

Deep FAIR – Knowledge Representation for Research Data about Complex Objects (Invited Talk)

Michael Kohlhase

FAU Erlangen-Nürnberg
michael.kohlhase@fau.de

Abstract

The publication, management, archiving, and re-use of research data (RD) is one of the current high-profile tasks in modern academia. Funding agencies mandate comprehensive research data practices, and the National Research Data Infrastructure (NFDI) supports the development of RD infrastructures with almost a billion Euro in the next decade in Germany alone.

The bulk of research data in the natural and engineering sciences essentially consists of "arrays of numbers", where the meaning of each data point can be uniformly described by its coordinates and a unit (e.g. for satellite images). For such data sets, the meaning can be captured by relatively simple metadata. This intuition is captured in the famous FAIR principles (Findability, Accessibility, Interoperability, and Reusability) which constitute the gold standard for RD in the current discussion.

In this talk we will look at an class of data that is often overlooked in the RD discussion: datasets of complex objects, such as the set of sculptures in a museum, the tabulation of all known elliptic curves, or the mathematical models used in a simulation. Such objects are complex in the sense that they are characterized by a large set of properties and relations to other objects. Such objects are best recorded by a formal description of their properties.

We will formulate a notion of "Deep FAIR Principles" for such object collections and look at the aspects of knowledge representation for the three examples above.