

THE RESEARCH DATA ALLIANCE COVID-19 DATA SHARING RECOMMENDATIONS AND GUIDELINES

Natalie Harrower | RDA COVID-19 WG Co-Chair, Director Digital Repository of Ireland [@natalieharrower](#)
Mark Leggott | RDA COVID-19 WG Co-Chair, Research Data Canada [@rdc_drc](#)



Agenda

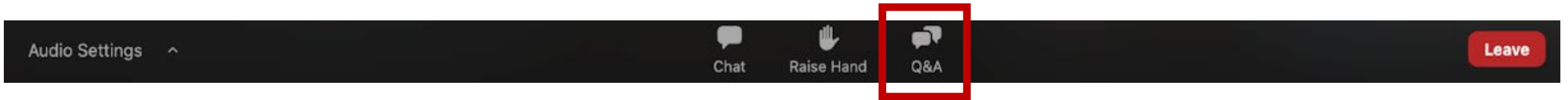
- > What is the RDA COVID-19 working group and why should we pay attention to its outputs?
- > Overview of the *RDA COVID-19 Recommendations and Guidelines on Data Sharing*
- > Learnings from the collaborative writing process
- > Example from a sub-group: 'Indigenous Data'
- > Q&A

<https://doi.org/10.15497/rda00052>



Questions & Answers

- > Please use the Q&A option to ask questions of the presenters. Questions will be monitored throughout the session and addressed at the end.
- > The Q&A option can be found at the bottom of your Zoom screen:



- > Please note that this **event is being recorded**, including questions and answers.

Background and Why

- > Request from the European Commission to the Research Data Alliance (RDA)
- > Working Group setup within weeks
 - 4 Research Areas, 4 Cross-cutting themes, each with Moderators
- > Structured through a series of teams
 - Co-Chairs, Co-Chairs + Moderators, 8 Themes, Editorial, Visualization, Zotero
- > April 1 - 30 June continual sprints, consultations, webinars,
 - 6 releases over 3 months
 - 143 pages in the end; 4-page Executive Summary, Infographic +
- > Exhausting and exhilarating!



Overview of the Guidelines



What are the Challenges Being Faced?



Critical Need for Rapid Data Sharing

Rapid massive research response with diverse outputs challenges **interoperability of data.**

A trade off between...



Timeliness

Precision



Lack of Harmonised Universal Standards and Context

Lack of pre-approved sharing agreements and archaic information systems hinder rapid threat detection and evidence-based response.

No universally adopted system or standard for



collecting



documenting



disseminating

COVID-19 research outputs.



Lack of documentation, context, and appropriate licensing challenges **reusability.**

What are the Objectives?



1.0
Clearly define detailed guidelines on data and software sharing for COVID-19 research.



1.1
Help stakeholders follow best practices to **maximise efficiency**.



1.2
Act as a **blueprint** for future emergencies to maximise the efficiency of their work.



2
Develop **recommendations** for funders and policymakers to maximise timely, quality data and software sharing and appropriate responses in health emergencies.





3
Address interests of researchers, policymakers, funders, publishers, and providers of data sharing infrastructures.

A Collaborative Cross-Disciplinary Effort

The work has been divided into **four research areas** with **four cross-cutting themes**.


The **guidelines** and recommendations listed here are highlights. Please find more detailed information in the [full-length publication](#).


 **Guidelines** - detailed practical advice aimed at researchers, data stewards, research software engineers, and public health officials.

 **Recommendations** - higher level generic advice aimed at policymakers, funders, publishers, and infrastructure providers.




CLINICAL


 Standardise terminologies, and find balance between timely data sharing and protecting privacy, confidentiality

 Organise data sharing and trial documents in trustworthy repositories




OMICS


 Select the best data formats and standards to fit the sub-discipline

 Promote use of domain-specific repositories to enable standardisation




EPIDEMIOLOGY


 Data models must include clinical data, disease milestones, indicators, reporting data, contact tracing and personal risk factors

 Incentivise publication of situational data, analytical models, scientific findings and reports



SOCIAL SCIENCES

 Enable interoperable cross-disciplinary, cross-cultural data use and collaboration


 Ensure robust funding streams for research aimed at understanding and managing the human aspects of the pandemic


Omics Example: 4.4.2 Guidelines for Host Genomics Data

Several different types of host genomics data are being collected for COVID-19 research. Some suitable repositories for these are:


1. **Gene expression data** should in general be retrieved from or deposited in the repositories listed below ([Blaxter et al., 2016](#)). To achieve load balancing, it is recommended to choose the respective regional repository. It should be noted that [INSDC](#) resources (i.e., [DDBJ](#), [ENA](#) and [NCBI](#)) synchronise most of their datasets daily².
 - 1.1. Transcriptomics of human subjects (requiring authorised access):
 - 1.1.1. [Database of Genotypes and Phenotypes](#) (dbGaP) ([Mailman et al., 2007](#))
 - 1.1.2. [European Genome-Phenome Archive](#) (EGA) ([Lappalainen et al., 2015](#)); the corresponding non-sensitive metadata will be available through EBI [ArrayExpress](#) ([Athar et al., 2019](#))
 - 1.1.3. [Japanese Genotype-phenotype Archive](#) (JGA) ([Kodama et al., 2015](#))
 - 1.2. Transcriptomics (from cell lines/animals):
 - 1.2.1. [ArrayExpress](#) ([Athar et al., 2019](#))
 - 1.2.2. [Gene Expression Omnibus](#) ([Barrett et al., 2013](#))
 - 1.2.3. [Genomic Expression Archive](#)
 - 1.3. Underlying reads can be retrieved from/will automatically be deposited to the corresponding read archive:
 - 1.3.1. [DDBJ Sequence Read Archive](#) (DRA) ([Kodama et al., 2012](#)), for submission documentation see [here](#)
 - 1.3.2. [European Nucleotide Archive](#) for submission documentation see [here](#)
 - 1.3.3. [NCBI Sequence Read Archive](#) (SRA) for submission documentation see [here](#)
 - 1.4. Microarray-based gene expression data:
 - 1.4.1. [ArrayExpress](#) ([Athar et al., 2019](#))
 - 1.4.2. [Gene Expression Omnibus](#) ([Barrett et al., 2013](#))


COMMUNITY

 Encourage public and patient involvement throughout data management lifecycle


 Balance between timely testing and contact tracing, emergency response, community safety, and individual privacy concerns


INDIGENOUS DATA GUIDELINES

 Indigenous governance of data collection, ownership, and sharing and use priorities is the central principle of Indigenous data sovereignty


 CARE Principles set minimum standards for collectors, users, and stewards of Indigenous data.


RESEARCH SOFTWARE

 Software used in data analysis must be able to reproduce results, if necessary

 Allocate financial resources to support development and maintenance of new research software

LEGAL AND ETHICAL CONSIDERATIONS

 Although the law provides the foundation for data handling, ethical frameworks should also inform expedited approval to maximise data use and sharing

 Expedite ethical review and approval for legal data sharing during a pandemic

Legal/Ethics Example: 10.4.5 Consent Guidelines

10.4.5 Consent

Consent is the act by which a participant, patient or data subject indicates that they permit something to happen to them, or to their data, which would otherwise not be able to happen. It covers a number of different specific contexts:

1. **Clinical:** a patient agrees to undergoing a procedure, including taking part in a trial;
2. **Data Protection:** a data subject agrees to personal data being processed for specified purposes;
3. **Research:** a participant agrees to take part in a research study or experiment.

In both cases, the informed consent sheets for clinical or research purposes would explicitly set out how data protection will be handled, as well as samples or biobanking, rights to self- images and others.

Giving consent should be informed (e.g. the individual knows what is going to happen and why), freely given (there is no coercion or similar motivation), given by somebody with capacity, unambiguous and auditable (the consent is recorded somewhere) (See also [Parra-Calderón, 2018](#)). Depending on the jurisdiction and the research domain, there may be an additional requirement to seek consent. This may include a representative community board as well as participants themselves.

Ideally, consent should be sought for collecting, processing, sharing and publishing data. However, there are other legal bases for processing personal data. Some specific examples from the European General Data Protection Regulation ([GDPR, 2016](#)) are described below. Our recommendation would therefore be as follows:

1. Where possible, use data where the data subject has provided a valid consent that includes or is compatible with intended use of the data and complies with the requirements on consent in the specific country or region.

Where these are not possible, there are other reasons why data may be used (see [Hallinan, 2020](#), [Ó](#)

Foundational Elements

What are the key recommendations?

The RDA COVID-19 Recommendations and Guidelines are aimed at developing a systematic approach for data sharing in public health emergencies that supports scientific research and policymaking, including an overarching framework, common tools and processes, and principles that can be embedded in research practice.

- 1** Coordinate cross-jurisdictional efforts to foster global **Open Science** through policy and investment.
- 2** Incentivise early publication and release of data and software outputs.
- 3** Invest in state of the art IT, data management systems **infrastructure, economies of scale, and people.**
- 4** Data, software and models should be **timely and FAIR: Findable, Accessible, Interoperable, Reusable.**
- 5** Require the use of **Data Management Plans.**
- 6** Use common domain-specific **metadata standards, and persistent identifiers.**
- 7** Provide **documentation** of context, methodologies used to define, construct, and compile data, data cleaning and quality checks, data imputation, and data provenance.
- 8** Use **Trustworthy Data Repositories** committed to the long-term preservation and sustained access to their data holdings.
- 9** Use common generic as well as domain-specific **metadata standards, and persistent identifiers.**
- 10** **Balance ethics and privacy,** taking into account public interests and benefits while addressing the health crisis.
- 11** Access should be as **open as possible** and as **closed as necessary.**
- 12** Seek **technical solutions** that ensure anonymisation, encryption, privacy protection, and de-identification to **increase trust** in data sharing.
- 13** Provide **legal frameworks that promote sharing** of surveillance data across jurisdictions and sectors.

Foundational Elements

With thanks to the team at CANARIE and Research Data Canada for the [Infographic](#)



RDA Community response

- > Call to action to create a fast track Working Group aimed at developing a system for data sharing in public health emergencies, specifically COVID-19
- > Around 600 RDA members and newcomers registered for the different groups
- > 165 active contributors to the documents
- > 6 Co-Chairs + Secretariat
- > Experts in different fields as group moderators
- > Regular calls and iterations
- > Weekly webinars, requests for comments
- > 5 releases produced (April – June 2020)
- > Final release – 30 June 2020

RDA's guiding principles:

- ✓ Openness
- ✓ Consensus
- ✓ Balance
- ✓ Harmonization
- ✓ Community-driven
- ✓ Non-profit and technology-neutral

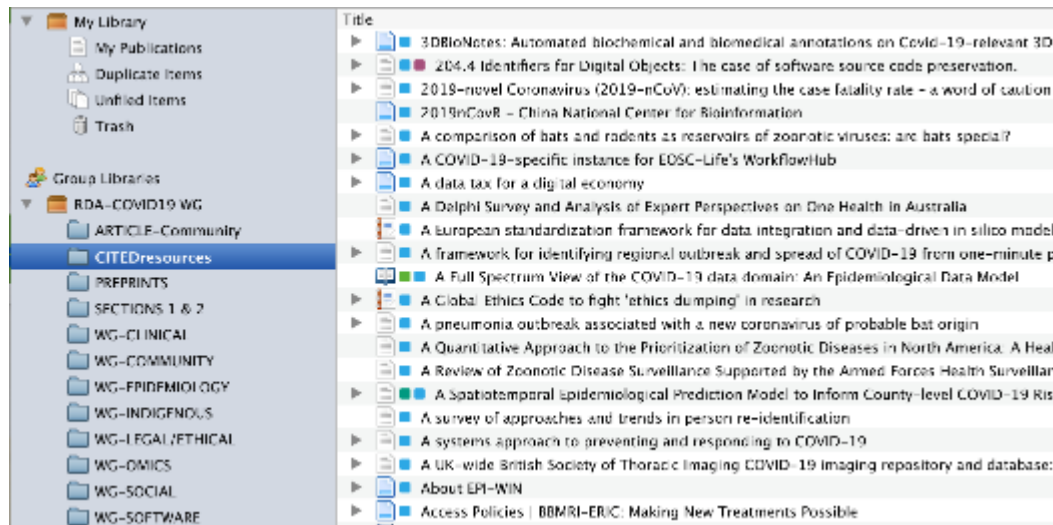
Indigenous Data sub-group Example

- > Group membership of a dozen
- > Produced the output at the final stages, in 2-3 weeks
- > Based on the CARE Principles previously written by the Global Indigenous Data Alliance (GIDA)
- > COLLECTIVE BENEFIT: “Data ecosystems shall be designed and function in ways that enable Indigenous Peoples to derive benefit from the data.”

Indigenous Data Guidelines	Indigenous data rights, priorities and interests must be recognised in all COVID-19 research and surveillance activities	Indigenous governance of data collection, ownership, sharing and use priorities is the central principle of Indigenous data sovereignty	<u>CARE Principles of Indigenous Data Governance</u> set minimum standards for collectors, users and stewards of data
-----------------------------------	--	---	---

RDA COVID-19 Zotero Library & Infographic

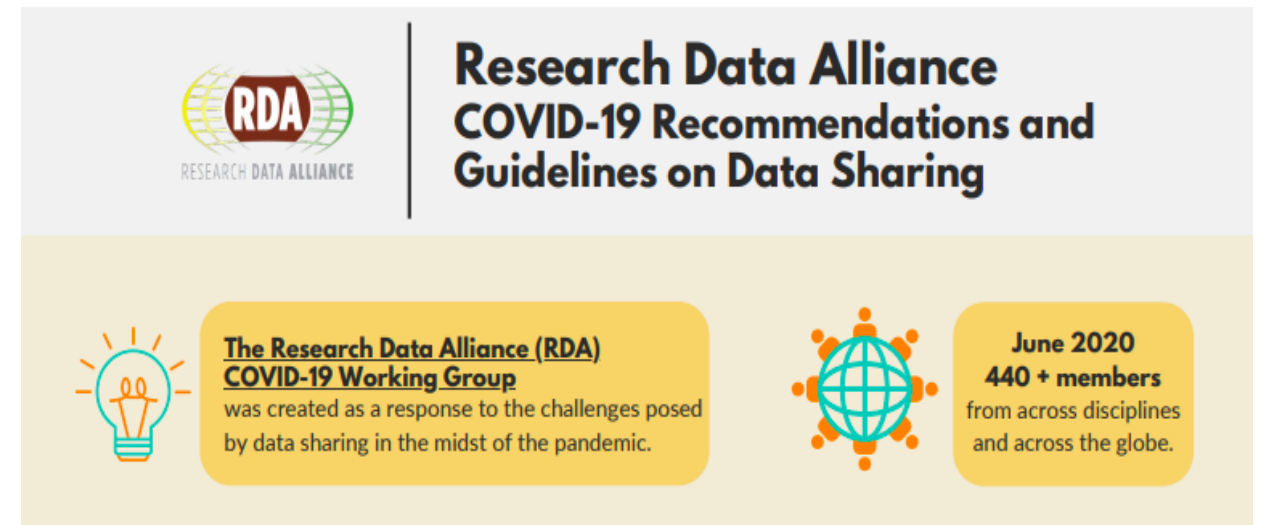
RDA COVID-19 Zotero Library



The screenshot shows a Zotero library interface. On the left, there is a sidebar with a tree view of folders. The main area displays a list of items with columns for 'Title' and 'Date Added'. The 'CITEDresources' folder is selected in the sidebar.

Title	Date Added
3DBioNotes: Automated biochemical and biomedical annotations on Covid-19-relevant 3D	
204.4 Identifiers for Digital Objects: The case of software source code preservation.	
2019-novel Coronavirus (2019-nCoV): estimating the case fatality rate - a word of caution	
2019nCovR - China National Center for Bioinformation	
A comparison of bats and rodents as reservoirs of zoonotic viruses: are bats special?	
A COVID-19-specific instance for EOSC-Life's WorkflowHub	
A data tax for a digital economy	
A Delphi Survey and Analysis of Expert Perspectives on One Health in Australia	
A European standardization framework for data integration and data-driven in silico model	
A framework for identifying regional outbreak and spread of COVID-19 from one-minute p	
A Full Spectrum View of the COVID-19 data domain: An Epidemiological Data Model	
A Global Ethics Code to fight 'ethics dumping' in research	
A pneumonia outbreak associated with a new coronavirus of probable bat origin	
A Quantitative Approach to the Prioritization of Zoonotic Diseases in North America: A Heal	
A Review of Zoonotic Disease Surveillance Supported by the Armed Forces Health Surveillar	
A spatiotemporal epidemiological Prediction Model to Inform County-level COVID-19 Ris	
A survey of approaches and trends in person re-identification	
A systems approach to preventing and responding to COVID-19	
A UK-wide British Society of Thoracic Imaging COVID-19 Imaging repository and database:	
About EPI-WIN	
Access Policies BBMRI-ERIC: Making New Treatments Possible	

RDA COVID-19 Infographic



The infographic features the RDA logo (Research Data Alliance) on the left. The main text reads: "Research Data Alliance COVID-19 Recommendations and Guidelines on Data Sharing". Below this, there are two yellow callout boxes. The first box, accompanied by a lightbulb icon, states: "The Research Data Alliance (RDA) COVID-19 Working Group was created as a response to the challenges posed by data sharing in the midst of the pandemic." The second box, accompanied by a globe icon, states: "June 2020 440+ members from across disciplines and across the globe."



RDA COVID-19 Data Stewardship Wizard

DS Wizard KNOWLEDGE MODEL TAGS PREVIEW Close

RDA Covid-19 Data Sharing Reco... Guidelines for Researchers

Expand all Collapse all

- RDA Covid-19 Data Sharing Recommendations: Omics
 - Guidelines for Researchers
 - Will you be re-using existing data?
 - Will you be collecting data?
 - No
 - Yes
 - Do you need generic guidance on metadata standards?
 - Do you need generic guidance on documenting data?
 - Do you need generic guidance on trustworthy repositories?
 - Do you need generic guidance on Ethics and Privacy?
 - Do you need generic guidance on Legal issues?
 - Will you be collecting clinical data?
 - Will you be collecting molecular biology data (omics)?
 - Yes
 - Do you need guidance for publishing Virus Genomics Data?
 - Do you need guidance for publishing Host Genomics Data?
 - Do you need guidance for publishing Structural Data?
 - Do you need guidance for publishing Proteomics Data?
 - Do you need guidance for publishing Metabolomics Data?
 - Do you need guidance for publishing Lipidomics Data
 - No
 - Will you be collecting epidemiological data?
 - Will you be collecting Social Science Data?
 - Will you be collecting data from Community Participation?
 - FAIR principles
 - Will you be processing data?
 - Will you be publishing/sharing data?
 - Recommendations for Policymakers
 - Recommendations for Funders
 - Recommendations for Publishers
 - Recommendations for Providers of Data Sharing Infrastructures

Chapter 0817e940 Delete

Title

Guidelines for Researchers

Text

Editor Preview

The COVID pandemic does not serve to remove the basic validity of the rights and interests on which these documents and principles are based. In other words, formal protocols for conducting research are required both during a pandemic and at other times, unless otherwise modified by the relevant authorities. The emergency does, however, mandate a reconsideration of the balance between these rights and interests - in particular between a research subject's right to privacy and the public interest in the outcome of research. In

You can use Markdown and see the result in the Preview tab.

Questions

- [Will you be re-using existing data?](#)
- [Will you be collecting data?](#)
- [Will you be processing data?](#)
- [Will you be publishing/sharing data?](#)

Ongoing Work and Future Steps

> Journal Articles and Endorsements

- 4-5 articles completed or in preparation by COVID-19 WG Members

> RDA Groups and RDA Plenary 16 Sessions

- Broader efforts under RDA WGs
- Infectious Disease BOF (goal to create a WG/CoP)
- Community Participation BoF / Citizen Science
- COVID-19 Epidemiology WG / Epidemiology WG

> Stakeholder support

- Adoption and implementation of the recommendations and guidelines;
- Policymakers, funders and publishers have a major influence on the behaviour of researchers and data stewards.

Value of RDA for COVID-19

Building the social and technical bridges to enable open sharing and re-use of data

RDA EU RDA US CONTACT US LOGIN REGISTRATION

O&A Members 63
Active Organisational & Affiliate members

MEMBERSHIP Members: 10966
Becoming a member of RDA is simple and open to both individuals and organisations
[Register now](#)

RDA Groups WG & IGs: 96
Discover what RDA Working and Interest Groups and all other Groups are up to and find out how to join them. [Explore Groups](#)

ABOUT RDA GET INVOLVED GROUPS RECOMMENDATIONS & OUTPUTS RDA FOR DISCIPLINES PLENARIES & EVENTS NEWS & MEDIA

The Value of RDA for COVID-19

Home » Get involved » The Value of RDA for... » The Value of RDA for COVID-19

13 July 2020 | 862 reads | Facebook | Twitter

Under public health emergencies, and particularly the COVID-19 pandemic, it is fundamental that data is shared in both a timely and an accurate manner. This coupled with the harmonisation of the many diverse data infrastructures is, now more than ever, imperative to share preliminary data and results early and often. It is clear that open research data is a key component to pandemic preparedness and response.

In late March, RDA received a direct request from one of its funders, the European Commission, to create global guidelines and recommendations for data sharing under COVID-19 circumstances. Over 600 data professionals and domain experts signed up and began work in early April 2020. They have produced a rich set of detailed guidelines to help researchers and data stewards follow best practices to maximise the efficiency of their work, and to act as a blueprint for future emergencies; coupled with recommendations to help policymakers and funders to maximise timely, quality data sharing and appropriate responses in such health emergencies.

On 30 June 2020, RDA published the final version of the RDA COVID-19 Recommendations and Guidelines on data sharing covering four research areas - clinical data, omics practices, epidemiology and social sciences - complemented by overarching areas focusing on legal and ethical considerations, research software, community participation and indigenous data.

The Outputs

The COVID-19 WG, from April 1st through June 30th, 2020, created more than five releases of the recommendations and guidelines, leading to the final endorsed version, "RDA COVID-19 Recommendations and Guidelines for Data Sharing," with ongoing efforts to add and review materials.

The Value of RDA for...

- COVID-19
- Funders
- Individuals
- Infrastructure Providers
- Libraries
- Organisations performing Research
- Regions
- Student/Early Career Programms
- The European Open Science Cloud (EOSC)
- Using RDA Community
- Request for Comments
- Call for Papers: Research Data Alliance
- Results Special Collection

The Outputs

The COVID-19 WG, from April 1st through June 30th, 2020, created more than five releases of the recommendations and guidelines, leading to the final endorsed version, "RDA COVID-19 Recommendations and Guidelines for Data Sharing," with ongoing efforts to add and review materials.

- RDA COVID-19 Recommendations and Guidelines for Data Sharing, final release, published 30 June 2020
- RDA COVID-19 Recommendations and Guidelines for Data Sharing Infographic
- RDA COVID-19 Guidelines and Recommendations - the prior 5 releases
- RDA-COVID19 WG Zotero Library

Citation: RDA COVID-19 Working Group. Recommendations and Guidelines on data sharing. Research Data Alliance, 2020. DOI: <https://doi.org/10.15497/rda00052>

Resources

- Final executive summary
- RDA COVID19 Press Release 30 June 2020 final - June 2020

Joint Statements

- RDA COVID-19 Recommendations and Guidelines for Data Sharing: How STM Publishers can Contribute (July 2020)
- GIDA-RDA COVID-19 Guidelines for Data Sharing Respecting Indigenous Data Sovereignty (July 2020)
- The Duty to Document does not Cease in a Crisis, it becomes more Essential (May 2020)
- Data Together COVID-19 Appeal And Actions (March 2020)

RDA FOR COVID-19 Events

A series of weekly "RDA COVID-19 Update Webinar" occurred almost every Tuesday between April and June 2020 and provided updates on the overarching COVID-19, Legal and Ethical, Research Software, Community Participation Working Groups, Indigenous Data contribution, and the four research themes (clinical, omics, epidemiology, social sciences), along with an opportunity for members to ask questions. Recordings and presentations from these sessions are posted on the **Events** meeting links.

Upcoming events include:

- RDA Ireland Meet The Experts Webinar - Data Sharing for COVID-19 Research: Recommendations and Guidelines from the RDA COVID-19 Working Group - 29 July 2020**

<https://www.rd-alliance.org/value-rda-covid-19-0>



...by the community for the community.

RDA

Email - enquiries@rd-alliance.org

Web - www.rd-alliance.org

Twitter - [@resdatall](https://twitter.com/resdatall)

LinkedIn - www.linkedin.com/in/ResearchDataAlliance

Slideshare - <http://www.slideshare.net/ResearchDataAlliance>

Natalie Harrower, Director, Digital Repository of Ireland [@natalieharrower](https://twitter.com/natalieharrower)

Mark Leggott, Executive Director, Research Data Canada [@rdc_drc](https://twitter.com/rdc_drc)



Thank you! Questions?