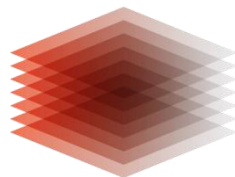

LEIBNIZ INFORMATION CENTRE
FOR SCIENCE AND TECHNOLOGY
UNIVERSITY LIBRARY

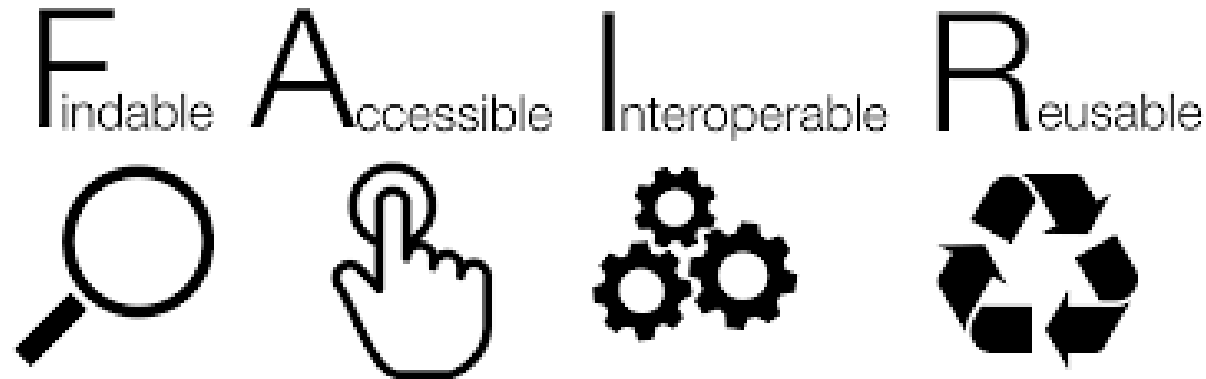


TIB

Semantic Research Data Management in the National Research Data Initiative (NFDI)

Jahrestagung der DPG, 28. September 2021

FAIR Data



Linked Data Principles

For implementing the FAIR principles

1. Use **URIs** to **identify** the “things” in your data
2. Use **http:// URIs** so people (and machines) can **look** them **up** on the web
3. When a URI is looked up, **return** a **description** of the thing **in** the W3C **Resource Description Format (RDF)**
4. Include **links to related things**

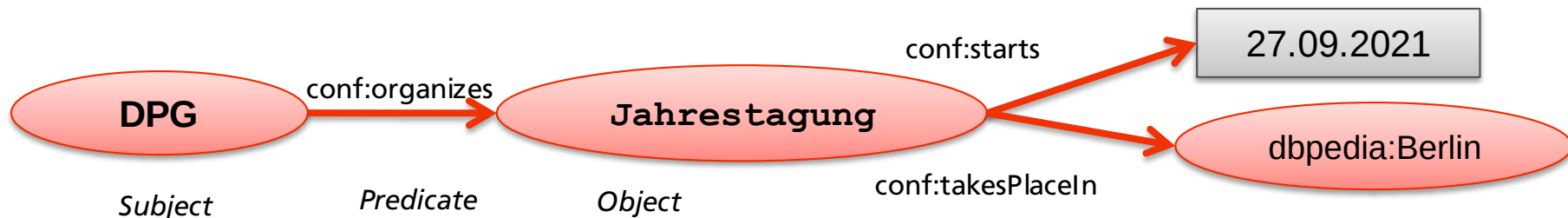
<http://www.w3.org/DesignIssues/LinkedData.html>

[1] Auer, Lehmann, Ngomo, Zaveri: **Introduction to Linked Data and Its Lifecycle on the Web.** [Reasoning Web 2013](#)



RDF & Linked Data in a Nutshell

1. Graph based RDF data model consisting of S-P-O statements (facts)



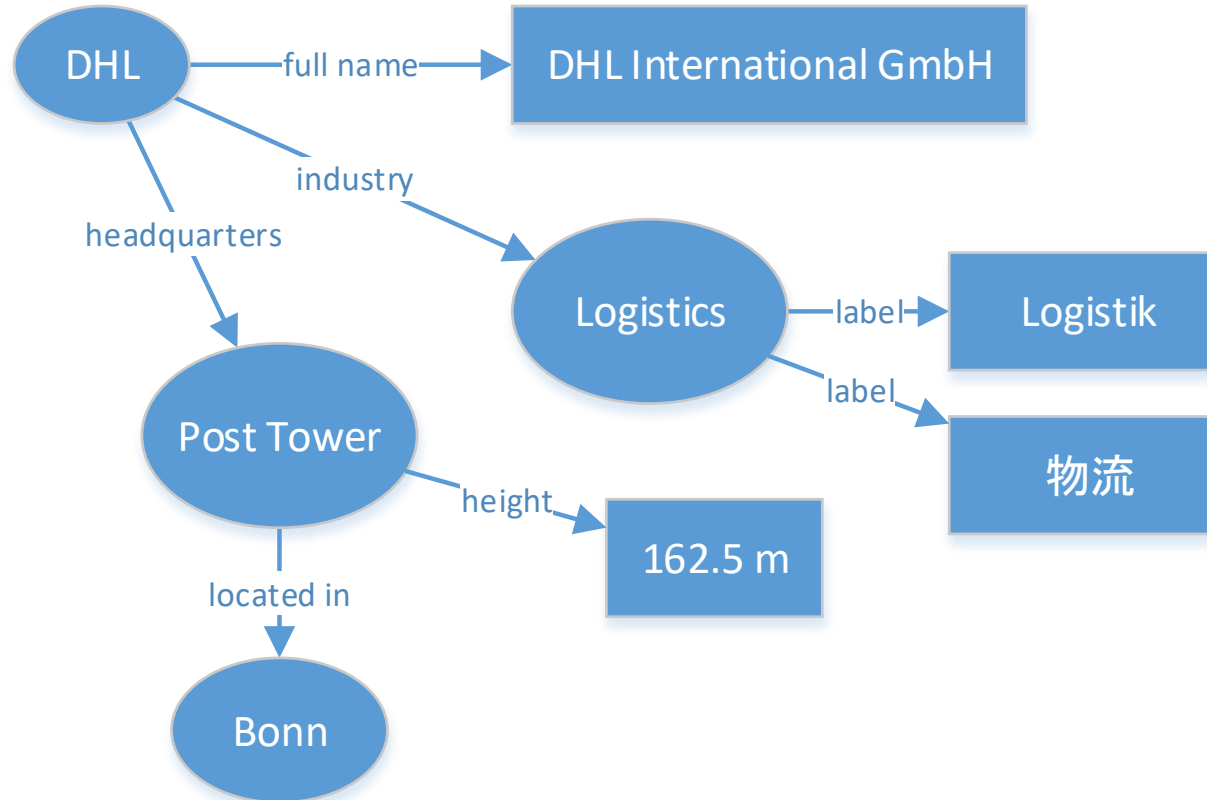
2. Serialised as RDF Triples:

DPG	conf:organizes	Jahrestagung .
Jahrestagung	conf:starts	"2021-09-27"^^xsd:date .
Jahrestagung	conf:takesPlaceAt	dbpedia:Berlin .

3. Publication under URL in Web, Intranet, Extranet

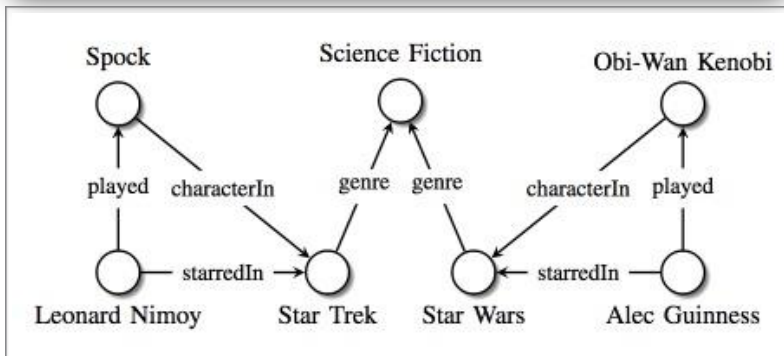
Linked Data

Creating Knowledge Graphs with RDF





The Knowledge Graph



GND	
Link zu diesem Datensatz	http://d-nb.info/gnd/1021356255
Typ	Person (piz)
Person	Appelbaum, Anne
Geschlecht	weiblich
Zeit	Lebensdaten: 1908-1998
Land	Deutschland (XA-DE); USA (XD-US)
Geografischer Bezug	Wirkungsort: New York, NY
Beruf(e)	Psychologin
Beziehungen zu Personen	Cassirer, Ernst (Vater)
Beteiligt an	1 Publikation 1. <i>Nachgelassene Manuskripte und Texte / Bd. 1. Zur Metaphysik der symbolischen Formen</i> 1995



Emerging Knowledge Graphs & Data Spaces

Search Engine Optimization & Web-Commerce

- Schema.org used by >20% of Web sites
- Major search engines exploit semantic descriptions



Pharma, Lifesciences

- Mature, comprehensive vocabularies and ontologies
- Billions of disease, drug, clinical trial descriptions

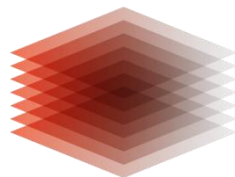


Digital Libraries

- Many established vocabularies (DublinCore, FRBR, EDM)
- Millions of aggregated from thousands of memory institutions in Europeana, German Digital Library



LEIBNIZ INFORMATION CENTRE
FOR SCIENCE AND TECHNOLOGY
UNIVERSITY LIBRARY



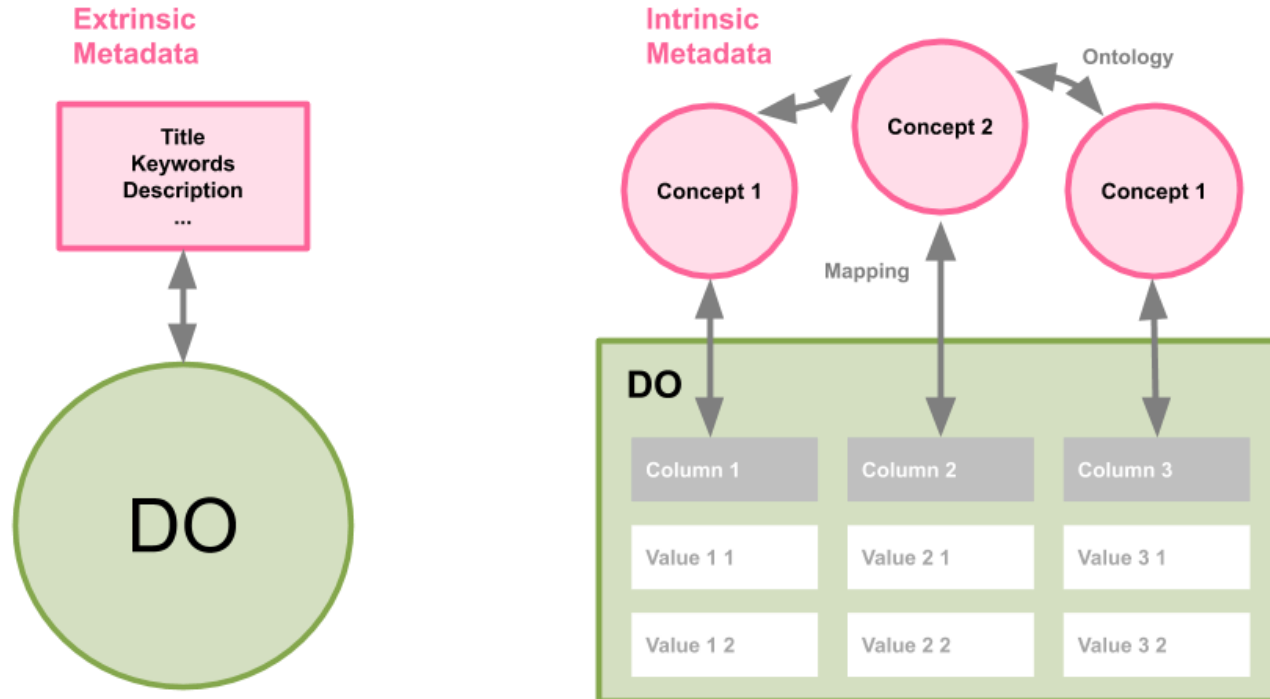
TIB

Realizing FAIR Research Data Management

Member of
Leibniz
Leibniz Association

From Extrinsic to Intrinsic Metadata

... by mapping raw data to shared conceptualizations (vocabularies/ontologies)



Ontology Layer

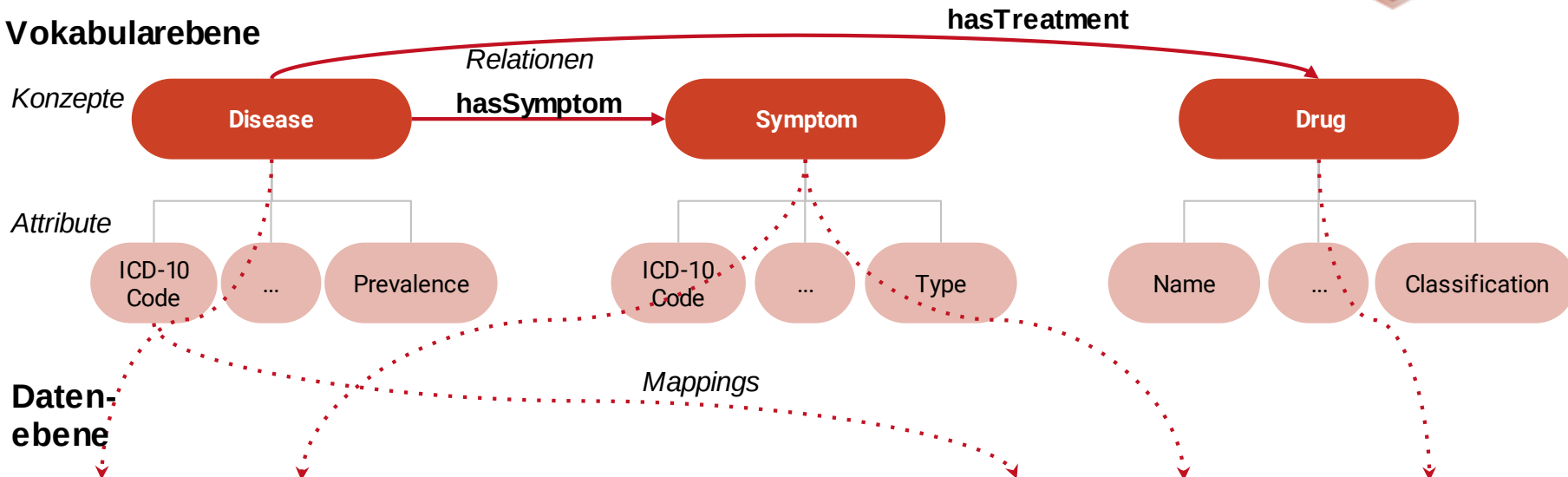
Collaborative Ontology
Development and
Publishing Environment –
TIB Terminology Service

Data Layer

Research Data Repository –
TIB Research Data Manager

Beispiel: Mapping von Forschungsdaten auf Vokabulare

Vokabularebene

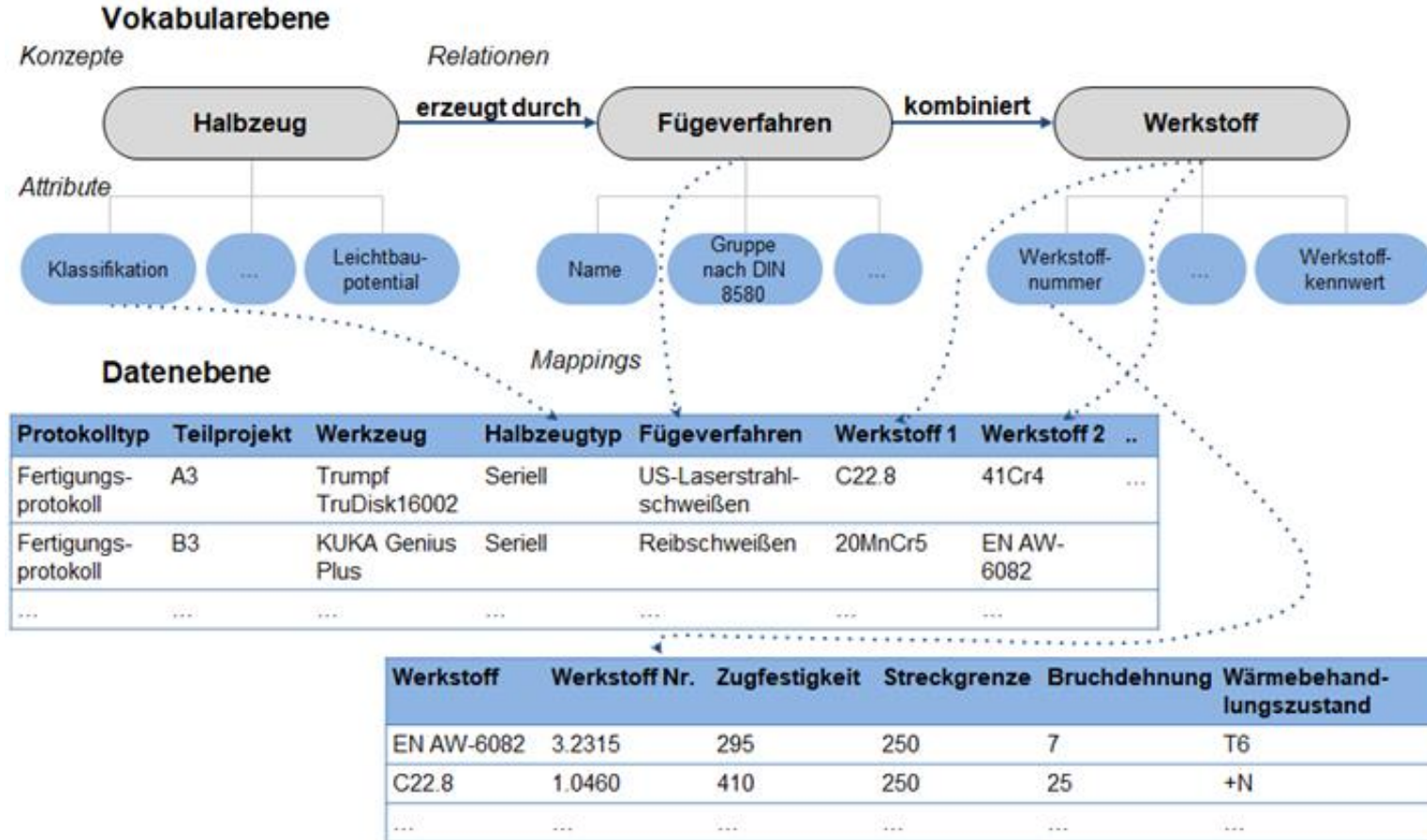


Daten-ebene

Krankheit	Symptom	Prävalenz
Grippe	Fieber	1000
Krebs	Blutung	30
...	...	...

Disease	ICD-10	Symptoms	Medication
Influenza	J10	Fever	Amantadin
Cancer	C00-C97	Bleeding	Chemotherapy
...	...	...	...

Beispiel: Semantische Forschungsdaten im Maschinenbau



TIB Terminology Service – offene CKAN Distribution für Forschungsdatenmanagement

**Collaborative Ontology Development and
Publishing Environment**

<https://service.tib.eu/ts4tib/ontologies>

<https://terminology.nfdi4chem.de/>

Welcome to the TIB Terminology Service

[Search](#)

Examples: [electric vehicle](#), [CHEBI:72695](#)

[Looking for a particular ontology?](#)

[TIB's Terminology Service](#)

With its new Terminology Service, [TIB – Leibniz Information Centre for Science and Technology and University Library](#) provides a single point of access to terminology from domains such as architecture, chemistry, computer science, mathematics and physics. The

[Community Terminology Services](#)

TIB maintains additional Terminology Services specific to several research communities as extensions to this service. [NFDI4Ing](#) and [NFDI4Chem](#) are particular examples of these extensions. It is planned to host more community specific Terminology Services as extensions of the central

[Report an Issue](#)

For feedback, enquiries or suggestions on our service or to request a new ontology please use our [GitHub issue tracker](#). For contacting us about further matters regarding the TIB Terminology Service, send an e-mail to information@tib.eu

Data Content

Updated 27 Sep 2021 20:08

- 43 ontologies
- 57,001 terms
- 6,536 properties
- 6,556 individuals

Tweets by [@TIBHannover](#)

 **TIB**
[@TIBHannover](#)

Gute Nachrichten von unseren Kolleg:innen von [@openaccessnet](#): Das Internetportal [open-access.network](#) wurde überarbeitet und bietet nun umfassende kostenlose Angebote zu den Themenschwerpunkten Informieren, Fortbilden und Vernetzen! [tib.eu/de/die-tib/neu](#)



Für freien Zugang zu Wi...
Überarbeitetes Internetpor...

Welcome to the NFDI4Chem Terminology Service

[Search](#)

Examples: [oxidation](#), [IAO_0020000](#)

[Looking for a particular ontology?](#)

About NFDI4ChemTS

The NFDI4Chem Terminology Service is a repository for chemistry and related ontologies providing a single point of access to the latest ontology versions. You can browse or search the ontologies and look into their terms and relations. The Terminology Service can be used either by humans through the website or by machines via the TS API. The NFDI4Chem Terminology Service is developed and maintained by TIB - Leibniz Information Centre for Science and Technology. It is part of the service portfolio of the NFDI4Chem consortium within the National Research Data Infrastructure.

Data Content

Updated 27 Sep 2021 20:08

- 19 ontologies
- 49,469 terms
- 2,311 properties
- 639 individuals

Tweets by [@Nfdi4Chem](#)



NFDI4Chem Retweeted



HerresPawlisLab

[@HerresLab](#)

Interested in the description of mixtures? We are happy to have Gerd Blanke and Leah McEwen in this week's Stammtisch [@Nfdi4Chem](#) talking about "MinChI – InChI for Mixtures"! 📺📅 Friday 24th september at 1.30 pm via Zoom! Please register here: nfdi4chem.de/index.php/even...



Sep 21, 2021

List of all ontologies in NFDI4ChemTS

Show

10

Search:

entries

Ontology Name	Short name	Description	Loaded	Action
Basic Formal Ontology	BFO	The upper level ontology upon which OBO Foundry ontologies are built.	Mon Sep 27 20:04:08 GMT 2021	Search Terms Properties Individuals Download
Bioinformatics operations, data types, formats, identifiers and topics	EDAM	EDAM is a simple ontology of well established, familiar concepts that are prevalent within bioinformatics, including types of data and data identifiers, data formats, operations and topics. EDAM provides a set of terms with synonyms and definitions - organised into an intuitive hierarchy for convenient use.	Mon Sep 27 20:05:55 GMT 2021	Search Terms Properties Individuals Download
CHEBI Integrated Role Ontology	CHIRO	CHEBI provides a distinct role hierarchy. Chemicals in the structural hierarchy are connected via a 'has role' relation. CHIRO provides links	Mon Sep 27 20:04:47 GMT	Search Terms

Ontologies are hierarchies of classes with associated properties

Molecular Interactions

A structured controlled vocabulary for the annotation of experiments concerned with protein-protein interactions. Developed by the HUPO Proteomics Standards Initiative.

[Terms](#)[Download](#)[Ontology Homepage](#)[Browse Terms](#)

└─ molecular interaction

└─ alias type

└─ attribute name

└─ biological role

└─ acceptor

└─ allosteric effector

└─ allosteric molecule

└─ ancillary

Ontology information

Ontology IRI: <http://purl.obolibrary.org/obo/mi.owl>

<http://purl.obolibrary.org/obo/mi.owl>

Version IRI: <http://purl.obolibrary.org/obo/mi/releases/2020-04-13/mi.owl>

<http://purl.obolibrary.org/obo/mi/releases/2020-04-13/mi.owl>

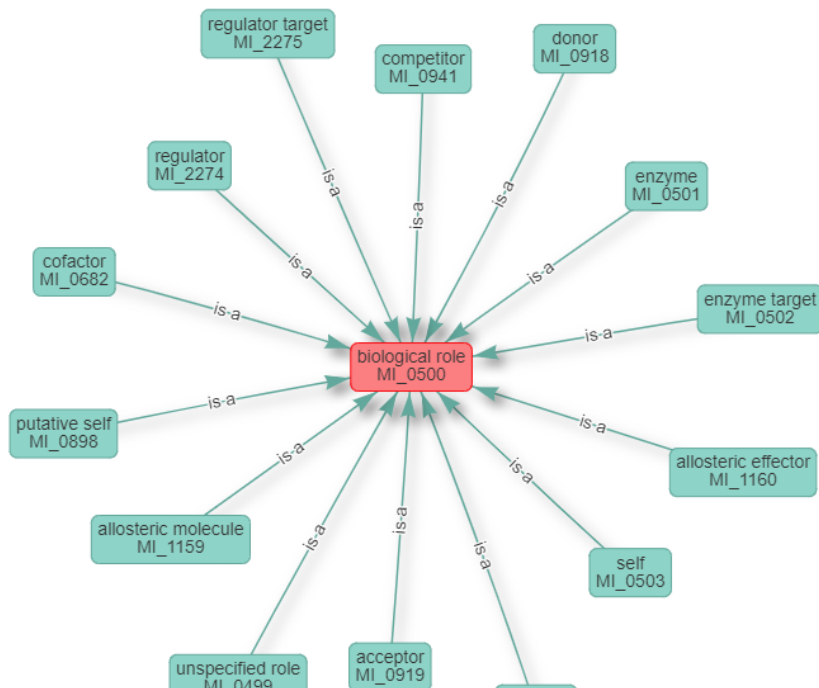
Ontology ID: mi

Version: 2020-04-13

Number of terms: 1566

Last loaded: Mon Sep 27 20:06:23 GMT 2021

Visualized term: biological role (http://purl.obolibrary.org/obo/MI_0500)



Legend

Relationship	Color	Visibility
Extended nodes (*)		-
is a		☒
part of		☐
Select/Deselect all		☐

List of extended nodes (*):

- biological role (MI_0500)



Classes ≡

Properties ≡

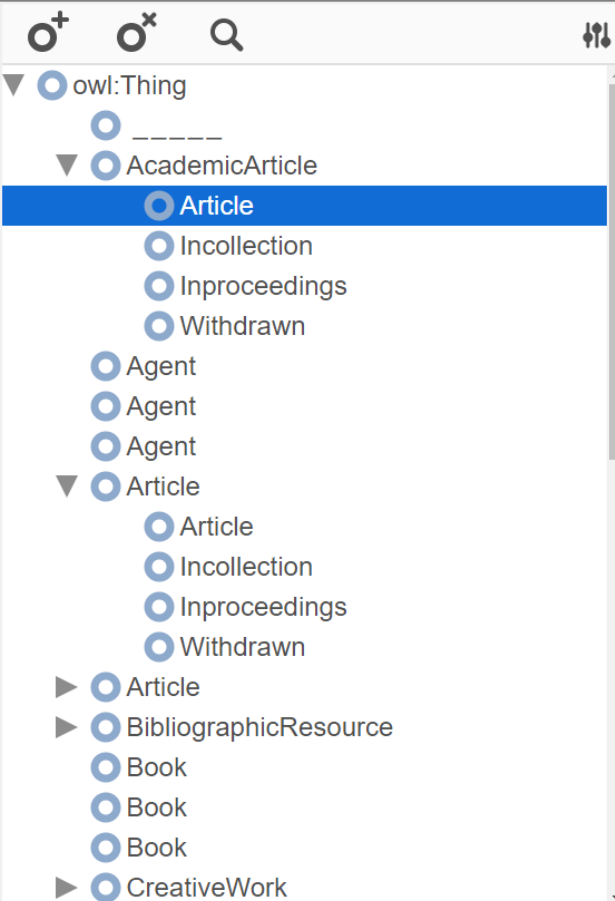
Individuals ≡

Comments ≡

Changes by Entity ≡

History ≡

Class Hierarchy



Class: Article

IRI
file:/srv/protege/webprotege-data/uploads/upload-2592441270055297140#Article

Annotations

rdfs:label	Article	en	✕
rdfs:comment	A journal article.	en	✕
Enter property	Enter value	lang	

Parents

Q13442814	✕
AcademicArticle	✕
Article	✕
Publication	✕
Article	✕
Enter a class name	

Relationships

Comments: A

Project Feed

TIB Data Manager – offene CKAN Distribution für Forschungsdatenmanagement

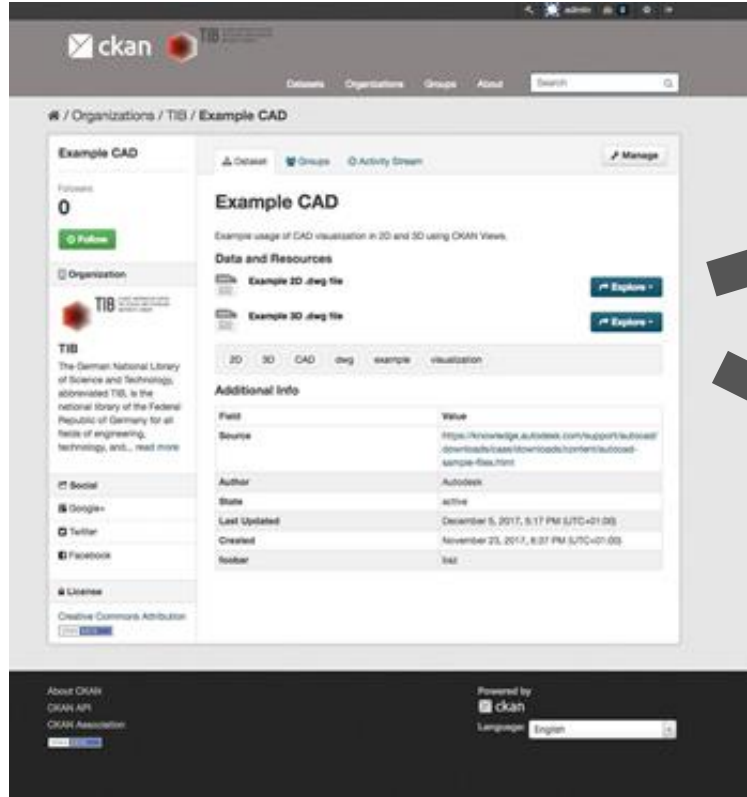
**Open-source Software Distribution der
renommierten CKAN Datenportal Plattform
für FDM**

CKAN: A Repository for Heterogeneous Data Collections

The screenshot shows the CKAN Datasets page. The header includes the CKAN logo, TIB logo, and navigation links: Datasets, Organizations, Groups, About. A search bar is present. The main content area is titled '4 / Datasets'. On the left, there are filters for Organizations (TIB (4)), Groups (None), Tags (2D (2), 3D (2), CAD (2), example (2), CA (1), Combustion (1), computer vision (1), dreg (1), EDTA (1), Experiment (1)), Show More Tags, Formats (TAR (1), vti2img4 (1)), and Licenses (Creative Commons At... (6)). The main list shows 4 datasets found. The first dataset is 'Jupyter notebooks' with a description: 'A collection of Jupyter Notebooks for science related projects LIQD Gravitation Wave Data Imagery Analysis 12 Steps to Navier-Stokes Computer Vision Machine Learning'. Other datasets listed are 'Example CAD', 'Pangaea CAD files', and 'Autocombustion reactions STFSO videos'. A callout box points to the '4 datasets found' text.

Data Collections with
different formats

CKAN: Different Views of the Same Data Collections



The screenshot shows the CKAN 2D view for the 'Example CAD' dataset. The page layout includes a sidebar with organization information (TIB), a main content area with a description, a 'Data and Resources' section listing two files ('Example 2D.dwg file' and 'Example 3D.dwg file'), and an 'Additional Info' table.

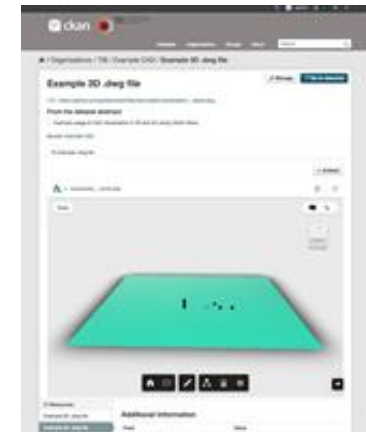
Field	Value
Source	https://knowledge.autodesk.com/support/tutorials/downloads/5445/downloads/tutorials/tutorials-sample-files.html
Author	Autodesk
Status	Active
Last Updated	December 5, 2017, 5:17 PM UTC+01:00
Created	November 23, 2017, 8:37 PM UTC+01:00
Isobar	Isat

2D
View



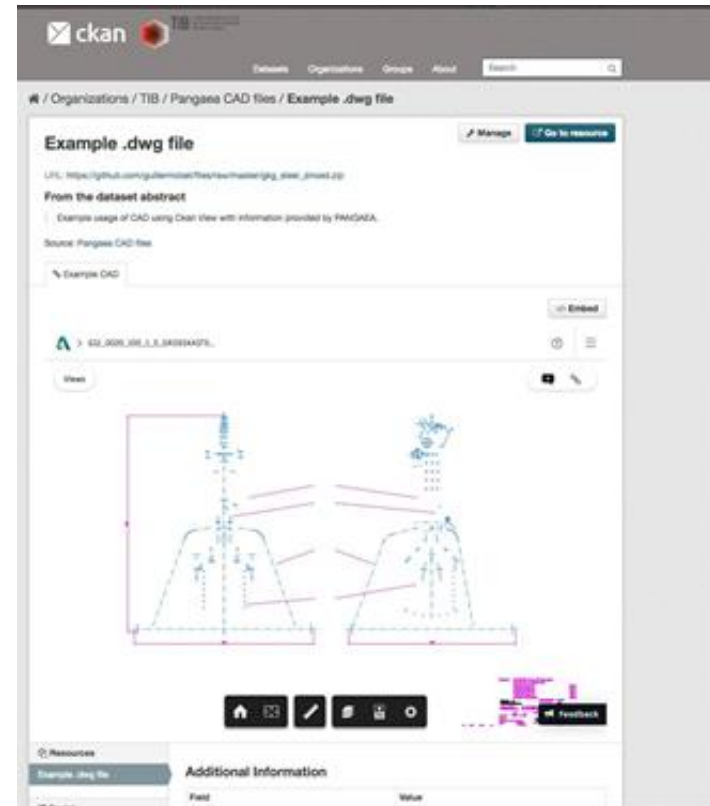
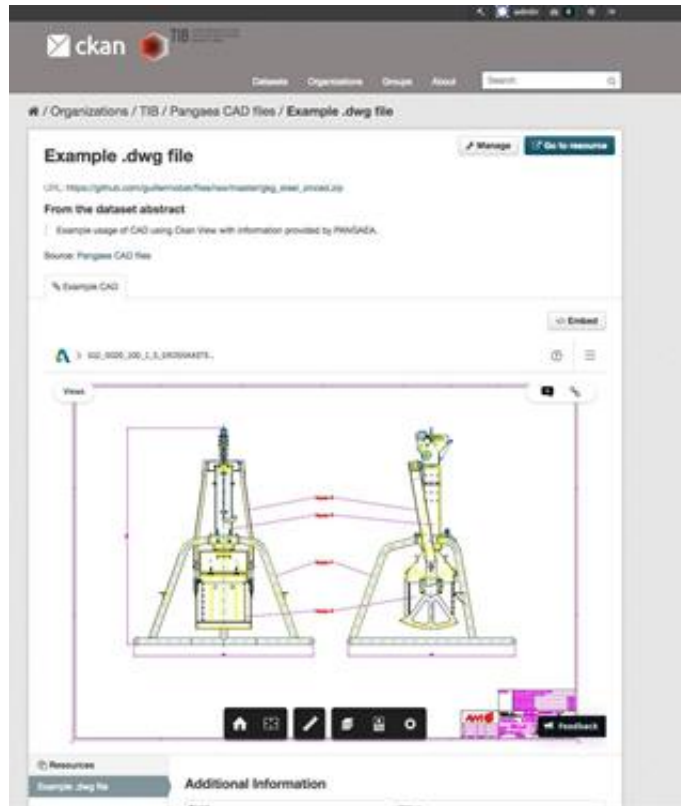
This screenshot shows the CKAN 2D view for the 'Example 3D.dwg file'. It displays a 2D wireframe model of a mechanical part, with a sidebar containing file details and a main area showing the model's 2D representation.

3D
View



This screenshot shows the CKAN 3D view for the 'Example 3D.dwg file'. It displays a 3D solid model of the same mechanical part, rendered in a light blue color, with a sidebar containing file details and a main area showing the 3D model.

CKAN: Visualizations of Data Collections using Auto CAD



CKAN: Playing a Video

ckan TIB

Datasets Organizations Groups About Search

/ Organizations / TIB / Autocombustion reactions ... / STF50 autocombustions with ...

STF50 autocombustions with varying Phi [Manage](#) [Go to resource](#)

URL: https://github.com/guilemofel/fresh/master/STF50_autocombustions_with_varying_phi_v2_HD.mp4

From the dataset abstract

Video about auto combustion reactions of STF50 with EDTA+CA: varying phi

Source: Autocombustion reactions STF50 video

Video

Embed

Autocombustion reactions of STF50 with EDTA+CA: varying ϕ

5:02 / 5:08

Resources

STF50 autocombustions ...

Social

Google+

Twitter

Facebook

Additional Information

Field	Value
Date last updated	December 1, 2017
Metadata last updated	unknown
Created	unknown
Format	video/mp4
License	Creative Commons Attribution



ckan TIB

Datasets Organizations Groups About Search

/ Organizations / TIB / Autocombustion reactions ... / STF50 autocombustions with ...

STF50 autocombustions with varying Phi [Manage](#) [Go to resource](#)

URL: https://github.com/guilemofel/fresh/master/STF50_autocombustions_with_varying_phi_v2_HD.mp4

From the dataset abstract

Video about auto combustion reactions of STF50 with EDTA+CA: varying phi

Source: Autocombustion reactions STF50 video

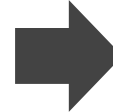
Video

Embed

Resources

STF50 autocombustions ...

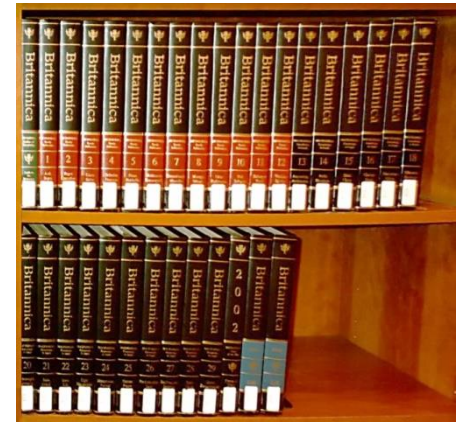
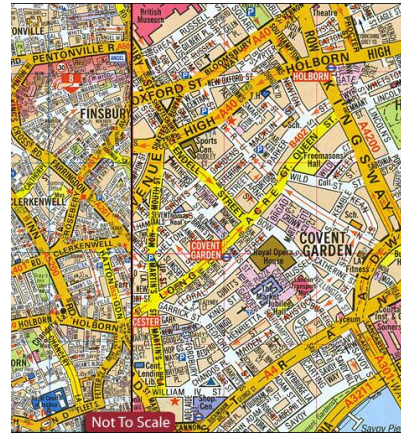
Additional Information



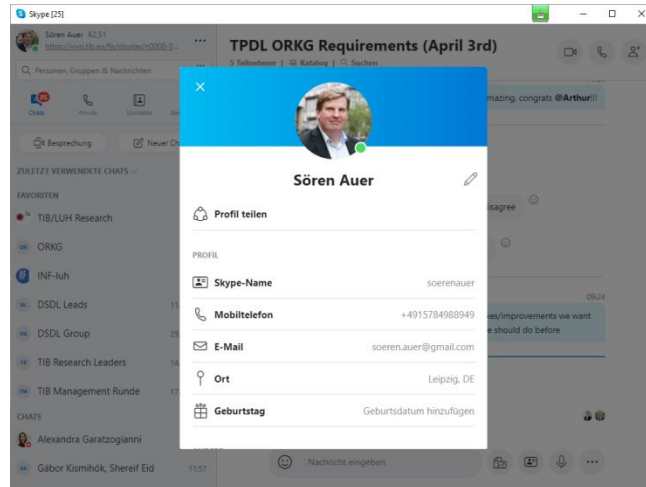
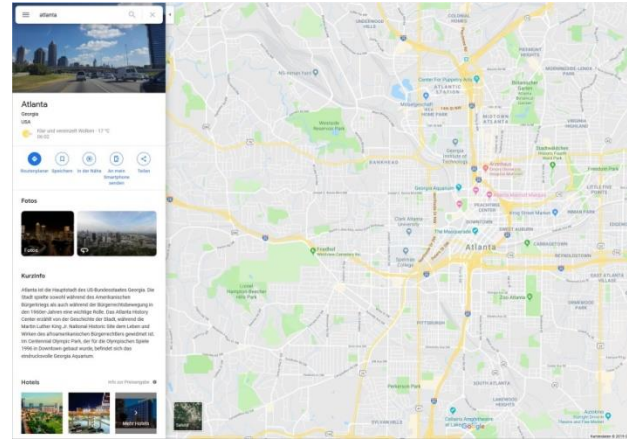
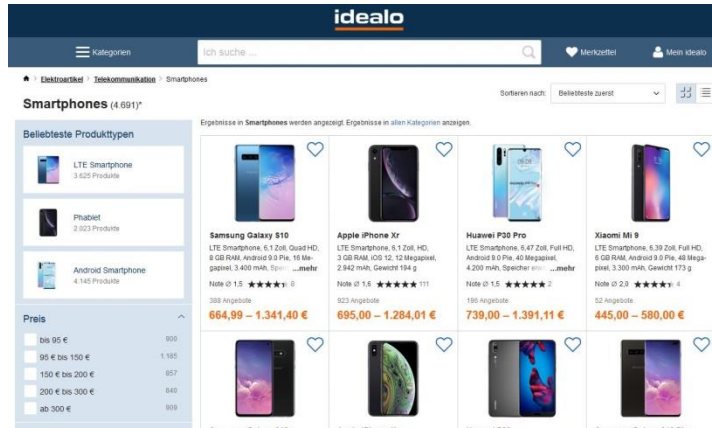
Organisation von Forschungsbeiträgen im Open Research Knowledge Graph

Wissensbasierte Informationsflüsse für die
Naturwissenschaften

How did information flows change in the digital era?



How does it work today?

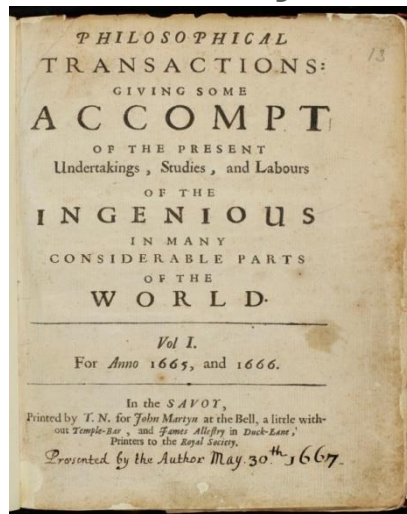


The World of Publishing & Communication has profoundly changed

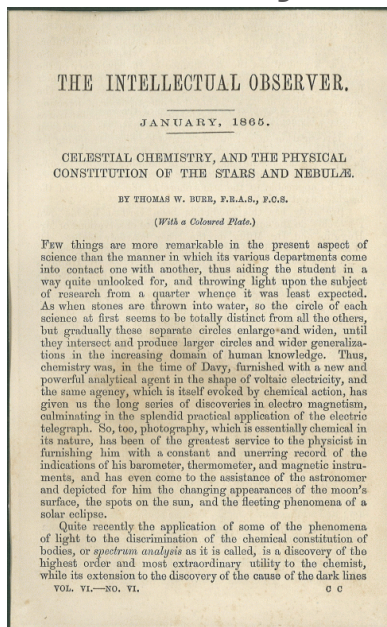
- **New means adapted to the new possibilities** were developed, e.g. „zooming“, dynamics
- **Business models** changed completely
- More focus on data, interlinking of **data / services and search** in the data
- Integration, **crowdsourcing**, **data curation** play an important role

Scholarly Communication has not changed (much)

17th century



19th century



20th century



21st century



Meanwhile other information intense domains were completely disrupted:

We need to rethink the way how research is represented and communicated

Challenges we are facing:

Digitalisation of Science

- Data integration and analysis
- Digital collaboration

Monopolisation by commercial actors

- Publisher look-in effects
- Maximization of profits ^[1]

Reproducibility Crisis

- Majority of experiments are hard or not reproducible ^[2]

Proliferation of publications

- Publication output doubled within a decade
- continues to rise ^[3]

Deficiency of Peer Review

- Deteriorating quality ^[4]
- Predatory publishing

[1] <http://thecostofknowledge.com>, <https://www.projekt-deal.de>

[2] M. Baker: *1,500 scientists lift the lid on reproducibility*, *Nature*, 2016.

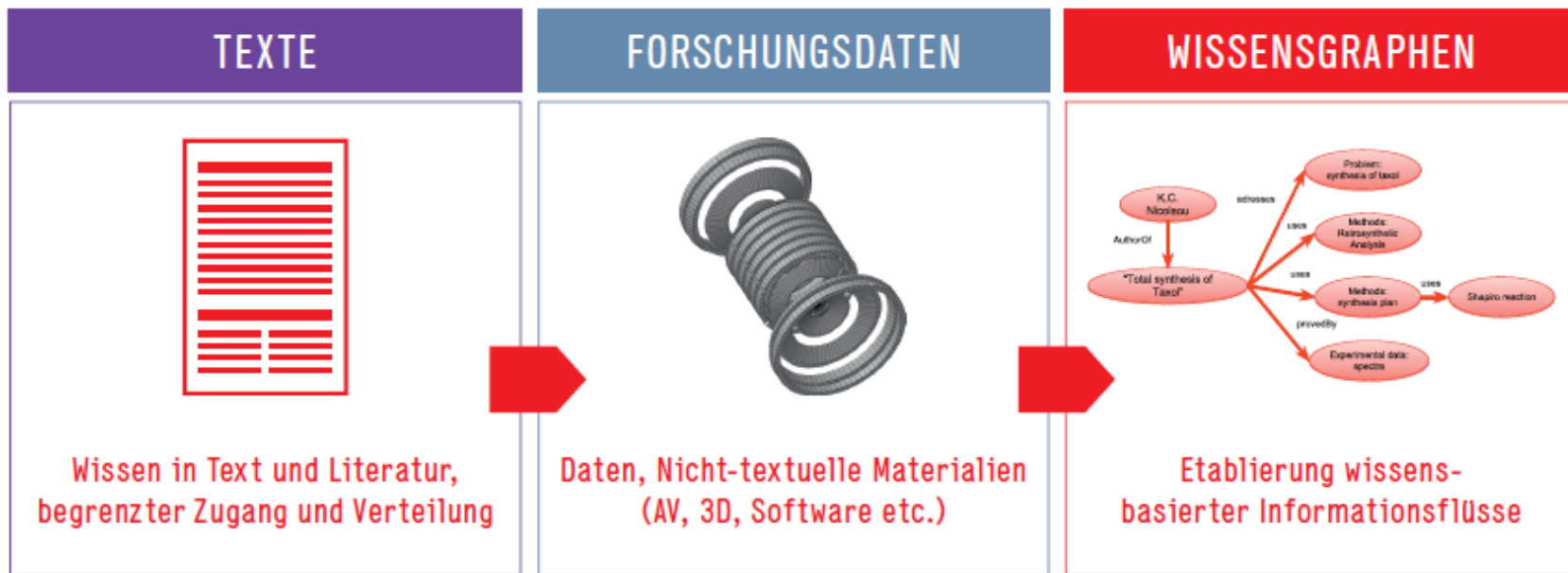
[3] *Science and Engineering Publication Output Trends*, National Science Foundation, 2018.

[4] J. Couzin-Frankel: *Secretive and Subjective, Peer Review Proves Resistant to Study*, *Science*, 2013.

Strategisches Ziel: Paradigmenwechsel im Informationsaustausch

Wissensbasierte Informationsflüsse in Wissenschaft und Technik

Herausforderungen: Digitalisierung der Wissenschaften, Monopolisierung durch kommerzieller Akteure, Proliferation von Publikationen, Reproducibility Crisis



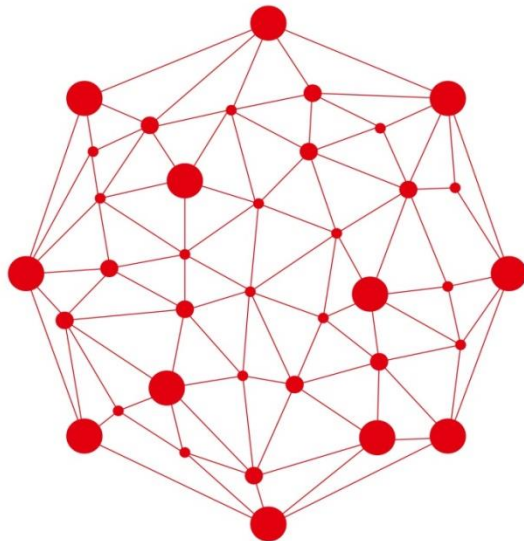
Open Research Knowledge Graph für die TIB-Fachdomänen

Übergreifende Konzepte

- Forschungsprobleme
- Definitionen
- Forschungsansätze
- Methoden

Artefakte

- Publikationen
- Daten
- Software
- Bild/Audio/Video
- Wissensgraphen / Ontologien



Open Research Knowledge Graph macht übergreifende und fachspezifische Konzepte **eindeutig identifizierbar** und **vernetzt diese semantisch** (mit klar beschriebenen Relationen) untereinander und mit relevanten weiteren Artefakten

Fachspezifische Konzepte

Mathematik

- Definitionen
- Theoreme
- Beweise
- Methoden
- ...

Physik

- Experimente
- Daten
- Modelle
- ...

Chemie

- Stoffe
- Strukturen
- Reaktionen
- ...

Informatik

- Konzepte
- Implementierungen
- Evaluationen
- ...

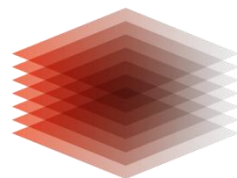
Technik

- Normen / Standards
- Prozesse
- Bauelemente
- Einheiten, Sensordaten

Architektur

- Normen / Regularien
- Bauelemente
- Modelle
- ...

LEIBNIZ INFORMATION CENTRE
FOR SCIENCE AND TECHNOLOGY
UNIVERSITY LIBRARY



TIB

Demo: Open Research Knowledge Graph

<https://orkg.org>

Properties	Estimation of the epidemic properties of the 2019 novel coronavirus: A mathematical modeling study Contribution 1 - 2020	Estimation of the epidemic properties of the 2019 novel coronavirus: A mathematical modeling study Contribution 2 - 2020	Estimation of the epidemic properties of the 2019 novel coronavirus: A mathematical modeling study Contribution 3 - 2020	Transmission potential of COVID-19 in Iran Contribution 1 - 2020	Transmission potential of COVID-19 in Iran Contribution 2 - 2020	Estimating the generation interval for COVID-19 based on symptom onset data Contribution 1 - 2020
Has research problem	COVID-19 reproductive number	COVID-19 reproductive number	COVID-19 reproductive number	COVID-19 reproductive number	COVID-19 reproductive number	COVID-19 reproductive number
Location	Wuhan City, China	Wuhan City, China	Wuhan City, China	Iran	Iran	Singapore
Study date	2020-01-10/2020-01-23	2020-01-23/2020-02-08	2020-01-10/2020-02-08	2020-02-19/2020-02-29	2020-02-19/2020-02-29	2020-01-21/2020-02-26
RO estimates (average)	4.38	3.41	3.39	3.6	3.58	1.27
95% confidence interval	3.63-5.13	3.16-3.65	3.09-3.70	3.2-4.2	1.29-8.46	1.19-1.36
Method	a weighted average of Exponential growth, Maximum likelihood, Sequential Bayesian, Time-dependent reproduction numbers, and SEIR model basic reproduction numbers by calculating weights from a Poisson loss function	a weighted average of Exponential growth, Maximum likelihood, Sequential Bayesian, Time-dependent reproduction numbers, and SEIR model basic reproduction numbers by calculating weights from a Poisson loss function	a weighted average of Exponential growth, Maximum likelihood, Sequential Bayesian, Time-dependent reproduction numbers, and SEIR model basic reproduction numbers by calculating weights from a Poisson loss function	generalized growth model	based on the calculation of the epidemic's doubling times: estimated epidemic doubling time of 1.20 (95% CI, 1.05, 1.44) days	generation interval

View paper

[View paper](#)
[Graph view](#)
[Edit](#)

Transmission interval estimates suggest pre-symptomatic spread of COVID-19

 March 2020

 Virology

 Lauren Tindale

 Michelle Coombe

 Jessica Stockdale

 Emma Garlock

 Wing Yin Venus Lau

 Manu Saraswat

 Yen-Hsiang Brian Lee

 Louxin Zhang

 Dongxuan Chen

 Jacco Wallinga

 Caroline Colijn

DOI: [10.1101/2020.03.20029983](#)

Share this paper: [f](#) [t](#)
Contribution 1

Contribution 2

Research problems

☐ Add to comparison

COVID-19 reproductive number

Contribution data

95% Confidence interval

1.45-2.48

Location

Singapore

Mean incubation period

7.1 (6.13, 8.25) days

Mean serial interval

4.56 (2.69, 6.42) days

R0 estimates (average)

1.97

Study date

2020-01-19/2020-02-26



Edit mode Every change you make is automatically saved

View paper

Graph view

Stop editing

March 2020

Virology

Lauren Tindale

Michelle Coombe

Jessica Stockdale

Emma Garlock

Wing Yin Venus Lau

Manu Saraswat

Yen-Hsiang Brian Lee

Louxin Zhang

Dongxuan Chen

Jacco Wallinga

Caroline Colijn

DOI: 10.1101/2020.03.03.20029983

Edit data

Share this paper:

Contribution 1

Contribution 2



Research problems

☐ Add to comparison

COVID-19 reproductive number ✕



Contribution data

95% Confidence interval

1.45-2.48



Location

Singapore



Mean incubation period

7.1 (6.13, 8.25) days

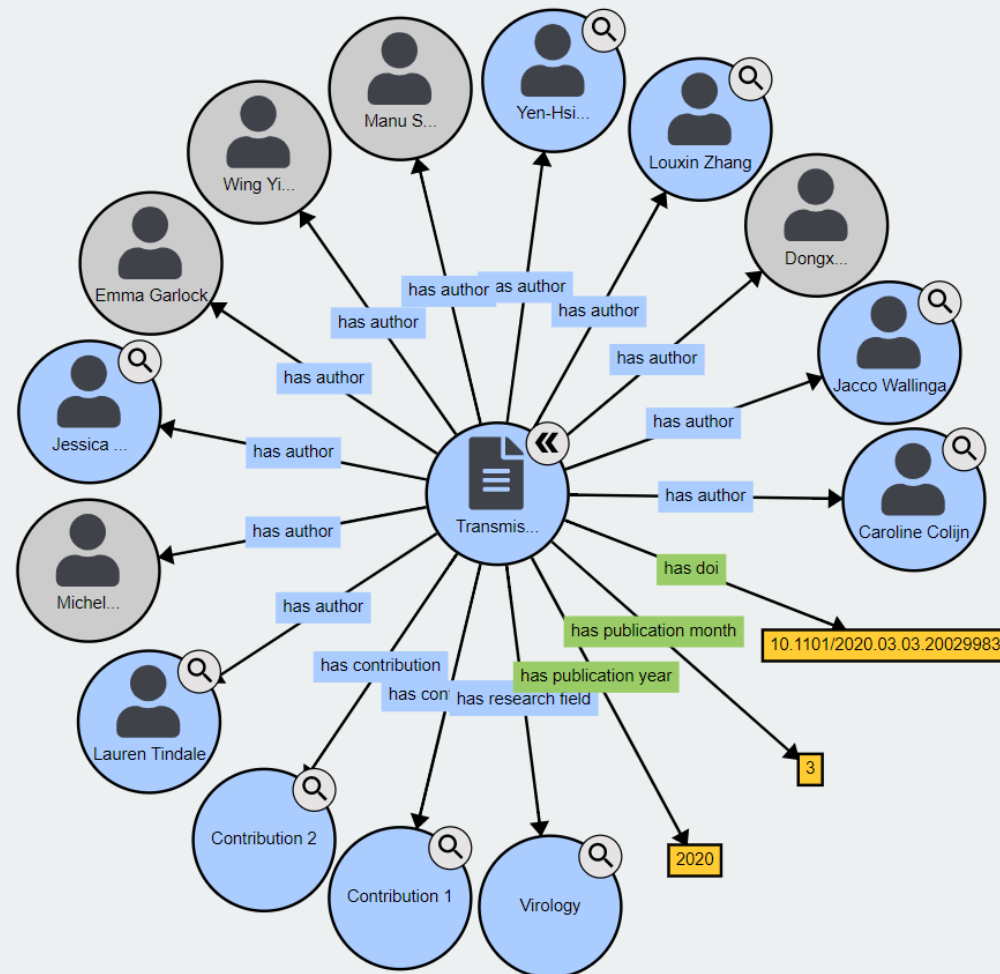


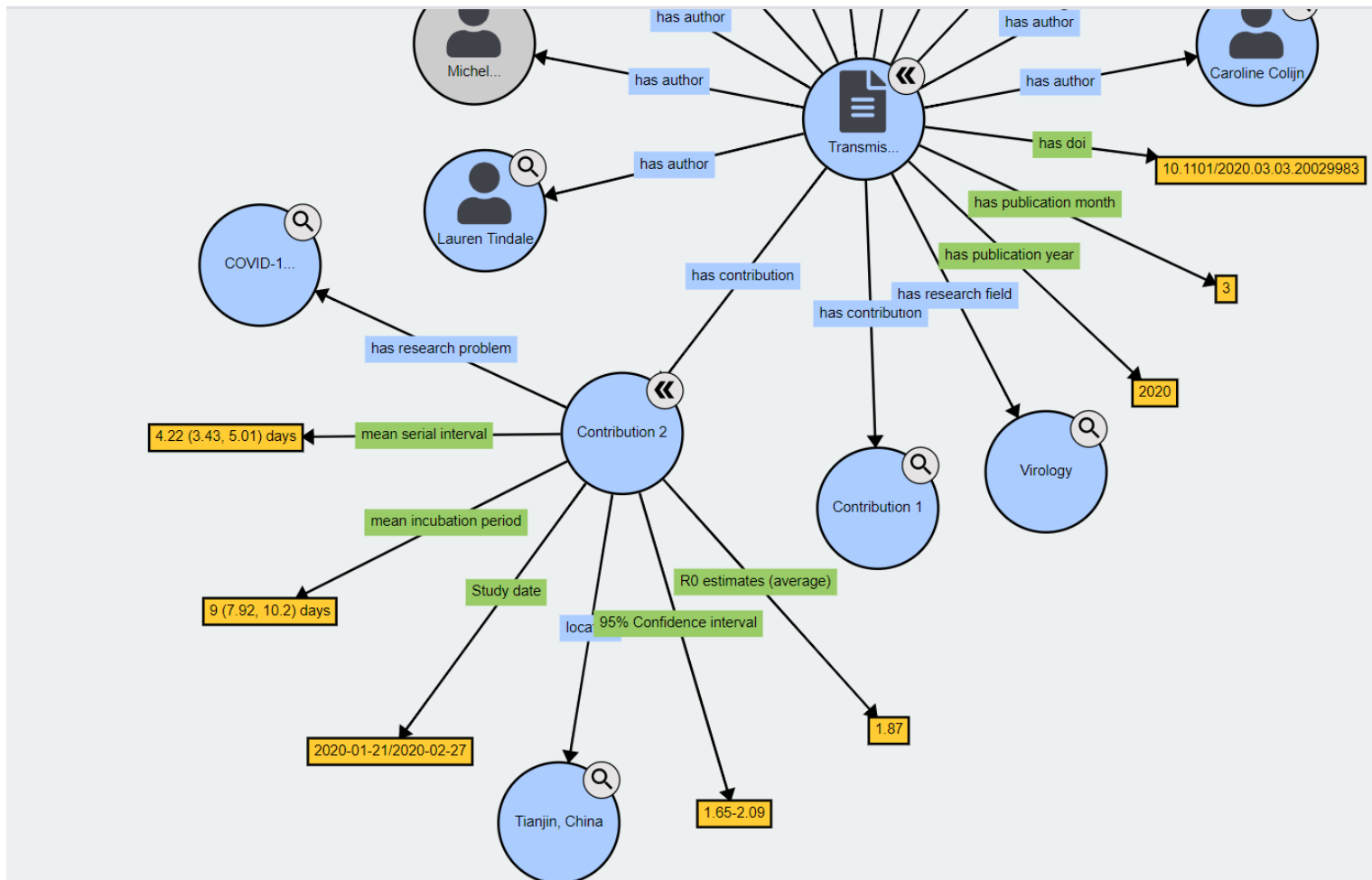
Edit value

Mean serial interval

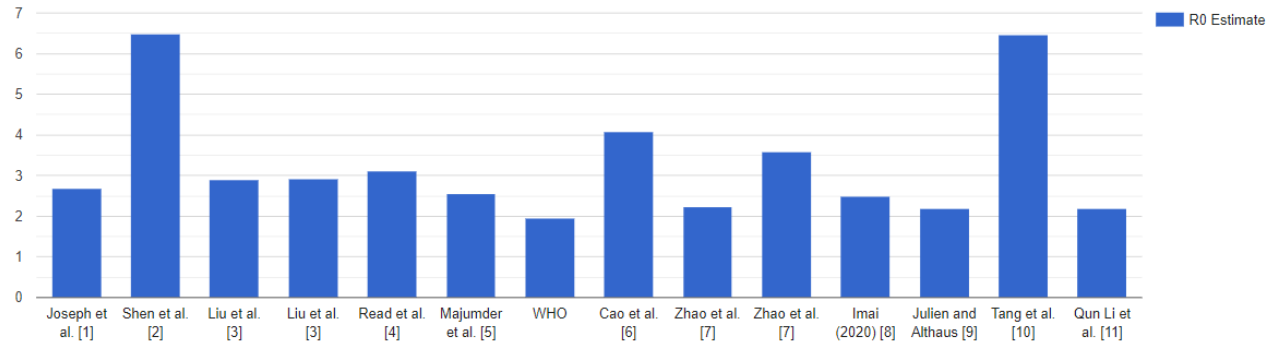
4.56 (2.69, 6.42) days



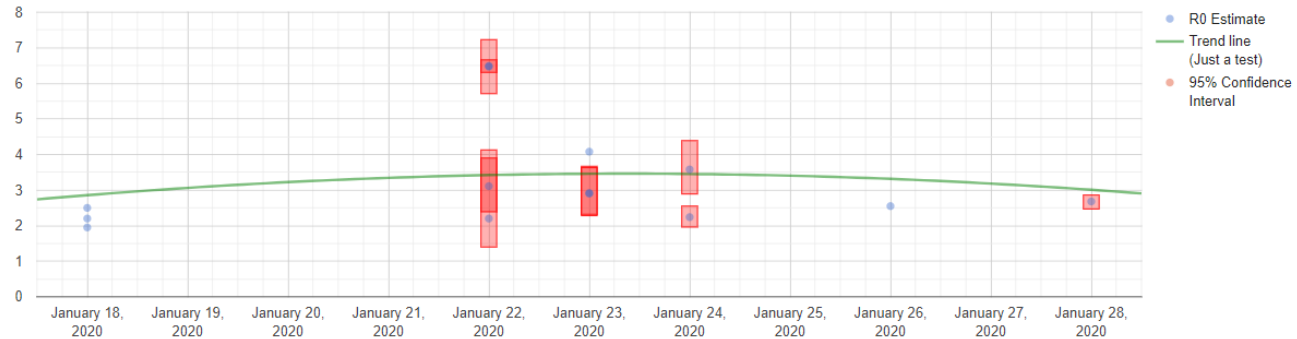




Bar Chart



Combo Chart



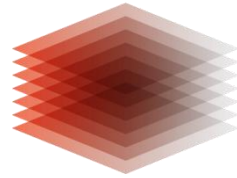
Zusammenfassung

Offene TIB Data Manager CKAN Distribution ermöglicht Realisierung föderierter organisations- oder fachspezifischer Portale zur Publikation von Forschungsdaten

Vokabulare, Taxonomien, Ontologien und Wissensgraphen ermöglichen die Etablierung eines gemeinsamen Verständnis der Forschungsdaten (→ Nachhaltigkeit, Wiederverwendbarkeit)

Mit dem geplanten Open Research Knowledge Graph können auch Forschungsbeiträge besser mit Forschungsdaten vernetzt werden

LEIBNIZ INFORMATION CENTRE
FOR SCIENCE AND TECHNOLOGY
UNIVERSITY LIBRARY



TIB



<https://de.linkedin.com/in/soerenauer>



<https://twitter.com/soerenauer>



https://www.xing.com/profile/Soeren_Auer



http://www.researchgate.net/profile/Soeren_Auer

Prof. Dr. Sören Auer



Leibniz
Universität
Hannover

TIB & Leibniz University of Hannover
Soeren.Auer@tib.eu

Member of

Leibniz
Leibniz Association