

Bibliographic data science workshop

Leo Lahti, D.Sc., University of Turku, Finland

Iiro Tiihonen, MA student, University of Helsinki, Finland

Mikko Tolonen, Prof. University of Helsinki, Finland

Metadata is an essential component in virtual research environments and infrastructures. Recent research has highlighted the research potential of metadata, both as a research object in itself and as complementary information supporting the analysis of broader data collections such as full texts or audiovisual material. The overall coverage of metadata collections can be considerable in terms of time, geography, and contents, and as such it can be viewed as Big Data or Linked Data in the humanities. Metadata is typically highly structured and standardized. This can facilitate data analysis and the sharing of best practices, research algorithms and workflows. The lack of dedicated data science tools and questions of representativity, completeness and quality have posed remarkable challenges for research use, however; new research is needed to harmonize, integrate, and analyze metadata collections in order to realize their overall research potential. We call this approach bibliographic data science.

This workshop will bring together researchers to critically evaluate the potential of metadata collections based on latest research and teaching in this area, considering the questions of appropriate modeling and research use. The workshop will consist of an introduction, interactive talks, and a concluding panel discussion. In particular, we invite presentations on the following aspects of bibliographic data science: (i) quantitative algorithms, ecosystems and tools; (ii) research cases; and (iii) education and open science. Suggestions in the form of short abstracts can be sent to Leo Lahti <leo.lahti@iki.fi>.