



Cross-Disciplinary Discovery in a Global Repository Network The Case of BASE – Bielefeld Academic Search Engine

Jochen Schirrwagen Bielefeld University Library https://orcid.org/0000-0002-0458-1004

Open Science Fair, Porto, 2019 Workshop on Data Discovery Across Disciplines



Search 147,483,777 documents from 7,383 source





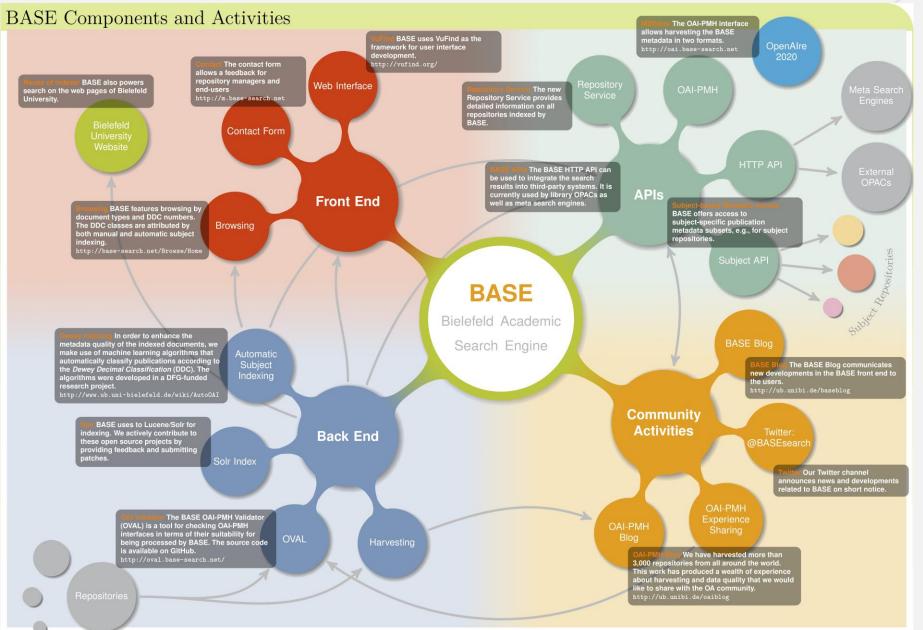
About BASE

- Officially started in 2004
- Aggregation from OAI-endpoints worldwide; all disciplines (repositories, journal platforms, digital collections, aggregators)
- ~70% Open Access from a total of ~147 million documents from 7,383 data sources (Sep-2019)
- Automatic classification using DDC
- OA status and License information processing
- OA Boosting
- ORCID member and BASE ORCID Search & Link Wizard



Universitätsbibliothek

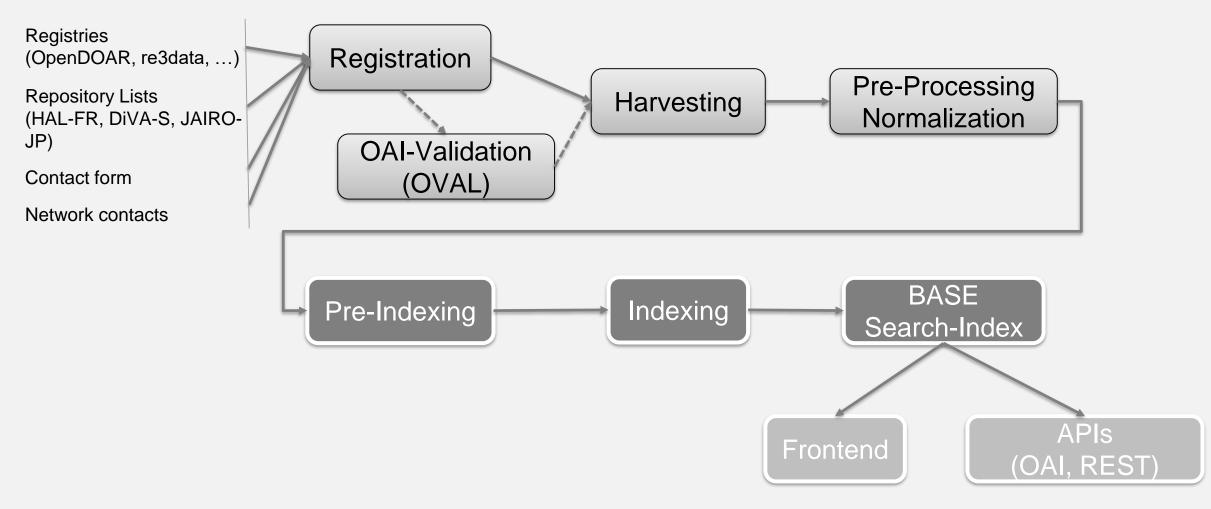








Aggregation and Data Workflow

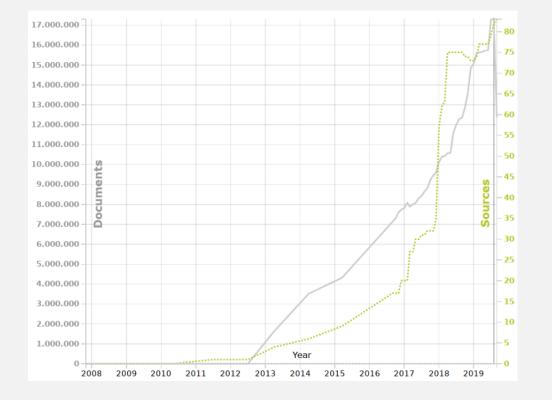




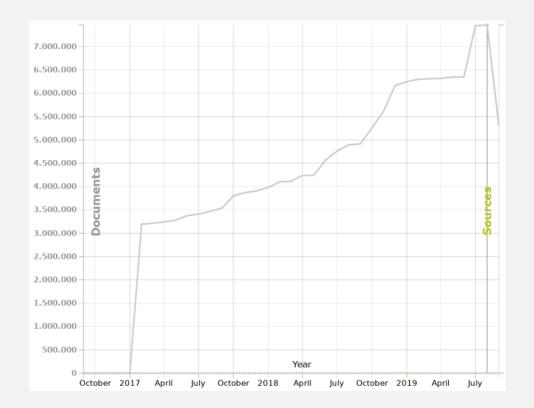


Integration of Data Repositories and Items in BASE

Number of Data Repositories incl. all items



Number of Research Data Items



Figures provided by Friedrich Summann, BASE





Discovery of Research Data – Search Interface

BASE			Login English 💟
Basic search Advanced search Browsing Search history			
Advanced Search	Document Type		
Entire Document 💟 poll data elections in Germany			
	Text		
Title	Book	Conference object	Patent
Author	Book part	Report	Thesis
Subject Headings	Journal/Newspaper	Review	Bachelor thesis
	Article contribution	Course material	Master thesis
	Other non-article	Lecture	Doctoral and
(Part of) URL		Manuscript	postdoctoral thesis
10 Hits pro page	Musical notation	Image/Video	Software
Access	Мар	Still image	Dataset
Open Access Income Access Unknown	Audio	Moving image/Video	Unknown
Linguistic tools	Terms of Re-use/Licences		
○ Verbatim search			
	Creative Commons		
Content sources	CC-BY	CC-BY-ND	CC-BY-NC-SA
Worldwide	CC-BY-SA	CC-BY-NC	CC-BY-NC-ND
	Public Domain		
Publication Year	CC0	Public Domain Mark (PD	PM)
Trans.	Software Licences		
From: To:	GPL	BSD	MIT





Discovery of Research Data – Search Results

BASE		Login English
Basic sear	ch Advanced search Browsing Search history	
Entire Document	poll data elections in Germany doctype:7	
	Verbatim search Additional word forms Multi-lingual search	
	Boost open access documents	
213 hits in 147,483,777 do	cuments	Sort Your Results
1. German Election P	anel Study, 1990 ; Version 1	Relevance
Author:	Kaase, M. [claim] ; Klingermann, HD. [claim] ; Kuechler, M. [claim] ; Pappi, F.U. [claim] ; Semetko, H.A. [claim] ; Forschungsgruppe Wahlen (Mannheim) [claim]	Refine Search Result
Description:	Data covering opinions about politicians, parties, and political issues were gathered for the first three waves of this panel study before the Bundestag election in 1990, and again for a fourth wave after the election. Topics include voting behavi	Author Construction Constructio
Publisher:	ICPSR - Interuniversity Consortium for Political and Social Research	Year of Publication
Year of Publication:	1994-03-10	Content Provider
Document Type:	Dataset ; survey data ; [Dataset]	Language
Content Provider:	da ra - Registration agency for social and economic data 🟮	Document Type
Detail View Email this Ad	d to Favorites Check in Google Scholar Export Record	Access
2. Volkskammerwahl Later 1991	- Ein Jahr danach 1991 ; Election for the East German Parliament - A Year	Terms of Re-use
Author:	Concret, Berlin [claim]	More Options
Description:	Changed living conditions and expectations after German-German unification, behavior at the polis 1990. Topics: Change of situation in life during the year; solution of urgent political tasks; reasons for economic condition and unemployment (scale	Search History Get RSS Feed
Publisher:	GESIS Data Archive	Set ATOM Feed
Year of Publication:	1991	Email this Search
Document Type:	Dataset ; [Dataset]	Save Search
Content Provider:	da ra - Registration agency for social and economic data 🟮	Browsing





Discovery of Research Data – Detailed View

BASE	Login English
Basic sear	ch Advanced search Browsing Search history
ome » Search » Detail Viev German Election Panel S	
Author:	Kaase, M. [claim] ; Klingermann, HD. [claim] ; Kuechler, M. [claim] ; Pappi, F.U. [claim] ; Semetko, H.A. [claim] ; Forschungsgruppe Wahlen (Mannheim) [claim]
Description:	Data covering opinions about politicians, parties, and political issues were gathered for the first three waves of this panel study before the Bundestag election in 1990, and again for a fourth wave after the election. Topics include voting behavior in the most recent general election, party preference, attitudes toward political candidates, issue relevance and competence of parties versus coalitions, projected results of this election, views regarding the respondent's personal economic situation, attitudes toward the 35-hour work week, opinions about the unification of Germany and the short- or long-term advantages or disadvantages of unification, interest in opinion polls and their possible influence on voting, and beliefs about television programming's neutrality in the election campaign. Other subjects covered were fear of losing one's job, interest in media coverage of political issues, preferred newspaper, political discussions among family members, education level attained, and religiosity. ; West German citizens over age 18 residing in private households. ; Stratified multistage random sample. ; This study was conducted in the former Federal Republic of Germany (West Germany) and does not include data from the former East Germany. The codebook is in German.
Publisher:	ICPSR - Interuniversity Consortium for Political and Social Research
Year of Publication:	1994-03-10
Document Type:	Dataset ; survey data ; [Dataset]
Language:	English
Subjects:	attitudes ; candidates ; economic conditions ; elections ; German reunification ; job security ; media coverage ; political attitudes ; political campaigns ; political interest ; political issues ; pelitical participation ; political parties ; public confidence ; public opinion ; voter attitudes ; voting behavior
Rights:	Not available ; This study is no longer distributed by ICPSR.
JRL:	https://doi.org/10.3000/ICF-SR00192.VI
Content Provider:	da/ra - Registration agency for social and economic data
	:• URL: https://www.da-ra.de/
	Continent: Europe
	Country: de Kumber of documents: 464,585
	Number of documents: 404,585 Open Access: 464,585 (100%)
	Type: Publications
	Source indexed in BASE since: 2019-01-22
	- BASE URL: https://www.base-search.net/Search/Results?q=dccoll:ftdara
Email this I Add to Eavorites	Check in Google Scholar Export Record





Opportunities and Limitations

- Global coverage of (OAI)sources
- Large amount of indexed items
- Cross-disciplinary search
- Common descriptive format (Dublin Core)

- Descriptive metadata quality varies across the data sources
- No publication dataset interlinking
- No links to discipline-specific metadata
- No (FAIR) indicators / metrics yet
- No optimization for research data search behaviour
- Many data repositories lack

standardized APIs (e.g. OAI-PMH)





Thank you for your attention!



base.ub@uni-bielefeld.de



@BASEsearch



https://base-search.net