

# Social Dynamics of Data Interoperability IG RDA18 Session Report

Distributed 2021-12-06

This report has been added as a record on Zenodo and is contained within the [Zenodo RDA Social Dynamics of Data Interoperability IG community](#).

## Table of Contents (subheadings are hyperlinked)

<b>Overview</b>	<b>2</b>
<b>Links</b>	<b>2</b>
<b>Introduction</b>	<b>2</b>
<b>Invited Presentation</b>	<b>3</b>
The social dynamics of data sharing: Steven McEachern	3
<b>Discussion: Developing a survey to examine the social dynamics that influence the implementation of the FAIR principles.</b>	<b>4</b>
<b>Next steps</b>	<b>6</b>

Contributors to this session report. Contributors are listed in alphabetical order using the [CRedit Contributor](#) framework.

<b>Name</b>	<b>CRedit</b>
Kristal Spreadborough	Writing – original draft, Writing – review & editing, Project administration, Conceptualization
Kheeran Dharmawardena	Writing – review & editing, Conceptualization

## Overview

The goal of this document is to provide a high level summary of the RDA18 session. It is designed to complement the [session recording](#) by highlighting the key themes and next steps from the session. The structure of this report reflects that of the [RDA18 session](#).

## Links

Zenodo community	<a href="https://zenodo.org/communities/rda-sensitive-data-ig">https://zenodo.org/communities/rda-sensitive-data-ig</a>
Zotero group	<a href="https://www.zotero.org/groups/4412096/rda_sddi_ig">https://www.zotero.org/groups/4412096/rda_sddi_ig</a>
Link to current charter	<a href="https://www.rd-alliance.org/group/social-dynamics-data-interoperability-ig/case-statement/social-dynamics-data-interoperability">https://www.rd-alliance.org/group/social-dynamics-data-interoperability-ig/case-statement/social-dynamics-data-interoperability</a>
Link to session recording	<a href="https://www.youtube.com/watch?v=aMqkqm0ggXo&amp;list=PL77Dq-HHs58ZC_0rzK7hZWGR8CMyKvi9c&amp;index=5">https://www.youtube.com/watch?v=aMqkqm0ggXo&amp;list=PL77Dq-HHs58ZC_0rzK7hZWGR8CMyKvi9c&amp;index=5</a>
Link to session page	<a href="https://www.rd-alliance.org/plenaries/rda-18th-plenary-meeting-virtual/10-key-social-challenges-sharing-data">https://www.rd-alliance.org/plenaries/rda-18th-plenary-meeting-virtual/10-key-social-challenges-sharing-data</a>
Notes from the session (including copy of the chat)	<a href="https://docs.google.com/document/d/1xuwo9Pc5K3kMZccwvhblflmeC-Wt2KcQdRe1iaV5xfl/edit?usp=sharing">https://docs.google.com/document/d/1xuwo9Pc5K3kMZccwvhblflmeC-Wt2KcQdRe1iaV5xfl/edit?usp=sharing</a>

## Introduction

Slides available here: <https://zenodo.org/record/5748372>

Introduction given by Kristal Spreadborough

Kristal gave a brief overview of the IG to date including key events ([slide 4](#)), an update on the charter and current themes the IG is addressing ([slides 5 - 6](#)), and current mechanisms for resource sharing ([slide 7](#)). Planned activity was introduced ([slide 8](#)).

Kristal introduced Dr. Steven McEachern who gave the invited presentation at the session ([slide 9](#)).

# Invited Presentation

## The social dynamics of data sharing: Steven McEachern

Slides available here: <https://zenodo.org/record/5748601>

Steve gave a 15 min presentation followed by a 10 min discussion. Themes from the presentation included:

- The supply of data - data custodians ([slides 3 - 4](#)): What are custodians' concerns. Steve discussed the incentives and risks associated with data sharing from the perspective of the data holders. Some elements include:
  - Privacy and confidentiality
  - Intended use
  - Intellectual property
- The demand for data - secondary users ([slides 5 - 6](#)): What do users expect? Steve discussed some the requirements of data users - the individuals or groups who would use the data shared in the above point.
- A model of data sharing - The Theory of Planned Behaviour ([slides 9 - 10](#)): Steve used this theory to explain how/why people might be encouraged to share data. Elements of this theory include:
  - Attitudes: attitudes towards the behaviour. E.g. I believe that data sharing is important
  - Norms: perceptions of the judgment of significant others. E.g. My colleagues would expect me to share my data
  - Perceived behavioural control: beliefs about control over one's ability to undertake the behaviour. E.g. I think it is relatively easy for me to share my data
  - Behavioural intention: the intent to perform the action. E.g. I intend to share the data I produced in my research project
  - Behaviour: the actual performing of the action. E.g. the deposit of the data in a repository
- Aligning supply and demand: Bridging the gap ([slides 11 - 20](#)) - Steve shared the results of some studies which examined data sharing practices, including those that considered this through the lens of the Theory of Planned Behaviour. Studies include ([full citations in the slides](#)):
  - McEachern & McDougall (2014)
  - Kim and Stanton (2015)
  - Zenk- Moltgen et al. (2018)
  - Melero and Navarro-Melina (2020)
- Steve concluded with prompts to consider when asking what is needed to encourage data sharing ([slide 21](#)).

The discussion addressed the following topics:

- Q: Is there conflicting data on whether social norms impact data sharing practices? A: it may be explained by domain norms - some domains tend to share data more than others.
- Q: If that data suggests that people who have more control over their data actually share less, could this be because people complete control with responsibility/risk? Data is not clear yet, institutional mandates, funder and journal policies could be impacting this.

## Discussion: Developing a survey to examine the social dynamics that influence the implementation of the FAIR principles.

Slides available here: <https://zenodo.org/record/5748372>

Discussion moderated by Kheeran Dharmawardena

Kheeran led a 45 minute discussion and collaboration activity focused on gathering key questions to inform the development of a survey investigating the social dynamics that influence the implication of the FAIR principles. We used the following prompts:

- What questions could we ask about the social aspects that foster or hinder making data Findable?
- What questions could we ask about the social aspects that foster or hinder making data Accessible?
- What questions could we ask about the social aspects that foster or hinder making data Interoperable?
- What questions could we ask about the social aspects that foster or hinder making data Reusable?

People contributed the following questions. [These are available as a csv here.](#)

### What questions could we ask about the social aspects that foster or hinder making data Findable?

Top

- 4 ■ How do we foster trust in data repositories? e.g. trust that researchers can maintain control over data, trust that the repository will be an enduring record.
- 4 ■ What are the disincentives and risks, real or perceived that prevent people from sharing data?
- 4 ■ Is data sharing through repositories valued in your organization?
- 4 ■ As researchers are usually busy and loaded with work, they would need specified worktime to share their data and make it findable.
- 4 ■ What are the demonstrable incentives for individuals to participate in data sharing?
- 3 ■ What are the perceived risks of making your data findable, if any?
- 2 ■ What incentives do you experience at your institution that encourage data sharing?
- 2 ■ Under what conditions are curation services required and what are the needs for those services?
- 2 ■ What do researchers consider 'findable' themselves, and what do they do to achieve that?
- 2 ■ How much "do it yourself" is required across different repos?
- 2 ■ What benefits from sharing data through repositories do you perceive?
- 2 ■ How secure do researchers feel in terms of their career situation?
- 2 ■ How do we create governance that is somewhat similar across repositories so that stakeholders know more or less what to expect when interacting with these repos?
- 2 ■ How should infrastructures that make data findable be governed?
- 2 ■ What level of "data findability" is normal in different disciplines, and different regions?
- 1 ■ What are the benefits of depositing large amounts of data?
- 1 ■ Is it useful to consider categories of repositories based on the service layer?
- 0 | how do people understand the principle of reuse?
- 0 | how do you define reuse?

### What questions could we ask about the social aspects that foster or hinder making data Accessible?

Top

- 2 ■ Do you want to know who uses your shared data? And do you want to have an effect on how it is used?
- 1 ■ Do you support researchers with metadata and data documentation?
- 1 ■ Does the institution offer guidelines and support for researchers working with qualitative data?
- 0 | How would you define interoperability?
- 0 | What governance arrangements do you have in place for managing who can access data through your service?

### What questions could we ask about the social aspects that foster or hinder making data Interoperable?

Top

- 7 ■ Do researchers understand what interoperability of data (on their field of science) mean?
- 4 ■ When making your data interoperable, who do you have in mind as the future user?
- 3 ■ How to ensure coordinated, interoperable Research Data Management at different national research institutes?
- 2 ■ Are there established standards or schemas for the types of data you create?
- 2 ■ What additional effort is involved in making your data interoperable?
- 2 ■ Do you understand what metadata means?
- 2 ■ Do researchers perceive that 'data sharing' is relevant to their work?
- 2 ■ What is your understanding of interoperability?
- 2 ■ How do you encourage people to plan for interoperability even if their data will not be "open"?
- 2 ■ Do you support researchers with data standards? data formats?
- 2 ■ What kind of data produced and shared by other researchers, would help you in your research?
- 2 ■ When you do share data, why do you do so? Conversely, if you don't why don't you?
- 2 ■ How important is interoperability of data to researchers using your service?
- 2 ■ Do you know what are open file formats?
- 2 ■ Do researchers know how to make it interoperable?
- 1 ■ How do you anticipate future users' needs? (any institutional support for instance?)
- 1 ■ What are more junior researchers being told by more senior ones about the risks and rewards of sharing?
- 0 | to what extent are researchers of different subject-specific fields locked in to proprietary services?
- 0 | What are some reasons that you might not use an open file format?

**What questions could we ask about the social aspects that foster or hinder making data Reusable?**

Top

- 1 Do data infrastructures have built-in visual indicators, labels to signal the degree of trustworthiness and operational readiness of datasets?
- 2 Do researchers have enough skills to reuse data available on a data platform?
- 3 Do research want their data to be reused?
- 4 how you encourage people to trust others data?
- 5 How much time do you spend on making your data re-usable?
- 6 Have opening your data increased the amount of collaboration/publications?
- 7 Does your institution have any policy regarding sharing and reuse of data
- 8 Are you afraid that someone would find mistakes in your data?
- 9 Are you afraid that someone would use your data without citing it?
- 10 If you would get direct benefits, i.e. money, would you make your data reusable?
- 11 Are the ways you measure variables understandable for future reuse?
- 12 what education is needed to understand how data can be reused (rather than just re gathered)?
- 13 For computational scientist: Do they know what a reproducible compendium/package is and should contain?

There was some overlap between these questions, so the co-chairs will curate a list of questions and share them at or before the Community Meeting in 2022.

## Next steps

1. Form small working groups to compile the questions into a survey which we can deploy in the RDA community - participants were asked to sign up on the Meeting Notes Document if they were interested in being involved. The co-chairs will use this to contact people.
2. Continue to cultivate discussion in this area.
3. Next community meeting scheduled for 1 Feb, 2022, 12:00:00 UTC

There is a rolling call for participation, please do get in touch!