Research Data Alliance: Understanding Big Data Analytics Applications in Earth Science

Morris Riedel (1), Rahul Ramachandran (2), and Peter Baumann (3)

(1) Juelich Supercomputing Centre, (2) NASA, (3) Jacobs University Bremen

The Research Data Alliance (RDA) enables data to be shared across barriers through focused working groups and interest groups, formed of experts from around the world – from academia, industry and government. Its Big Data Analytics (BDA) interest groups seeks to develop community based recommendations on feasible data analytics approaches to address scientific community needs of utilizing large quantities of data. BDA seeks to analyze different scientific domain applications (e.g. earth science use cases) and their potential use of various big data analytics techniques. These techniques reach from hardware deployment models up to various different algorithms (e.g. machine learning algorithms such as support vector machines for classification). A systematic classification of feasible combinations of analysis algorithms, analytical tools, data and resource characteristics and scientific queries will be covered in these recommendations. This contribution will outline initial parts of such a classification and recommendations in the specific context of the field of Earth Sciences. Given lessons learned and experiences are based on a survey of use cases and also providing insights in a few use cases in detail.