong](#), chaired the panel where [Kirstie Whitaker](#), [John Hammersley](#), [Martin Donnelly](#), [Caroline Jay](#) and [James Baker](#) facilitated the continuation of the conversation started by the keynotes around culture change and productivity.

The organisers expected the productivity debate to expand around practices and softwares that facilitate or deter personal and team productivity – e.g. continuous integration –, but the debate showed that, as Neil Chue Hong summarised, "productivity is defined by you as the person and is something personal to the way that you want to achieve something rather than being something defined as part of a much wider 'this is what the world needs to do'[sic]. The audience contributed with some interesting examples of different perspectives on productivity; for example, not worrying about food because the company where you work provides free catering and the fact that development tools only really work if they are an agreed choice amongst project team members.

## Q&A on Sustainable Software Practice

Neil Chue Hong, Institute Director, chaired this Q&A panel consisting of [Daniel S. Katz](#), [Aly s Brett](#), [Carina Haupt](#), [Louise Brown](#) and [Joseph Parker](#). The panellists explored their personal experience leading Research Software Engineers teams or other groups/organisations that provide training and guidance on sustainable software practice. Carina reminded the audience of Kirstie Whitaker's take-home message: every little helps because incremental progress is still progress.

The panellists mentioned the importance of highlighting success stories – e.g. The Carpentries long-term assessment results (Jordan et al. 2017) showed that people are doing things differently and being more productive. The importance of listening to your users was also highlighted by more than one panellist, focussing on the fact that listening could save many hours of software development.

## Lightning talks

CW18's lightning talks offered a quick (read two minutes) presentation covering a topic from someone very closely involved with the work. Lightning talks are an excellent way to seed potential collaborators and introduce areas of interest to those attending. With 90 attending, it's not possible or desirable that everyone presents a lightning talk. However, we had 32 lightning talks in total. There were an array of talks from pet software projects, simple rules, life experience to future events – all very much in line with the workshop themes. A [scan through the slides](#) (Software Sustainability Institute 2018) is a great way to get up to speed with what was presented.

## Discussion Sessions

These sessions give attendees a chance to discuss matters of interest within the scope of the workshop themes. The possible topics are collected during registration and during the workshop to allow both long-standing and nascent ideas to surface. If questions from previous CW's appear, we add them and the associated blog(s) to the CW FAQ page. Table 1 has the topics available for discussion at CW18. Groups of four to six people had 45 minutes to discuss their topics and then 45 minutes to write a blog post about it. This method of creating outputs from the discussion sessions is called speed blogging.

Table 1.

Topics for discussion during Collaborations Workshop 2018's Discussion Session.

Topic	Sub Topic
Culture Change	What should a culture change case study include?
	How do we influence more senior staff to promote culture change?

Topic	Sub Topic
	<p>How do we develop opportunities for those who are sympathetic to culture change to get more involved?</p> <p>How do you build a community of practice in a domain or across institutions?</p> <p>How do you build a community of practice at an institution (e.g. HPC users)?</p> <p>What type of training do you need to be better at culture change?</p> <p>How do you support giving researchers the time and means to do software development well?</p> <p>How to encourage better software practice (code sharing, documentation) from your colleagues and/or students?</p> <p>Which tools promote better culture in how research software is developed?</p> <p>What kind of data and analysis can help us inform research policy?</p> <p>How to persuade beginners from less traditionally computational research areas that software is important?</p>
Productivity	<p>How to make the most of RSEs' expertise in particular fields?</p> <p>How to manage expectations while juggling different projects with overlapping deadlines and competing resources?</p> <p>How the use of tools for improving productivity and quality in software engineering can be promoted in the wider community?</p> <p>How to handle big data in a productive way?</p> <p>What should a Bot that helps manage pull requests act?</p>
Capability, Careers and Diversity	<p>What progress has there been in RSE career paths in the last year and what challenges remain?</p> <p>How to promote software work in the career path of non-RSEs?</p> <p>RSE negotiating with HR: strategies and case studies</p> <p>How different levels of RSE interact/coordinate with each other?</p> <p>How to balance developing software to support research and participating in research more directly?</p> <p>How to accelerate the process for recognition of RSE roles based in departments?</p> <p>Strategies for establishing national / institutional / departmental RSE groups</p> <p>Incentives for trainers in the style of the Carpentries</p>
Software Sustainability	<p>Sharing Reproducible Research: minimum requirements and desired features</p> <p>Good practices and tools for software management, including plans</p> <p>Good practices and tools for testing scientific code and continuous integration</p>

Topic	Sub Topic
	How to increase the visibility of software and repositories?
	Software Sustainability in Galleries, Libraries, Archives, and Museums
	Where to host standard examples across disciplines?
	Will supporting Machine Learning add to software technical debt?
Others	Minimal requirements for archiving code and software shared with research: journal, author and reuser perspectives
	What are the most popular research software tool categories in computational research? What is missing?
	Developing new software vs discover existing software tools? Research tool directories
	Designing and integrating sustainable software curricula for beginners into research methodology programmes
	Getting external contributors in (open source) research software

Four speed blogs were published as outputs of the discussion sessions: "[Sharing reproducible research - minimum requirements and desirable features](#)", "[Establishing RSE groups at different levels](#)", "[Five failed tests for scientific software](#)" and "[Archiving code and software shared with research: journal, author and re-user perspectives](#)".

## Collaborative Ideas

The collaborative ideas session at CW18 brought together researchers software engineers and people who work in similar domains. We created the groups, and each group had an hour to identify a problem they faced in their work or domains that could use better software and then outline a proposed solution, preferably something they could make progress on in the 24 hours of the Hack Day. There are strong networking and social aspects to the Collaborative Ideas sessions: they are much more about coming up with a Hack Day idea as getting to know and working with people from different institutions, focus and career stages. Table 2 has the list of collaborative ideas that were submitted.

Table 2.

Ideas generated during Collaborations Workshop 2018's Collaborative Ideas session.

Discoverability of software for potential reuse or collaboration is an issue that is common across research domains

Providing evidence for software and data citations in published papers

Can we work together and crowdsource to provide some standard responses and examples?

Automated Software-Assisted Provisioning

Documentation of research software, its use and connections to datasets and narrative research outputs

Interactively explore the similarities and differences between how you can accomplish the same task in different languages
Understand, translate, or even refactor Fortran code
Simplifying the use of JupyterLab as a teaching tool
Code review across projects for software sustainability
Teaching programming to researchers across all research domains
Education and career choices in the absence of relevant or up-to-date information
Finding the right library to fit your project
Time to download the entire dataset and pre-process it locally

After the collaborative ideas session, the ideas are printed, and attendees get to vote on their preferred ideas. The winners were the "Providing evidence for software and data citations in published papers". Runner-ups were "Simplifying the use of JupyterLab as a teaching tool" (second place), "Automated Software-Assisted Provisioning" (third place). Each member of one of the winning teams was awarded an Amazon voucher.

## Mini-workshops & Demos

Mini-workshops and demo sessions are 30-minute parallel sessions that allow more depth on a particular topic, tool, or technique. CW18 hosted mini-workshops and demos on Overleaf, citation file format, scientific code manifesto, (research) software engineering initiatives and testing. In addition to that and in the spirit of being an unconference, [James Baker](#) decided to chair an additional session to look at the newly released [JupyterLab](#).

The full list of mini-workshops and demos is in Table 3.

Table 3. Mini-workshops and Demos Sessions hosted at Collaborations Workshop 2018	
Title	Authors
Overleaf as a collaborative tool for scientific writing	Vincent Knight (Cardiff University)
Making research software easily citable with the Citation File Format	Stephan Druskat (Humboldt-Universität zu Berlin)
Code is Science – open-source scientific code manifesto	Yo Yehudi (University of Cambridge)
Python testing with pytest	Matt Williams (University of Bristol)

The Software Engineering Initiative of the German Aerospace Center	Carina Haupt (German Aerospace Center) and Tobias Schlauch (German Aerospace Center)
A Registry for Research Software Repositories	Alexander Struck (Humboldt-Universitaet zu Berlin)
Building Effective, Sustainable, Research Software Communities	Jeremy Cohen (Imperial College London)

## Hack Day

The CW18 Hack Day (HD) started off on the evening of 27th March with an introduction to the format of this session and an excellent talk by the Institute's Research Software Group lead, [Steve Crouch](#), on technical choices (and how to also have fun). In total, 13 ideas were pitched, see Table 4. The pitches are a way to sell the idea to the audience to let them know what they can work on and which team they can choose to join. By the next morning, we had eight confirmed teams. The judges visited each team in the morning and the afternoon to evaluate their progress and especially assess their teamwork.

Table 4.

Ideas pitched for Collaborations Workshop 2018's Hack Day.

Choose a Visualisation *
Align the Citation File Format and CodeMeta *
A Research Software Dashboard to Help Research Software Grow
The Carpentries platform to support translations *
Parson's problems for mortals
Productivity Approximator *
Keeping up with the Joneses: Providing evidence for software and data citations in published papers *
Electronic lab notebooks done the FOSS way
Code is Science Manifesto *
Geospatial metadata sharing *
Streamlining JupyterLab as a teaching tool
Automated Network Map construction *
Promoting Code Review Within a Research Software Community: Matchmaking and Gamification

\* Ideas developed by one of the teams.

One of the wonderful aspects of the HD is that people have a chance to try something new, learn something new, practise collaborative work, and get something tangible done in a

short space of time often within a new environment and language. Having worked effectively with a new set of people often translates into future collaborations.

At the end of the Hackday, each team had five minutes to explain how they met the judging criteria and show the features and functionality of their new project. Was the entry novel? Was it potentially impactful? Were best practice in infrastructure used? Were the demo and presentation informative? Was their transparency in what had been done and what was the future potential?

It was a challenge for the judges as they were choosing among excellent entries, and all teams put in masterful efforts. The winners were "Productivity Approximator" – each team member won a Raspberry Pi 3 kit + Google AIY Voice Kit + sensor packs full size. Runner-ups were "Keeping up with the Joneses: Providing evidence for software and data citations in published papers" (second place), "Code is Science Manifesto" (third place) and "Carpentry Translation Infrastructure (CTI)" (fourth place).

## Social Programme

In addition to the workshop dinner at Jurys Inn Cardiff, the Institute also organised a few walks around the centre of Cardiff offering a change of scene for those attending and gave them an opportunity to explore local culture.



Figure 5. [doi](#)

Photo group of Collaborations Workshop 2018.

The city centre tour before the workshop dinner was lead by [Nikoleta E. Glynatsi](#), who also facilitated the visit to [Bute Park](#) on the morning of the second day, and to [Alexandra Gardens](#) on the afternoon of the 27th March.

The Collaborations Workshops social programme is beneficial for attendees, see Fig. 5, as it allows them to explore the city which they would otherwise miss while networking or working hard at the event. They get to know more about the place from our local contacts; thus getting a deeper insight into the places where CW takes place.

## Conclusions and Take-Home Messages

Collaborations Workshop 2018 was a great success, described by attendees as "a great opportunity to learn about programming for research" and "the best ideas-sharing workshop around". At the end of three immersive days with many conversations about best practice and tooling, it was clear that there is now an increased focus on interoperability and human factors in software sustainability. Many discussions concluded with the idea that making tools work together would increase productivity, that culture change is inherently about people, and that if we get the human factors right (e.g. providing resources, training and support to allow people to advocate for better software practice), then the change around research software is more likely to be successful.

## Acknowledgements

Many people came together to offer advice, resources and help running the workshop in addition to the Institute staff. Without them (and the rest of the CW delegates), the Collaborations Workshop would not be able to take place.

We extend our thanks to everyone who contributed to making CW18 a successful event:

- [Our sponsors: figshare](#) and [Overleaf](#)
- Our attendees
- [The School of Mathematics, Cardiff University](#) staff
- [Our steering committee](#)
- Our organising committee, events team and [staff](#)
- [Our funders](#): UK Engineering and Physical Sciences Research Council (EPSRC) through grant EP/H043160/1 and EPSRC, BBSRC and ESRC Grant EP/N006410/1.

## References

- Hammersley J (2018) Making things easier to do can have a bigger impact than you might imagine. In: Software Sustainability Institute Collaborations Workshop 2018. <https://doi.org/10.6084/m9.figshare.6081275.v1>
- Jordan K, Teal T, Becker E, Word K (2017) Data Carpentry - Our long-term assessment results are in! <https://datacarpentry.org/blog/2017/10/long-term-survey-results>. Accessed on: 2018-9-25.
- Katz D (2018) Research Software Sustainability: WSSSPE & URSSI. In: Software Sustainability Institute Collaborations Workshop 2018. <https://doi.org/10.6084/m9.figshare.6081248.v1>
- Penfold N (2018) Sustainability in research communication and innovation. In: Software Sustainability Institute Collaborations Workshop 2018. <https://doi.org/10.6084/m9.figshare.6033749.v2>
- Software Sustainability Institute (2018) Collaborations Workshop 2018 - Lightning Talks. <https://doi.org/10.6084/m9.figshare.c.4244906.v1>
- Upson M (2017) Data in government - Reproducible Analytical Pipelines. <https://dataingovernment.blog.gov.uk/2017/03/27/reproducible-analytical-pipeline/>. Accessed on: 2018-9-26.
- Upson M (2018) Effecting Culture Change in a large organisation. In: Software Sustainability Institute Collaborations Workshop 2018. <https://doi.org/10.6084/m9.figshare.6081263.v1>
- Whitaker K (2018) How far we'll go. In: Software Sustainability Institute Collaborations Workshop 2018. <https://doi.org/10.6084/m9.figshare.6025652.v2>