***RDA Interest Group Draft Charter***

***Name of Proposed Interest Group (IG):*** ***Weather, climate and air quality***

***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):*

In a field where the amount and the data sources increase exponentially, it becomes critical to find efficient solutions at a community level to address problems such as storage, indexing, metadata, sharing, or analysis. The global objectives of this Interest Group will be to explore and discuss the challenges for management, efficient analysis and dissemination of large and diverse datasets generated by the weather, climate and air quality communities.

These communities follow one of the main ideas of data sharing of RDA as most (but unfortunately not all) of the datasets used are completely open and freely available. Most of the concepts and problems that this IG aims to address correspond to classic RDA issues (metadata, data standards, efficient sharing, etc...). However, as these communities already have rules and policies structured in terms of standards and data portals, they need some advanced community specific solution to these issues.

These communities, which are not (yet) adequately represented in RDA, are one of the biggest data producers in terms of volume (for example, the CMIP6 project is planned to generate few hundreds PB of data and this is only one example of a project for one community), have a range of users well identified (including both the public and private sectors as weather and climate sensitivity is very democratic) and, therefore, can bring an important added value to the work of the RDA community.

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

With the increase of the data volumes and sources, scientists and users working with weather, climate and air quality data find themselves fighting with issues like having data “stuck” locally and, hence, difficult to share among institutions, data repositories being too big to be indexed/explored, or the software suggested to access and manipulate the data demanding unrealistically large amounts of memory to efficiently compute diagnostics, metrics and products (an example are post-processing operations of large climate and air quality simulations sometimes having a wallclock time longer than the simulation itself). This IG should discuss and propose solutions to help tackling these issues, working with the organizations dealing with these issues and/or those that host or generate these data. Collaborations and close links with the ESGF[[1]](#footnote-1) (“an international effort led by the American Department of Energy (DOE), and co-funded by NASA[[2]](#footnote-2), NOAA[[3]](#footnote-3), NSF[[4]](#footnote-4) and other international laboratories to facilitate advancements in Earth System Sciences and data dissemination”), international air quality and climate data providers such as the Copernicus programme[[5]](#footnote-5) or the ENES infrastructure[[6]](#footnote-6) (network in which several of the IG members’ participate) will ensure the work of this IG will be well integrated in the global framework of the Earth Sciences data community. As an example for the air quality component, the SDS WAS[[7]](#footnote-7) (managed by the BSC in collaboration with AEMET[[8]](#footnote-8)) being a WMO[[9]](#footnote-9) international framework, it will ensure a strong link with the world meteorological organization and, consequently, strengthens the links between this IG and other national meteorological agencies.

***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.):*

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

The main goals of the group will be:

* Encourage weather, climate and air quality communities to enforce the data curation by converging to common ontologies and metadata, defining or proposing standards in collaboration with the already working RDA Metadata IG
* Work on open data / open science initiative fostering the adoption of technologies (i. e. PIDs generation) which facilitate experiments reproducibility
* Involve communities like ENES (Earth System modelling), ESIWACE[[10]](#footnote-10) (Centre of Excellence in Simulation of Weather and Climate) and ENVRI[[11]](#footnote-11) (environmental research)
* Participate in Data Sharing Infrastructure like EUDAT[[12]](#footnote-12) and the already mentioned ESGF

One of the objectives of this IG is also to use the political influence of RDA to push for solutions at global (community) level.

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

The target audience of this IG is of course representatives of the different entities of the weather, climate and air quality communities. An example of the targeted audience can be found in the list of participants to the Birds of a Feather session where the creation of this IG was discussed at the last RDA plenary in Tokyo (April 1-3 2016). This session was attended by people from the community with different profiles: from software and data engineers to pure scientists and decision makers.

Recommendations from and collaboration with other RDA IG/WG such as the Array Database Assessment (several persons are members of both the “Weather, climate and air quality” IG and the Array Database Assessment WG) to get inputs on how to deal with new technologies or with Geospatial, Global Water Information IGs to get experience of domain specific data management.

***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.):*

The outcomes can be summarized as:

* identification of the main issues and bottlenecks the community is confronted to in terms of data sharing and management
* proposal of enhancements to the data centers and researches that can improve the shareability and usability of the data
* proposal for efficient data analysis tools enhancements

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

A mailing list has been setup so that people from the IG can interact. To subscribe, interested parties can send an email with the subject “Subscribe” to earthsciences-rda-ig@bsc.es. Apart from these discussions by email, regular (every 3 months) teleconferences to follow-up the advancements of the work and/or discussions will be set up by the IG coordinators.

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

After the first six months of “official life” of the IG, a list of relevant actors from the community that have an involvement in the IG can be provided (MS1: definition of the ecosystem and state of the art targeted in the IG). After the first 12 months, a list of issues and specific problems, using illustrative examples from the community, will be provided, leading to an eventual WG born from this IG (MS2: definition of the precise problems of the community and list of actors to solve them)

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

|  |  |  |  |
| --- | --- | --- | --- |
| FIRST NAME | LAST NAME | EMAIL | TITLE |
| Enrique | Arias | [Enrique.arias@uclm.es](mailto:Enrique.arias@uclm.es) | Pr |
| Francesco | Benincasa | francesco.benincasa@bsc.es | Mr |
| Gary | Berg-Cross | [gbergcross@gmail.com](mailto:gbergcross@gmail.com) | Dr |
| Timea | Biro | t.biro@trust-itservices.com | Ms |
| Idir | Bouarar | [Idir.bouarar@mpimet.mpg.de](mailto:Idir.bouarar@mpimet.mpg.de) | Mr |
| Pierre-Antoine | Bretonnière | pierre-antoine.bretonniere@bsc.es | Mr |
| Maria | Brovelli | maria.brovelli@polimi.it | Prof |
| Reinhard | Budich | [Reinhard.budich@mpimet.mpg.de](mailto:Reinhard.budich@mpimet.mpg.de) | Mr |
| Emmanouil | Chaniotakis | [chaniotakis@certh.gr](mailto:chaniotakis@certh.gr) | Mr |
| Antonio | Cofiño | [Antonio.cofino@unican.es](mailto:Antonio.cofino@unican.es) | Pr |
| Simon | Cox | [Simon.cox@csiro.au](mailto:Simon.cox@csiro.au) | Dr |
| Sabine | Darras | Sabine.darras@obs-mip.fr | Ms |
| Francisco | Doblas-Reyes | [Francisco.doblas-reyes@bsc.es](mailto:Francisco.doblas-reyes@bsc.es) | Dr |
| Matteo | De Felice | [Matteo.defelice@enea.it](mailto:Matteo.defelice@enea.it) | Dr |
| Daniel | Field | [Daniel.field@atos.net](mailto:Daniel.field@atos.net) | Mr |
| Mike | Fortun | profmikefortun@gmail.com | Prof |
| Fabrizio | Gagliardi | [Fabrizio.gagliardi@bsc.es](mailto:Fabrizio.gagliardi@bsc.es) | Dr |
| Siddeswara | Guru | [siddeswara.guru@csiro.au](http://www.the-email-archive.com/email-data-siddeswara-guru-csiro-au-26325564.htm) | Dr |
| Maggie | Hellström | M.Hellstroem@gsi.de | Dr |
| Stephanie | Kethers | Stephanie.kethers@ands.org.au | Dr |
| Kwo-Sen | Kuo | [kuo@bayesics.com](mailto:kuo@bayesics.com) | Dr |
| Fernando | Martin | [Fernando.martin@ciemat.es](mailto:Fernando.martin@ciemat.es) | Pr |
| Yasuhiro | Murayama | [murayama@nict.go.jp](mailto:murayama@nict.go.jp) | Prof |
| Ben | Postance | B.Postance@lboro.ac.uk | Mr |
| Matthew | Mayernik | mayernik@ucar.edu | Dr |
| Alessandro | Spinuso | [Alessandro.spinuso@knmi.nl](mailto:Alessandro.spinuso@knmi.nl) | Mr |
| Tobias | Weigel | weigel@dkrz.de | Dr |
| Jens | Weismueller | [Jens.weismueller@lrz.de](mailto:Jens.weismueller@lrz.de) | Mr |
| Malcolm | Wolski | m.wolski@griffith.edu.au | Mr |

1. <http://esgf.llnl.gov/> [↑](#footnote-ref-1)
2. [www.nasa.gov/](http://www.nasa.gov/) [↑](#footnote-ref-2)
3. [www.noaa.gov/](http://www.noaa.gov/) [↑](#footnote-ref-3)
4. [www.nsf.gov/](http://www.nsf.gov/) [↑](#footnote-ref-4)
5. <http://www.copernicus.eu/main/overview> [↑](#footnote-ref-5)
6. <https://verc.enes.org/> [↑](#footnote-ref-6)
7. <http://sds-was.aemet.es/> [↑](#footnote-ref-7)
8. <http://www.aemet.es/en/portada> [↑](#footnote-ref-8)
9. [www.wmo.int](http://www.wmo.int) [↑](#footnote-ref-9)
10. <https://www.esiwace.eu/> [↑](#footnote-ref-10)
11. <http://envri.eu/> [↑](#footnote-ref-11)
12. <https://www.eudat.eu/> [↑](#footnote-ref-12)