**BASE (Bielefeld Academic Search Engine)**

- a Next Generation Multi-Level Repository-based Service Provider

base.ub@uni-bielefeld.de - Bielefeld University Library

**Introduction**

BASE (short for Bielefeld Academic Search Engine, [http://www.base-search.net](http://www.base-search.net)) is one of the largest search engines for academic open access content. BASE is harvesting, aggregating, normalising and enriching OAI metadata from repositories all over the world, and is acting as an OAI service and data provider at the same time. Bielefeld University Library has been operating and developing BASE since 2004 with a small but efficient project staff. BASE has switched to free and open source software (Lucene Solr, VuFind) in 2011. Bielefeld UL’s expertise in search engine technology and OAI metadata harvesting gained through BASE prepared the ground for several projects, project participations and invitations to expert committees. Today BASE plays an outstanding role within the global repository landscape and network based on sharing and exchanging services, information and metadata.

**BASE Components and Activities**

In order to enhance the metadata quality of the indexed documents, we make use of machine learning algorithms that automatically identify publications according to the Dewey Decimal Classification (DDC). The algorithms were developed in a DFG-funded research project.

The DDC classes are attributed by Dewey Indexing In order to enhance the automatic classification of publications according to Dewey Decimal Classification (DDC). The algorithms were developed in a DFG-funded research project.

BASE uses Lucene/Solr for indexing. We actively contribute to these open source projects by providing feedback and submitting patches.

The BASE-OAI-PMH Validator (OVAL) is a tool for checking OAI-PMH interfaces in terms of their suitability for being processed by BASE. The source code is available on GitHub. [http://oval.base-search.net/](http://oval.base-search.net/)

The BASE-OAI-PMH Validator (OVAL) is a tool for checking OAI-PMH interfaces in terms of their suitability for being processed by BASE. The source code is available on GitHub.

The OAI-PMH interface allows harvesting the BASE repositories in two formats. [http://oval.base-search.net/](http://oval.base-search.net/)

OAI-PMH Blog

**BASE Fact Sheet**

- > 90 million documents indexed
- from > 4,195 repositories
- from > 100 countries
- metadata aggregation and enrichment
- multilingual search
- OAI interface, search APIs

The global repository network as indexed by BASE