

Business and Economics Research Data Center

BERD@BW

Baden-Württemberg

<https://www.berd-bw.de>

Business, Economic and Related Data

BERD@NFDI

National Research Data Infrastructure

<https://www.berd-nfdi.de>

BERD: The knowledge graph of German companies

Dr. Renat Shigapov (Data Scientist) & Dr. Irene Schumm (Head of Research Data Center)

Mannheim University Library, 23.02.2021

Motivation: German company data is fragmented over many providers, registers and time spans

Modern



Historical



Problems

- ☹ No unique identifier over providers
- ☹ No unique identifier over time

Consequences

- ☹ Limited Accessibility & Researchability
- ☹ Example: Named entity linking on texts and tables

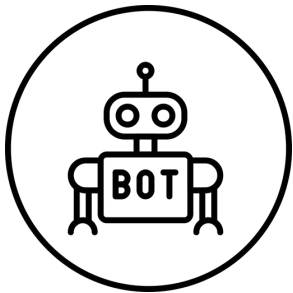
Solution

- ☺ BERD knowledge graph of German companies

Phase 1: BERD knowledge graph construction & testing

BERD ontology and BERD data

Python +
WikidataIntegrator
(a wrapper of the
Wikibase API)



SPARQL request



company	state	city	address	Search
LEI	ISIN or WKN	Creditreform-ID	Handelsregister	
min. employees	max. employees	legal form	SIC industry	

UniCredit Bank AG

Copy	CSV	Excel	PDF	Print	JSON
Property		Value			
also known as		Bayerische Hypo- und Vereinsbank Aktiengesellschaft			
BIC code		HVBWDEM1XXX, HYVEDEMMXXX			
Crefo number		315034553			
Handelsregister: court		Amtsgericht München			
Handelsregister: court, code & number		Amtsgericht München HRB 42148			

FRONT-END

- ✓ Python
- ✓ Flask
- ✓ JavaScript

BACK-END

- ✓ Wikibase

SPARQL response

Search BERD@BW

UniCredit Bank AG (Q4222)

Item Discussion

No description defined

Bayerische Hypo- und Vereinsbank Aktiengesellschaft

In more languages

Language	Label	Description	Also known as
English	UniCredit Bank AG	No description defined	Bayerische Hypo- und Vereinsba...

Statements

BIC code

HVBWDEM1XXX

1 reference

LEI

2ZCNR88UK83OBTEK2170

+ add reference

HYVEDEMMXXX

1 reference

+ add value

Crefo number

315034553

1 reference

+ add value

legal form

Aktiengesellschaft

1 reference

+ add value

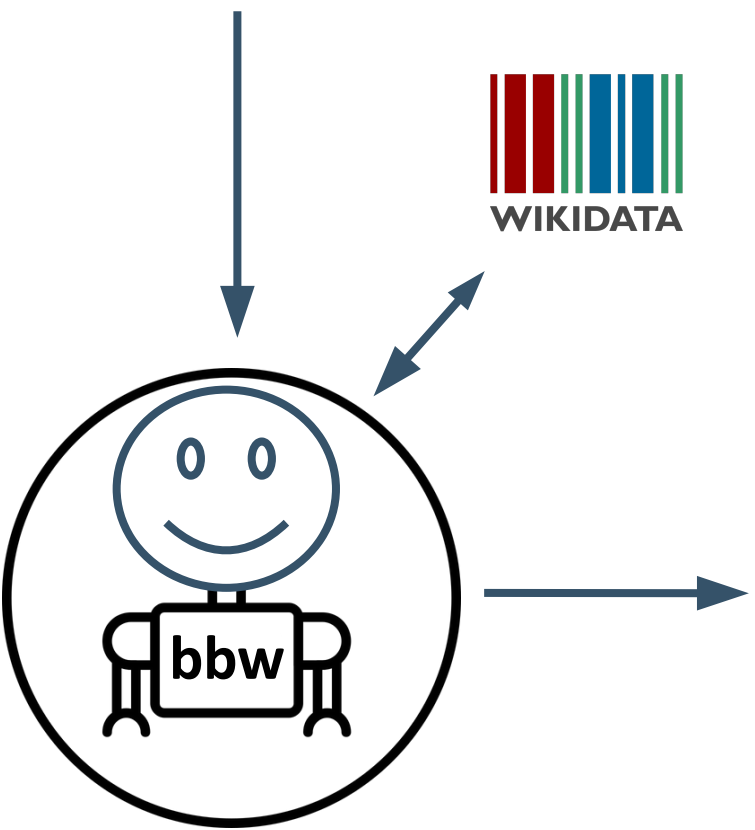


Phase 2: Tabular data annotation for ontology learning, data integration and matching services

1	Deutsche Telekom	mobile phone industry	Bonn	1994-01-01	https://www.telekom.com/
2	Lufthansa	air transport	Cologne	1953-01-01	https://www.lufthansa.com
3	SAP SE	software development	Walldorf	1972-04-01	https://www.sap.com



<https://github.com/UB-Mannheim/bbw>



1	Deutsche Telekom	mobile phone industry	Bonn	1994-01-01	https://www.telekom.com/
2	Lufthansa	air transport	Cologne	1953-01-01	https://www.lufthansa.com
3	SAP SE	software development	Walldorf	1972-04-01	https://www.sap.com
property		industry	headquarters location	inception	official website
type	business	industry	urban municipality of Germany		
datatype		WikibaseItem	WikibaseItem	Time	Url

Phase 3: Speeding up and automating knowledge graph construction with RaiseWikibase

A wrapper of the Wikibase API

- ☹️ Create and edit 1 million entities per 2-12 days.
- ☹️ Adding the templates, modules and texts takes time.

RaiseWikibase (a wrapper of MariaDB)

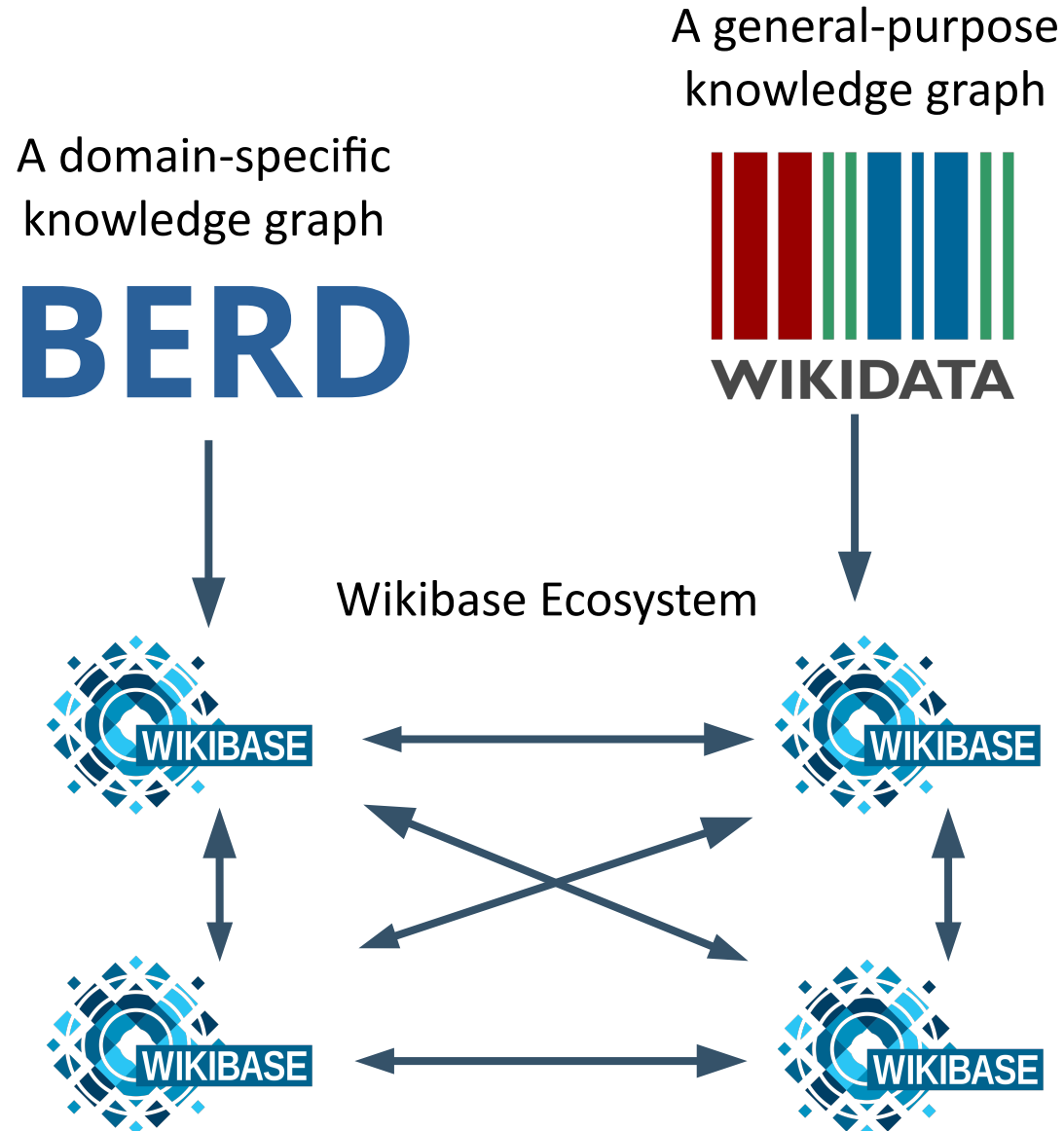
- 😊 Create and edit 1 million entities per hour.
- 😊 The templates, modules and texts can be easily and quickly uploaded.

50 times speed up!



Soon at ESWC 2021 and <https://github.com/UB-Mannheim/RaiseWikibase>

Summary



Takeaway 1: Wikibase is an excellent software for research data management!

Takeaway 2: Let's create more tools that will make Wikibase even better!

Takeaway 3: BERD knowledge graph for accessibility and researchability of **B**usiness, **E**conomic and **R**elated **D**ata.



<https://github.com/shigapov>
Renat.Shigapov@bib.uni-mannheim.de
Irene.Schumm@bib.uni-mannheim.de

