RDA Sensitive Data IG: Towards a common understanding of sensitive data across regions and disciplines

More information and contact details are available at https://www.rd-alliance.org/groups/sensitive-data-interest-group

The RDA Sensitive Data Interest Group aims to promote the FAIR principles and reproducible research, while drawing attention to the unique challenges and opportunities associated with sensitive data and exploring strategies to mitigate risks.

Sensitive data level: how, where and why?

This poster is presented in tandem with the Sensitive Data IG session at RDA Plenary 18 (which shares the same name). In previous meetings, the Sensitive Data IG has scoped the interest and needs of the RDA community. At this meeting, we focus on specific challenges and opportunities associated with working with Sensitive Data. This poster explores one of these:

- how do we understand sensitive data definitions in different regions and disciplines, and
- how might we develop a shared language around sensitive data. For example, developing an understanding of different community agreed vocabularies and how these relate to each other.

The first step will be to **identify together common terminology on sensitive data usable across disciplinaries**.

This poster complements the RDA Sensitive Data Interest Group session.

It is developed by the Sensitive Data IG co-chairs and is an opportunity for us to convey our ideas and seek input from the community in different fora.

You can join us!

The future work for sensitive data terminology building consist of three actions.

First, case studies of different types of sensitive data must be identified. These will include, for example, case studies from the humanities, medicine, biodiversity, and military cases.

Second, different classification systems for the level of sensitive data must be shown and compared. This should explore, for example, what constitutes a high level of sensitive data in one classification system compared to another.

Third, examples of how sensitive data is managed across different regions must be detailed. For example, what protocols govern the use of sensitive data in a national context compared to the European context?

By presenting these examples, we aim to:

1. promote wider discussion on how sensitive data is currently defined.
2. draw out how context can impact whether data is considered sensitive over time

In doing so, we take steps towards developing a **common language** for discussing sensitive data classifications and levels across contexts/disciplines/localities.

Methodology/poll for case studies to identify types of sensitive data

- Six open ended questions during the session calling for sensitive data examples concerning military topics / health data / biodiversity - environment risk / personal data / industrial data / other
- Sharing specific case studies, including horror / success stories.
- Open call for helping to build a shared terminology around sensitive data, drawing on your examples, experiences, or case studies.

Methodology/poll for identification and comparison of classification systems for the level of sensitive data

- In your discipline (to be defined by participants), what are the relevant terms to describe sensitivity?
- In your region / place / country (to be defined by participants), what are the official classifications for sensitive data?

Methodology/poll to describe how sensitive data are managed across regions

- In your region / place / country (to be defined by participants), what are the relevant governance policies or protocols for sensitive data?
- Within the region, does governance differ by discipline or industry?
- By what process are definitions of sensitive data re-assessed or updated?

General objectives of the Sensitive Data Interest Group

The IG objectives will be informed by the community. We take the following objectives as a starting point:

- Develop a shared understanding and refined definition of sensitive data.
- Define various levels of sensitivity for data and develop tools to assess this sensitivity.
- Develop an understanding of how sensitivity relates to openness.
- Identify different consent models.
- Identify types of sensitive data holdings and resources across various domains.
- Identify existing data definitions and standards for different types of sensitive data.
- Identify challenges in collecting, using, archiving and sharing sensitive data.
- Engage with key stakeholders working in the area of sensitive data management/analytics.
- Identify existing solutions for sensitive data collection, analysis, storage and dissemination.
- Identify differences in how sensitive data is managed between groups and regions.

Take home message

The Interest Group is focusing on developing an understanding of the terminology around Sensitive Data.

We seek a **diversity of perspectives across disciplines and regions**.

We warmly welcome new and existing members to join this conversation at RDA18 and beyond.

POSTER LEAD: Romain David (European Research Infrastructure on Highly Pathogenic Agents [ERINHA], 0000-0003-4073-7456)
CO-AUTHORS (in alphabetical order): Dharma Akmon (University of Michigan, 0000-0002-1359-0586); Nichola Burton (Australian Research Data Commons, 0000-0003-4470-4846); Sarah C. Davidson (Max Planck Institute of Animal Behavior, The Ohio State University, 0000-0002-2766-9201); Steven McEachern (Australian Data Archive, 0000-0001-7848-4912); Kristan Kang (Australian Research Data Commons,0000-0002-2057-1033); Aleksandra Michalewicz (University of Melbourne, 0000-0002-7328-2470); Priyanka Pillai (University of Melbourne, 0000-0002-3768-8895); Kristal Spreadborough (University of Melbourne, 0000-0002-7022-3213); Frankie Stevens (Australia's Academic and Research Network, 0000-0002-2556-473X).
DOI: 10.5281/zenodo.5663084
Presented at Research Data Alliance 18th Plenary - remotely, November 2021.

CC BY