

23rd Plenary Meeting Collaborative Notes

Group(s) name(s) organising the session: *BoF - Stefanie Kethers*

Session link: [Agricultural Research Data Integrity](#)

Session scheduled date/time/breakout session: *Wednesday, 13 November, Breakout 3, 08:00 – 09:30 CST*

Venue: *Edificio de Ingeniería, Classroom 206*

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.

Summarise the session in three sentences:
This session introduced Agricultural Data Integrity as a potential topic for a Working Group within RDA. After the presentation of our problem statement and a few lightning talks to set the scene, the second half of the session started to clarify the scope of the proposed group. This included discussion of what we mean by “Agricultural Data” and “Agricultural Data Integrity”, as well as a brief discussion of use cases and key parties to be involved.
Key outcomes/actions/takeaways
1. <i>There was interest in forming a WG on this topic, under the umbrella of the RDA Improving Global Agricultural Data (IGAD) CoP.</i> 2. <i>We need to provide clear definitions of “Agricultural Data” and “Agricultural Data Integrity”.</i> 3. <i>We need to capture use cases in agriculture data integrity which will be utilised in a definition and guidance document for policy and designing technical solutions.</i>
Synergies and/or possible collaborations identified with RDA groups and other groups:
• RDA Improving Global Agricultural Data (IGAD) Community of Practice • RDA Ethics in Agricultural Data WG
Highlight text that will be used in social media mentions (please make sure the text is clear and appropriate for public consumption & comprehension)
Direct link to Group home page

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
Stefanie Kethers	Australian Research Data Commons	Melbourne, Australia	stefanie.kethers@ardc.edu.au
Cynthia Sims Parr	USDA-ARS	Vermont, US	cynthia.parr@usda.gov
David Molik	United States Department of Agriculture, Agriculture Research Service (USDA-ARS)	Manhattan, Kansas, USA	david.molik@usda.gov david.molik@gmail.com
Whitney Zatzkin	Bioeconomy Information Sharing and Analysis Center	Washington, DC USA isac.bio	whitney@isac.bio
Ryan O'Connor	RDA Europe (RDA TIGER!)	Cork, Ireland	ryan.oconnor@rda-foundation.org
Anne-Sophie Bage	INRAE	France	anne-sophie.bage@inrae.fr
Limor Peer	Yale University	USA	limor.peer@yale.edu
Clement Jonquet	INRAE	Montpellier, France	Clement Jonquet
Trish Radotic	ARDC, RDA	Perth, Western Australia	trish.radotic@ardc.edu.au
Hassan Oukhouya	Laboratory LMSA, Department of Mathematics, Faculty of Science, Mohammed V University in Rabat	Morocco	hassan_oukhouya@um5.ac.ma

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

Slide deck:

https://docs.google.com/presentation/d/18YMEav7RXBpiJEAXg3puVeCqacO8LUOL/edit#slide=id.g31186e39355_0_0

Agenda

- 1) Welcome, introductions: Stefanie Kethers
- 2) Problem Statement: Whitney Zatzkin
 - We need to consider what actions we are taking to protect our data from contamination during our work and analyses.
 - What criterion for trustworthiness must be met before data is shared through this public data initiative?
 - Who pushes the button to upload new datasets to the world's leading open science datasets? What is the screening process for someone asking to use the data? When is it verified for duplicate submissions?
 - What protects data integrity (accuracy, completeness, quality)?
 - [More than 10,000 peer-reviewed publications](#) were removed in 2023 due to concerns with their data. We are leveraging algorithms on our data across the global economy and must make those decisions with safe, secure data that is saying what we believe it is saying.
- 3) Lightning talks
 - David Molik - Use Case
 - Chris Baker (pre-recorded video) - Use Case
 - Stefanie Kethers - International Data Spaces (skipped due to lack of time)
 - Cynthia Parr - Agricultural Community of Practice
 - For those interested in the four Research Agencies of the USDA:
<https://www.ree.usda.gov/agencies-and-offices>
 - Action: Join the Ethics in Agricultural Data WG
<https://www.rd-alliance.org/groups/ethics-in-agricultural-ag-data-wg/activity/>
 - Q&A
- 4) Discussion

A few questions to get us started - please feel free to add your own

 - Can you think of other use cases to consider?

- Centralised Repositories
- Farm generated data
- What are the socio-technical requirements resulting from these use cases?
- Some level of technical know-how
- Are you aware of any existing solutions that might help address these requirements?

5) Wrap-up and next steps

What does Agricultural Data Mean?

How do you track it before ag data gets shared? How do you ensure the data provided matches the metadata provided.

Data as an asset, currency - import and export law enforcement perspective; this forum might be a good place to develop and discuss these ideas.

Put risk in the problem statements, asking the relevant questions asking if this data is safe to use and to what extent.

Who might need to be recruited and included in this effort. Eg. including manufacturers for example, look at bringing one kind of group like this (industry partners)

Equipment manufacturers

Ag Researchers (government and academic)

Policy

Develop a list of relevant use cases.

Look at other interested people who can help to drive these WGs, from at least two regions outside of US and Europe.

Defining Ag Data Integrity Word Cloud

Trust in the data

Knowing of the provenance of the data from the inception of the data

Sensitivity of the data

Potential of the data for misuse

Impact of the data

Potentialities of the data (in general)

Avoiding the commoditization of data

Risk

Inherent risk

Cost

Defining socio-technical aspects of Ag Data Integrity

Perception of data integrity

Has to be seen for what it is

Defining the work

Capture use cases in agriculture data integrity which will be utilised in a definition and guidance document for policy and designing technical solutions.