Carnegie Mellon’s first Open Science Symposium – Themes about research data and their reuse

Huajin Wang, Melanie Gainey, Ana Van Gulick
University Libraries, Carnegie Mellon University, Pittsburgh, PA, USA

Background

In October 2018, Carnegie Mellon University (Pittsburgh, PA, USA) hosted its first Open Science Symposium: [https://events.mcs.cmu.edu/oss2018/](https://events.mcs.cmu.edu/oss2018/). The two-day event was a collaboration between data specialists in the University Libraries and life sciences faculty and was funded by a grant for life sciences from the DSF Charitable Foundation. The goal of the event is to build awareness and support for the adoption of open research practices, and encourage innovative ideas about data sharing in the life sciences and related disciplines. Presentations from the event have been documented on Open Science Framework: [https://osf.io/54gue/](https://osf.io/54gue/).

Recurring Themes

- Incentives and standards for data sharing need to come from the ground-up within the disciplines.
- New collaborative methods for data collection and analysis are reshaping the way that science is done.
- Open data ≠ reusable data. Reproducibility and reusability make data valuable, not just open.

Emerging Hot Topics

- Open hardware and software
- Large scale team science / observatory science
- Open framework for citizen science
- National-wide data collection and curation:
  - NCI cancer research commons
  - NCBI Short Read Archive
  - ENCODE DCC
- Different flavors of reproducibility
- Reproducible computing on HPC
- Evolving platforms by publishers:
  - eLife reproducible articles
  - Pre-printing: Pros and cons

Invited talks + Panels (18 speakers):
- Open Science in Research
- Open Data and Reproducibility
- Open Tools and Platforms
- Open Access

Scientific Speed Dating + Reception

Workshops:
- Code Ocean
- Bioconductor
- KiltHub Deposit-a-thon
- BenchSci
- Pittsburgh Supercomputing Center
- ENCODE DCC
- protocols.io

Sessions

DAY 1

- Invited talks + Panels (18 speakers):
  - Open Science in Research
  - Open Data and Reproducibility
  - Open Tools and Platforms
  - Open Access

DAY 2

- Invited talks + Panels (18 speakers):
  - Open Science in Research
  - Open Data and Reproducibility
  - Open Tools and Platforms
  - Open Access

Emerging Hot Topics

- Open hardware and software
- Large scale team science / observatory science
- Open framework for citizen science
- National-wide data collection and curation:
  - NCI cancer research commons
  - NCBI Short Read Archive
  - ENCODE DCC
- Different flavors of reproducibility
- Reproducible computing on HPC
- Evolving platforms by publishers:
  - eLife reproducible articles
  - Pre-printing: Pros and cons

Incentives and standards for data sharing need to come from the ground-up within the disciplines.

Recurring Themes

- New collaborative methods for data collection and analysis are reshaping the way that science is done.
- Open data ≠ reusable data. Reproducibility and reusability make data valuable, not just open.

We thank all speakers for their enthusiasm and contribution; DSF Charitable Foundation and industry sponsors for financial support; Mellon College of Science and University Libraries for administrative support, and Dr. Eric Yttri for co-organizing.