Contribution ID: **48** Contribution code: **2-35**   Type: **Poster**

Using Schema-based Metadata for Image Labels accessed with FAIR Digital Objects

Scientific image data sets can be continuously enriched by labels describing new features which are relevant for some specific task. This process can be automated by means of Machine Learning (ML) techniques. Although such an approach shows clear advantages, especially when it is applied to large datasets, it also poses an important challenge:

Relabeling image data sets curated by different scientists, in order to collectively use them for ML, requires a common agreement on the labels which can be used. This can be achieved thanks to the use of a standardized way to describe the label information: a metadata schema including vocabularies. Furthermore, machine-actionable decisions on the label information for relabeling can be enabled by the representation of images and schema-based metadata as FAIR Digital Objects (DOs).

We introduce a metadata schema including vocabularies to describe ML image data represented as FAIR DOs that can be accessed for relabeling. The specifications of the metadata schema are presented. The relevance of a standardized metadata description including vocabularies for relabeling ML image data is emphasized. It is shown how the metadata is accessed with FAIR DOs and how vocabularies support automated relabeling. This contribution supplements the content of "FAIR DO Application Case for Composing Machine Learning Training Data"with a focus on the semantic aspects for relabeling.

Please assign your poster to one of the following keywords.

Processes/Policies

Please assign yourself (presenting author) to one of the stakeholders.

Scientist/ Data Producer

Please specify "other" (stakeholder)

In addition please add keywords.

FAIR, Machine Learning, semantics

Primary authors: BLUMENROEHR, Nicolas (Karlsruhe Institute of Technology, Steinbuch Centre for Computing); AVERSA, Rossella (Karlsruhe Institute of Technology)

Presenter: BLUMENROEHR, Nicolas (Karlsruhe Institute of Technology, Steinbuch Centre for Computing)

Session Classification: Postersession II

Track Classification: Postersession