



**INTERNATIONAL
DATA WEEK**
13-16 October 2025
Brisbane, Australia



RDA 25th Plenary Meeting [part of IDW 2025] Collaborative Notes

Group(s) name(s) / session title	(Joint session) Sensitive Data Interest Group, Professionalising Data Stewardship IG - <i>Data management terminologies: from challenges to community recommendations</i>
Session link for more information	https://www.rd-alliance.org/groups/sensitive-data-interest-group/plenary-participation
Session programme time	13:30 – 15:00 AEST 03:30 – 05:00 UTC 05:30 – 07:00 CET RDA 25 rd Plenary Meeting Breakout Session 5 Venue: Plaza Terrace Room
Presentation slides	P 03. Thursday 16 October - Plaza Terrace Room - 1330 - 1500.pptx Poster: P RDAPlenary25_QofDM_PosterRD_ASB_V1.3_05102025_17_24CET... Paper in preparation: = Draft_Term4DM_V1.2_15092025_16.15CET



Session summary (to be filled by group chairs):

We will use the content in the table below to highlight your work to the RDA community in a report by the Technical Advisory Board ([RDA P24 report example](#)). **Please complete ALL fields below by Friday, 31 October, to be included in the report** and other communication. We also ask you to update your group webpage with the key takeaways from your post-plenary session.

1.	Describe what your session is about in three sentences maximum: <i>Example: This session was about ... The key points the session covered were to finalise / update / discover / identify ect.</i>
2.	Highlight a maximum of five key outcomes/actions/takeaways of your session: <i>Example:</i> 1. Need to learn more about how... 2. Review tools to identify next... 3. Discussion about upcoming... 4. Identify future collaboration with... 5. Follow up with new members...



**INTERNATIONAL
DATA WEEK**
13-16 October 2025
Brisbane, Australia



3. What collaborations or synergies did you discover between your group and other RDA group(s) or external (outside of the RDA) organisations?

[Note this information is used for creation of an interactive map ([RDA P24 map example](#)) included in the post-plenary report. It is also used to enable further development of collaboration / cross-fertilisation knowledge exchange activities within RDA]
(Example: please use the most relevant group(s) from <https://www.rd-alliance.org/group-directory/> or include a URL to an external organisation)



Attendee Check-in

New to the RDA? Check out on where to start [here](#)

Get involved in the [RDA Community](#)

This meeting will take place in accordance with the [RDA Code of Conduct](#)

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

Full Name	Affiliation	Location	Email
Catherine Gonzalez	RWTH/NFDI4Microb iota	Aachen, Germany	c.gonzalez@itc.rwth-aachen.de
Liise Lehtsalu	RDA Europe	Italy	liise.lehtsalu@rda-foundation.org
Federico Grasso Toro	University of Bern	Bern, Switzerland	federico.grasso@unibe.ch
Antje Manske	GESIS / Base4NFDI	Mannheim, Germany	antje.manske@gesis.org
Linlin Zhao	Deakin University	Melbourne, Australia	l.zhao@deakin.edu.au
susanne den boer	university of copenhagen	copenhagen, denmark	susanne.denboer@adm.ku.dk
Sebastian Lopez Marcano	University of Queensland	Brisbane, Australia	s.lopezmarcano@uq.edu.au
Rasha Abu Safieh	University of Otago	Dunedin, New Zealand	Rashaabusafieh.alfar@otago.ac.nz
Olga Churakova	University of Bern	Bern, Switzerland	olga.churakova@unibe.ch



John Graybeal	GO FAIR US, San Diego Supercomputer Center	United States	jbgraybeal@ucsd.edu
Nathalie Gaudreault	Allen Institute, Seattle,	United States	nathalieg@alleninstitute.org
Moazzam S.B. Syed	TU Dublin	Ireland	moazzam.syed@tudublin.ie
Romain David	ERINHA	France / Belgium	romain.david@erinha.eu
Anne-Sophie	INRAE	France	anne-sophie.bage@inrae.fr
Alison Specht	TERN, University of Queensland	Australia	a.specht@uq.edu.au

Despite an increase in the availability of proposed definitions for many data ecosystem tasks, there remains an urgent need to harmonise data management terminologies that support FAIR principles and actively improve the quality of data management.

In this session, we use a five-level severity scale (from minor disruption to research termination), inspired by established risk assessment frameworks (likelihood × impact).

The approach aims to link terminological obstacles directly to consequences on projects such as time, cost, quality, compliance, and reproducibility.

Meeting agenda

This session is structured around **three interactive components**:

1. Introduction & Survey (Mentimeter)

- Participant input on terminological obstacles they have experienced.
- Scores of the severity of obstacles (Levels 1 – 5) in participants' own contexts..

2. Round Table

- Produce comparable insights across disciplines while capturing local nuance.
- Discuss how obstacles appear across different domains and communities.
- Compare severity ratings across contexts and examples.

3. Discussion

- Identify stakeholders needed to address specific obstacles.
- Connect challenges to possible mitigation strategies.
- Shape next steps for the collaborative paper and checklist development.



**INTERNATIONAL
DATA WEEK**
13-16 October 2025
Brisbane, Australia



Wrap-up (5 min): Summarise findings and explain how they will feed into the P25 paper sprint and future RDA recommendations (all participants are invited to be co-authors).

Expected Outcomes

- List of **key terminology obstacles** with highest research impact.
- Identification of **priority areas for harmonisation** and governance.
- Inputs for the **collaborative RDA paper** and **terminology checklist**.
- Commitments or suggestions for **community actions** post-session.

Discussion Prompts

- Compare your **severity ratings** across disciplines and contexts.
- Share real-world examples of term conflicts (e.g., “Data Owner” vs “Data Steward”).
- Identify **cross-cutting patterns** – are some issues universal (e.g. polysemy), while others are domain-specific?
- Discuss how to move from **identifying issues** → to **aligning, embedding, and sustaining** common solutions.

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

The session focused on the standardisation of data management terminologies, with an emphasis on FAIR principles and **the enhancement of data management quality**. Ambiguous or overlapping terms—such as *Data Steward* vs. *Data Custodian*, *Dissemination* vs. *Archiving*, or *Repository* vs. *Data Lake*—often lead to confusion, inconsistent practices, and reduced interoperability.

The group is working on a paper building on a collaborative effort initiated at the RDA 23rd Plenary (P23) to be consolidated here, to identify, debate, and refine terminological challenges. The aim is not to produce a definitive glossary but to document areas of persistent ambiguity, highlight their practical implications, and propose community-driven pathways for action. Specifically, our objectives are to: (i) analyse key areas of terminological ambiguity that affect the quality of data management, (ii) synthesise collective insights from RDA plenaries and related projects into a coherent framing, and (iii) propose actionable recommendations—including a practical checklist—to guide research communities in adopting, aligning, and sustaining shared terminologies. By doing so, we seek to reframe terminology as a shared asset rather than a barrier to trust, interoperability, and professionalisation in research data management.

Aligning and clarifying terminology

We started by referring to an existing set of examples of terminological confusion by Specht et al. (2024) <https://doi.org/10.1371/journal.pone.0311967>

term	definition	example(s) from Specht et al. 2024
Polysemic	(1 term or concept, multiple meanings), a form of ambiguity [86 p.89]	experiment - run
Linguistic confusion	due to (apparent) multiple terms for a meaning (i.e. synonyms)	prediction - estimation; census tract - census sector - block



Fuzziness	in language or terminology matches (existing and evolving nuances), a form of ambiguity [87]	treated -non-treated; control - reference
Idiomatic	language terms adopted from another language with different contexts, a form of ambiguity [86 p.7]	mirror - control
Disciplinary domains (communities of practice)	different disciplines assign different meanings to the same term [88,89]	ecoregion; survey - census; wave-round-repeat instance
Country of practice or origin	may assign different meanings to the same term [90]	survey - census
Vagueness	involves uncertainty about the actual meanings of particular terms, e.g. administrative units differ in name and nature from region to region and country to country [86 p.83, 91]	village - town; municipality -city
Lexical gap	New concepts (words, phrases) that don't exist in the target language [92]	digital presence; checklist; data management plan
Adoption challenges	a 'new' term needs to have community endorsement for it to be recognised and used and this has not happened yet [93,94]	in-country - out-of-country; checklist - computational notebook

The references cited are listed [below](#)

This was then followed by a mentimeter survey of the participants, some results of which are shared here. 29 people responded to the mentimeter questions.

Responses were put into context for later cross correlation. Country of origin, discipline, work practice were explored.

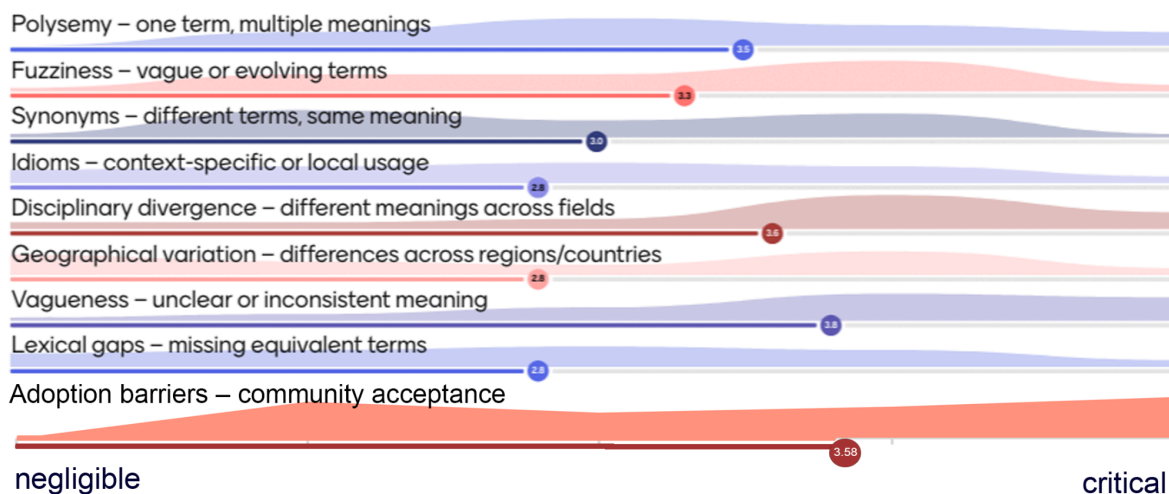
A critical question was to tell us what problems arise from the use or understanding of data management terms?



A series of questions then explored the severity of effect terminological confusion had on their workflow. The three highest ranking sources of terminological confusion were polysemy, vagueness and adoption barriers.



**INTERNATIONAL
DATA WEEK**
13-16 October 2025
Brisbane, Australia



Respondents gave us a snapshot of terms they had come across in their work that had created problems.

Decolonisation. Different meanings due to disciplinary or cultural context

Observation vs Measurement, Process vs Procedure, Quality control vs Data Validation

data v metadata: these are terms that have contextual meaning in a particular situation, but generally their def'ns are meaningless and are recursive. In the end their usefulness depends on local def's

any new technology buzzword. AI-ready vs ML-ready vs LLM-ready.

FAIR vs interoperable -- is there anything in FAIR that isn't a subsumed element of interoperability?

any term used in modeling, examples: sensor vs instrument vs platform (e.g. SensorML; data collection vs data set vs data record vs data value; work v instance v item in BibFrame

Responsible vs accountable

For example in land cover classification, the term "forest" is different definitions, eg. numbers of trees, canopy cover, diameter of tree trunks, geographical extent - width at least 10 meters,

metrics versus measurements

An example was additionally provided via these collaborative notes:

The word 'data sharing' is problematic. Data stewards understand this as data publication with a wider audience after project end. Researchers understand this as data collaboration with their collaboration partners (closed) during the project.

The survey is to be refined and repeated to explore relationships and contextual influences.

References

86. Kroeger PR. Analyzing Meaning: an introduction to semantics and pragmatics: Second corrected and slightly revised edition / Volume 0. [Place of publication not identified]: Language Science Press; 2019.
87. Zhang Q. Fuzziness - vagueness - generality - ambiguity. Journal of Pragmatics [Internet]. 1998 [cited 2024 Jan 14];29(1):13–3 Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0378216697000143>
88. Madin J, Bowers S, Schildhauer M, Krivov S, Pennington D, Villa F. An ontology for describing and synthesizing ecological observation data. Ecological Informatics [Internet]. 2007 [cited 2024 May 6];2(3):279–96. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S1574954107000362>
89. Vanderbilt KL, Lin CC, Lu SS, Kassim AR, He H, Guo X, et al. Fostering ecological data sharing: collaborations in the International Long Term Ecological Research Network. Ecosphere [Internet]. 2015 [cited 2022 Mar 29];6(10):art204. Available from: <http://doi.wiley.com/10.1890/ES14-00281>
90. Vanderbilt KL, Blankman D, Guo X, He H, Lin CC, Lu SS, et al. A multilingual metadata catalog for the ILTER: Issues and approaches. Ecological Informatics [Internet]. 2010 [cited 2023 Oct 11];5(3):187–93. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S1574954110000336>



**INTERNATIONAL
DATA WEEK**
13-16 October 2025
Brisbane, Australia



91. Kennedy C. 23. Ambiguity and vagueness: An overview. In: Maienborn C, Heusinger KV, Portner P, editors. *Semantics* [Internet]. Berlin, Boston: DE GRUYTER MOUTON; 2011 [cited 2024 Jan 15]. Available from: <https://www.degruyter.com/document/doi/10.1515/9783110226614.507/html>
92. Tabrizi HH, Pezeshki M. Strategies Used in Translation of Scientific Texts to Cope with Lexical Gaps (Case of Biomass Gasification and Pyrolysis Book). *TPLS* [Internet]. 2015 May 31 [cited 2024 Mar 15];5(6):1173. Available from: <http://www.academypublication.com/issues2/tpls/vol05/06/07.pdf>
93. Ullman RE, Enloe Y. Accelerating Technology Adoption Through Community Endorsement. In: Di L, Ramapriyan HK, editors. *Standard-Based Data and Information Systems for Earth Observation* [Internet]. Berlin, Heidelberg: Springer Berlin Heidelberg; 2010 [cited 2024 Jan 19]. p. 227–48. Available from: http://link.springer.com/10.1007/978-3-540-88264-0_13
94. Borycz J, Olendorf R, Specht A, Grant B, Crowston K, Tenopir C, et al. Perceived benefits of open data are improving but scientists still lack resources, skills, and rewards. *Humanit Soc Sci Commun* [Internet]. 2023 Jun 20 [cited 2023 Jun 22];10(1):339. Available from: <https://www.nature.com/articles/s41599-023-01831-7>