

## 23rd Plenary Meeting Collaborative Notes

**Group(s) name(s) organising the session:** *BoF - Hugh Shanahan*

**Session link:** [Advancing Data Skills: Expanding the CODATA-RDA SoRDS Curriculum](#)

**Session scheduled date/time/breakout session:** *Tuesday, 12 November, Breakout 2, 14:30 – 16:00 CST*

**Venue:** *Edificio de Ingeniería, Classroom 205*

### Session summary (for Group co-chairs)

We will use the content in the table below to highlight your work to the RDA community as a report organised by the Technical Advisory Board and to a wider audience through English & Spanish social media mentions.

Please complete ALL fields below by **Friday, 29th November, close of business** to be included in the report & social media activities.

<b>Summarise the session in three sentences:</b>
<b>Key outcomes/actions/takeaways</b>
<ol style="list-style-type: none"> <li>1. <i>Example</i></li> <li>2. <i>Example</i></li> <li>3. <i>Example</i></li> </ol>
<b>Synergies and/or possible collaborations identified with RDA groups and other groups:</b>
<b>Highlight text that will be used in social media mentions (please make sure the text is clear and appropriate for public consumption &amp; comprehension)</b>
<b>Direct link to Group home page</b>
<a href="http://datascienceschools.org/">http://datascienceschools.org/</a>  <a href="https://www.rd-alliance.org/groups/codatarda-research-data-science-schools-low-and-middle-income-countries">https://www.rd-alliance.org/groups/codatarda-research-data-science-schools-low-and-middle-income-countries</a>

Get involved in [RDA Community](#)

Check out [P23 programme sessions](#)


This meeting will take place according to the [RDA Code of Conduct](#)

### Attendee Check-in

Please complete this table to indicate your attendance (add rows as needed):

Full Name	Affiliation	Location	Email
Martin Weise	TU Wien	Austria	martin.weise@tuwien.ac.at
Katarzyna Biernacka	TU Berlin	Germany	katarzyna.biernacka@tu-berlin.de
Magdalena Szuflińska-Zurawska	Gdansk University of Technology	Poland	Magdalena.szuflińska@p.g.edu.pl
Rob Quick	Indiana University	US	rquick@iu.edu
Flo Grattarola	Czech University of Life Sciences Prague / Biodiversidata	Czech Republic / Uruguay	grattarola@fzp.czu.cz
Ana Lorena Rojas-Sabogal	Pontificia Universidad Javeriana	Colombia	rojas-ana@javeriana.edu.co
Marcela Alfaro Cordoba	UC Santa Cruz	USA	macordob@ucsc.edu
Daniel Manrique-Castano	Digital Research Alliance of Canada	Canada	daniel.manrique-castano@alliancecan.ca

**Collaborative Session Notes** (To be used by participants and chairs during the session)

Link to the Slides:  BoF P23.pdf

Link to curriculum materials currently being taught in Costa Rica:

<https://github.com/CODATA-RDA-DataScienceSchools/Materials/tree/master/docs/DataHeredia2024>

## Agenda:

1. Welcome and Introduction (**10 minutes**)
2. Justification for Curriculum Expansion (**20 minutes**)
3. Discussion: New Topics and Modules (**30 minutes**)
4. Formation of the New Working Group (**10 minutes**)
5. Q&A and Open Discussion (**10 minutes**)
6. Summary and Next Steps (**5 minutes**)

## Program for the School:

Day	Module
1	Opening + Open and Responsible Science + Research Data Management
2	Software Carpentry - Unix + Software Carpentry - Git
3	R/Python
4	R/Python + Author Carpentry
5	RDM Lab + Open and Responsible Science
6	Data Visualization + Information Security
7	Data Analysis (Statistical Analysis and Intro to Machine Learning)
8	Data Analysis (Neural Networks)
9	Data Analysis (Deep Learning) + Intro to Computational Infrastructures
10**	Computational Infrastructures + Closing

\*\* This is a half day with closing after lunch

## Questions

- Are there any specific skills missing or underrepresented in the current curriculum?
  - Need to clarify goals in order to better understand what is missing
  - To teach the importance of data in machine learning need more “basics” - data cleaning, data analytics etc
  - FAIRification can sometimes mean that we forget about data quality and the work that needs to be done. Could reach out to RDA groups wrt support
  - “Low computing technologies”
  - Get more education designers and pedagogy people involved
  - Get connection to additional materials to enable “deep dives”
- Which topics or skills do you find most valuable, and which ones would you deprioritize if additional time were needed for new content?
  - Do some asynchronous teaching to free up time
  - Less about DMPs and more about data management - be topical for the lived experiences of the students
  - More time for OpenRefine instead of DMPs
  - Data visualisation could be shortened - do something on holistically managing data through a pipeline. I.e. have a dirty data set that is cleaned and then visualised

3. Are there particular areas that you feel should have dedicated modules?
  - Integrate training on accessibility into the modules - perhaps have a separate group working on accessibility within the RDA
  - The image of the curriculum isn't not clear so I don't understand where it starts, and I don't see any of the last part of the data lifecycle as to what to do with the data when the research finishes. Topics like publication data, preserve, metadata, repositories, data articles, and anonymization.
  
4. "How do you see the role of domain-specific applications (e.g., social sciences, climate science) in a general data science curriculum?"
  - Sensitive data
    - Someone commented that there are a lot of different jurisdictions in regards to sensitive data, we can't teach everything.
    - Could make a "tool box" that is domain specific
  
5. Would an additional week for specialized, advanced workshops be beneficial to cover these new topics in depth? Should we mix everything into one single week?
  -
  
6. How could we balance foundational skills with more advanced, topic-specific modules to serve both newcomers and advanced participants?
  
  
7. Should we consider modular, flexible tracks within the program, allowing participants to focus on their areas of interest
  - a. This could help address question #6
  
8. Are there other groups (e.g., industry partners, academic institutions) that we should consult to make the curriculum as comprehensive and relevant as possible?
  - a. Digital Library Federation
  - b. Research Data Access and Preservation (RDAP) Association
  - c. Schools of information science
    - i. increasingly these schools bring together Library Science and Computer Science.

