Research Data Alliance
COVID-19 Recommendations and Guidelines on Data Sharing

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.

June 2020
440 + members from across disciplines and across the globe.

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.

Global Effort to Raise the Bar for Data Sharing

- 117 cross-sectoral signatories to the Wellcome Trust statement in January 2020.
- Agreement by 30 leading publishers on immediate open access to COVID-19 publications and underlying data.

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**: **F**indable, **A**ccessible, **I**nteroperable, **R**eusable.
5. Require the use of **Data Management Plans**.
6. Use common generic as well as 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. **Expedite** article and data review processes, **prioritising** and **fast-tracking data** at all stages.
10. **Balance ethics and privacy,** taking into account individual versus public interests, and community benefit and value, 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.

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

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

- Read the full-length publication at: doi.org/10.15497/rda00052
- Visit the RDA COVID-19 Working Group Outputs page for updates from the RDA community, a Zotero library of references, and a series of guides and navigational tools to further facilitate access.

Infographic designed by CANARIE