Opening reproducible research (o2r)

Markus Konkol\textsuperscript{1,*} (m.konkol@uni-muenster.de), Daniel Nüst\textsuperscript{1,*} (daniel.nuest@uni-muenster.de), Marc Schutzeichel\textsuperscript{2} (m.schutzeichel@uni-muenster.de), Edzer Pebesma\textsuperscript{1} (edzer.pebesma@uni-muenster.de), Christian Kray\textsuperscript{1} (c.kray@uni-muenster.de), Holger Przibytzin\textsuperscript{2} (holger.przibytzin@uni-muenster.de), Jörg Lorenz\textsuperscript{2} (joerg.lorenz@uni-muenster.de)

\textsuperscript{1} Institute for Geoinformatics, Münster
\textsuperscript{2} University and State Library, Münster
* Shared co-first authorship

The opening reproducible research project (o2r) addresses computational reproducible research in geosciences. By computational reproducibility we mean that researchers are able to rerun the computational steps described in a scientific publication in order to confirm the results reported in the paper. o2r supports scientists in submitting and executing reproducible research by developing the executable research compendium (ERC). An ERC not only allows to reproduce the results reported in a paper but also facilitates access to and reuse of research components. An ERC is created by the author of an article and composed of all components that underlie the research results, i.e. data, software, documentation, and user interface (UI) bindings. Data contains all inputs for the analysis. Software comprises the runtime environment and code created by a researcher and all underlying libraries or tools to reproduce the analysis in form of a Docker container. Documentation comprises both instructions, such as a README file, and the actual scientific publication, for example, in PDF format, any supplemental records, and metadata in standardized formats. UI bindings provide linkage between research components and user interface widgets. They can be used to attach suitable UI widgets to static diagrams in order to make them interactive, for example, a slider that allows to change the parameter shown in a diagram. They are also an entry point and implicit documentation for human readers.

o2r focuses on Open Science since creating an ERC requires that datasets and software are not restricted by their license. Research that is based upon closed-source software and unlicensed data is consequently out of scope.

In our project we first aim at specifying what must be included in an ERC so that the results are reproducible (software). Packaging libraries, software versions, and dependencies is an important challenge. We will then design and evaluate user interfaces that support authors and readers alike. Authors need guidance while creating an ERC, i.e. building a valid compendium, filling out a minimal set of metadata, and creating ui bindings in order to document computations and make static figures interactive. Readers will be provided with an interface that facilitates access to and reuse of research components. Users are able to drill deeper into the data and the analysis with the help of figures enhanced with interactive widgets. Executable research compendia support the trend towards Open Science. They make research more transparent, understandable, and reusable and ensure reproducibility for the future, because they are self-contained and archivable.

Although ERCs constitute a new form of publishing research results, they do not require the scientific landscape to change fundamentally as they can be integrated into the existing review and publication cycle.