

# Opening Reproducible Research

Daniel Nüst\*, Markus Konkol\*, Edzer Pebesma\*, Christian Kray\*, Stephanie Klötgen°, Marc Schutzeichel°, Jörg Lorenz°, Holger Przibytzin°, Dirk Kussmann°

\* Institute for Geoinformatics (ifgi) and University ° Regional Library Münster (ULB) Contact: daniel.nuest@uni-muenster.de

## Reproducible research is not so hard. Why do so few researchers do it?

Researchers find it difficult to reproduce a computational analysis conducted some years ago. Beyond the mere text of a publication, data or code might not be accessible or functional. Used base software or libraries have been updated or discontinued. Creating a reproducible workflow in the first place is a burden. The reward for publishing reproducibly is low.

## How does Opening Reproducible Research improve the situation?

- Develop a standard for packaging all components of reproducible research in an executable research compendium
- Understand desired and potential interactions with digital research objects, then implement and evaluate them
- Integrate reproducible research packages into the existing review and publication cycle, and archival platforms
- Provide an online platform and offline tools to create and manage research compendia

## The project

The DFG-funded project *Opening Reproducible Research* aims to improve the exchange and archiving of geoscience research results that are published over the Internet, by facilitating access to them and by simplifying their reuse in the form of a research compendium.

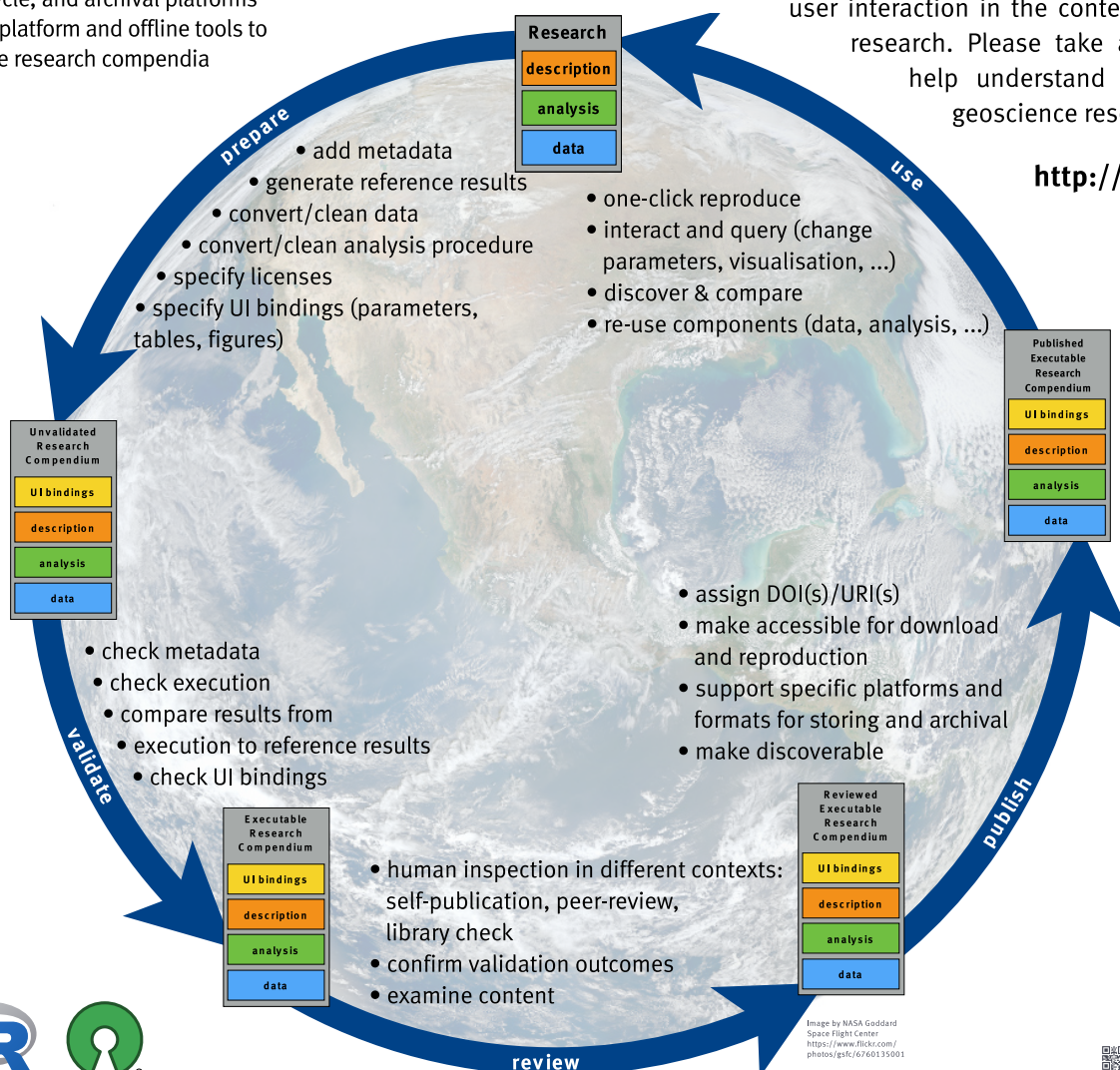
The project focuses on the publication process and the interaction with digital research objects as technical challenges to increase publication of open reproducible research.

Combining these goals will set a clear innovation in the area of reproducibility, which is at the heart of the scientific method.

## Join our survey

We run an online questionnaire investigating user interaction in the context of reproducible research. Please take a few minutes to help understand reproducibility in geoscience research.

<http://o2r.info/survey>



## Technologies



## Executable research compendium

The executable research compendium provides a formal structure for packaging all elements of a reproducible research workflow. It has three parts: (i) the researcher's workspace with all files (scripts, configuration, text, documentation) used for an analysis and the final publication, (ii) a snapshot of all software used (libraries, tools, base software, OS), and (iii) metadata to denote core contents. The software environment is available both in an executable form (Docker image) and as a complete machine readable recipe (Dockerfile). The metadata provide the basis for tools that execute research compendia and interact with them. It also clearly defines licenses.

## Integration with publication cycle and digital preservation

Scientific journals have well-established review and publication workflows, which are used by scientists. *Opening Reproducible Research* supports scientists by integrating executable research compendia into the review processes of existing journal management systems. The digital preservation community creates standards and tools for long term archival of digital assets. Executable research compendia follow these standards by extending the BagIt standard and integrating extensive standardized metadata documents.

## Interacting with research compendia

An executable research compendium has a number of interaction and query possibilities ranging from manipulating input parameters or variables to using components for new research. Dedicated user interfaces support the interaction with and publication process of research compendia. We thus focus on reducing the effort required to produce, discover, and reuse an executable research compendium by developing intuitive user interfaces.

Cursor image by Lordalpha1, drawer icon by Sebastian Gabriel, gear icon by McGalloway; from Wikimedia Commons