

For more information see the [RDA Interest Group Template 14 April 2015 ksf 0](#)

RDA Interest Group Draft Charter Template

Name of Proposed Interest Group: Sensitive Data Interest Group

RDA site: <https://www.rd-alliance.org/groups/sensitive-data-interest-group>

1. Introduction (A brief articulation of what issues the IG will address, how this IG is aligned with the RDA mission, and how this IG would be a value-added contribution to the RDA community)

Sensitive Data: A working definition of sensitive data is: Information that is regulated by law due to possible risk for plants, animals, individuals and/or communities and for public and private organisations. Sensitive personal data include information related to racial or ethnic origin, political opinions, religious or philosophical beliefs, trade union membership and data concerning the health or sex life of an individual. These data that could be identifiable and potentially cause harm through their disclosure. For local and government authorities, sensitive data is related to security (political, diplomatic, military data, biohazard concerns, etc.), environmental risks (nuclear or other sensitive installations, for example) or environmental preservation (habitats, protected fauna or flora, in particular). The sensitive data of a private body concerns in particular strategic elements or elements likely to jeopardise its competitiveness.

Adapted from: David et al., 2020, "Templates for FAIRness evaluation criteria - RDA-SHARC IG" <https://zenodo.org/record/3922069#.YCU7ehKq2w>

A range of disciplines collect data which are potentially sensitive, presenting serious barriers to reuse and reproducibility. There are a number of barriers which need to be overcome before sensitive data can be utilised safely and to its best advantage. One major challenge is that not all sensitive data is alike, with significant disciplinary variation in how sensitive data is defined, linked, managed, stored, and reused. Additionally, common approaches to working with, sharing and managing data are not always appropriate for sensitive data. For example, sensitive data exposes the different perspectives underlying the FAIR and CARE principles. Further, sensitive data requires careful stewarding such that it can be disseminated in an ethically and culturally appropriate way. Nonetheless, sensitive data has significant potential to be utilised in the conduct of novel and impactful work. Therefore, it is essential that a set of community standards and best practices be developed for sensitive data usage and management.

Issues the IG will address

In addition to issues identified by the RDA community as this IG develops, we envisage this IG will address the following issues:

1. Data carries with it different levels of sensitivity depending on its context (e.g., research discipline, who the data is about, what the data is being used for). However, it is not always clear how we should assess data for sensitivities in different contexts. A resource is needed for those working with data to allow them to make informed decisions about data sensitivity and, consequently, data governance, management, and usage.
2. Sensitive data is often identified. However, re-identification can be possible and can cause serious harm. Resources are needed on mechanisms of reidentification and the different risks for different types of sensitive data.
3. Data that has been labeled sensitive is often not shared beyond the team that collected/created this data. This means that data collection is sometimes duplicated, and is a challenge for reproducible research. More ethically and culturally safe sharing of sensitive data may also enhance the robustness of research design and development. Resources are needed which provide information for those working with sensitive data with information about how that data can be shared and reused in a safe and ethical manner.
4. At times there is a duality between sharing and reusing data in general, and for stewarding data in culturally and ethically appropriate ways. This duality is exacerbated in the context of sensitive data due to lower rates of data sharing, and increased potential for harm. Guidelines are needed for balancing principles of data sharing and reuse (e.g., FAIR) with ethically and culturally appropriate principles (e.g., CARE) specifically in the context of sensitive data.
5. Consent is a major consideration when sharing any data, especially sensitive data. However, informed consent can be challenging to obtain, especially when reusing data. This is sometimes a barrier to sharing sensitive data. Guidelines are needed that explore consent models, especially post-hoc consent, for governing the primary and secondary use of sensitive data.

How this IG is aligned with the RDA mission

The RDA Vision: This IG aligns with the RDA vision because it will develop mechanisms for the responsible reuse of sensitive data - a data source that is both extremely valuable but which also carries many ethical and cultural considerations. Sensitive data will play an increasingly significant role in addressing the grand challenges of the 21st century, such as issues of social and environmental justice. Indeed, the benefits and potential harms of sensitive data are increasingly being discussed in public forums as corporations and private companies leverage such data for profit. As mechanisms for sensitive data reuse become widely available (such as through the work of this IG), new innovation and invention will be fostered through the reuse of sensitive data. This IG has participants from University and non-University sectors, which strongly positions the IG to engage with all the variety of stakeholders.

The RDA Mission: This IG aligns with the RDA mission as it develops guidelines for the technical components of working with sensitive data, and for addressing the social aspects of working with sensitive data including fostering discussion around the cultural and ethical considerations of data reuse. This IG is well positioned to meet these challenges given the diverse backgrounds of the initial members. The connection between the technical aspects of working with sensitive

data (such as secure virtual environments) and the ethical and cultural aspects (such as consent, disciplinary perspectives and norms, and CARE principles) is a key point of interest for this IG.

How this IG would be a value-added contribution to the RDA community

Sensitive Data is ubiquitous. However, its context varies. For this reason, this IG complements the work of a range of existing IGs and WGs, including:

- [COVID-19 Data Working Group](#)
- [Raising FAIRness in Health Data and Health Research Performing Organisations \(HRPOs\) Working Group](#)
- [Reproducible Health Data Services Working Group](#)

- [Biodiversity Data Integration](#) Interest Group
- [Education and Training on Handling of Research Data](#) Interest Group
- [Ethics and Social Aspects of Data](#) Interest Group
- [Health Data](#) Interest Group
- [International Indigenous Data Sovereignty](#) Interest Group
- [Social Dynamics of Data Interoperability](#) Interest Group
- [Social Sciences](#) Interest Group

- Infectious Diseases Community of Practice (forthcoming)

The aims of the Sensitive Data IG is to provide a space to focus explicitly on sensitive data. While the scope is interdisciplinary, this IG focuses on sensitive data types. Our planned activities will compliment the above IGs as we address sensitive data in domain specific terms (e.g., sensitive data in the health domains) as well as in general terms (e.g., systems for sharing sensitive data). The Sensitive Data IG already has members from a number of the above IGs, which will aid us in coordinating our activities with these groups. The Sensitive Data co-chairs are collectively members of over 20 RDA groups.

All members of the Sensitive Data IG are also active members of the RDA community. We will draw on this to ensure that our efforts take account of previous work in the RDA, and to ensure that our group remains up-to-date on RDA activities.

2. User scenario(s) or use case(s) the IG wishes to address (what triggered the desire for this IG in the first place):

We identified the following key reasons for forming this IG. We envisage that additional use cases will be developed through working with the RDA community following endorsement.

1. There are a lack of guidelines for working with sensitive data both within and between disciplines/research areas. One reason for this is because sensitive data varies between contexts (e.g., between disciplines). To develop a cohesive but also targeted set of

guidelines, a group is needed which comprises members of a range of disciplines with a shared interest in sensitive data.

2. There is a need for a framework which considers the ethical and cultural aspects of sensitive data, alongside the technical aspects. Individuals may want to share their sensitive data and may have conducted all the necessary ethical/cultural safe guards. However, they may lack an understanding of how this can be achieved with the technical resources available to them, what repository or sharing mechanism can handle such data, and how best to access persistent IDs which allow them to track the use of their data. Conversely, individuals may have the ideal technological solution for sharing without an understanding of the ethical/cultural considerations. A group is needed to facilitate a dialogue between the ethical/cultural and technical aspects of sensitive data sharing, and to produce tangible outputs which progress this discussion.
3. There is a general consensus that sensitive data is highly valuable but that it is not being utilised to its full potential. While there is a range of anecdotal support for this claim, a body of work is needed which explores and documents the state of sensitive data primary and secondary usage, and which examines the underlying causes of sensitive data reuse practices within and between disciplines.
4. There is a recognition that there are a number of stakeholders with respect to sensitive data assets, and that each stakeholder has different requirements, needs, expectations, and terminology (e.g., in the case of health data, government, hospitals, researchers, community members). A group is needed which can synthesise the main expectations of different stakeholders to develop resources of individuals and organisations to use when engaging with, sharing, and accessing sensitive data (i.e., a resource for a shared language between stakeholders).
5. There is a need for adequate and specialised resourcing and infrastructure to manage, work with, and share sensitive data. Different data types require different solutions for management, analysis, and sharing. While a range of solutions are available for these different data types, their suitability for sensitive data is not always clear. Work is required to assess solutions for different sensitive data types specifically.
6. Our era is experiencing the most brutal collapse in biodiversity that the earth has known. Yet biodiversity produces many ecosystem services, and resources. However, species and habitat diversity is undermined by many human activities. The preservation of both fragile and overly coveted species and resources makes the publication of their geolocation sensitive. Other data concerning the characteristics of certain pathogens have also proven to be sensitive.
7. The humanities and social science disciplines likewise require clear guidance regarding collection, use and reuse of sensitive data. This may encompass specific ethical considerations pertaining to data collection (e.g., balancing FAIR v CARE principles), research data collection methods when working with vulnerable individuals or communities often on sensitive topics, the joining of disparate datasets, and considerations of how long such data should be retained, and where.

3. Objectives *(A specific set of focus areas for discussion, including use cases that pointed to the need for the IG in the first place. Articulate how this group is different from other current activities inside or outside of RDA.):*

1. Using the definition presented at the top of this document as a starting point, develop a shared understanding and refined definition of sensitive data.
2. Define various levels of “sensitivity” for data.
3. Data should be as open as possible and as closed as necessary. Within this context, develop an understanding of how sensitivity relates to openness.
4. Identify different consent models.
5. Identify types of sensitive data holdings and resources across various domains.
6. Identify existing data definitions and standards for different types of sensitive data.
7. Identify challenges in collecting, using and sharing sensitive data.
8. Engage with key stakeholders working in the area of sensitive data management/analytics .
9. Identify existing solutions for sensitive data collection, analysis, storage and dissemination.
10. Identify differences in how sensitive data is managed between groups and regions.

4. Participation *(Address which communities will be involved, what skills or knowledge they should have, and how will you engage these communities. Also address how this group proposes to coordinate its activity with relevant related groups.):*
Use these people to help grow the case studies

While the interested participants in this interest group are currently mostly from Australia, we have been working to establish this group as part of a global Community of Practice. We are currently developing a strategy to achieve international engagement.

To further this effort, the group has seen the recent addition of chairs from Europe and the USA to the group. The Social Science Interest Group, which comprises a broad international membership base and chairs from Norway, USA and Australia, also has formal participation in the Sensitive Data interest group.

The next phase of this engagement strategy will be through specific engagement with RDA groups and other stakeholders covering a range of domains and geographic regions. Specific stakeholders to be approached are still to be determined, but will be drawn from the these target groups These include:

- RDA Interest Groups: Social Science IG (established), International Indigenous Data Sovereignty IG (Initial approach made, pending response), Ethics and Social Aspects of Data IG, RDA-COVID19 WG (and the various sub-groups), Reproducible Health Data Services WG, Epidemiology common standard for surveillance data reporting WG,

Domain Repositories IG, Health Data Interest Group, RDA/NISO Privacy Implications of Research Data Sets IG, Virtual Research Environment IG, Social Dynamics of Data Interoperability IG

- Communities outside of RDA: Relevant domain and discipline communities, eg. The SSHOC and EOSC work programs around sensitive data, US and Canadian networks of Research Data Centres, International and Regional Statistical Agencies (WHO, UNStat, Eurostat, National Statistical Offices), (HEALTH DATA EXAMPLE COMMUNITIES???)

5. Outcomes (Discuss what the IG intends to accomplish. Include examples of WG topics or supporting IG-level outputs that might lead to WGs later on.):

1. To identify the key expectations of the community and use these to refine the IG's objectives.
2. List different types of data across disciplines such as health, social sciences, etc and how different levels of sensitivities apply to those types of data.
3. Identify best practices in sensitive data management across multiple regions, domains and disciplines and how to adapt the best practices.
4. Engage with relevant RDA IGs, WGs and CoPs to identify priorities in the area of sensitive data management.
5. Gather common guidelines and recommendations for working with sensitive data in different disciplines and in different regions.
6. Catalogue of ethical, philosophical and cultural principles that underpin the use of sensitive data assets.

6. Mechanism (Describe how often your group will meet and how will you maintain momentum between Plenaries.):

The IG will meet every 3 - 4 weeks via Zoom. Meeting times will be alternated to accommodate as many time zones as possible. Google Docs will be used to develop shared documentation. Email will be used to communicate about meetings and tasks requiring follow-up between meetings. The current chairs/members of the IG are already successfully using this system to meet and maintain momentum.

The IG will also meet regularly at Plenaries as an opportunity to workshop new ideas with the RDA community and foster new engagements. The group will also establish an informal communication channel through Slack, or a similar platform, to allow for ongoing conversation. The group will also organise webinars and information sessions between Plenaries to share ideas and for group members to stay in touch with the activities of the group. The IG will also use our RDA page to share documents and communicate regularly with the RDA community.

7. Timeline (Describe draft milestones and goals for the first 12 months.):

Initial activities: The group met for the first time as a Birds of a Feather session at RDA 16. Following this, a core group of interested members met to begin drafting the group charter. This group also submitted a proposal for an IG session for RDA 17. The group will send the draft charter for initial TAB review and community consultation in the lead up to RDA 17. The draft charter will also be sent for feedback specifically to members who have joined the IG page and who attended the BoF session at RDA 16. The draft charter and TAB/community/group feedback will be discussed at the RDA 17 session. Following this, the revised charter will be submitted for formal endorsement.

First 12 months: Once the IG is formally endorsed, we will undertake the following activities in the first 12 months:

1. Formally launch the IG - update our RDA IG site, call for additional co-chairs, share the approved charter with group members, establish a regular meeting time, establish RDA mailing list for the IG.
2. Engage in group consultation to identify the main themes of interest and develop a strategy for establishing working groups/task forces to address these.
3. Engage with stakeholders for feedback on key sensitive data issues and to develop the IGs networks within and outside of RDA.
4. Invite existing RDA IGs identified in section 4 above to provide feedback on, and participate in, working groups/task forces themes.
5. Presentation of webinar/workshop to workshop working group/task force topics and open the working groups/task forces topics for group comment through interactive platforms like Google Docs.
6. Formalise the working groups/task forces, share the goals of the working groups/task forces with the group and RDA more broadly to increase participation, prepare for RDA18 as an opportunity to share progress of the IG and working groups/task forces.
7. Prepare reports and outputs from the working groups/task forces, share reports with the community, present a webinar/workshop to share the outputs with the community.
8. Hold an IG meeting to assess the progress from the preceding 12 months and determine the next steps for working groups/task forces.

8. Potential Group Members (Include proposed chairs/initial leadership and all members who have expressed interest.)

- People interested in leadership:

FIRST NAME	LAST NAME	EMAIL	TITLE/AFFILIATION
Kristal	Spreadborough	kristal.spreadborough@unimelb.edu.au	University of Melbourne, Research Data Specialist

Aleks	Michalewicz	aleksm@unimelb.edu.au	University of Melbourne, Research Data Specialist
Priyanka	Pillai	priyanka.pillai@unimelb.edu.au	University of Melbourne, Research Data Specialist
Nichola	Burton	nichola.burton@ardc.edu.au	ARDC, Data Technologist
Keith	Russell	keith.russell@ardc.edu.au	ARDC, Manager (Engagements)
Stefanie	Kethers	stefanie.kethers@ardc.edu.au	ARDC, RDA Director of Operations
Steven	Mceachern	steven.mceachern@anu.edu.au	steven.mceachern@anu.edu.au Australian Data Archive, Director
Romain DAVID	European Research Infrastructure on Highly Pathogenic Agents	Romain.david@erinha.eu	Data manager, Research fellow European Research Infrastructure on Highly Pathogenic Agents
Dharma	Akmon	dharmrae@umich.edu	Director of Project Management and User Support Assistant Research Scientist Inter-university Consortium for Political and Social Research University of Michigan

- People who have joined on the Sensitive Data RDA IG so far

Name	Country
Frankie Stevens	Australia
Vince Bayrd	United States
Bénédicte Madon	France
Tiiu Tarkpea	Estonia

Lars Eklund	Sweden
Kristan Kang	Australia
Amy Nurnberger	United States
Su Nee Goh	Singapore
Robert Pocklington	Australia
Kristan Kang	Australia
Genevieve Rosewall	Australia
Graham Smith	United Kingdom

- People who attended the BoF expressed interest in participating following the BoF:

Name	Affiliation and role	Email	Interested in participating further?
Marjolaine Rivest-Beauregard	McGill University, MSc student	marjolaine.rivest-beauregard@mail.mcgill.ca	Yes
Kiera McNeice	Cambridge University Press, Research Data Manager	kmcneice@cambridge.org	Yes
Matthew Viljoen	EGI Foundation, Service Delivery and information security lead	matthew.viljoen@egi.eu	Yes
Stephanie Thompson	Research Data Management, University of Birmingham	s.e.m.thompson@bham.ac.uk	Yes

Y. G. Rancourt	Portage Network, Curation Officer	yvette.rancourt@carl-abrc.ca	Yes
Thea Lindquist	University of Colorado Boulder, Center for Research Data and Digital Scholarship, Executive Director	thea.lindquist@colorado.edu	
Briana Ezray	Penn State University, Research Data Librarian - STEM	bde125@psu.edu	Yes
Gen Rosewall	Agile Business Analyst, AARNet	gen.rosewall@aarnet.edu.au	Yes
Becca Wilson	University of Liverpool, UK ; Research Fellow	becca.wilson@liverpool.ac.uk	Yes
Karen Thompson	University of Melbourne	karen.thompson@unimelb.edu.au	Yes
Jeaneth Machicao	Universidade de São Paulo / Research fellow	machicao@usp.br	
Jules Sekedoua KOUADIO	Gustave Eiffel University	jules.kouadio@univ-eiffel.fr	yes
Mahamat Abdelkerim Issa	Institut national de recherche scientifique (INRS), Québec, CA, Phd. Student	Mahamat_Abelkerim.Issa@ete.inrs.ca	Yes

Erin Clary	Portage, Canadian Association of Research Libraries	erin.clary@carl-abrc.ca	Yes
Kylie Burgess	Research Data Lead, University of New England	kburge22@une.edu.au	Yes