FAIR Research Objects - CREDiT and Attribuion

Ruth Duerr[1]

[1]Ronin Institute for Independent Scholarship

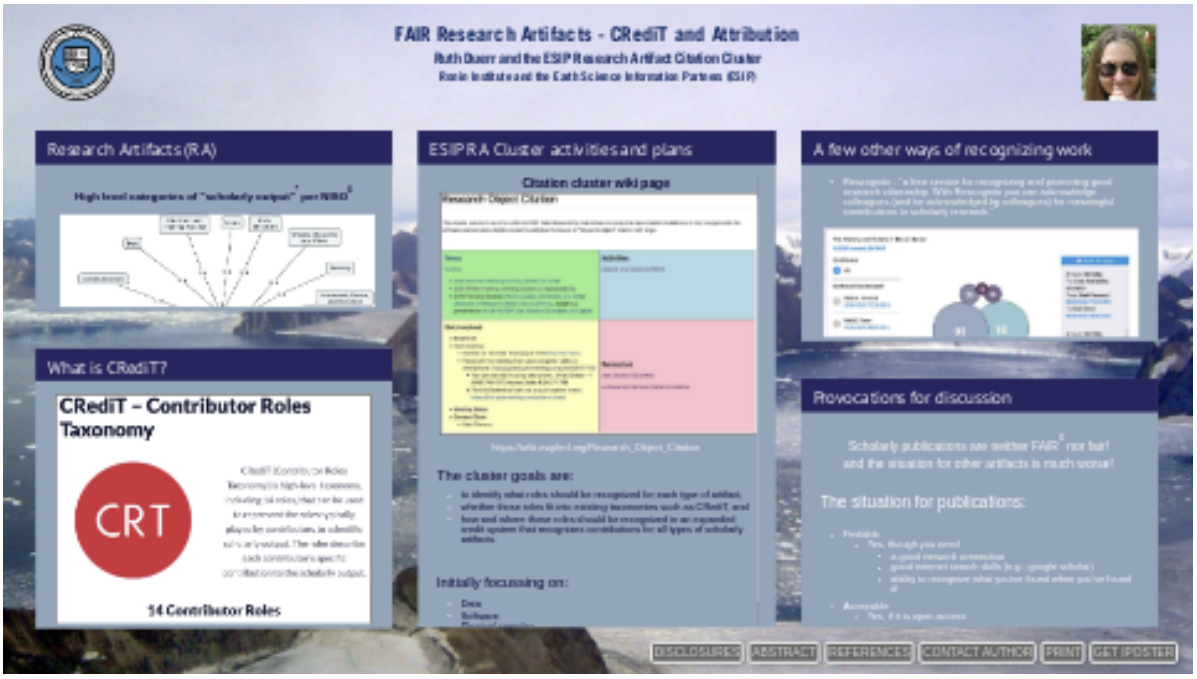
November 23, 2022

Abstract

As noted in the session description any research object produced should be made 'FAIR'; but 'FAIR' objects require attribution. What does that mean for the vast variety of research objects? What roles need to be acknowledged, how and where? These are questions the Federation of Earth Science Data Partners (ESIP) Research Object Citation cluster have been addressing. In particular, we have been applying the lens of the CREDiT taxonomy roles in these discussions. Here our conclusions to date will be described.

Posted on 23 Nov 2022 — CC-BY 4.0 — https://doi.org/10.1002/essoar.10505384.1 — This a preprint and has not been peer reviewed. Data may be preliminary.

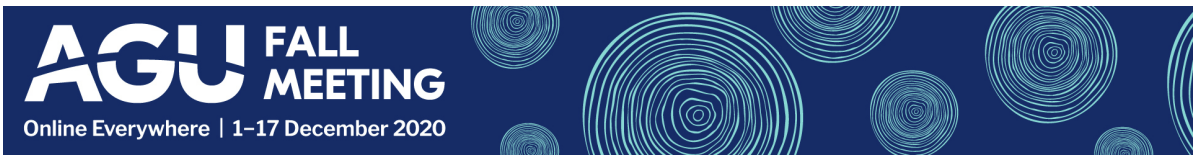
FAIR Research Artifacts - CRediT and Attribution



Ruth Duerr and the ESIP Research Artifact Citation Cluster

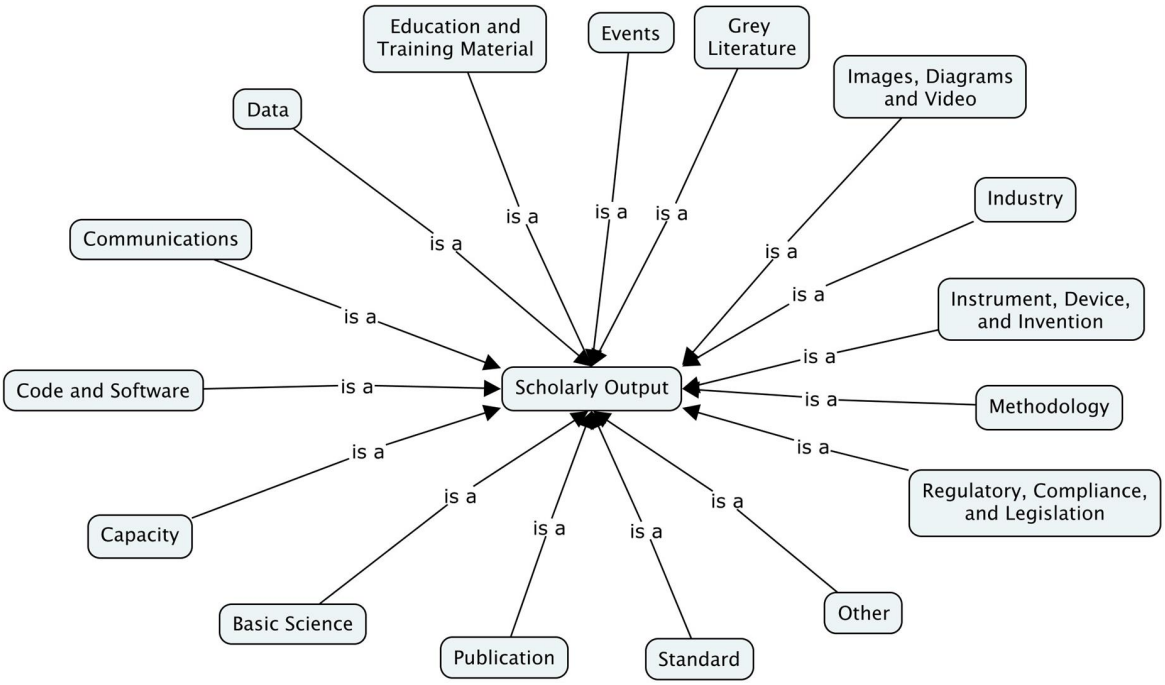
Ronin Institute and the Earth Science Information Partners (ESIP)

PRESENTED AT:



RESEARCH ARTIFACTS (RA)

High level categories of "scholarly output"* per NISO[4]



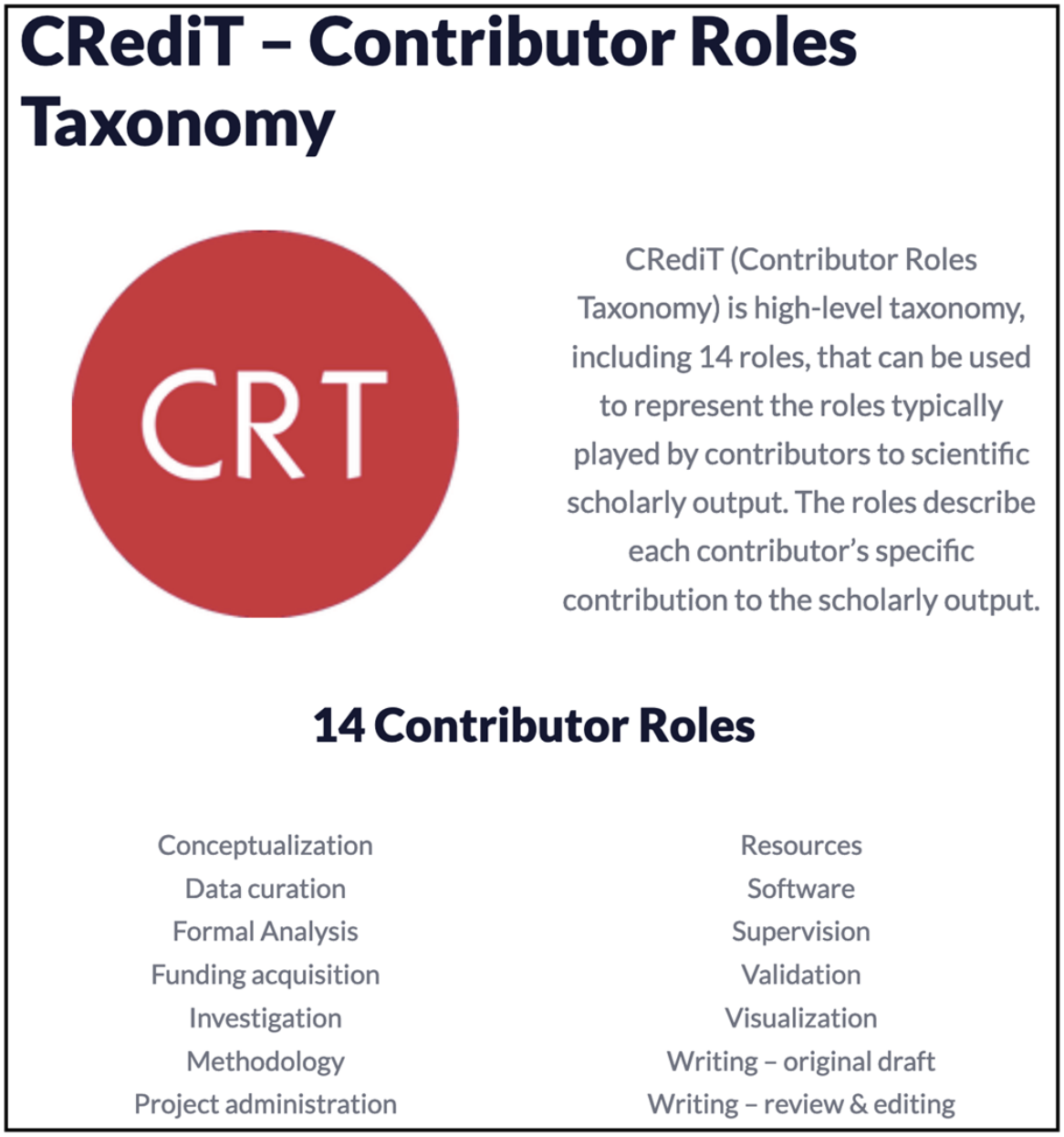
It should be noted that most of these high level categories have several to many subcategories!

Which of these categories actually have the highest impact on the world?

- Industry?
- Regulatory, Compliance, and Legislation?
- Instrument, Device, and Invention?
- Education and Training Material?
- Other?

* Scholarly output are research artifacts!

WHAT IS CREDIT?

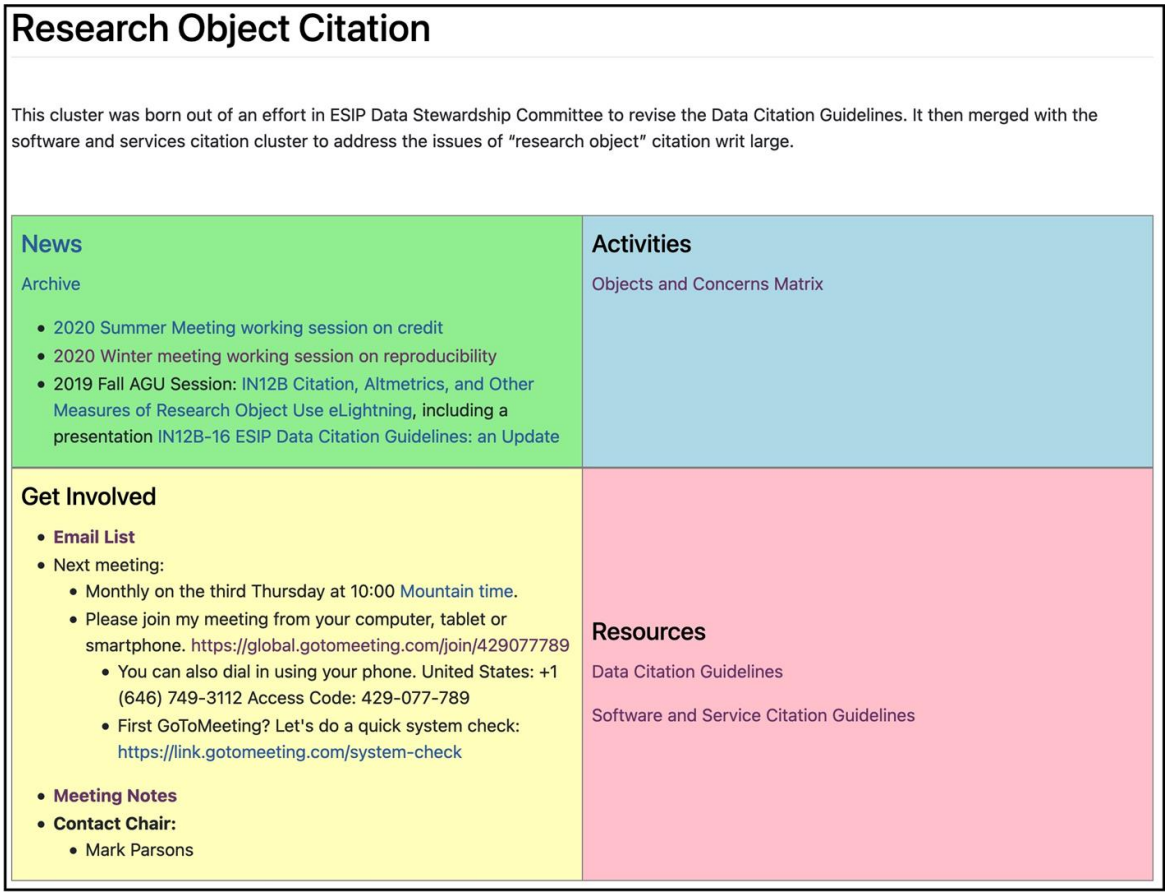


Concerns about CRediT[5]

- Very publication centric!
- Ignores the broader impacts of research
- Can exacerbate negative power dynamics

ESIP RA CLUSTER ACTIVITIES AND PLANS

Citation cluster wiki page (https://wiki.esipfed.org/Research_Object_Citation)



Research Object Citation

This cluster was born out of an effort in ESIP Data Stewardship Committee to revise the Data Citation Guidelines. It then merged with the software and services citation cluster to address the issues of "research object" citation writ large.

News

Archive

- 2020 Summer Meeting working session on credit
- 2020 Winter meeting working session on reproducibility
- 2019 Fall AGU Session: IN12B Citation, Altmetrics, and Other Measures of Research Object Use eLightning, including a presentation IN12B-16 ESIP Data Citation Guidelines: an Update

Activities

Objects and Concerns Matrix

Get Involved

- **Email List**
- Next meeting:
  - Monthly on the third Thursday at 10:00 Mountain time.
  - Please join my meeting from your computer, tablet or smartphone. https://global.gotomeeting.com/join/429077789
    - You can also dial in using your phone. United States: +1 (646) 749-3112 Access Code: 429-077-789
    - First GoToMeeting? Let's do a quick system check: https://link.gotomeeting.com/system-check
- **Meeting Notes**
- **Contact Chair:**
  - Mark Parsons

Resources

Data Citation Guidelines

Software and Service Citation Guidelines

https://wiki.esipfed.org/Research_Object_Citation (https://wiki.esipfed.org/Research_Object_Citation)

The cluster goals are:

- to identify what roles should be recognized for each type of artifact,
- whether those roles fit into existing taxonomies such as CRediT, and
- how and where those roles should be recognized in an expanded credit system that recognizes contributions for all types of scholarly artifacts.

Initially focussing on:

- Data
- Software
- Physical samples
- Semantic resources
- Educational resources

Methods:

- Monthly telecons
  - 3rd Thursday at 12 pm Eastern time

- See the ESIP calendar (https://www.esipfed.org/get-involved/telecon-calendar) for telecon details
- See the cluster wiki (https://wiki.esipfed.org/Research_Object_Citation) for documentation and other links
- Sign up for cluster email (https://lists.esipfed.org/mailman/listinfo/esip-citationguidelines)

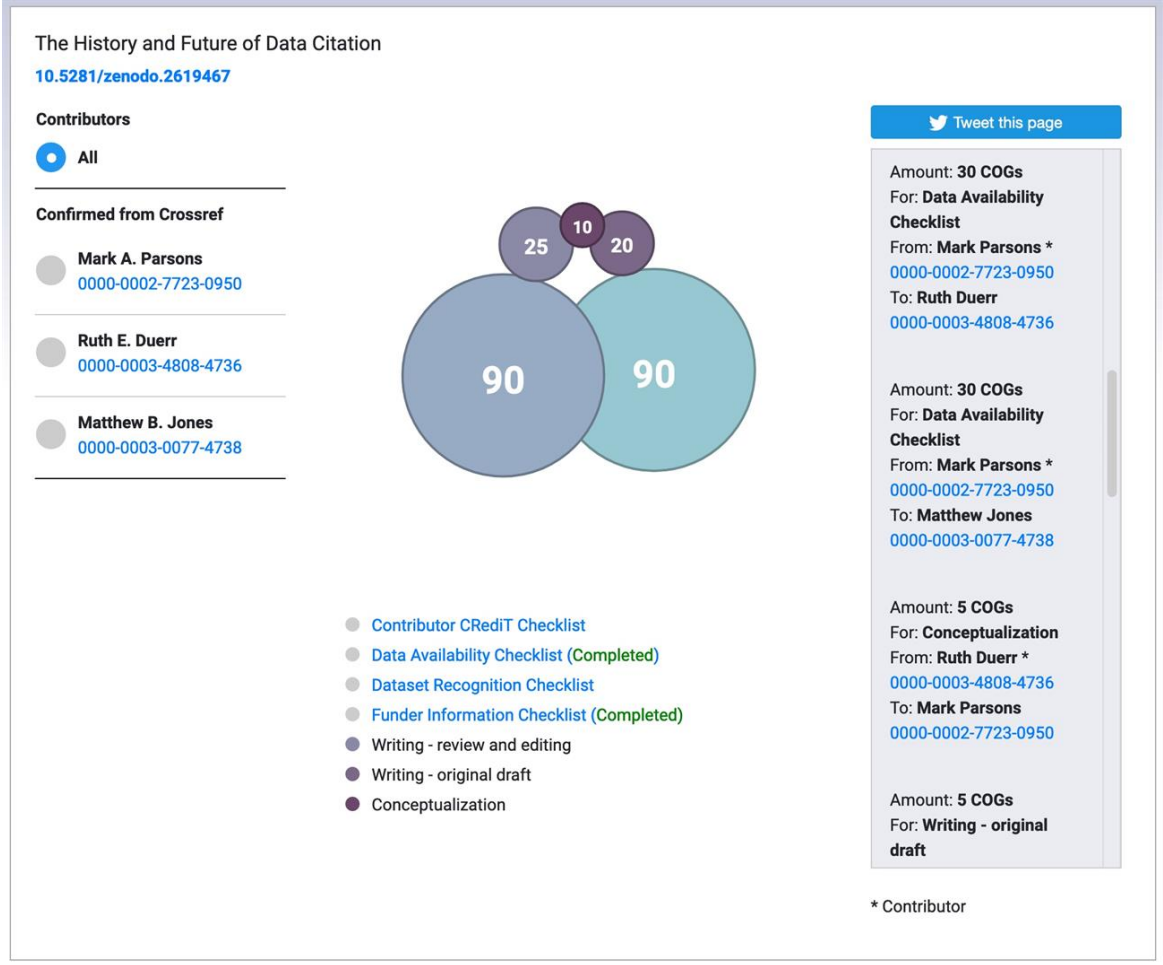
- Twice a year sessions at the ESIP summer/winter meetings
  - ESIP winter 2021 meeting will be held at the end of January
  - Schedule is not officially published yet but registration is open (https://2021esipwintermeeting.sched.com/info)

Activities

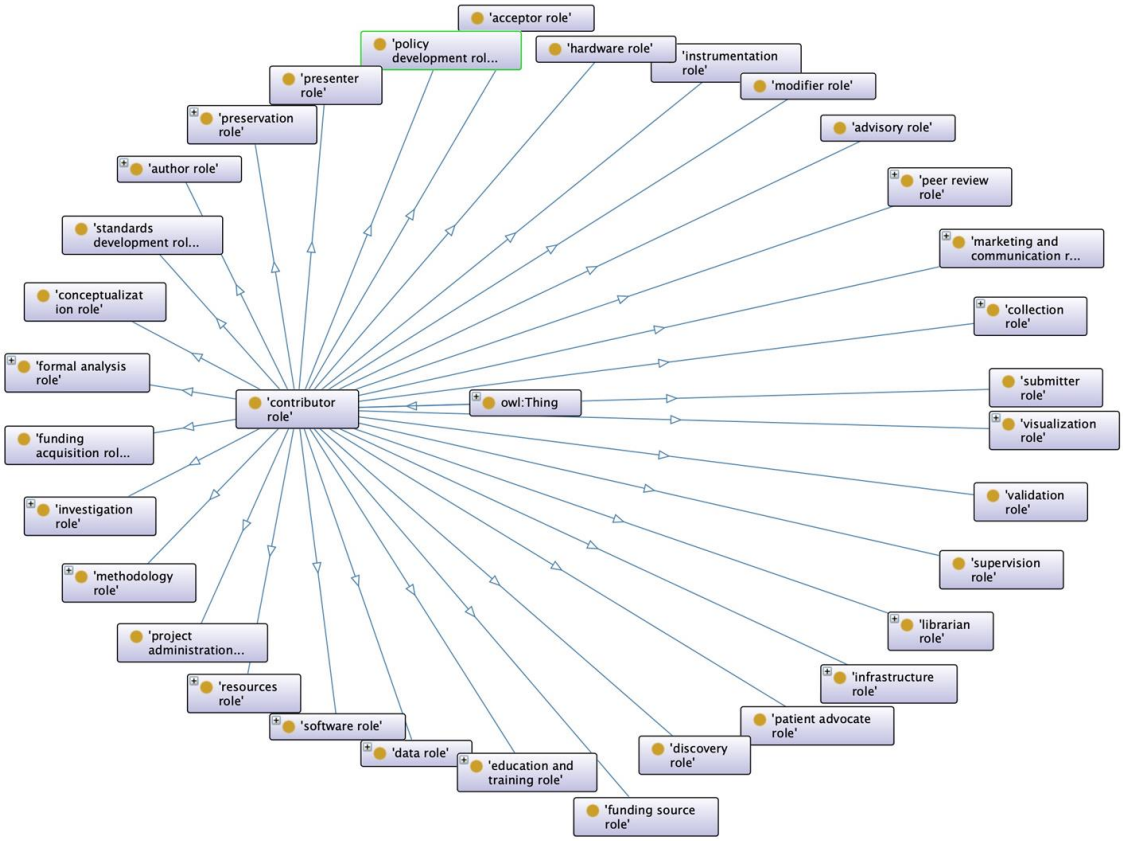
- Initially looked at reproducibility - specifically when or under what circumstances is it necessary to identify an object to enable reproducibility?
- Currently examining the issue of credit using CRediT and ISO 19115 role codes
- Contemplating writing a paper summarizing our results

A FEW OTHER WAYS OF RECOGNIZING WORK

- Rescognito - "a free service for recognizing and promoting good research citizenship. With Rescognito you can acknowledge colleagues (and be acknowledged by colleagues) for meaningful contributions to scholarly research."[1]



- Contributor Roles Ontology[2]



- ISO Role codes[3]

CI_RoleCode

ISO 19115		Added in ISO 19115-1	
+ resourceProvider + custodian + owner + user + distributor + originator	+ pointOfContact + principalInvestigator + processor + publisher + author	+ sponsor + coAuthor + collaborator + editor + mediator	+ rightsHolder + contributor + funder + stakeholder

- Software roles [4]

Designer	Debugger	Maintainer
Coder	Architect	Documentor
Tester	User support	Manager

PROVOCATIONS FOR DISCUSSION

Scholarly publications are neither FAIR[6] nor fair!

and the situation for other artifacts is much worse!

The situation for publications:

- Findable
  - Yes, though you need
    - a good network connection
    - good internet search skills (e.g., google scholar)
    - ability to recognize what you've found when you've found it!
- Accessible
  - Yes, if it is open access
  - Yes, if not open access it is only accessible if you
    - have lots of ready cash
    - work at an organization with a substantial journal budget
    - are willing to pirate a copy
- Interoperable
  - Yes, if available in JATS or similar format
  - No for printed or pdf copies (just try extracting the text out of a bunch of pdf journal articles from 10-20 years of the same journal!)
- Reusable
  - Yes, if it is open access
  - Yes, for fair use
  - No, for most publications with copyright
- fair
  - You mean the public paid for government sponsored research, yet can't generally access the results they paid for?
  - Is it really necessary to write your paper in your disciplines equivalent of Greek?

DISCLOSURES

This work is based on joint efforts of the ESIP Research Artifact Citation Cluster with support from the National Aeronautics and Space Administration (NASA), National Oceanic and Atmospheric Administration (NOAA) and the United States Geologic Survey (USGS).

Transcript of talk:

Hello, my name is Ruth Duerr from the Ronin Institute and today I am presenting some of my own thoughts and much of the work of the ESIP Research Artifact Citation Cluster.

But first, since I seem to be getting deep into the semantic world, let's talk about a few terms!

I don't think I need to say anything about FAIR - Findable, Accessible, Interoperable and Reusable.

But what is a Research Artifact? Well the first thing to note is that this talk originally was entitled FAIR Research Objects not FAIR Research Artifacts! Why was the name changed? Because of concern within the ESIP group (I'll be talking about them more later), concern about the existing definitions of "research object" out there in places like researchobject.org. So we, the cluster, have tentatively decided to talk about artifacts not objects.

But in researching the literature for this talk I ran across the NISO concept of "scholarly output", which for the purposes of this talk I am going to equate to "research artifacts".

This is a concept map of the high level categories of "research artifacts" from the National Information Standards Office. Among the several categories that should be familiar to you (data, software, publications, etc.) there are many things that you may not think of as artifacts - for example, laws, regulations, inventions, whole industries; yet clearly these can be legitimate outputs or artifacts of scholarship!

Moreover, I'd like you to think about which of these actually have the highest real-world impact. I submit that it isn't publications!

Another term in the title of this talk is the term CRediT, which stands for the Contributor Roles Taxonomy. CRediT started off being very publication centric, though its scope is broadening now to some extent. But the question is, whether it properly accounts for all of the contributor roles possible with all of the kinds of scholarly output in NISO's list?

Moreover, there has been some concern within the ESIP cluster about the implicit assumptions underlying some of the terms used - its an issue of power dynamics and inequities.

But CRediT isn't the only way of recognizing work, Rescognito is different in that it is a peer recognition system. Here you see a paper that Mark, Matt, and I worked on and you can see that not only does Rescognito use terms from its own list of roles, but it also uses CRediT terms, something called a data availability checklist, and other checklists. I can imagine adding lists for all kinds of artifacts/outputs.

Below that we have the Contributor Roles Ontology, the ISO Role Codes and a set of Software roles. There is a lot going on!

But enough about terminology! What is ESIP and this Cluster up to?

Well… ESIP has been involved in citation for years. So far it has endorsed guidelines for data and software. But the folks involved recognize that there are many other kinds of research artifacts, and so now attention has turned to

- Identifying what roles should be recognized for each type of artifact
- Whether those roles fit into existing taxonomies, etc. such as CRediT
- And how those roles should be recognized in an expanded credit system

We are initially focussing on data, software, physical samples, semantic resources, and educational resources primarily because that's the expertise within the group, which you should read as an invitation to come and join us, especially if there are other kinds artifacts that you are interested in!

Towards that end, there are links on the poster to our email list, wiki, and calendar, so please join us!

Lastly, there is supposed to be time for discussion at the end of this session, so I thought I would wear my Ronin hat and be provocative by showing that Scholarly publications are neither FAIR nor fair, but at least they are visible. Most of the other types of research artifacts are neither recognized or visible. How scientific is that?

ABSTRACT

As noted in the session description any research object produced should be made 'FAIR'; but 'FAIR' objects require attribution. What does that mean for the vast variety of research objects? What roles need to be acknowledged, how and where? These are questions the Federation of Earth Science Data Partners (ESIP) Research Object Citation cluster have been addressing. In particular, we have been applying the lens of the CREDiT taxonomy roles in these discussions. Here our conclusions to date will be described.

REFERENCES

References

1. Architecting Attribution. (2019). Contributor Role Ontology (Version 1.0) [Collaborative data model]. Project retrieved from https://github.com/data2health/contributor-role-ontology. Ontology retrieved from: https://zenodo.org/record/2593739#.XTZWoZNKgl4
2. Ted Habermann, John Kozimor, and Ross Bagwell. (2016, October 18). *Individuals, Organizations, and Roles*. https://wiki.earthdata.nasa.gov/display/NASAISO/Individuals+Organizations+and+Roles (https://wiki.earthdata.nasa.gov/display/NASAISO/Individuals%2C+Organizations%2C+and+Roles)
3. Pierre Alliez, Roberto Di Cosmo, Benjamin Guedj, Alain Girault, Mohand-Said Hacid, et al.. Attributing and Referencing (Research) Software: Best Practices and Outlook from Inria. Computing in Science and Engineering, Institute of Electrical and Electronics Engineers, 2019, pp.1-14.10.1109/MCSE.2019.2949413 . hal-02135891v2
4. NISO RP-25-2016 Outputs of the NISO Alternative Assessment Metrics Project, available at https://www.niso.org/publications/rp-25-2016-altmetrics
5. “CRediT – Contributor Roles Taxonomy.” n.d. CASRAI (blog). Accessed 30 January 2020. https://casrai.org/credit/
6. Wilkinson, M.D., Dumontier, M., Aalbersberg, I.J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.W., da Silva Santos, L.B., Bourne, P.E. and Bouwman, J., 2016. The FAIR Guiding Principles for scientific data management and stewardship. Scientific data, 3(1), pp.1-9.