

STELLA at DIPF

Integration of additional data-providers

Daniel Schiffner
Tamara Heck

DIPF provides and maintains several information retrieval infrastructures

FIS Bildung Literaturdatenbank (German Education Index)

- Indexing by 30 partners
- Recently hit 1.000.000 entries

peDOCS

- Open Access repository with ~ 20.000 publications

[Deutscher Bildungsserver](#) (German Education Server)

- Access to documents
- Background information regarding the German education system
- Also gives international context

[Forschungsdatenzentrum Bildung FDZ](#) (Research Data Centre for Education)

- Part of German Network of Educational Research Data
- Participates in the NFDI

Those four infrastructures are accessible via

[Fachportal Pädagogik](#) (German Education Portal)

- Meta search as well includes documents from ERIC, Library of Congress and more

Also:

[IWWB](#)

- Portal for finding vocational training courses
- current project aims at implementing new features
- e.g. testing of chat bots

Many partners collaborate in several infrastructures

- 30 at FIS Bildung Literaturdatenbank
 - 80+ Partners at IWWB
- Harmonization of interests and capacities have to be considered

Research on and for infrastructures commit to

supporting open science practices of others via

- Open access repositories
- (Open) data centers

follow open science practices

- Transparency of infrastructure and processes (e.g. metadata)
 - E.g. implementing infrastructure as “citable data”
- Reproducible data
- Collaboration and participation, as well with research-extern groups like infrastructure users

DIPF aims at improving its infrastructures via user studies

- Non-reactive methods like logfiles, click rate
- Reactive methods like online surveys, workshops, usability and search tests

Possible questions related to retrieval relevance assessment

- How is the completeness of metadata influencing relevance decision?
 - Background: diverse metadata due to different sources (meta search Fachportal and IWWB) and different indexing providers (e.g. FIS Bildung)
- How do the searches and relevance assessments differ for the same resources (e.g. data) in different frontends?
 - E.g. data from FDZ is searchable via FDZ portal and via Fachportal

Due to the infrastructure and exchanges in the backend:

- Creating snapshots (partially distributed Systems)

Maintain reproducibility, i.e. as data sets are updated, studies need to consider changes

- Defining and interpreting click rate as infrastructures differ and DIPF has diverse user groups like researcher and school teachers, educational practitioners

(Research) Question

- How to provide the data in a reproducible way?
 - Researchers need to know which data set they use and how updates might influence outcomes
- How to feedback information created by the living lab participants?
 - DIPF wants to know outcomes that can improve their systems, i.e. researchers need to make their data and results transparent, and transfer them to the data providing infrastructures
- How to provide “new methods” to the running system with live users?
 - How to roll-out (and roll-back) a method to the live-system?
 - How to detect bad ranking behavior?
 - What “influence” do data providers have and how should they use it?

Thank You for Your attention