

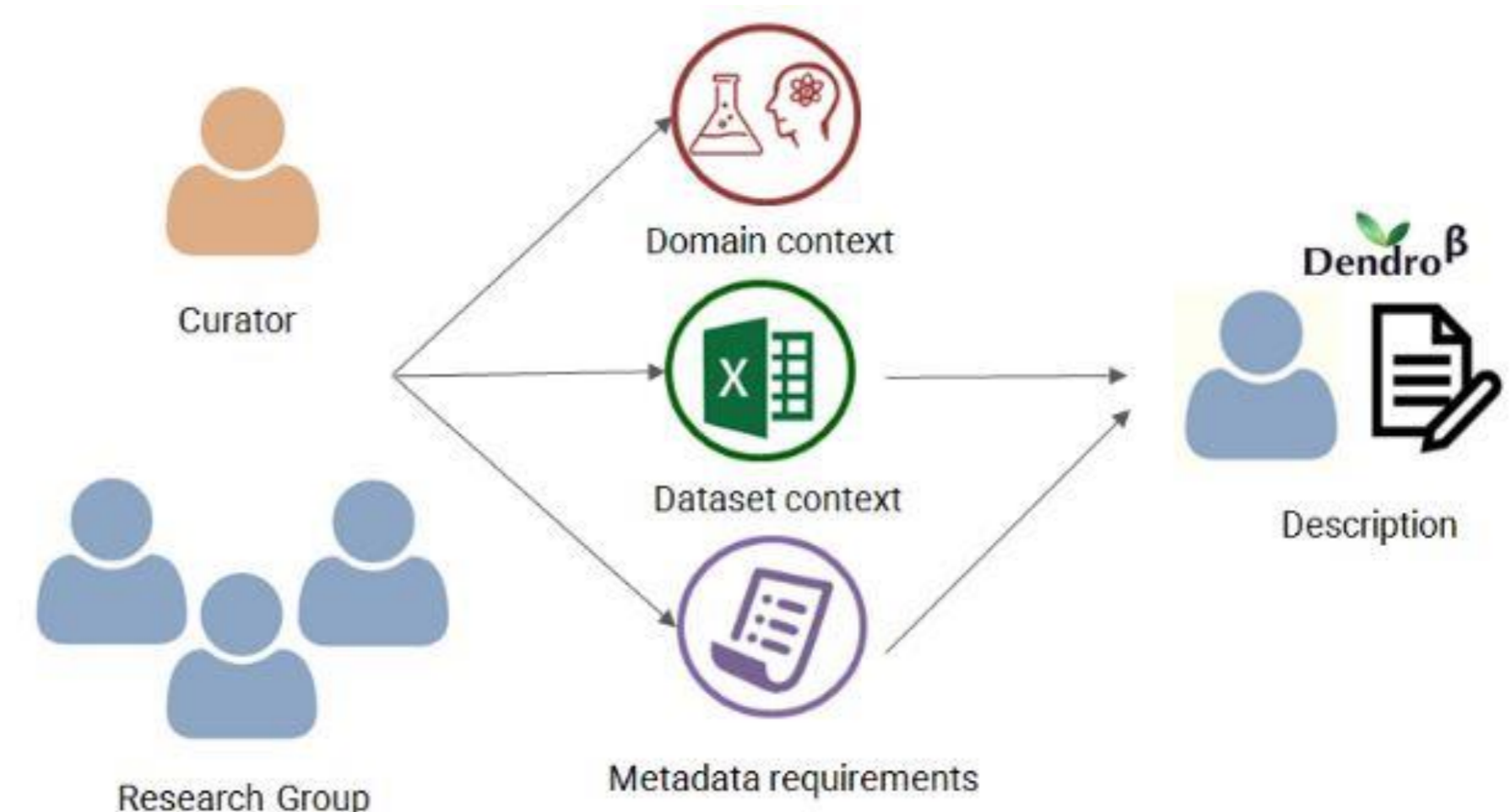
A DATA CURATOR'S WORKFLOW TO ENGAGE RESEARCHERS IN RDM: RESULTS FROM 13 MULTI-DOMAIN DATA DESCRIPTION SESSIONS

João Aguiar Castro / Cristina Ribeiro
INESC TEC, FEUP

BACKGROUND

- Engage researchers in RDM through data description using domain specific metadata models;
- 13 data description sessions conducted at University of Porto, between January 2018 and September 2019, with researchers from different domains;
- Data description was performed in Dendro, a staging RDM platform;
- Participants completed a questionnaire to measure their attitude towards data description.

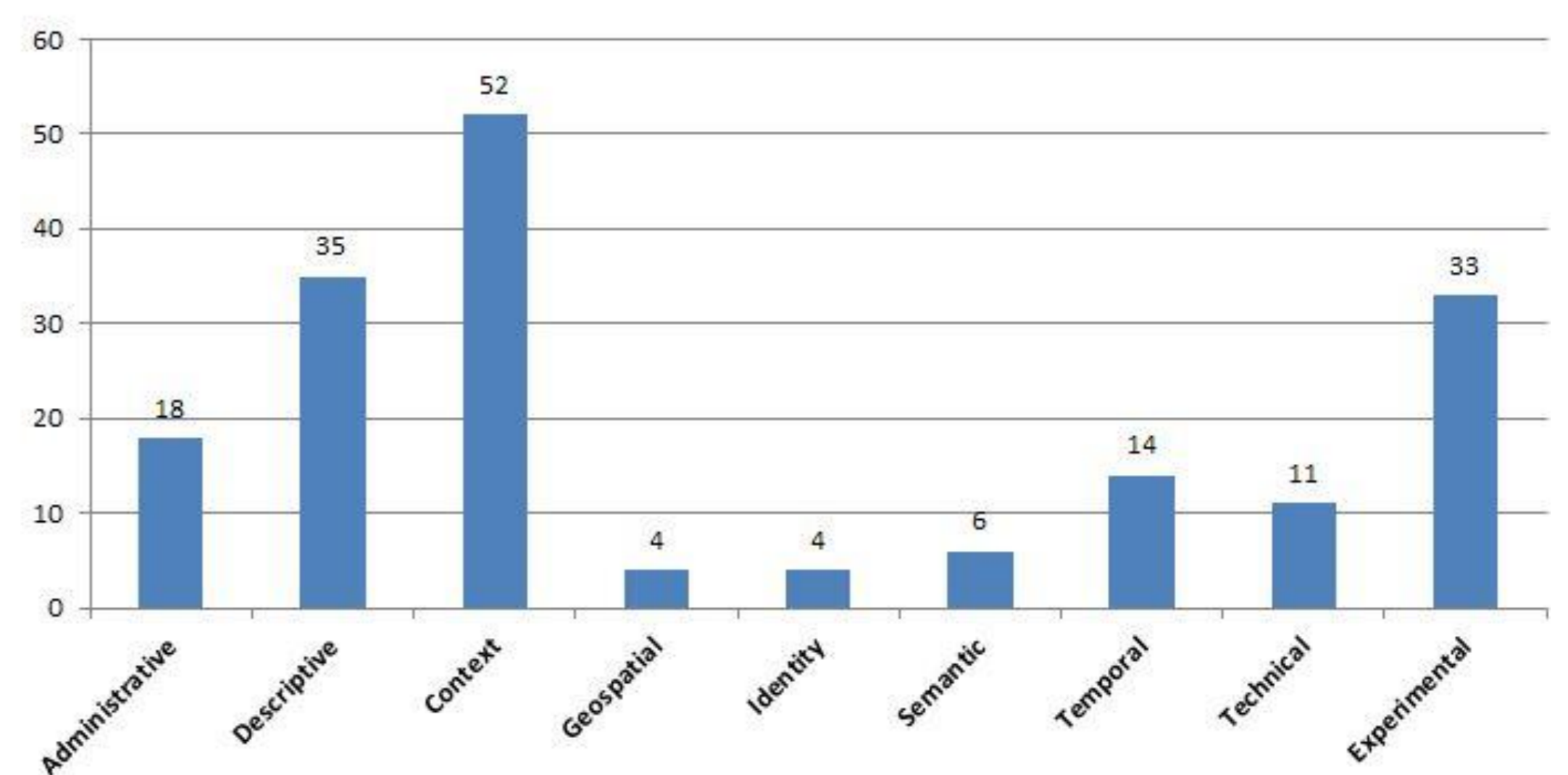
DATA CURATOR'S WORKFLOW



DESCRIPTORS COUNT

- 179 descriptors filled in (total 351 minutes); 89 unique descriptors (59 only once);
- 15 most used descriptors represented 50% of the total metadata;
- Abstract, Sampling Procedure* and *Data Collection Methodology* were the most used descriptors (8 times each);
- Longest session: 75 minutes to fill in 21 descriptors (Session 3: Consumption Sociology);
- Shortest session: 10 minutes to fill in 13 descriptors (Session 10: Sustainable Chemistry);
- Average of 13,7 descriptors in 27 minutes**

METADATA CATEGORIES FREQUENCY



METADATA QUALITY

Good

Session 1: Family Psychology
Session 2: Clinical Psychology
Session 3: Consumption Sociology
Session 4: Structural Adhesive Joints
Session 12: Health

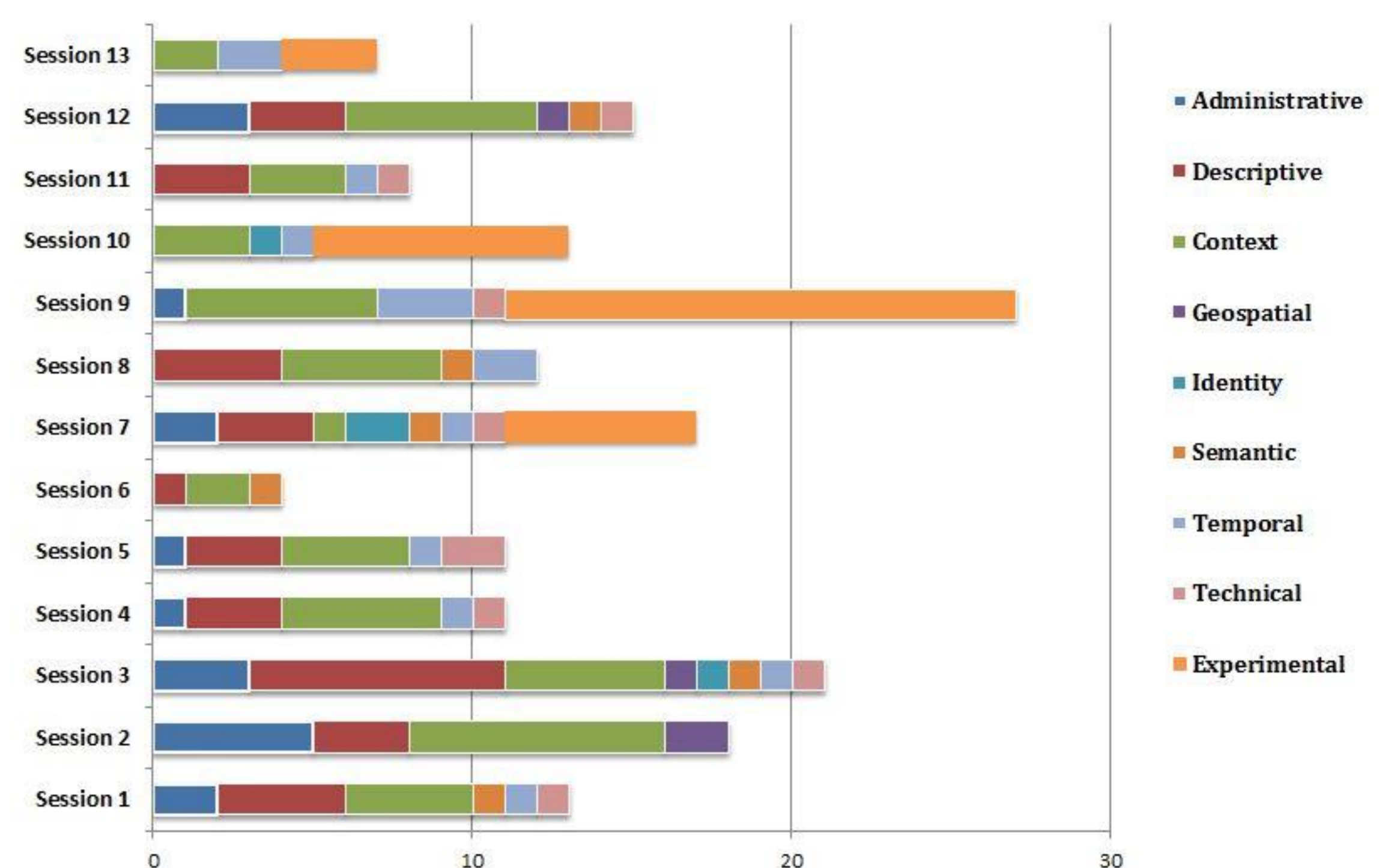
Satisfactory

Session 4: Services
Session 8: Organizations Sociology
Session 9: Magnetic Materials
Session 10: Sustainable Chemistry

Poor

Session 5: Nutrition
Session 6: Work Psychology
Session 11: Cultural Studies
Session 13: Magnetic Dynamics

METADATA CATEGORIES DISTRIBUTION BY SESSION



RESEARCHER'S PERSPECTIVES

- Researchers have shown an aptitude to fill in scientific descriptors;
- Development of knowledge and skills in metadata production by the researchers;
- Metadata creation can be a realistic task for the researchers;

yet... those who spent more than 20 minutes describing data considered the activity time-consuming;

Institutional support is essential for the adoption of RDM tools

DATA DESCRIPTIONS IS AN (...) ACTIVITY

