



## Studiengang in Bibliotheks- und Informationswissenschaft 2019-2021

Hindermann, Max (2021). Machine indexing of institutional repositories: indexing Edoc with Annif as proof of concept

### **Abstract**

The potential of some institutional repositories is hampered by the lack of subject indexing. I argue that this situation can be amended with the help of machine indexing. As proof of concept I show that Edoc — the University of Basel's repository — can be successfully indexed using the Annif-client Python library. In order to do so, I assess the performance of hundreds of Annif configurations in assigning subject terms to a sample data set against a gold standard constructed from cleaned and reconciled author keywords.