Results of the un-deRSE 2023 Breakout Session "RSE Educational Resource Discoverability"

Toby Hodges[1], Christian Knüpfer[2], Michele Mesiti[3], David Pape[4], Philipp Matthias Schäfer[2], Tobias Schlauch[5], and Philipp S. Sommer [*][6]

[1]The Carpentries
[2]Friedrich Schiller University Jena
[3]Karlsruhe Institute of Technology
[4]Helmholtz-Zentrum Dresden-Rossendorf
[5]German Aerospace Center
[6]Helmholtz-Zentrum Hereon

28. September 2023

During the breakout session, we, the participants, discussed three questions and answered them in the form of collections of notes organized on pin boards.

After a short introduction of the issue under discussion, we split the group in two halves and discussed the questions

- What standards, tools, and infrastructure exists?

and

- What information are we interested in when looking for resources?

(both in the context of research software engineering (RSE) educational resources (ER) discoverability). Each half worked on both questions, then we rejoined and summarized the results.

---

[*]This list of authors includes those persons that were present at the workshop and indicated their wish to be listed. There were others who contributed as well.

Finally, building on our these results, we discussed the third question:

- What is missing and how could we go about creating it?

The rest of this document consists of one section per question discussed, presenting our results, which are mostly collections of keywords and references, occasionally with comments.

What standards, tools, and infrastructure exists?

Answers to this question were not restricted to existing standards, tools, and infrastructure that are directly applicable to RSE ER, but rather anything that we might build upon or draw inspiration from was allowed. Figure 1 shows a cropped photograph of the pin board for this question at the end of the session.
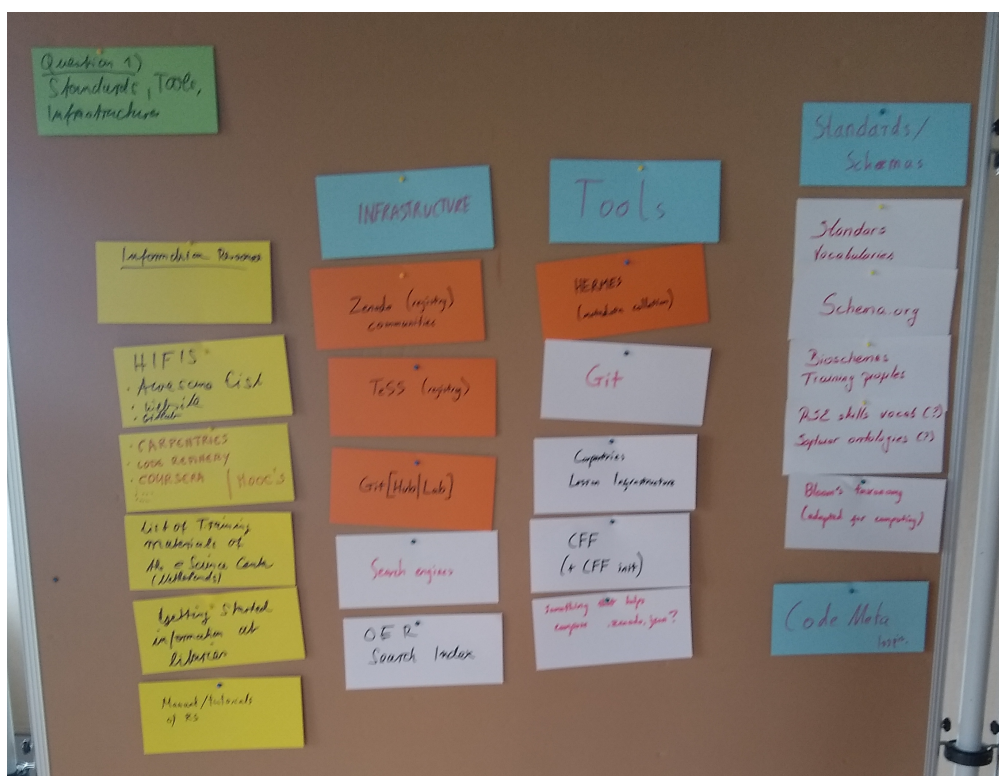

Figure 1: Pin Board with our answers to "What standards, tools, and infrastructure exists?"

In addition to standards, tools, and infrastructure, we added a category, information resources, under which we collected current documents that list and link to or host existing ER:

- HIFIS [7]
- The Carpentries [5, 8, 11, 12]
- CodeRefinery [1]
- Many MOOC-platforms offer courses on topics relevant to RSE
- eScienceCenter - Training materials [18]
- getting started information of libraries
- Manuals/Tutorials of RStudio [13]

We collected the following notes on existing standards:

- Standard Vocabularies
- Schema.org
- Bioschemas' Training profiles [3]
- RSE skill vocabularies
- Software ontologies
- Bloom's taxonomy adapted for computing
- CodeMeta [17]

We collected the following notes on existing tools:

- Hermes [9], which is already being adapted to generate metadata for educational resources
- Git, as a provider of metadata on contributors
- Carpentries Lesson Infrastructure [16], which does/will generate and incorporate metadata using Hermes in the compiled educational resources
- CFF [4, 6] with CFF init [14] as inspiration

We collected the following notes on existing infrastructure:

- Zenodo communities
- TeSS [2, 15]
- Git[Hub|Lab]
- Search engines
- OERSI.org [10], an existing registry

What information are we interested in when looking for resources?

Figure 2 shows a cropped photograph of the pin board for this question at the end of the session.

We collected desired information in three categories: that which teachers are interested in when searching, that which learners are interested in, and that which both are inter-

ested in. We briefly thought about a fourth category covering providers (authors and maintainers) of material, but as they are usually not searching for that material in these roles, they should be interested in providing metadata that satisfies the other groups interests.

For both roles, teacher and learner, we expect the following information to be of interest:

- topic
- link to the resource itself
- up-to-dateness
- references/links to related resources (prerequisites/follow-ups)
- language
- duration
- is-part-of relationship (episode of a course, for example)
- target audience: required level of expertise
- target audience: domain
- technical requirements
- peer review information (who/when)
- format/medium (video, course notes, slides, etc.)
- outcomes: skills taught
- suitability for self-learning/teaching
- measure of quality/user rating

Only for teachers, we expect the following information to be of interest:

- license regarding use, reuse, and republishing the resource
- active maintenance
- references to technical support by IT, the platform, etc.

The information that we think only learners are interested in, mostly revolves around courses that use the resource:

- events, that teach based on the resources
- contact information
- references to learner support during and after courses
- mode of courses (online, onsite, hybrid)
- cost of courses
- schedule of courses

What is missing and how could we go about creating it?

Figure 3 shows a cropped photograph of the pin board for this question at the end of the session.

For next steps or their desired outcomes we collected the following list:

- work on RSE learning vocabulary
- metadata standard for RSE ER with adoption in the community
- training for applying and creating metadata to ensure and establish good practices
- advocating use of metadata standard
- tooling to generate metadata for RSE ER (an extension of Hermes[9] has already been worked on)
- extend or create an index or a registry in which the RSE community collects metadata on ER

References

[1]   *All CodeRefinery lessons.* CodeRefinery. URL: https://coderefinery.org/lessons/from-coderefinery/ (visited on 10/17/2023).

[2]   Niall Beard et al. "TeSS: a platform for discovering life-science training opportunities". In: *Bioinformatics* 36.10 (2020), pp. 3290–3291.

[3]   Leyla Jael Castro et al. "Bioschemas training profiles: A set of specifications for standardizing training information to facilitate the discovery of training programs and resources". In: *PLOS Computational Biology* 19.6 (2023), e1011120.

[4]   *Citation File Format (CFF).* URL: https://citation-file-format.github.io/ (visited on 10/17/2023).

[5]   *Community Developed Lessons.* The Carpentries. URL: https://carpentries.org/community-lessons/ (visited on 10/17/2023).

[6]   Stephan Druskat et al. "Citation File Format 1.2. 0". In: (2021).

[7]   *Education and Training.* Helmholz Federated IT Servies. URL: https://www.hifis.net/services/software/training.html (visited on 10/17/2023).

[8]   *Lessons.* The Carpentries. URL: https://datacarpentry.org/lessons/ (visited on 10/17/2023).

[9]   Michael Meinel et al. *hermes.* Version proof-of-concept.

[10]   *Open Educational Resources Search Index.* Hochschulbibliothekszentrum NRW and Technische Informationsbibliothek. URL: https://oersi.org/ (visited on 10/17/2023).

[11]   *Our Lessons.* The Carpentries. URL: https://software-carpentry.org/lessons/ (visited on 10/17/2023).

[12]   *Our lessons.* The Carpentries. URL: https://librarycarpentry.org/lessons/ (visited on 10/17/2023).

[13]   *Product Guides.* Posit Software, PBC. URL: https://docs.posit.co/guides/ (visited on 10/17/2023).

[14] Jurriaan H. Spaaks et al. *cffinit*. Version 2.3.1. Aug. 2023. URL: `https://github.com/citation-file-format/cff-initializer-javascript`.

[15] *TeSS (Training eSupport System)*. Training eSupport System. URL: `https://tess.elixir-europe.org/` (visited on 10/17/2023).

[16] *The Carpentries Workbench*. The Carpentries. URL: `https://carpentries.github.io/workbench/` (visited on 10/17/2023).

[17] *The CodeMeta Project*. URL: `https://codemeta.github.io/` (visited on 10/17/2023).

[18] *Training materials*. netherlands eScienceCenter. URL: `https://www.esciencecenter.nl/training-materials/` (visited on 10/17/2023).
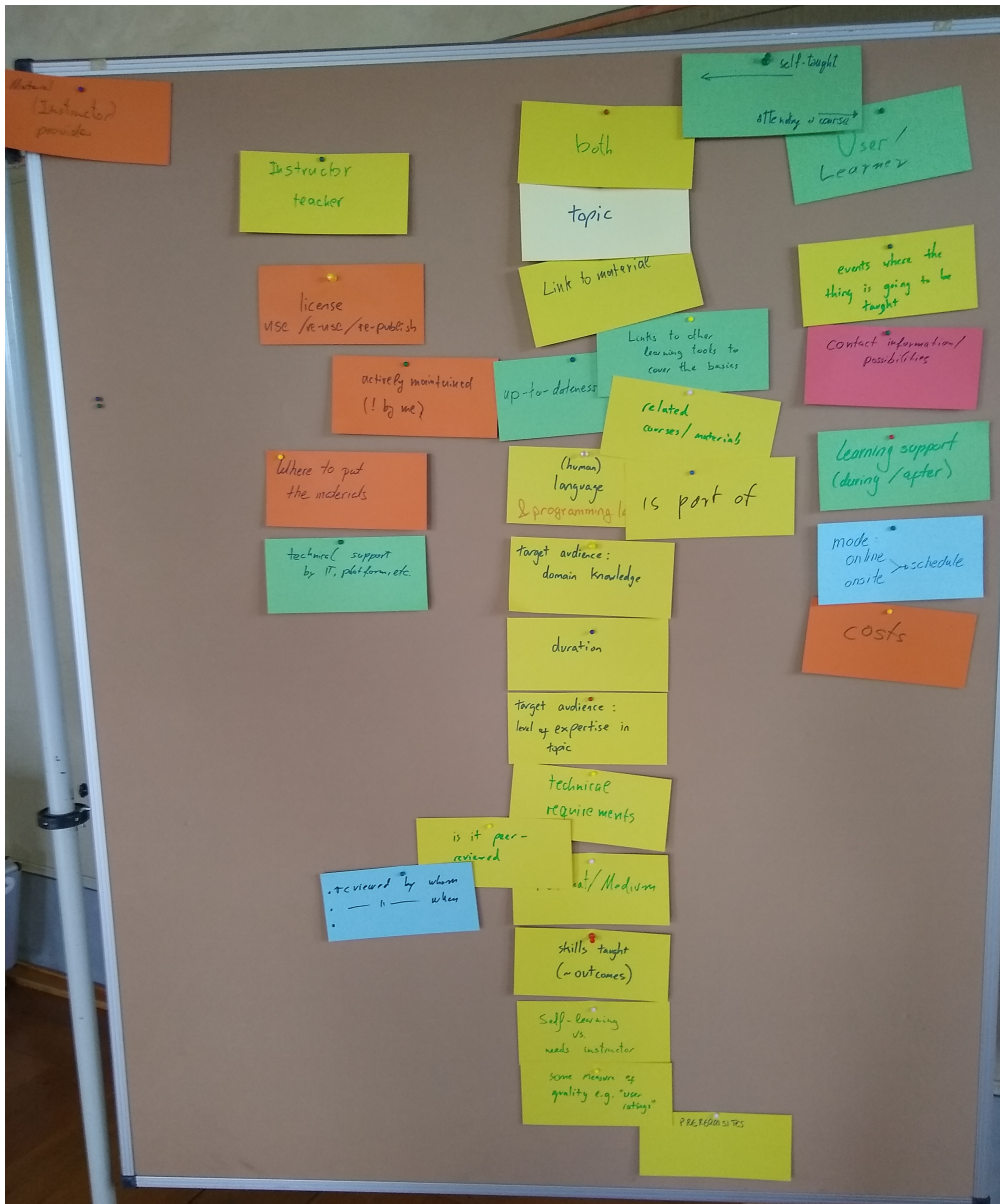
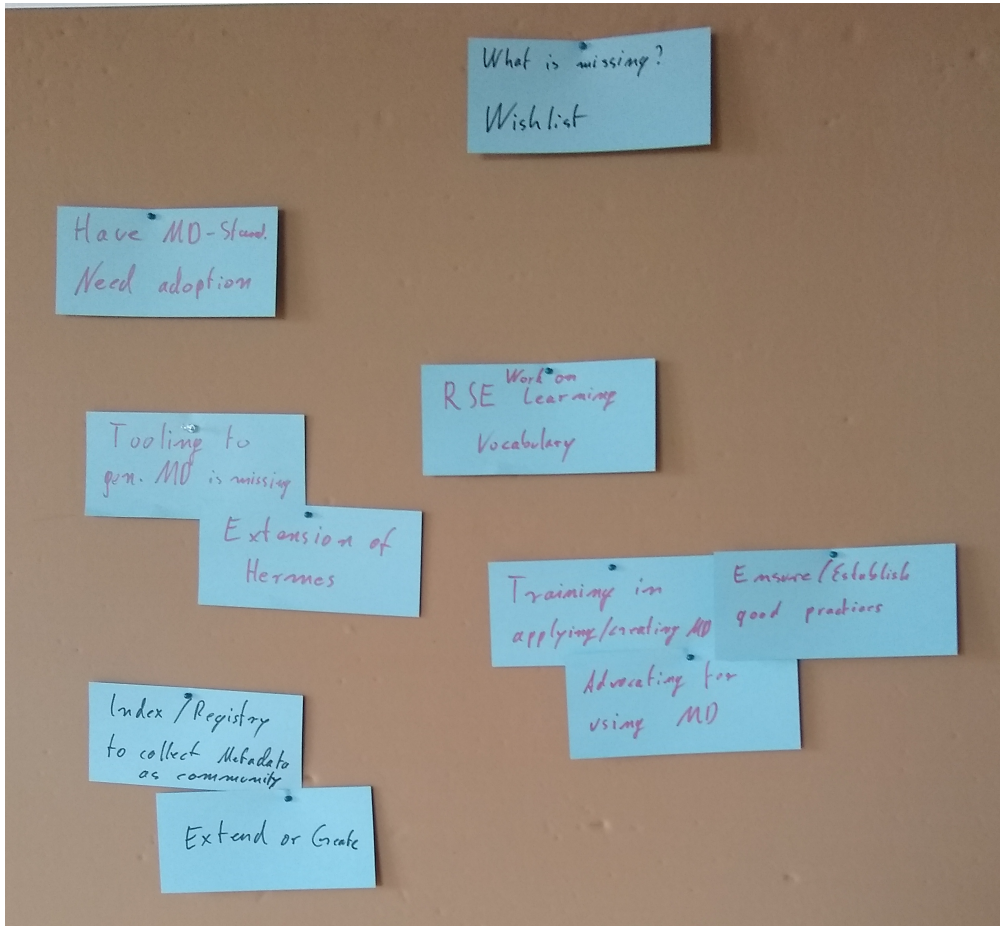Figure 2: Pin Board with our answers to "What information are we interested in when looking for resources?"

Figure 3: Pin Board with our answers to "What is missing and how could we go about creating it?"