World Data System (WDS) – European Geosciences Union (EGU) Training Workshop Survey Report

Final report based on the responses received by 19 December 2019.

Executive Summary

The [WDS–EGU Workshop 'Data Curation and Management: Current Achievements and Future Challenges'](#) was held on 6–8 November 2019 in Paris, France. This document is a summary report based on the 14 survey responses received by 19 December 2019.

The survey was sent to ensure that:

- The Workshop covered topics relevant to participants.
- The level of information was appropriate.
- The participants not only learned useful information, but enjoyed the experience.

All respondent comments are included without revision. One respondent skipped answering all questions regarding the level of information presented, as well as the yes/no question in the WDS introduction and ECR Network sections, which resulted in at least one 'not answered' response in many cases.

Six Key Points From the Survey

1. Participants were mostly satisfied with the Workshop, including the communication prior to the Workshop. They appreciated the efforts needed to organize such a Workshop,as well as the training itself.
2. Participants enjoyed the interactive session (World café) the most, and more hands-on exercises are wanted in addition to this. In particular, there was a desire to have hands-on exercises in 'How to exchange data', 'Ontologies and knowledge networks', and 'Cloud platforms in science'.
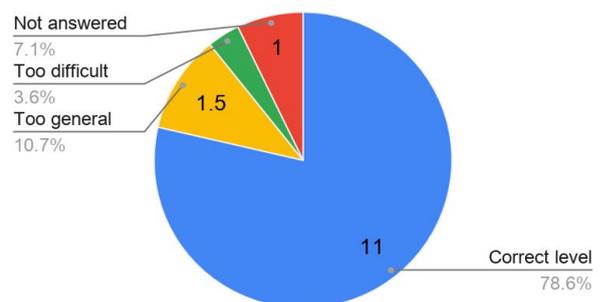3. Participants were at very different levels for some of the topics.

4. A need was expressed to inform participants of related, basic material to read prior to the Workshop such that the training is understood more effectively, beneficially, and in depth.
5. The length of the Workshop was too short to cover everything, and some sessions were completed too quickly and overloaded with information.
6. The order of sessions may need to be reconsidered; especially, the 'heavier' sessions should take place in the morning.

Results

General Feedback

All 14 respondents were satisfied with logistical aspects such as the application process, communication and practical details, travel and accommodation, and so on. Also, all 14 respondents thought that the Workshop overall matched with their expectations. Regarding the level of the information presented, some respondents selected two responses from

Level of the information presented

multiple choice, which are counted as 0.5 each. Nine respondents felt the level of the information presented during the Workshop was 'at the correct level', yet three respondents felt it was the combination of 'at the correct level' and 'too general'. One respondent felt it was the combination of 'at the correct level' and 'too difficult'.
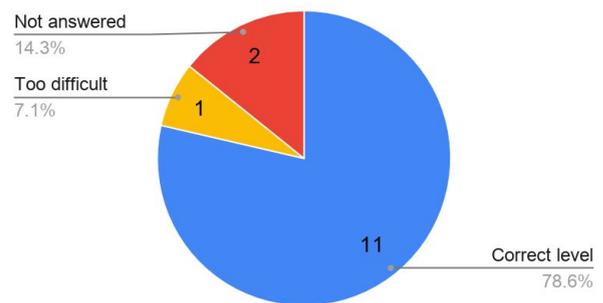
Specific comments from respondents included:

● The workshop and accommodation was very well organised. Communication with the organisers before the workshop was very supportive and helpful.
● The exchange with other participants was especially useful to me. The program provided enough time and incentives for that.
● The time was too short for the kind of topics that were to be covered, with too much of presentations rather than practical and discussion sessions.
● I don't think anything was problematic. Perhaps the presentations on technical aspects of data such as ontologies and cloud computing could be shifted earlier in the day. As a lot of vocabulary and concepts are introduced in these sessions, it might be useful to grasp them when we are not too tired ;)

- I'd suggest to have a lectures about 'heavy' topics as Big Data and Ontology in the morning sessions. In the afternoon, it would be nice to have more practical exercises, interactive sessions and group work. It would be really nice if we could do some ontologies together.
- It would be better if some basic information about certain topics is distributed in advance.

Module 1: What are Research Data?

11 respondents said the information presented during this session was 'at the correct level'. One respondent said it was 'too difficult', and two did not answer.
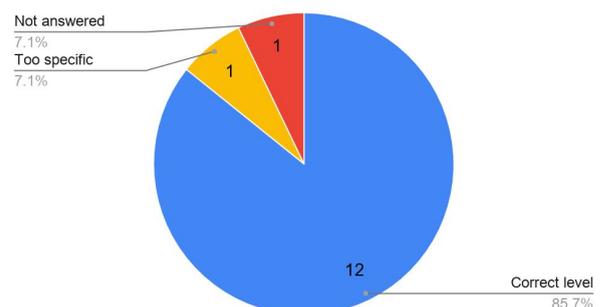
Specific comment from respondents included:

Level of the information presented

Not answered
14.3%

Too difficult
7.1%

2

1

11

Correct level
78.6%

- The presenter explained every concept clearly and gave practical examples. I found it very interesting and useful.
- Understanding the intrinsic value of data. Gaining an understanding of the data across type, domain, granularity, completeness. I think it helps to reflect on the type of research data generated.
- It was a very good and broad introduction to data and data management concepts. The distinction between different types of data and data products and considerations about metadata were specifically useful for my current project

Module 2: Research Data Management

12 respondents said the information presented during this session was 'at the correct level', one respondent said it was 'too specific', and one did not answer.

Specific comments from respondents included:

Level of the information presented

Not answered
7.1%

Too specific
7.1%

1

1

12

Correct level
85.7%

- The presentation was very interesting and the presenter really used very simple terms that were easy to understand.

- Very well organised and thorough presentation. The parts about day-to-day data management strategies, metadata standards and lists of repositories were specially useful and interesting for me.
- I really liked how life cycle models were used to create awareness for managing research data. Also, I was pleasantly surprised that openness was discussed so prominently.
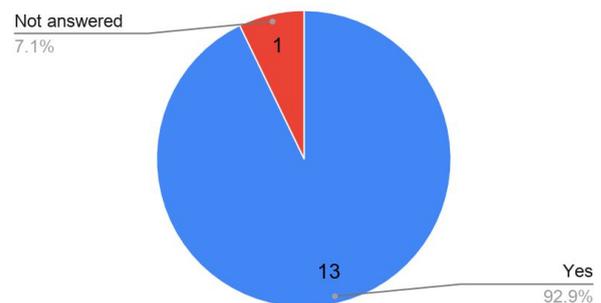
Module 3: World Café

Except for one who did not answer, all respondents stated that this session was useful.

Specific comments from respondents included:

Useful?

Not answered
7.1%

1

13

Yes
92.9%

- It was great to hear about the others as it makes you realize how much you have in common regardless of the discipline/institution.
- This activity was very important because it was an icebreaker activity for the participants.
- Four questions ranging from the challenges we face with respect to data management and expectations from the workshop. The best part of the exercise was to get up and start reflecting on how we search and manage data. Learning from peers. The best part was reflecting back on these questions at the end of the workshop. Useful workshop tool!
- Maybe a better organisation of the exercise: a bit more time, everybody should speak and bring on some points from personal experience.
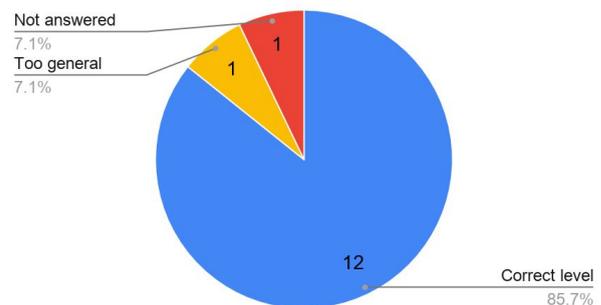
Module 4: Open Science and Data From the Perspective of a Practicing Scientist

12 respondents said the information presented during this session was 'at the correct level'. One respondent said it was 'too general', and one did not answer.

Specific comments from respondents included:

- Again a very insightful session. I was happy to hear about researchers' needs concerning infrastructures and services.

Level of the information presented

Not answered
7.1%
Too general
7.1%

1

1

12

Correct level
85.7%

- Presentation had a very good flow and covered a large amount of information. The parts about effect of publishing in open access journals and reporting failure were very interesting. Also I learned about pre-print archives.
- I would have liked some more practical things on this. It is still not clear to me how to make open science a reality.
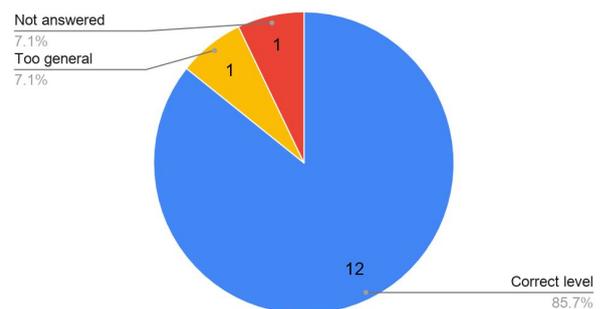
Module 5: Data Management Plans

12 respondents said the information presented during this session was 'at the correct level', and also 12 respondents thought the individual activity and debrief was interesting/useful. One respondent said it was 'too general', and it was not interesting/useful. One respondent did not answer.
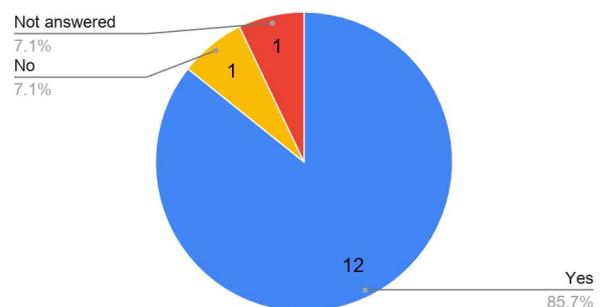
Specific comments from respondents included:

- Great and useful. I'd love to go through not only the DMP template, but to see real example. It would be nice if we could go through one very well written DMP (+ another one not so well written). With some real DMP we could have a discussion, what is well written, what isn't etc. It's always easier to learn from existing DMP, which has been written for a certain project.

Level of the information presented

Not answered
7.1%
Too general
7.1%
1
1
12
Correct level
85.7%

Interesting/useful?

Not answered
7.1%
No
7.1%
1
1
12
Yes
85.7%

- I was pleasantly surprised by the focus on DMPs in this workshops. I was already familiar with DMPs, but I think ECRs can usefully apply what they learned in this session.
- Writing a DMP for a research project highlighted many unnoticed aspects of dealing with data. I really liked the individual assistance by the organisers. Sharing experiences with other participants was also useful; not only to see that we share many common challenges, but also in regard to designing better research data support services.
- The time accorded to the session was insufficient, and I think if we had done it in groups, it would have been more interesting.
- An idea for the data management plan. Form as many groups as "teachers" (better if the component of the group share the same field as much as possible) and make them work

together in less time but more focused. The whole exercise took two days and it was very loose.
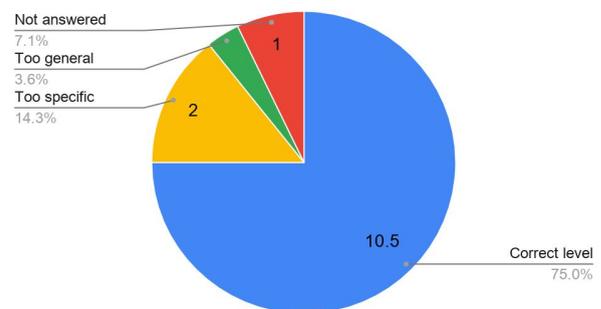
Module 6: How to Exchange Data

One respondent selected two responses from the multiple choice, which are counted as 0.5 each.

Ten respondents felt the level of the information presented during the session was 'at the correct level', while two felt it was 'too specific', and one 'too general'. One respondent felt it was a combination of 'at the correct level' and 'too general'.

Level of the information presented

Not answered 7.1%
Too general 3.6%
Too specific 14.3%
1
2
10.5
Correct level 75.0%

Specific comments from respondents included:

- This session provided many useful strategies. Also, a lot of useful tools were mentioned that I plan to check and use!
- There are different types of data, quantitative and qualitative. This lecture only focused on quantitative data. It would be great to hear more about how can qualitative data could be shared.
- Maybe a hands-on session on preparing an example data set for publication and submitting it to a repository would be useful here.
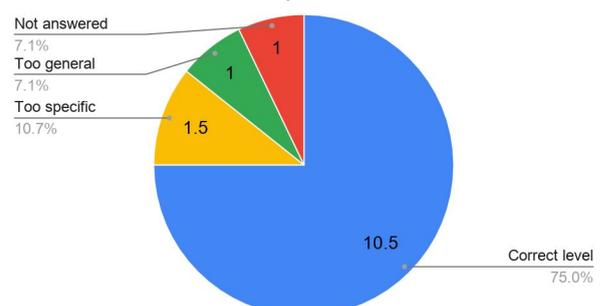- I would add tools to share data to make it a bit more practical/useful in the long term.

Module 7: Copyright and Plagiarism

One respondent selected two responses from the multiple choice, which are counted as 0.5 each.

Ten respondents felt the level of the information presented during the session was 'at the correct level', yet one felt it was 'too general'. One respondent felt it was the combination of 'at the correct level' and 'too specific'. One person did not answer.

Level of the information presented

Not answered 7.1%
Too general 7.1%
Too specific 10.7%
1
1
1.5
10.5
Correct level 75.0%

Specific comments from respondents included:

- Legal issues are a known barrier to data sharing, so I was very happy to hear more about intellectual property rights and licensing!
- Conceptual aspect of copyright and plagiarism became very clear. This is something we always face in research carrier and therefore it was highly useful session.
- There should have been more discussion on IPR on the data acquired by a group of people. Something from legal point of view.
- I would have liked information on choosing appropriate licenses for data sets, although this has to be decided for every data set individually. But some very general issues to consider would have been nice (for example: when can it be useful to choose an NC or ND license etc.)
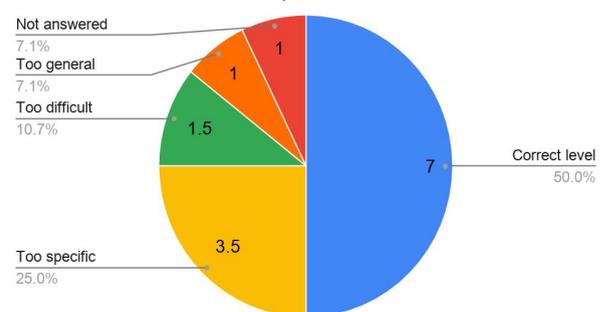
Module 8: Big Data

One respondent selected two responses from the multiple choice, which are counted as 0.5 each.

Seven respondents felt the level of the information presented during the session was 'at the correct level', with one selecting it was 'too general' and one 'too difficult'. One respondent felt it was a combination of 'too difficult' and 'too specific'. One person did not answer.

Level of the information presented

Not answered 7.1% — 1
Too general 7.1% — 1
Too difficult 10.7% — 1.5
Too specific 25.0% — 3.5
Correct level 50.0% — 7

Specific comments from respondents included:

- This was very interesting but at some times rather difficult to follow for a non specialist
- There were a lot of technical information but they were presented with simple examples that made them easy to grasp.
- This session provided a great overview of the relationship between available data and research methods.
- The presentation was too long with very congested slides and too much material presented within a very short time without much explanation.
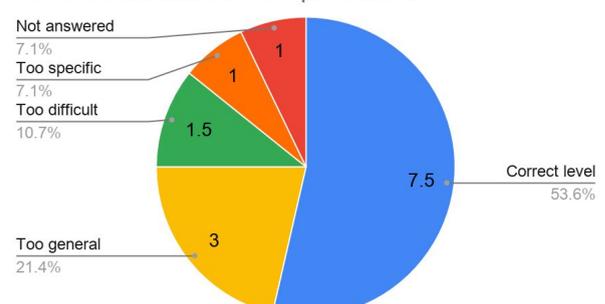
Module 9: Ontologies and Knowledge Networks

Three respondents selected two responses from the multiple choice, which are counted as 0.5 each.

Level of the information presented

Not answered 7.1% — 1
Too specific 7.1% — 1
Too difficult 10.7% — 1.5
Too general 21.4% — 3
Correct level 53.6% — 7.5

Seven respondents felt the level of the information presented during the session was 'at the correct level', yet two respondents felt it was 'too general' and one 'too difficult'. Two respondents give the combination of answers that the session was 'at the correct level' and 'too general', and one that it was 'at the correct level' and 'too difficult'. One person did not answer.

Specific comments from respondents included:

- Was hurriedly presented, the slides were very congested and busy.
- I think this was a rather interesting introduction to the powerful world of semantic web. Although the talk included ontologies that have been put in use for data interoperability but I still miss the practicality of data interoperability.
- Was a bit too general and for somebody who is new to the topic a bit more information on the background would have been useful.
- The related basic material would have been circulated prior to the workshop. This would have been more beneficial for the participants.
- I liked the examples of discipline specific vocabularies. Finding appropriate vocabularies can be difficult to find.
- I'd love to have an exercise to make some ontologies together, then it would be easier to understand. When I was a student, we made some ontologies in GIS. I remember, at that lecture I understood why we need ontologies to make databases.
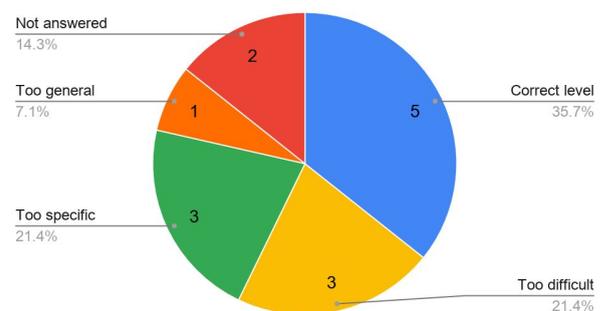
Module 10: Cloud Platforms in Science

Five respondents said the information presented during this session was 'at the correct level'. In contrast, three respondents said it was 'too difficult' and another three 'too specific', while one respondent selected 'too general'. Two respondents did not answer.

Specific comments from respondents included:

- I liked that this session showed the variety of cloud computing and storing services.
- Again, this session was very well structured. Although it was quite technical, I could easily follow.
- Interesting, yet challenging to grasp everything. I think it was important to have this presentation to introduce the new developments in data storage and management but it would have been best to shift the talk earlier in the day..
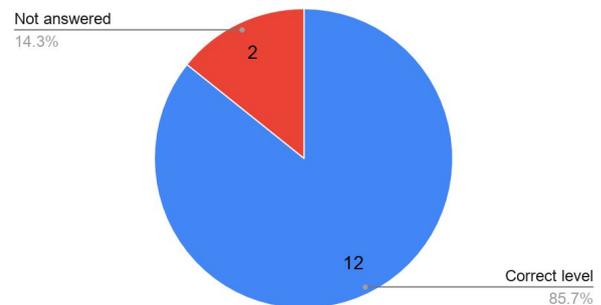
● I hope to learn some useful tools for managing data in clouds, e.g. where and which software, and some hands-on practice.

Module 11: Data Repositories and Their Certification

12 respondents said the information presented during this session was 'at correct level'. Two respondents did not answer.
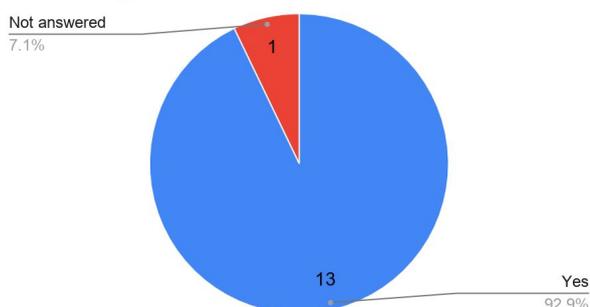
Specific comments from respondents included:

Level of the information presented

Not answered
14.3%

2

12

Correct level
85.7%

● Too much information within a very short time.
● Same as cloud platforms. I suggest to do a quick exercise.
● Before WDS-EGU workshop I didn't know that there are certificates for data repositories at all. Everything was well explained. Thank you!
● The part about the roles and responsibilities of a data repository (beyond just storing data) was very informative for me
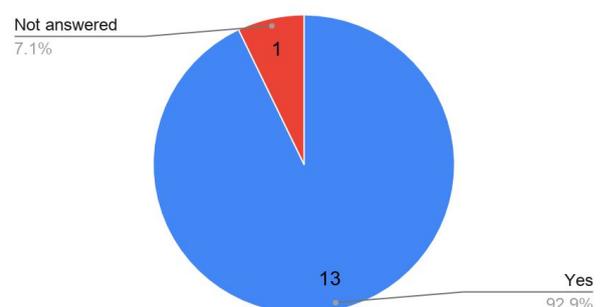
Module 12: The Roles and Responsibilities of the World Data System

Except for one person who did not answer, all respondents felt that the session was interesting/useful and that they understood the roles of WDS.

Interesting/useful?

Not answered
7.1%

1

13

Yes
92.9%

Understand the role of WDS?

Not answered
7.1%

1

13

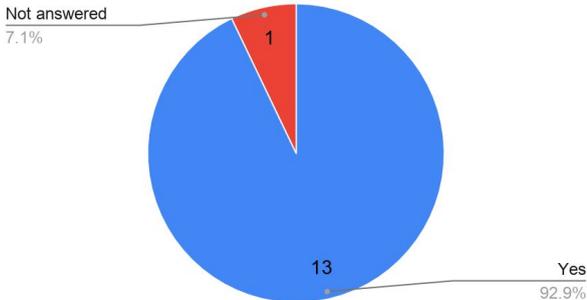Yes
92.9%

Specific comments from respondents included:

● At the moment I don't have questions. But I'm pretty sure that I will have some in the future and I know which people I can contact.
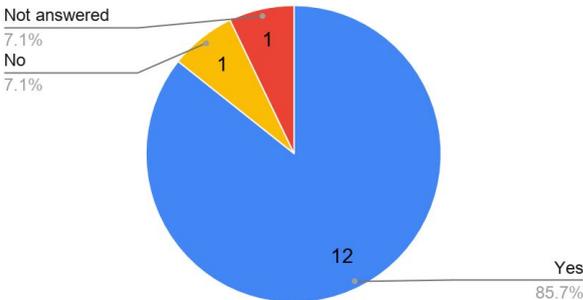● How to take forward the mission of WDS to the rest of the world, especially to India.

Module 13: WDS Early Career Researcher and Scientists Network

All respondents felt the session was interesting/useful, except for one who did not answer. The majority of the respondents intend to join and contribute to the WDS ECR Network.

Specific comments from respondents included:

- If it is possible, I would like to stay informed on the network's activities. Because I am still working on my master's degree, I don't think I could contribute much at the moment. But I would definitely be open to join once I am finished!
- I already joined!
- I am not sure how I could contribute to the Network.

—(End of Document)—