

Consortium of scientific software registries and repositories

April 18, 2024 Call

[Previous call notes](#)

Attendees (16:00 UTC call x)

Alain Monteil, INRIA: HAL		Kristin Vanderbilt, IMCR	
Alejandra Gonzalez-Beltran, Science and Technology Facilities Council, UK Research and Innovation		Lorraine Hwang, Computational Infrastructure for Geodynamics	
Alexandros Ioannidis, CERN: Zenodo	x	Maaïke de Jong, Research Software Directory	
Allen Lee, ASU, CoMSES Net	x	Matt Jones, NCEAS	
Ana Trisovic, Dataverse, Harvard		Michael R. Crusoe, CWL ; Debian-Med	
Anita Bandrowski, SciCrunch , UCSD		Mike Hucka, Caltech, SBML	
Bruce Wilson, ORNL DAAC		Morane Gruenpeter, Software Heritage	
Bryce Mecum, NCEAS, UC Santa Barbara: CodeMeta		Moritz Schubotz, swMATH	
Carly Robinson, DOE-OSTI		Olaf Teschke, swMATH	
Carole Goble, University of Manchester, WorkflowHub		Paula Martinez, ReSA, ARDC,	
Celine Sarr, DDMoRe repository		Alice Allen, ASCL	x
Colin Smith, Environmental Data Initiative		Perrine Paul-Gilloteaux, BioImage Information Index	
Daniel Garijo, Ontosoft, MINT	x	Pete Meyer, SBGrid and BioGrids	
David Kennedy, Neuroimaging Informatics Tools and Resources Collaboratory NITRC.org		Peter Teuben, UMD, ASCL	
David Long, (old EiC for RSCL) Yang Du (new RSCL EiC)		Piotr Sliz, SBGrid and BioGrids	
Harry Bhadeshia, Materials Algorithms Project (MAP)		Natalie Barnett, DOE-OSTI	x
Hervé Ménager, Institut Pasteur: ELIXIR bio.tools		Shelley Stall, AGU	
Jeanette M. Sperhac, GHub		Susan Borda, University of Michigan Library	
Joy Ku, SimTK		Ted Carnevale, Neuroscience Department, Yale: ModelDB	

Jurriaan Spaaks, Netherlands eScience Center: Research Software Directory		Tom Morrell, Caltech	x
Kay Graf, OSSR	X	Tom Pollard, MIT: PhysioNet	
Vanessa Sochat, Lawrence Livermore National Lab: The Research Software Encyclopedia and working on Research Software Registry (as an OCI registry)		Alexander Struck (de-RSE, Cluster of Excellence "Matters of Activity")	
Jason Maassen, Netherlands eScience Center: Research Software Directory		Alex Wade Digital Science / Dimensions	
		Thomas Pronk. Amsterdam UMC, NL-RSE, OSNL Open Research Software Advisory Panel	
Maxence Azzouz-Thuderoz, swMATH	X	Philip Durbin, Dataverse, Harvard	x

Guests:

- Rebecca Ringuette (heliophysics at NASA); rebecca.ringuette@nasa.gov
- <https://science.data.nasa.gov/news/software-workshop-2024>

Apologies: Alain

Agenda

● Introductions for new attendees

- Rebecca is interested in controlled vocabularies for programming languages and other scientific software metadata. Examples:
 - <https://bio-tools.github.io/biotoolsSchema/> des nouvelles d'un collègue qui a eu un malaise et a été transporté à l'hôpital. Cela montre que vous vous souciez de son bien-être, et cela peut également être une manière de recueillir des informations sur la situation pour organiser le travail en conséquence.
 - <https://wikidata.org>
 - https://github.com/github-linguist/linguist/blob/master/lib/linguist/language_s.yml
 - At comses we've been treating it like a user entered folksonomy with predictably messy / emergent results

● Recent events:

- CoS: Katz, D. S., Gruenpeter, M., Vrouwenvelder, K., & Garijo, D. (2024, March 21). Progress in including research software in the scholarly communication ecosystem. Year of Open Science Culminating Conference. Zenodo. <https://doi.org/10.5281/zenodo.10849689>
- The 2024 Software Heritage Symposium: <https://www.softwareheritage.org/2024/03/12/software-heritage-symposium-summit-2024-website/>
- [Mining Software Registries](#) (Lisbon, April 15-16)

- **CodeMeta Updates and Crosswalk Check In**
 - Updates from Moranne?
- **Best Practices Self Assessment**
 - [2023-9 Best Practices Self Assessment](#)
 - A few new updates, yay!
 - **PENDING TO DO:** Make the comparison and get mean scores to see if (in general) adoption is more widespread.
- **In-person Workshop (Spring 2025) Planning**
 - Target 75k\$ from Sloan.
 - 25 ppl? (maybe more?)
 - Tentative goals of the workshop (need to cull / refine)
 - Finalize ongoing assessments on adoption of our published best practices for scientific software registries and repositories
 - Assess and identify focal areas for how these best practices should continue to evolve to improve the general state of published scientific and research software for transparency, reuse, interoperability, and reusability and consider publishing an open access article “Revisiting best practices ...”
 - Improve registry and repository metadata interoperability
 - Identify and gather requirements and use cases to guide the development of cyberinfrastructure and tools that reduce the friction of adoption of these best practices, e.g., the Invenio GitHub Archiver¹ or Software Metadata Extraction Framework²
 - Use these requirements to develop software libraries and GitHub apps that assist scientific researchers in publishing their code according to the FAIR4RS Principles and other emerging FAIR efforts like FAIR4ML, etc.
 - Alice: Codemeta harvester so we can search all the registries?
 - Consider hosting a mini symposium to reflect on the state of research software and panel discussion with clear focus (volunteer Dan Katz to be the keynote speaker?)
 - AGU Conference Center in DC
<https://www.agu.org/building/pages/reserve>
 - Archived previous workshop:
<https://asclnet.github.io/SWRegistryWorkshop/>
- **NIH Research Software Best Practices RFI**
<https://grants.nih.gov/grants/guide/notice-files/NOT-OD-24-005.html>
 - **PENDING TO DO:** Alice: will submit to Zenodo
 - Where is the copy that went to NIH?
 - Please post link **HERE:**

¹ <https://github.com/caltechlibrary/iga>

² <https://github.com/KnowledgeCaptureAndDiscovery/somef>

- **PENDING TO DO:** Add it on our website.

Presentation:

N/A

Future Presentations:

To Be Scheduled:

- Tom Pollard: PhysioNet and code review
- Allen: Make Models FAIR (<https://tobefair.org>)
- ~~Melissa Harrison~~: JATS
 - Melissa left her position. Maybe Anita knows someone to present.
- Measuring impact for software (usage, citations) - Supporting software sustainability in any way
- Software dependencies (e.g., what is done in GitHub).
- How arXiv works
- IEEE Code Ocean
- FAIR for AI
- Impact of schema.org metadata in broader research graphs - what's the impact?
- Cite Software initiative and RESA
- RDA FAIR4ML -> Dan, Fotis, Leyla-Jael?
- Google scholar and software indexing
(e.g., https://scholar.google.com/scholar?hl=en&as_sdt=0%2C3&q=comses)

AOB/Open discussion

-

Action Items

- Alice: Alphabetize the table to include new participants.