

Business and Economics Research Data Center

BERD@BW

Baden-Württemberg

Der BERD@BW-Wissensgraph für offene Firmendaten mit Wikibase

FAIR knowledge-graph-based research data management

Dr. Renat Shigapov, Dr. Irene Schumm, Markus Herklotz & Lars Oberländer

Mannheim University Library, 04.03.2021

Outline

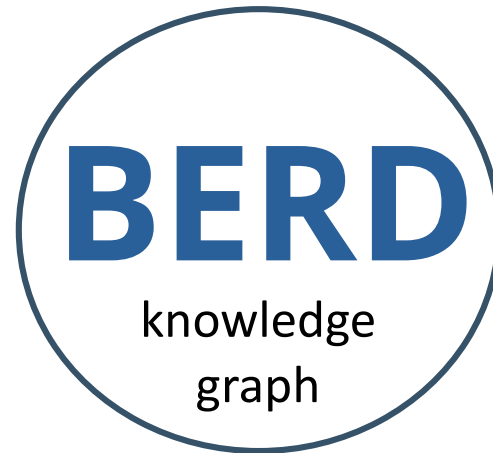
Motivation

Part 1: Creating

Part 2: Scaling

Part 3: Using

Conclusions



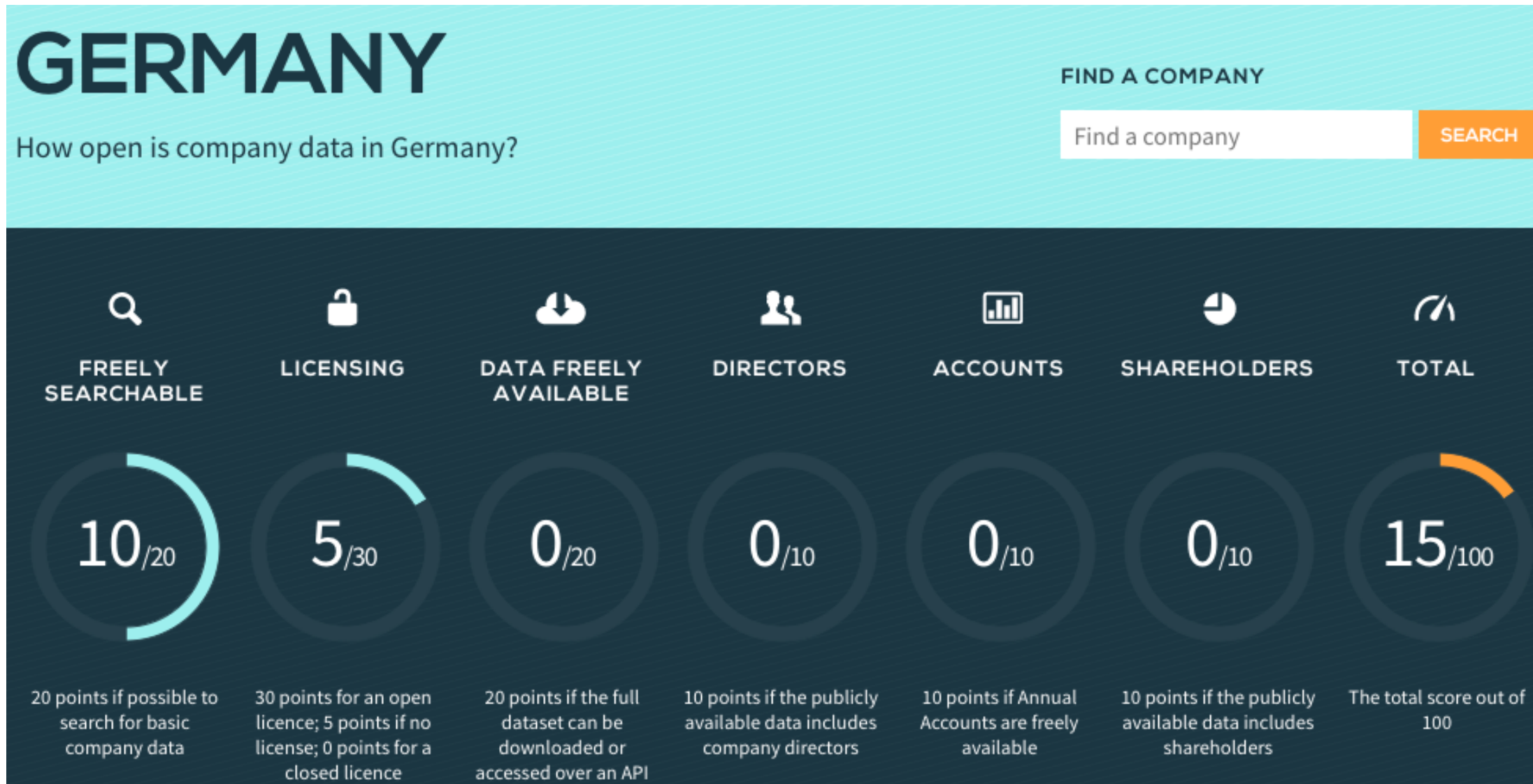
instance of



15

100

The open company data index for Germany is 15/100

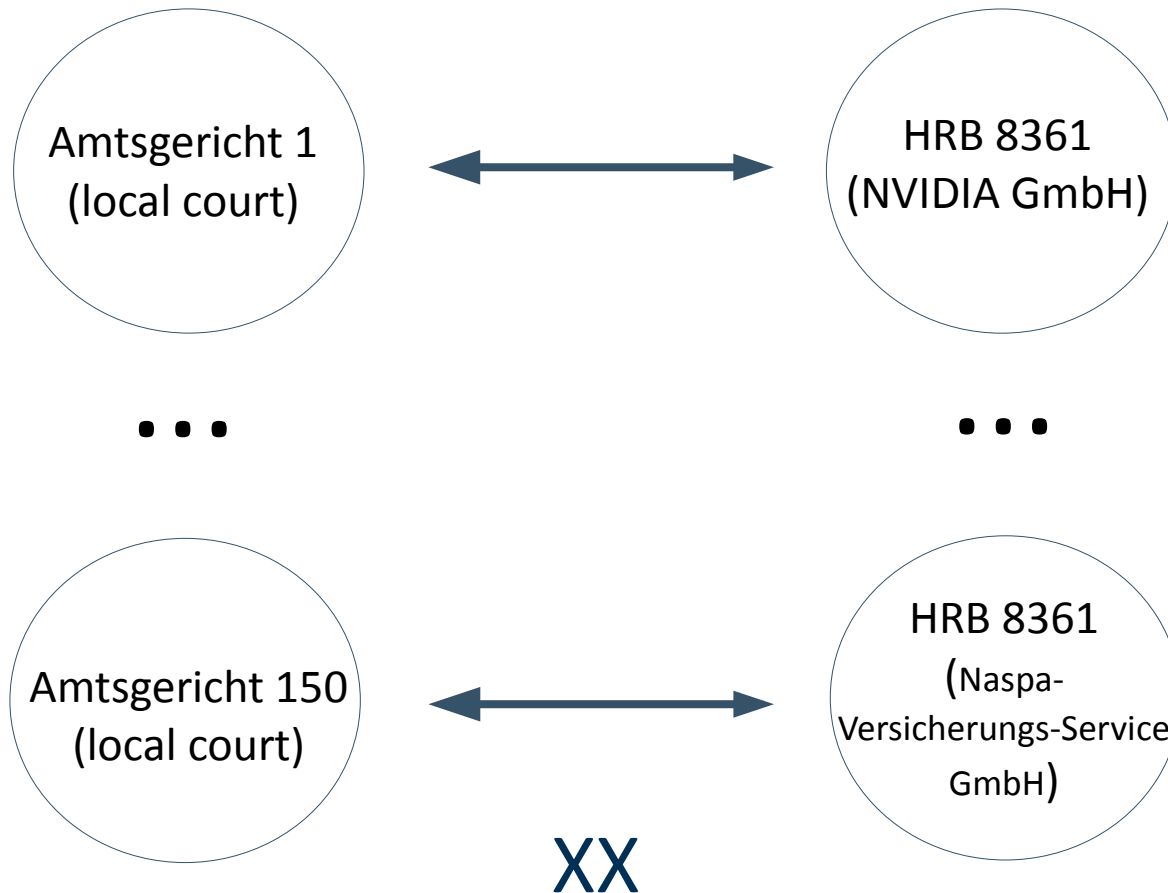


Source: <http://registries.opencorporates.com/jurisdiction/de>

The curse of German company data

**Many
registration authorities**

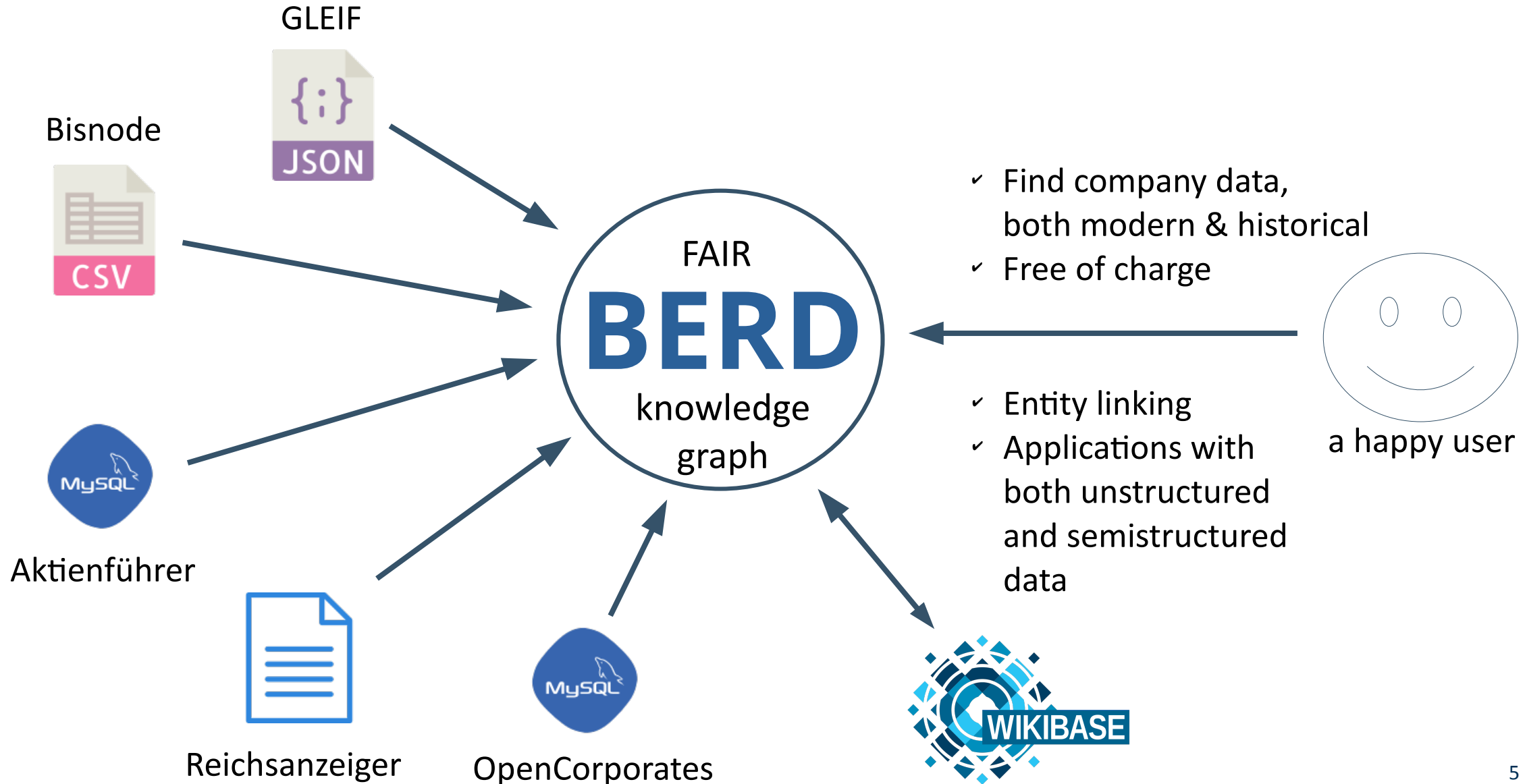
**Non-unique
company identifiers**



The historical & modern
German company data
ARE NOT:

- ☹ Findable
- ☹ Accessible
- ☹ Interoperable
- ☹ Reusable

Solution: BERD knowledge graph of German companies



Outline

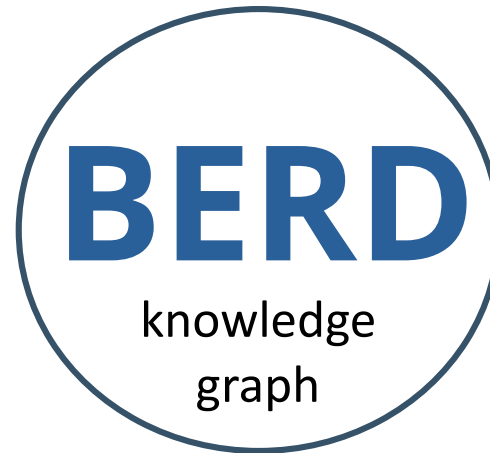
Motivation

Part 1: Creating

Part 2: Scaling

Part 3: Using

Conclusions



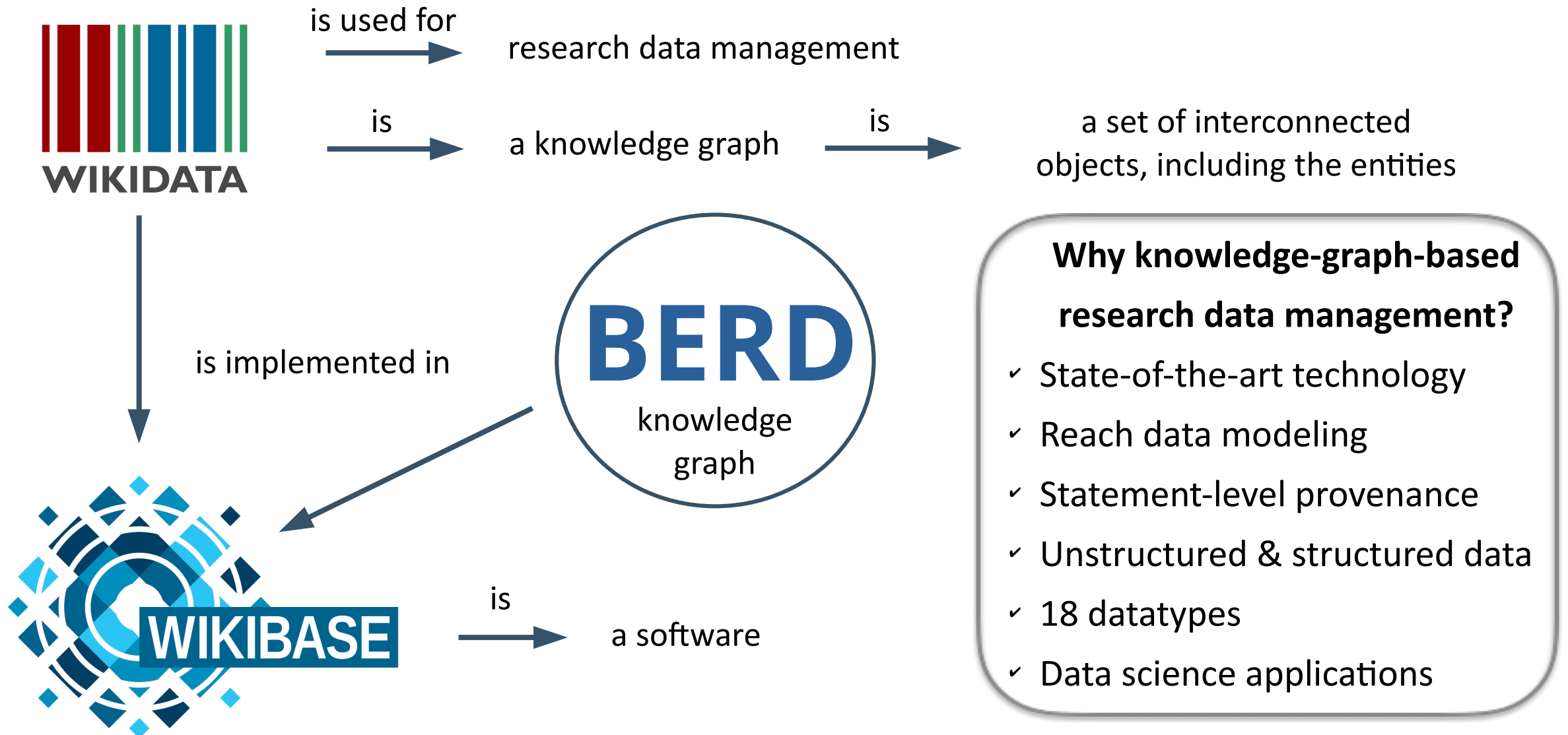
instance of



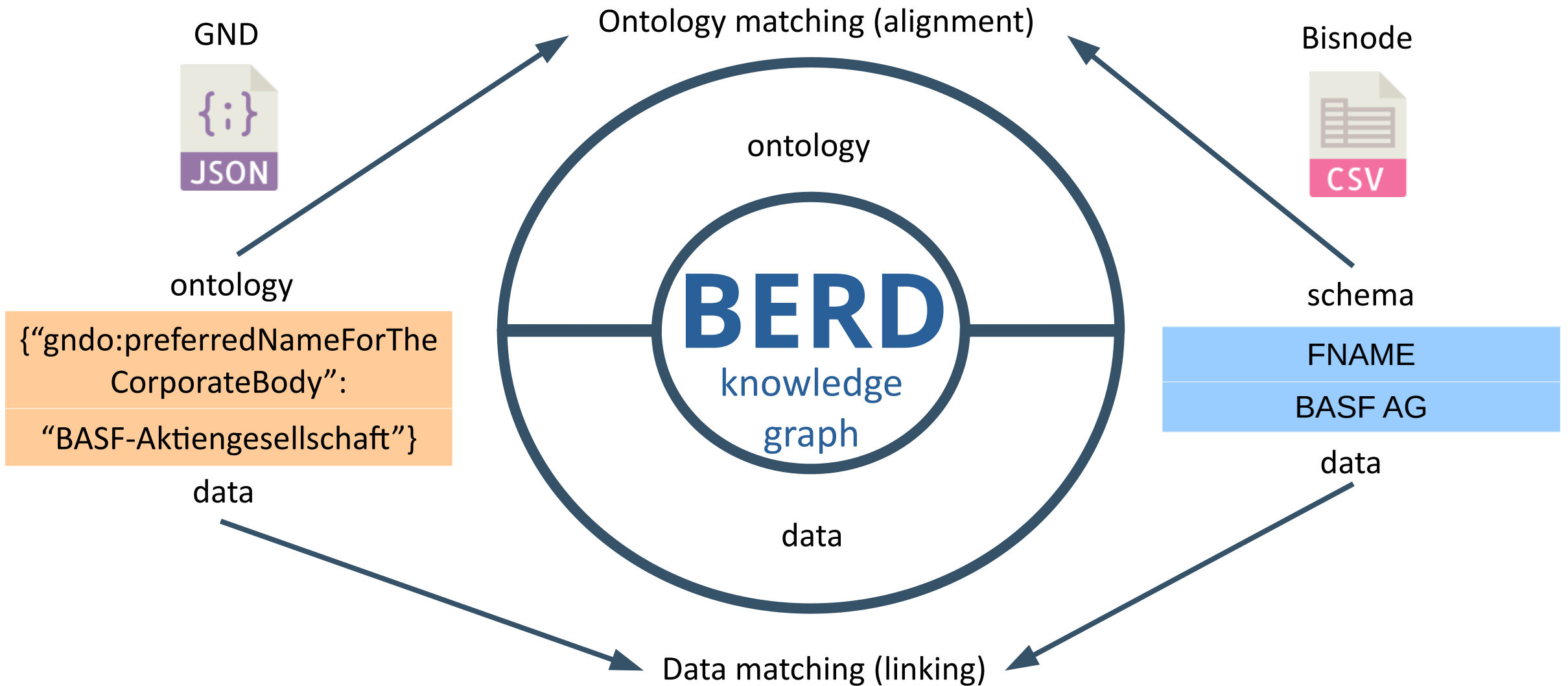
15

100

Wikibase, Wikidata & knowledge graphs

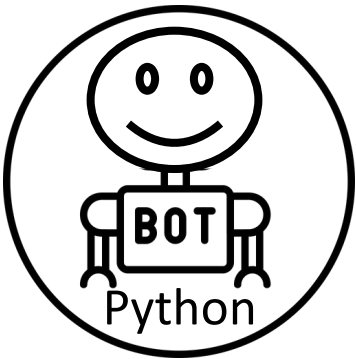


Knowledge graph = ontology + data

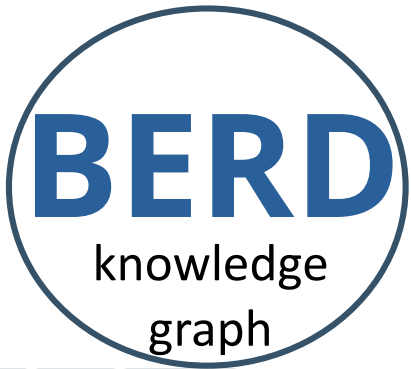


Filling BERD ontology and BERD data via Wikibase API

Company page
RDF export



Wikibase API



SPARQL endpoint
Data analysis and export

Search BERD@BW

Navigation

Main page

Recent changes

Random page

Help about MediaWiki

Wiki tools

Special pages

Concept URI

MGW Gesellschaft für Geld- und Wertpapiervermittlung mbH (Q20000)

Item

Discussion

No description defined

MGW Gesellschaft für Geld- und Wertpapiervermittlung mbH

In more languages

Configure

Language	Label	Description	Also known as
English	MGW Gesellschaft für Geld- und Wertpapiervermittlung mbH	No description defined	MGW Gesellschaft für Geld- und ...

Statements

Crafo number

326021821

1 reference

legal form

Gesellschaft mit beschränkter Haftung

1 reference

headquarters postal code

82041

1 reference

instance of

organization

1 reference

query.BERD-BW.de

Examples

Help

More tools

1 #Local courts sorted by the number of companies registered in the court

2 #defaultView:BubbleChart

3 SELECT ?Court (COUNT(?Court) AS ?NumberOfCompanies) WHERE {

4 ?company wdt:P3 wd:Q2;

5 wdt:P9 ?Court.

6 SERVICE wikibase:label { bd:serviceParam wikibase:language "en". }

7 }

8 GROUP BY ?Court

9 ORDER BY DESC (?NumberOfCompanies)

10 LIMIT 10

Bubble chart

Outline

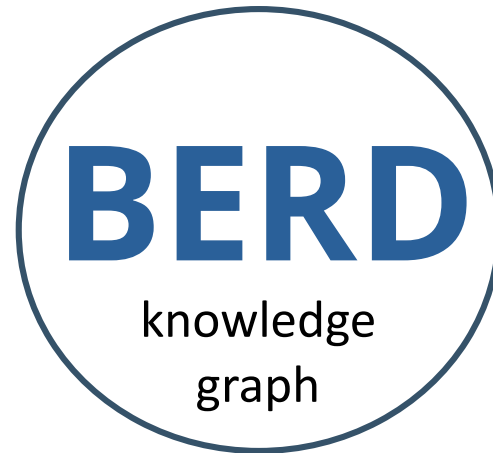
Motivation

Part 1: Creating

Part 2: Scaling

Part 3: Using

Conclusions



instance of



15

100

Towards a large-scale knowledge graph



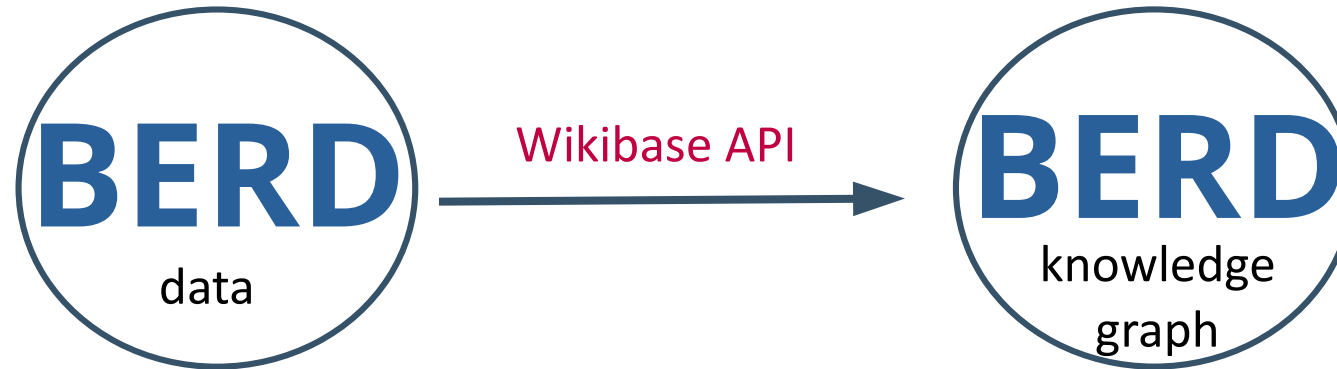
10000-100000 entities

1000000-10000000 entities

- More ontology alignment
- More data matching
- More data integration
- More data filling

A small problem with big data

A million edits of entities per 2-12 days



A million edits of entities per 1-2 hours

Towards the large-scale BERD knowledge graph

Search BERD@BW

Navigation

[Main page](#)
[Query service](#)
[Recent changes](#)
[Random page](#)
[Help about MediaWiki](#)
[New item](#)
[New property](#)
[BERD@BW](#)
[BERD@NFDI](#)

Wiki tools

[Special pages](#)
[Concept URI](#)

WSM Bau Leipzig GmbH & Co. KG (Q1000000)

- Item
- [Discussion](#)

No description defined

▾ In more languages

Configure

Language	Label	Description	Also known as
English	WSM Bau Leipzig GmbH & Co. KG	No description defined	

Statements

OpenCorporates ID

de/U1308_HRA17733

retrieved

4 November 2018

▸ 1 reference

street address

Hildegardstraße 40, 04315 Leipzig. (English)

retrieved

4 November 2018

▸ 1 reference

Native company number

Leipzig HRA 17733

retrieved

4 November 2018

▸ 1 reference

Outline

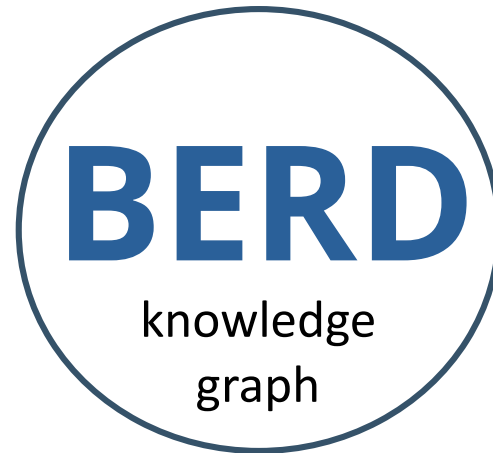
Motivation

Part 1: Creating

Part 2: Scaling

Part 3: Using

Conclusions



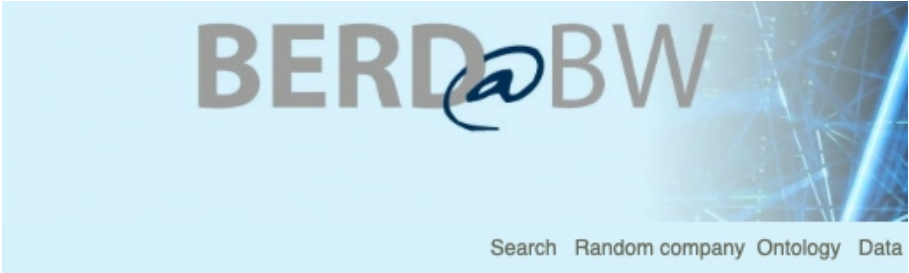
instance of



15

100

Wikibase as a backend for search services



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

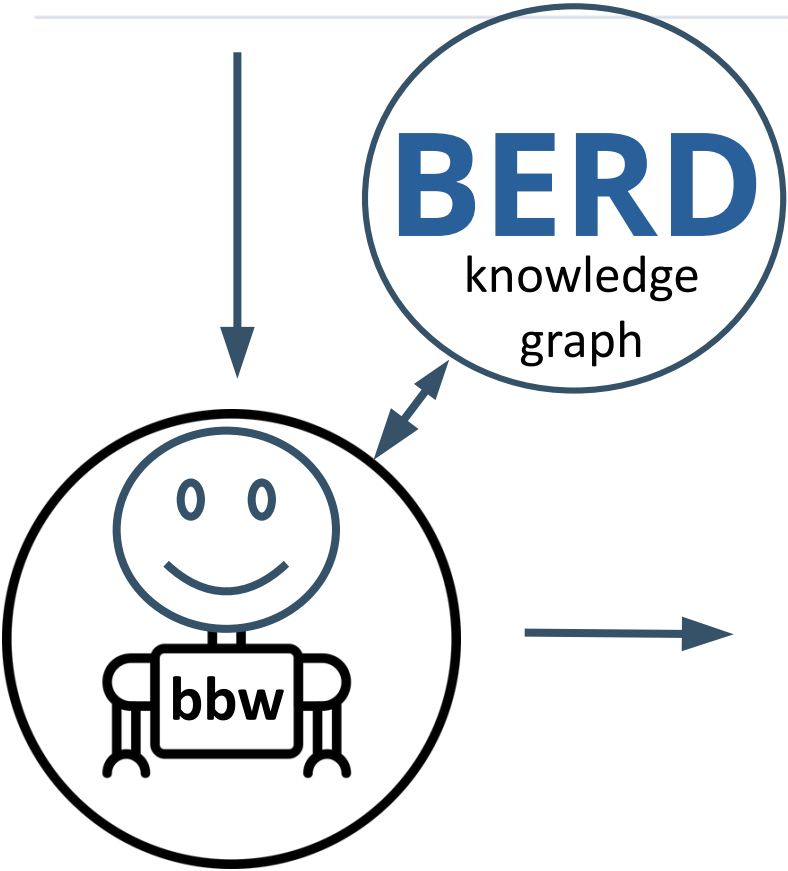
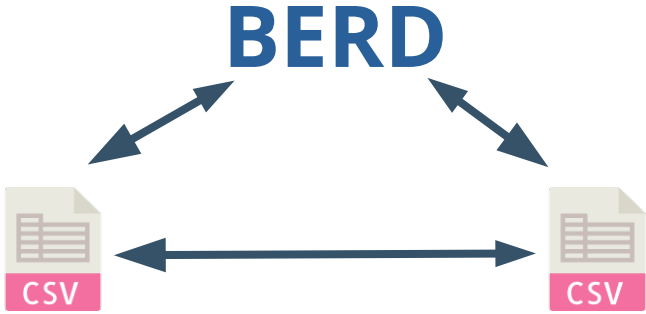
Graubner Industrie-Beratung

Copy CSV Excel PDF Print JSON

Property	Value
also known as	Graubner Industrie-Beratung GmbH
Crefo number	318147083
Handelsregister: court	Amtsgericht Stuttgart
Handelsregister: court, code & number	Amtsgericht Stuttgart HRB 331527
headquarters address	Schwimmbadstr. 26
headquarters city	Bad Herrenalb
headquarters fax number	(07083) 524914

Semantic annotation, entity linking & data integration

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



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

Outline

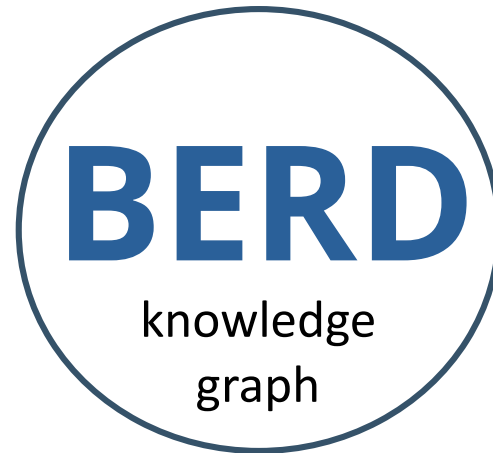
Motivation

Part 1: Creating

Part 2: Scaling

Part 3: Using

Conclusions



instance of



15

100

Towards the open
company data index = 100

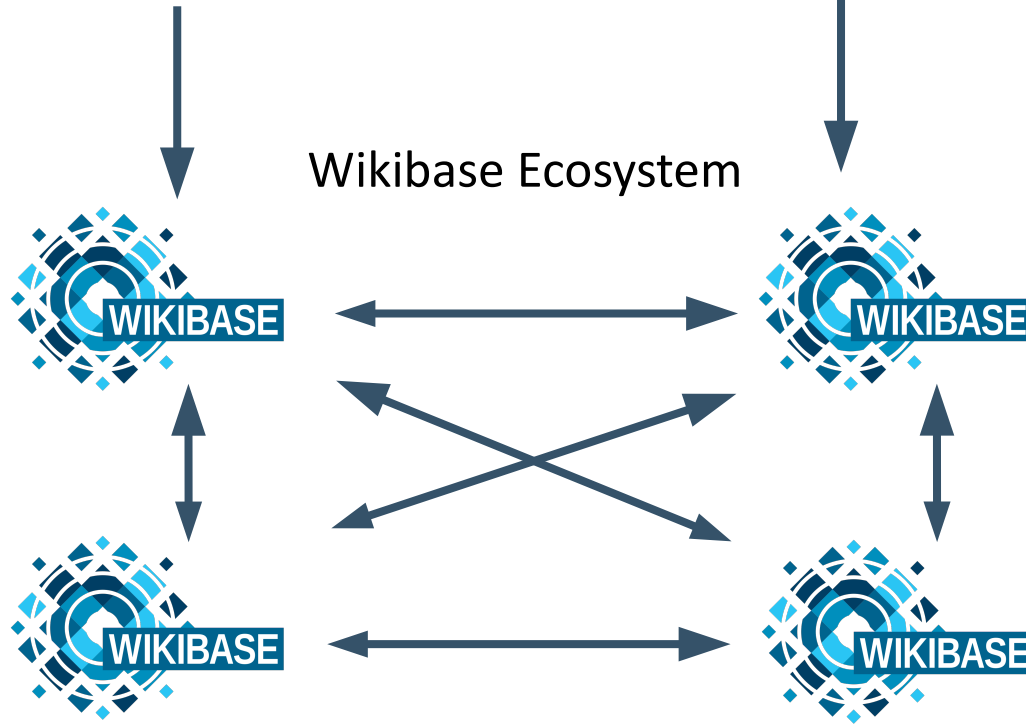


A general-purpose
knowledge graph



A domain-specific
knowledge graph
of German companies

BERD



Conclusions

- ☺ FAIR BERD knowledge graph
- ☺ State of the art research data management
- ☺ FAIR **B**usiness, **E**conomic and **R**elated **D**ata

FAIR BERD software:
<https://www.berd-bw.de/services-tools>