FAIR Scientific Workflows: Status, Challenges and Research Opportunities

Sarah Cohen-Boulakia

Professor
Université Paris-Saclay

Biography

Sarah Cohen-Boulakia is a full Professor at the Laboratoire de Recherche en Informatique at Universite Paris-Sud/Paris-Saclay. She holds a Ph.D. in Computer Science and a habilitation from the same University. She has been working for fifteen years in multi-disciplinary groups involving computer scientists and biologists of various domains. She spent two-years as a postdoctoral researcher at the University of Pennsylvania, USA and 18 months at the Institute of Computational Biology (IBC) of Montpellier, France. Dr. Cohen-Boulakia's research interests include provenance and design of scientific workflows, reproducibility of scientific experiments, integration, querying and ranking in the context of biological and biomedical databases. She currently co-animates a National working group on reproducibility of scientific experiments and she is involved in the European Research Infrastructure ELIXIR (https ://www.elixir-europe.org/).

Summary

Computational workflows describe the complex multi-step methods that are used for data collection, data preparation, analytics, predictive modelling, and simulation that lead to new data products. They can inherently contribute to the FAIR data principles: by processing data according to established metadata; by creating metadata themselves during the processing of data; and by tracking and recording data provenance. These properties aid data quality assessment and contribute to secondary data usage. Moreover, workflows are digital objects in their own right. In this talk we will discuss FAIR principles for computational scientific workflows. We will highlight remaining challenges to address the workflows specific nature in terms of their composition of executable software steps, their provenance, and their development.