Open Science & Digital Innovation
Splinter session

EOP-S Working Group

EO Science Strategy Workshop, 19-20 June Bruges, BE

20/06/2023

- Open Science and Digital Innovation Splinter Session

- Engaging discussion, thanks to all participants.

- We had 7 questions to guide the discussions

- Plus, a 10-min opportunity at the end to explore any other comments/questions

Q1. Open Science principles should be applied by default to all ESA projects (exceptions apply). What is currently closed that can be made open?

- Code produced under ESA funded activity should be available to the wider community

- Reusable, open source CCI algorithms could foster
  - Transparency and provenance across the digital value chain, traceability is kept between data and products
  - Reusability and definition of new products

- Challenges to be addressed:
  - Education and guideline needs on how to write open reusable source code
  - Additional effort needs to be financially accounted for, i.e. funding for in-depth documentation, testing infrastructure, for long term maintenance
  - Collaboration between groups and knowledge exchange needs to be encouraged to fill education gaps
  - Access to computation facilities is an issue
  - Data should be offered in a form that can be readily consumed
  - Dedicated support for software refactoring from closed source to open source needed

Q2. If tomorrow ESA decides to provide open access to all project outcomes, what would be your thoughts?

- Idea supported by the audience
- Should be a standard procedure
- ESA should provide examples/guidelines on how to comply with the requirements
- Embargo period should be a requirement
- ESA should provide a platform with sharing codes, results, etc to guide through the procedure.

Q3. What are the main technological advances that have high potential to innovate in the next 4-5 years? How do you see them adopted in the field of Earth Observation?

- Focus on workflows and reproducibility across platforms, i.e., System of System approach
- Organise big data in a way that can be used by algorithms – AI and Analysis Ready Data
- Need for AI benchmark training datasets for Earth Observation projects
- AI generated data to help the understanding of physical processes with higher resolution and high frequency
- Digital Twins, computing acceleration and AI needed - to capture all the physical processes
- IoT -> Future connectivity and real time mission tasking
- Ensure Traceability and Trust across the digital value chain

Q4. Projects tend to end when funding ends. How can we ensure long term engagement of scientific research into operational services? What setup would allow our science community to be part of the support and evolution of operational services?

- Create different reward systems, to facilitate continuous development and operations (e.g., EUMETSAT)
- Project setup: not to separate the scientific development from the operational maintenance
- Design a methodology to engage both scientists and operational providers to ensure long-term sustainability and maintenance of scientific products.
- Engage with final users to implement the production of scientific products as operational services.
- Communicate clearly on project outcomes, science study/ pre-operational/ operations

Q5. How would you like scientific activities to be communicated by ESA?

- Improve Web Discoverability (of ESA data, products, instruments, etc.)
- Communicate not only about ESA-funded science, but also about non-ESA funded science that uses ESA data
- Continuity of outreach and communication programmes and activities is key
- Develop and provide open access to more tools for scientific visualisation
- Develop a gallery of reusable examples of how to best communicate research results

Q6. What are the organizations ESA should partner with to ensure maximising quality of outputs? Could you facilitate such partnerships?

Partnerships with the European Open Science Cloud Services,

- There is a need to link different data stores and facilities.
- There is a need to make EO data available to their communities

Joining EGI Federation (federation of computing and storage resource providers)

Q7. What are the metrics we should use to demonstrate impact? How can we track status of adoption and performance/impact in the long-term?

- Currently metrics describe use of the data through the number of publications, number of products
- Need for acknowledging that a series of activities have an impact, not just a single activity or a publication
- Emphasize metrics representing usability and diversity of data and software , i.e. through:
  - Numbers of collaborators from different groups
  - Use in different scientific fields
  - Citizen collaborators involved
- Leave an open text for projects to describe what they think their impact is

Other questions / comments?

- AI has been hyped, we do good and bad with it
- When using AI as a method, we need to account for the full provenance chain. This includes the understanding of EO data, processing algorithms and uncertainty in output.

- We need to reach a point to be able to check "what is clear " and "what is not?" For instance, there are waves in the atmosphere, which cannot be measured directly in observations, but they can be modelled. There are uncertainties in the output. So, it's important to have traceability and visibility in such an approach.

- Our communities are continuously stretched - we have many satellites and have not been able to sufficiently use all these data.
    - How to take better benefit of the observing capabilities that we already have?
    - How can we enhance the existing capabilities we have, e.g., Sentinel-1 or Sentinel-2
    - ESA Programme on larger scale to look at the exploitation of existing data