



Contribution ID: 26

Type: Poster

Machine Learning based Similarity Models for Research Dataset Discovery Systems

This poster is focussing on improving the process of research dataset discovery by developing machine learning models that can estimate the similarity between two datasets. In addition to traditional keyword-based search methods, the relevance of a dataset can also be determined by its similarity to existing, relevant datasets. The proposed models incorporate metadata about the datasets as well as the scientific publications that cite and use them, known as the “context”. The evaluation of these models shows that considering the context of a dataset leads to more accurate estimates of dataset similarity.

Primary authors: GEBREYOHANNES, Aberham (Bonn University); Dr ZERR, Sergej (Hochschulrechenzentrum, Bonn University)

Presenter: Dr ZERR, Sergej (Hochschulrechenzentrum, Bonn University)