Housekeeping

- Collaborative notes [https://tinyurl.com/6tj2tu8](https://tinyurl.com/6tj2tu8)
- These slides [https://tinyurl.com/3cmwzjhc](https://tinyurl.com/3cmwzjhc)

- Meeting etiquette
  - Add your name to the participants list
  - Add your questions in the chat
  - Raise your hand if you wish to speak
  - Please be aware that the session is being recorded and will be made publicly available
<table>
<thead>
<tr>
<th>time</th>
<th>presentations</th>
<th>lead</th>
</tr>
</thead>
<tbody>
<tr>
<td>00.00</td>
<td>Short <em>introduction</em> to the group and its aims (5 minutes)</td>
<td>Morane</td>
</tr>
<tr>
<td></td>
<td>Overview of other <em>ongoing working groups</em> related to software in RDA (5 mins)</td>
<td></td>
</tr>
<tr>
<td>00.10</td>
<td>Presentation of the <em>EOSC SIRS report</em>: Scholarly infrastructures for research software (25 mins)</td>
<td>Roberto</td>
</tr>
<tr>
<td>00.35</td>
<td>Update of the activity in the <em>FORCE11 Software Citation Implementation Working Group</em> including the ongoing task forces (CodeMeta, journals, repositories...)(15 mins)</td>
<td>Neil</td>
</tr>
<tr>
<td>00.50</td>
<td>A <em>mini-workshop</em> session on code <em>quality</em> assessment (30 mins) Conclusion of the workshop activity (5 mins)</td>
<td>Neil</td>
</tr>
<tr>
<td>01.25</td>
<td><em>Next steps</em> (5 minutes)</td>
<td>Morane</td>
</tr>
</tbody>
</table>
The Software Source Code IG

Co-chairs:
- Neil Chue Hong
- Julia Collins
- Roberto Di Cosmo
- Stepping down: Mingfang Wu

TAB Liaison: Gretchen Greene

VP17 coordinator: Morane Gruenpeter

Objectives:
A forum for discussing research software inside RDA
- issues on management, sharing, discovery, archival and provenance of software source code.
- It will pay special attention to source code that generates research data and plays an important role in scientific publications.

https://www.rd-alliance.org/groups/software-source-code-ig
Why software source code?

“Source code provides a view into the mind of the designer.”
Len Shustek, Computer History Museum

“Go to the code!”

Three pillars of Open Science, Software Heritage CC-By 4.0
2019
Chronology

**BOF RDA P9**, Barcelona April 2017
- motivations => 60 participants

**RDA P10**, Montreal September 2017
- motivations, survey of ontologies, metadata use cases

**RDA P11**, Berlin March 2018
- started the idea for a dedicated identification WG

**RDA P13**, Philadelphia April 2019
- FAIR for Software Source Code and launch of the SCID WG

**FORCE2019**, Edinburgh October 2019
- full day hackathon on Research Software (motivated the CodeMeta generator)

**RDA VP15**, Australia March 2020
- Open discussion about the creation of a new group, the FAIR4RS WG (which was launched in June 2020)

**RDA VP16**, Costa Rica November 2020
- Existing efforts and practices in Academia

**RDA VP17!!!**
Software at RDA and in academia

● Related groups:
  ○ RDA, ReSA and FORCE11 [FAIR for Research Software Working Group](https://www.rd-alliance.org/groups/FAIR-for-Research-Software) (FAIR4RS WG)
    ■ Welcome to join the work defining FAIR principles for research software
  ○ RDA & FORCE11 [Software Source Code Identification WG](https://www.rd-alliance.org/groups/Software-Source-Code-Identification-WG) (SCID IG)
    ■ Output published in September 2020
  ○ FORCE11 [Software Citation Implementation Working Group](https://www.rd-alliance.org/groups/Software-Citation-Implementation-WG) (SCIWG)
    ■ Ongoing WG about software citation

● Related software sessions during VP17:
  ○ [WG - CURE-FAIR: Progress Update](https://www.rd-alliance.org/events/WG-CURE-FAIR-Progress-Update)
A joint RDA Working Group, FORCE11 Working Group, and Research Software Alliance (ReSA) Taskforce.

Coordinating of a range of existing community-led discussions on:
- How to define and effectively apply FAIR principles to research software,
- How to achieve adoption of these principles.

https://www.rd-alliance.org/group/fair-4-research-software-fair4rs-wg/
Ice-breaker question 1

Why are you interested in Software Source Code?

https://tinyurl.com/6tj2tu8
SIRS report presentation
By Roberto Di Cosmo

Ice-breaker question 2

Are you using a scholarly infrastructure for Source Code?

https://tinyurl.com/6tj2tu8
FORCE11
Software Citation Implementation WG update
By Neil Chue Hong
FORCE11 Software Citation Implementation Working Group
(co-chairs: N. Chue Hong, M. Fenner, D. S. Katz)

Following-on from FORCE11 Software Citation Working Group and the Software Citation Principles it developed

Objective: Produce concrete guidelines for software citation, and implement them within the scholarly research community (software developers, repositories and registries, journals and conference and publishers, indexers, institutions)

A community with monthly calls to discuss challenges and progress in implementing software citation, with task forces for

- **Guidance** - developing documents for developers, authors, and reviewer
- **Journals** - coordinating editors and publishers to simplify and implement guidance
- **Repositories** - developing best practices document for handling software
- **CodeMeta** - standardizing metadata for software, moving towards merging into schema.org
Software Citation Principles

Software is a critical part of modern research.

1. Importance
2. Credit and Attribution
3. Unique Identification
4. Persistence
5. Accessibility
6. Specificity


Image courtesy of DataCite
The journey so far...

2015
Software Citation WG started
~55 members (researchers, developers, publishers, repositories, librarians)
Reviewed existing community practices & developed use cases

2015
Software Citation Principles published
Started with data citation principles, updated based on software use cases and related work, working group discussions, community feedback
Software citation principles published after community review: 10.7717/peerj-cs.86

2016
Software Citation Implementation WG started
Group set up to:
1. endorse the principles
2. develop sets of guidelines for implementing the principles
3. help implement the principles
4. test specific implementations of the principles.

2016
Software Citation Implementation WG started
Wrote document to identify software citation implementation challenges: arXiv 1905.08674

2017/18
Task forces publish resources
Codemeta Task Force provides recommendations for schema changes
Repositories Task Force runs workshop to identify best practices
Journals Task Force starts adoption process with journals and publishers to promote The Importance of Software Citation: 10.12688/f1000research.26932.1

2019/20
Adoption of software citation increases?
How does software citation fit with other related work:
FAIR for Research Software
Open Research / Open Science
Reproducibility
Software catalogs
Objective:
Develop guidance for different stakeholders to help implement software citation, principally authors of research articles seeking to cite software correctly and developers of software looking to make their software easier to cite.

Using checklists as an effective way of ensuring consistency and completeness

- See The Checklist Manifesto: How to Get Things Right by Atul Gawande for examples

Milestones and activity:
- Developed and published guidance for paper authors and software developers
  - Software Citation Checklist for Authors
  - Software Citation Checklist for Developers
- Developed Software Citation Primer used as basis for paper by Journals Task Force
- Started Software Citation Checklist for paper reviewers, on hold pending work from Journals Task Force and Code Review WG

Guidance Task Force on hiatus while Journals Task Force provides feedback on guidance
Objective:
Bring together representatives of Research Software Registries and Repositories to discuss and improve practices

Milestones and activity:
- Held two-day workshop (funded by the Sloan Foundation) bringing together managers and editors of scientific software registries and repositories
- Task force completed work and spun out new coalition of registries and repositories to build a community of practice (consortium meets monthly): [https://github.com/scicodes](https://github.com/scicodes)
**Objective:**

Work with organizations that publish journals, proceedings, monographs to improve how software is cited in their works and the scholarly processing ecosystem.

Journals (and conferences) need guidance on what authors should submit regarding the software they use.

**Milestones and activity:**

- **Paper** (co-written with ~20 publishers) providing guidance to journals & conferences published.
- **Publicized in Scholarly Kitchen** and many talks.

**Next steps:**

- Help produce versions with examples and citation styles appropriate for intended audiences.
- Work on what happens after article is submitted – how citations are processed and indexed, working with JATS4R WG.
Objective:
Understand metadata needed to describe software. Update CodeMeta and express all codemeta properties using schema.org

Milestones and activity:
- Reviewed CodeMeta schema and opened issues in the CodeMeta repository with proposals for updating schema
- Validate & integrate proposals in the next CodeMeta release (v3)
- Preparing formal proposal to schema.org
- Planning a webinar in September to showcase existing tooling for generating and using codemeta files.
  - E.g. CodeMeta generator from Software Heritage
    - Hosted version
    - Github repository
Next steps for Software Citation Implementation

- Raised the profile of software citation with many stakeholder groups
  - Lots of good work done, starting to have effects
- Communities need to produce their own versions of the generic guidance and promote and incentivize the use of software citation
- Tracking effects is a research challenge
  - Can we determine if software citation increases over time?
- Other technical and social challenges remain
  - Swimming in a sea of identifier types
  - Where and how metadata should be stored
- And citation is just a part of a larger FAIR, open, and reproducible goal
Mini-workshop: code quality assessment
What is code quality?

Software quality: does the software serve its purpose well?

Code quality: does the code serve its purpose well?
- Readable
- Modifiable

Based on:
https://blog.ploeh.dk/2019/03/04/code-quality-is-not-software-quality/
How do we assess code quality?

We’ll answer a series of questions in parallel, via the collaborative notes document, and reflect on our answers.

1. What is your definition of code quality?

2. What do you look for when assessing the readability of a piece of software source code?

3. What do you look for when assessing whether software source code is modifiable / reusable?

This will help the Code Review Community WG develop guidance for code review during development, and at the time of publication - contact Holly Meunier to get involved in the group.

Collaborative notes: https://tinyurl.com/6tj2tu8
4. Who should/can ensure the **software quality (functionality)**, **code quality (readability)** or the **curation quality (FAIRness?)** of the software? Which roles are responsible?

5. Which type of quality can be ensured by **repository managers**? What should repository managers look for when source code is **deposited** to ensure it is well **curated**?

6. Do we need to **keep humans in the loop**? Can quality be assessed automatically? If so, which type of quality and why?

Collaborative notes: [https://tinyurl.com/6tj2tu8](https://tinyurl.com/6tj2tu8)
Software curation on hal.archives-ouvertes.fr

Detecting extraneous or abusive content (illegal or harassing),

Verifying consistency between the metadata and the software source code itself, and verifying due credit to authors

Completing or correcting the deposit metadata if needed.
Next steps for the SSC IG

Open questions (to answer on the notes):

- What **themes** would you like to discuss during the next plenaries?

- We are looking for **contributions** to the next session? (call for short presentations)

Join us on the mailing list:

https://www.rd-alliance.org/groups/software-source-code-ig
Thanks for joining