

**Slido.com**

**Enter code:**

**unconfercece**



**Q&A and panel**

+

+

+

+

+

# **CLARIN perspective**

**Daan Broeder, CLARIN**

# CLARIN RDA involvement

Daan Broeder – CLARIN ERIC

Data together unconference session, EOSC Symposium

Madrid, Sept 2023



# Introducing CLARIN



## The research infrastructure for language as social and cultural data

CLARIN is a digital infrastructure offering data, tools and services to support research based on language resources.

- CLARIN is an ERIC since 2012
- Distributed infrastructure, coordinating and working with the long term national CLARIN projects that maintain CLARIN centers
- B, C centers providing metadata and resources
- Knowledge centers providing expertise
- Technical infrastructure aimed at coherent discovery and access of resources Virtual Language Observatory, but also data management tools CLARIN LR Switchboard, Virtual Collection Registry
- Collaboration with other SSH infrastructures: DARIAH, CESSDA (SSHOC) and wider in EOSC projects

# RDA Background

- CLARIN has been involved in RDA activities from the start, also inspired by strong CLARIN affiliated people and organisations
- More general data management topics:
  - PIDs, DTR, Certification, FIM, DataFabric, Dynamic Citations, ...
- compared to thematic ones:
  - [Linguistics Data IG](#), outcome: [recommendations for citation of research data in linguistics](#)
- CLARIN RDA interests are mainly linked to core business publishing, discovery and facilitating access to language data
- Note there is no overall coordination wrt. RDA participation, but
  - RDA participation often based on personal interest
  - For many CLARIN is just one of the collaborations & projects they are involved in
- For CLARIN, funded projects and the related expert discussions, have to take priority wrt involvement. Lack of people!

# Perception of RDA within CLARIN

- RDA awareness by management, developers, researchers
  - Except by management awareness seems limited, also (hearsay) surveying ESFRIs, RDA does not play a big role
  - RDA doesn't appear often in discussions, some interesting discussions moved elsewhere eg. DataFabric & FDO
- Usefulness as a discussion & networking platform
  - This is currently rated as the most valuable aspects of RDA, **a free and open global platform**, to discuss technical and non-technical aspects of data infrastructure without direct project/funding repercussions
  - Including the networking
  - Some group events seem only specific project/policy promotion, but is this bad?
  - But not always clear what is going on, what are new activities etc. You have to be already involved to know
- RDA direct outcomes
  - Lack of direction and progress, overlapping activities
  - Difficult to quantify output and alignment
- Other fora, eg. national and project collaborations such as the SSH SSHOC collaboration are more visible and get more attention

# Improve RI involvement, some ideas

Some communities and infra's seem more involved (BioMed, OpenAIRE,..): from a need or opportunity, their activities in RDA seem more supported and coordinated, no need for action

But how could RIs as CLARIN profit and contribute more

- In general there is a lack of expert staff to participate in all projects and discussions. Too many similar events, and we have to prioritize for funded projects and other
  - RDA to have synergy with topics in other but funded projects EC, thematic clusters
  - Anchoring explicit RDA involvement plans in project proposals ie. have a RDAP next to a DMP
- RDA as a standardisation instrument is maybe not insufficiently exploited, but compared with eg. ISO it misses some things as an administrative/editorial infrastructure, but that is also a different business model
- Note there is a contradiction in the mentioned criticisms:
  - perceived lack of direction and sluggishness could be amended at the cost of inclusivity ie. be more selective about topics old-ideal was for it to be like IETF
  - RDA to be solely considered as a network for (technical) information exchange, -> try to get rid of the project/organisational promotion, but that also can be perceived as a unnecessary barrier