| | |
|---|---|
| **Title** | From peer-reviewed to peer-reproduced: a role for research objects in scholarly publishing |
| **Author** | *Alejandra Gonzalez-Beltran* [1], Peter Li [2], Jun Zhao [3], Maria Susana Avila-Garcia [4], Marco Roos [5], Mark Thompson [5], Eelke van der Horst [5], Rajaram Kaliyaperumal [5], Ruibang Luo [6], Tin-Lap Lee [7], Tak-wah Lam [6], Scott C. Edmunds [2], Susanna-Assunta Sansone [1], Philippe Rocca-Serra [1] |
| **Affiliation** | [1] Oxford e-Research Centre, University of Oxford, UK [2] GigaScience, BGI HK Research Institute, Hong Kong [3] InfoLab21, Lancaster University, UK [4] Nuffield Department of Medicine, University of Oxford, UK [5] Department of Human Genetics, Leiden University Medical Center, The Netherlands [6] HKU-BGI Bioinformatics Algorithms and Core Technology Research Laboratory & Department of Computer Science, University of Hong Kong, Pokfulam, Hong Kong [7] School of Biomedical Sciences and CUHK-BGI Innovation Institute of Trans-omics, The Chinese University of Hong Kong, Shatin, Hong Kong |
| **Contact URL** | alejandra.gonzalezbeltran@oerc.ox.ac.uk<br>http://isa-tools.github.io/soapdenovo2/<br>https://github.com/ISA-tools/soapdenovo2 |

The reproducibility of science in the digital age is attracting a lot of attention and concerns from the scientific community, where studies have shown the inability to reproduce results due to a variety of reasons, ranging from unavailability of the data to lack of proper descriptions of the experimental steps.

Multiple research object models have been proposed to describe different aspects of the research process. Investigation/Study/Assay (ISA) is a widely used general-purpose metadata tracking framework with an associated suite of open-source software, which offers a rich description of the experiment's hypotheses and design, investigators involved, experimental factors, protocols applied. The information is organised in a three-level hierarchy where 'Investigation' provides the project context for a 'Study' (a research question), which itself contains one or more 'Assays' (taking analytical measurements and key data processing and analysis steps). Nanopublication (NP) is a research object model which enables specific scientific assertions, such as the conclusions of an experiment, to be annotated with supporting evidence, published and cited. Lastly, the Research Object (RO) is a model that enables the aggregation of the digital resources contributing to findings of computational research, including results, data and software, as citable compound digital objects.

For computational reproducibility, platforms such as Taverna and Galaxy are popular and efficient ways to represent the data analysis steps in the form of reusable workflows, where the data transformations can be specified and executed in an automatic way.

In this presentation, we will address the question of whether such research object models and workflow representation frameworks can be used to assist in the peer review process, by facilitating evaluation of the accuracy of the information provided by scientific articles with respect to their repeatability.

Our case study is based on an article on a genome assembler algorithm published in GigaScience, but due to the proven used of the respective research object models in their respective communities, we argue that the combination of models and workflow system will improve the scholarly publishing process, making science peer-reproduced.