Nils C. Hanwahr

Big Data Environments: Database Infrastructures, Data-intensive Knowledge, and Epistemology in a new Ecological Epic

This dissertation project asks how expert knowledge about nature is mediated by Big Data scientific infrastructures. Claims of a new "4th Paradigm" of scientific research are evaluated and questioned by contextualizing them in institutional discourses on epistemology and the history of database technology.

The development of Big Data infrastructures since the introduction of relational databases in 1970 is outlined by tracing trajectories and entanglements in case studies from the Earth Sciences and Life Sciences.

The aim of the study is to derive recommendations for policy design on the interception of data-intensive research and evidence-based science policy. An initial chapter outlines the discourse of claims since the rise of the term 'Big Data' around the year 2000.

A following chapter question these claims by contrasting them with the history of database technologies, narrated along the story of Jim Gray, database pioneer and inventor of transaction processing.

Case studies on the Ocean Observatories Initiative and the National Ecological Observatory Network provide an attempt to historicize current infrastructure projects that will crucially frame our view of the environment.

A final normative chapter will draw conclusions from the analysis and critique of the alleged 4th Paradigm and consider ways to promote reproducibility, transparency, and accountability in data-intensive scientific infrastructures.