“Weather, climate and air quality” IG session

9th RDA Plenary Meeting, Barcelona, 5-7 April 2017

research data sharing without barriers
rd-alliance.org

Pierre-Antoine Bretonnière, Francesco Benincasa
Barcelona Supercomputing Center

Agenda

- 9.00 - 9.10: **Introduction**
- 9.10 - 9.40: Presentation of participants
- 9.40 - 10.00: Review of IG plans
- 10.00 - 10.30: Definition of activities schedule for next year

Introduction of the IG objectives

- Recent and exponential increase of data volume and diversity leads to new problems in the earth sciences community.

- Need for new data solutions to avoid:
  - Having data “stuck” locally and difficult to share among institutions.
  - Too complex data indexing and discovery of big data repositories.
  - Undocumented datasets or lacking metadata.

Introduction of the IG objectives

- Some of these problems are common to other communities (data transfers, metadata, pids,...) but some others (ontologies, data organization, dynamic datasets,...) are very **specific to the weather, climate and air quality community**.

- **Strong pressure** from a **large user community** (urban air quality, climate-change adaptation, industry sensitive to environmental pressures).

Introduction of the IG objectives

- Use RDA infrastructure to gather people from our communities but also data specialists and try to find answers to these data challenges.

- RDA should be seen as a unique opportunity for sharing data knowledge with other communities and push for political decisions in terms of data standards.

- Foster collaboration with other RDA IG/WG.

Introduction of the IG objectives

- In this Research Data Alliance Interest Group, we want to explore and discuss the challenges for the use and **efficient analysis** of **large** and diverse **datasets** from the **climate, weather and air quality** communities.

- Based on a collaboration between several **research** meteorological and European climate institutes and taking into account input from the **private** (renewable energy, satellite and agriculture sectors for example) **and public sectors**, this IG will suggest **practical and applicable solutions** for Big Data issues encountered by these communities, both at **technological and policy** level.

General data sharing and management challenges

- Volume
  - Physical storage
  - Discovery, indexing, ...
- Variety
  - Data sources
  - Data formats
  - Data access protocols
- Velocity
  - Processing tools (Diagnostics, analytics and visualization)
- Veracity
  - Metadata

Specific data sharing and management challenges

- Community specific data formats
- Community specific or project specific conventions (CMOR, CF, …)
- Data analysis close to the storage close to the HPC (bring the compute to the data)
- Data sharing tools, infrastructures, portals (ESGF, Copernicus, EUDAT, …)
- Something else?

Use cases

- Proposed solution to volume problem: a case of earth sciences workflow (Francesco Benincasa)
- MPI meteorological and trace gas distribution campaign (Martin Kunz)
- Proposed solution for metadata and provenance (Pierre-Antoine Bretonnière)

Discussion and next steps

- Are you aware of **other RDA IG or WG** we should/could collaborate with and how?
- Do you see any overlaps with other IG or WG?

https://rd-alliance.org/groups

- Apart from the ones mentioned previously, do you think we should add **other objectives** to this IG?
- How do you see the **collaboration with the industries**/private sector?
- What added value could the **SMEs** bring to this IG?

Discussion and next steps

- Should we consider officially involving **other entities or consortia** (IS-ENES, EUDAT, ESGF)?

- How do the **earth observations and satellites** communities deal with their data (ESA, EUMETSAT)?

- Ideas of **solutions**?

Proposed milestones

- MS1: definition of the ecosystem and state of the art
- MS2: definition of the precise problems of the community and list of actors to solve them

Actions

- For those interested, subscribe to RDA IG page: https://www.rd-alliance.org/groups/weather-climate-and-air-quality
- Definition of the roadmap of the IG
- Co-chairs?