LEIBNIZ INFORMATION CENTRE
FOR SCIENCE AND TECHNOLOGY
UNIVERSITY LIBRARY

TIB

The Open Research Knowledge Graph and its Use for Plasma Physics

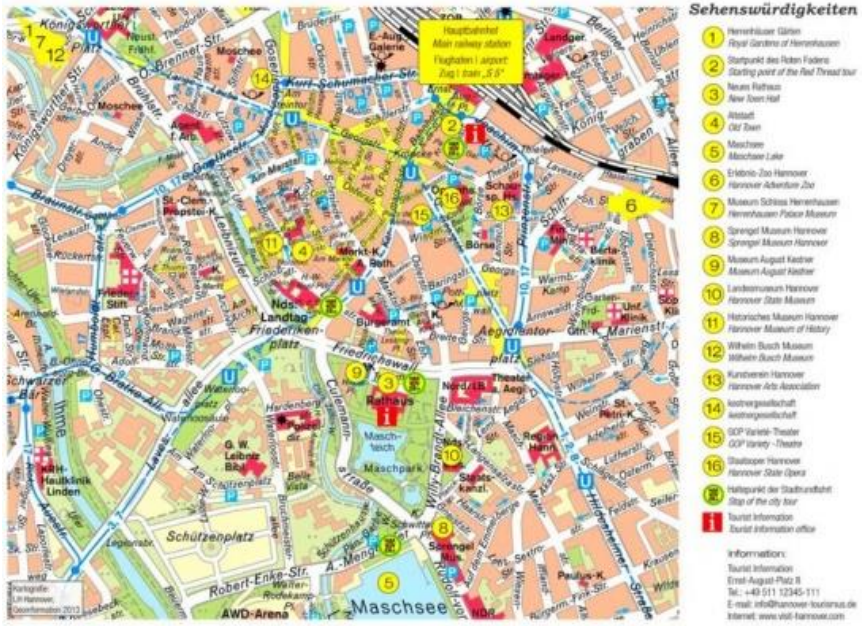
Int. Workshop on FAIR Data in Plasma Science
May 18

Leibniz Association

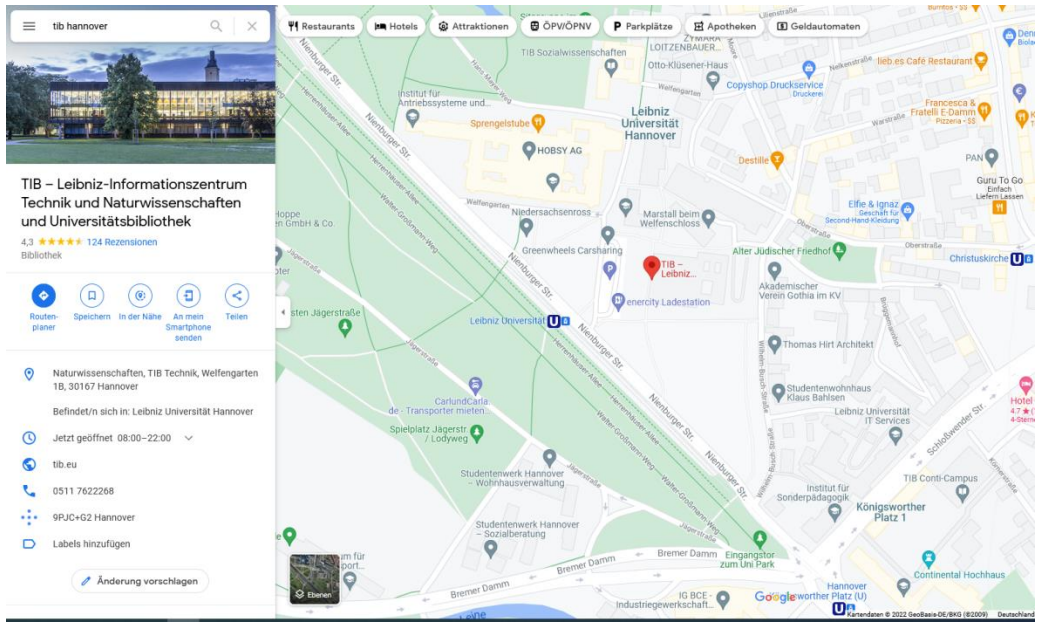
Digitalization in everyday life

Navigation to TIB in Hannover

50 years ago



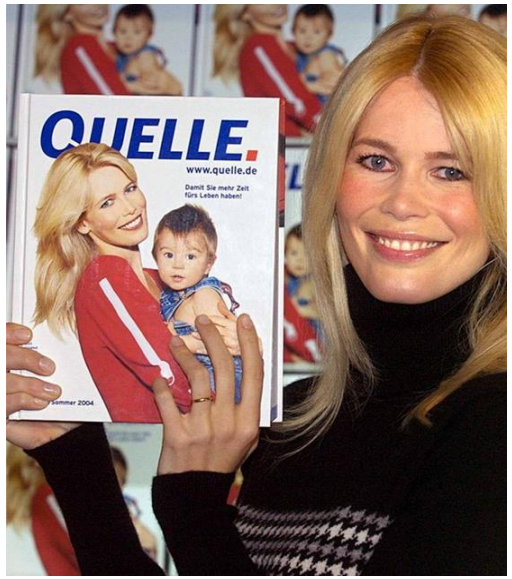
Now



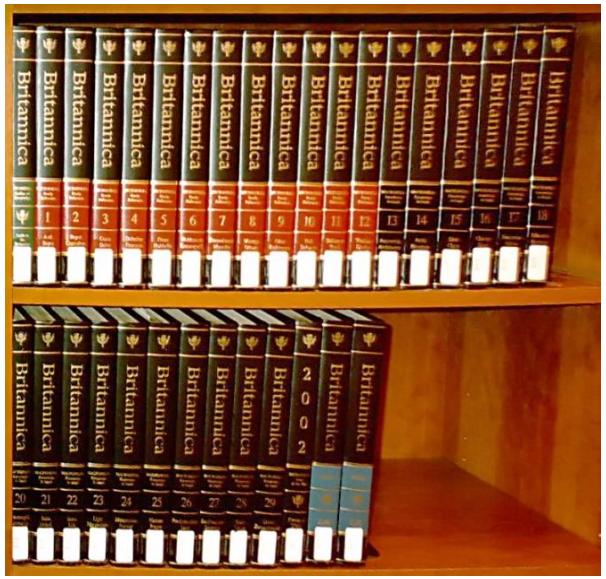
+ New Features:
  - → Zoom in
  - → Traffic jam warning
  - → Opening hours
  - → Interesting places around

Similar in other domains….

Who still remembers?



Mail order catalogs



Encyclopedias

Phone books

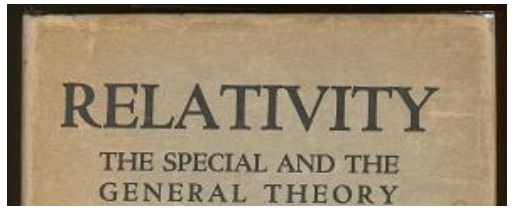
Whole industries got disrupted and our lives were significantly changed

What about Science?

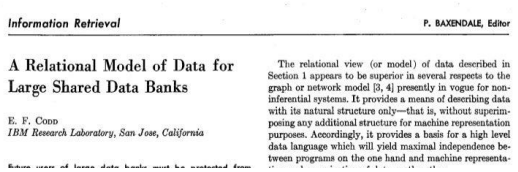
Over 300 years ago


ACTA
ERUDITORUM
ANNO M DC LXXXIV
publicata,

100 years ago


RELATIVITY
THE SPECIAL AND THE
GENERAL THEORY

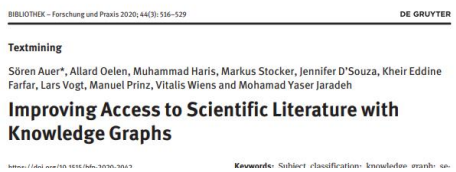
20 years ago


Information Retrieval P. BAXENDALE, Editor

A Relational Model of Data for Large Shared Data Banks

E. F. Codd
IBM Research Laboratory, San Jose, California

Today

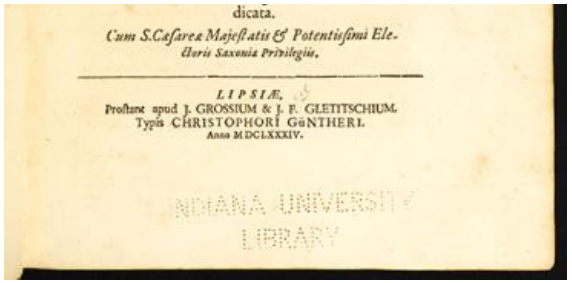

BIBLIOTHEK – Forschung und Praxis 2020; 44(3): 516–529 DE GRUYTER

Textmining

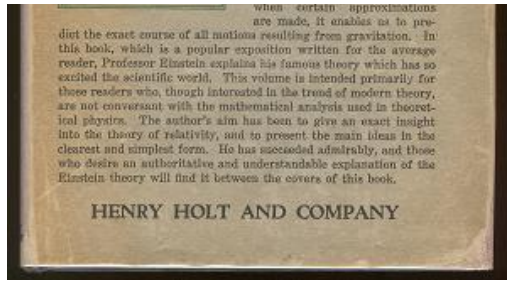
Sören Auer*, Allard Oelen, Muhammad Haris, Markus Stocker, Jennifer D'Souza, Kheir Eddine Farfar, Lars Vogt, Manuel Prinz, Vitalis Wiens and Mohamad Yaser Jaradeh

Improving Access to Scientific Literature with Knowledge Graphs

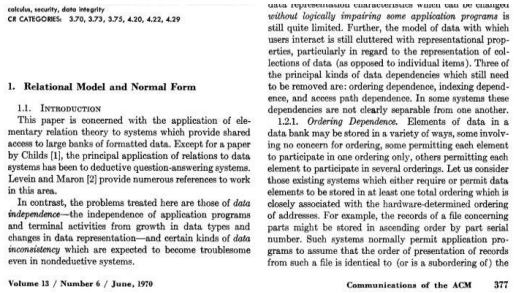
Science does not harvest the full potential of digitalization


dicata.
Cum S.Cæsareæ Majestatis & Potentissimi Electoris Saxoniæ Privilegio.

LIPSIÆ,
Prostant apud J. GROSSIUM & J. F. GLEITTSCHIUM.
Typis CHRISTOPHORI GüNTHERI.
Anno MDCLXXXIV.


when certain approximations are made, it enables us to predict the exact course of all motions resulting from gravitation. In this book, which is a popular exposition written for the average reader, Professor Einstein explains his famous theory which has so excited the scientific world. This volume is intended primarily for those readers who, though interested in the trend of modern theory, are not conversant with the mathematical analysis used in theoretical physics. The author's aim has been to give an exact insight into the theory of relativity, and to present the main ideas in the clearest and simplest form. He has succeeded admirably, and those who desire an authoritative and understandable explanation of the Einstein theory will find it between the covers of this book.

HENRY HOLT AND COMPANY


calculus, security, data integrity
CR CATEGORIES: 3.70, 3.73, 3.75, 4.20, 4.22, 4.29

1. Relational Model and Normal Form

1.1. Introduction

This paper is concerned with the application of elementary relation theory to systems which provide shared access to large banks of formatted data. Except for a paper by Childs [1], the principal application of relations to data systems has been to deductive question-answering systems. Levein and Maron [2] provide numerous references to work in this area.

In contrast, the problems treated here are those of *data independence*—the independence of application programs and terminal activities from growth in data types and changes in data representation—and certain kinds of *data inconsistency* which are expected to become troublesome even in nondeductive systems.

Volume 13 / Number 6 / June, 1970 Communications of the ACM 377


knowledge graph such as the ORKG can be used to give a condensed overview on the state-of-the-art addressing a particular research quest, for example as a tabular comparison of contributions according to various characteristics of the approaches. Further possible intuitive access interfaces to such scholarly knowledge graphs include domain-specific (chart) visualizations or answering of natural language questions.

*Corresponding author: Prof. Dr. Sören Auer, auer@tib.eu
Allard Oelen, allard.oelen@tib.eu
Muhammad Haris, muhammad.haris@tib.eu
Dr. Markus Stocker, markus.stocker@tib.eu
Dr. Jennifer D'Souza, jennifer.dsouza@tib.eu
Kheir Eddine Farfar, kheir.farfar@tib.eu
Lars Vogt, Lars.vogt@tib.eu
Manuel Prinz, manuel.prinz@tib.eu
Vitalis Wiens, vitalis.wiens@tib.eu
Mohamad Yaser Jaradeh, yaser.jaradeh@tib.eu

manuellen (crowd/expert sourcing) und (halb-)automatisierten Techniken ein. Nur mit einer solchen Kombination aus menschlicher und maschineller Intelligenz können wir die erforderliche Qualität der Darstellung erreichen, um neuartige Explorations- und Unterstützungsdienste für Forscher zu ermöglichen. Im Ergebnis kann ein Wissensgraph wie der ORKG verwendet werden, um einen komprimierten Überblick über den Stand der Technik in Bezug auf eine bestimmte Forschungsaufgabe zu geben, z. B. als tabellarischer Vergleich der Beiträge nach verschiedenen Merkmalen der Ansätze. Weitere mögliche intuitive Nutzungsschnittstellen zu solchen wissenschaftlichen Wissensgraphen sind domänenspezifische Visualisierungen oder die Beantwortung natürlichsprachlicher Fragen mittels Question Answering.

Schlüsselwörter: Sacherschließung; Wissensgraph; Semantic Web; Crowdsourcing; Text Mining

Not much has changed!

A Consequence of Document Centered Information Flows: The Publication Flood

- ~ 2.5 Mio new publications per year
- Researchers lack overview, even in small fields
- Loss of knowledge
- Answering questions is like looking for a needle in the haystack

An Example – CRISPR

Specific research questions:

- Who applied CRISPR to butterflies?
- How to apply CRISPR with minimal costs?
- How do different genome editing techniques compare?

Further Challenges of Document-Orientation

Reproducibility Crisis

ELSEVIER

Monopolization of commercial actors

Deficiency of Peer-Review

Lack of machine assistance

Predatory Publishing

Time to Rethink Scholarly Communication!

The solution is not „better pdfs“…

*“The lightbulb was **not** invented by improving the candle.”*
Oren Harari

Digitalization is **more** than just Digitization!
Current and future scientific challenges can not be tackled with an outdated communication system.

Digitalize Knowledge, Not Documents!

The Open Research Knowledge Graph

ORKG

As the name already suggests, ORKG is a **knowledge graph**.

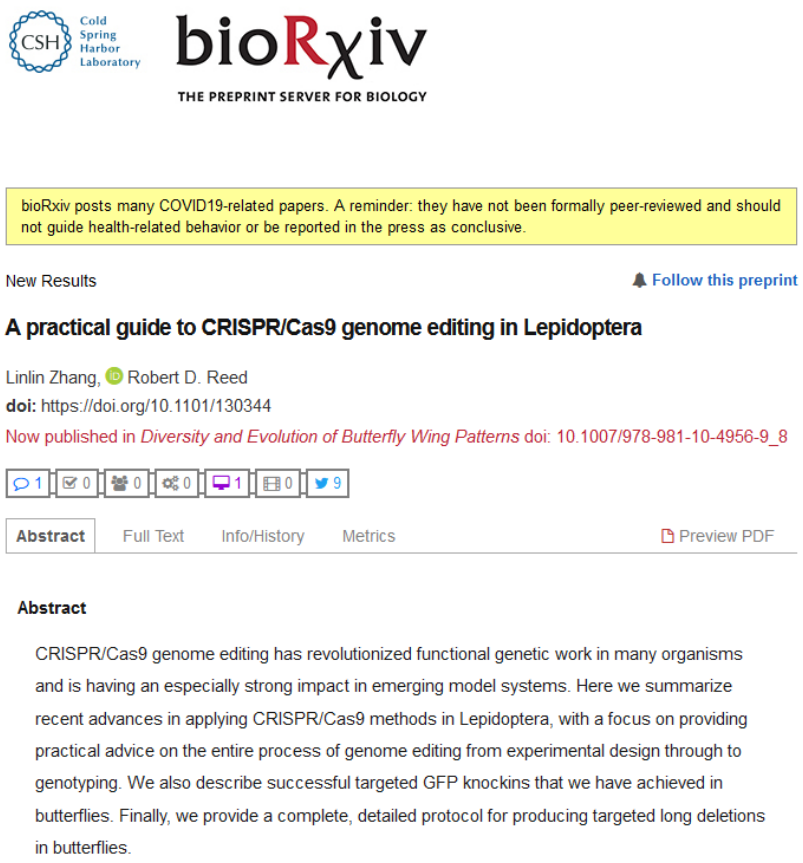
Knowledge Graphs are widely used in industry…

Why not use them for (open) science as well?

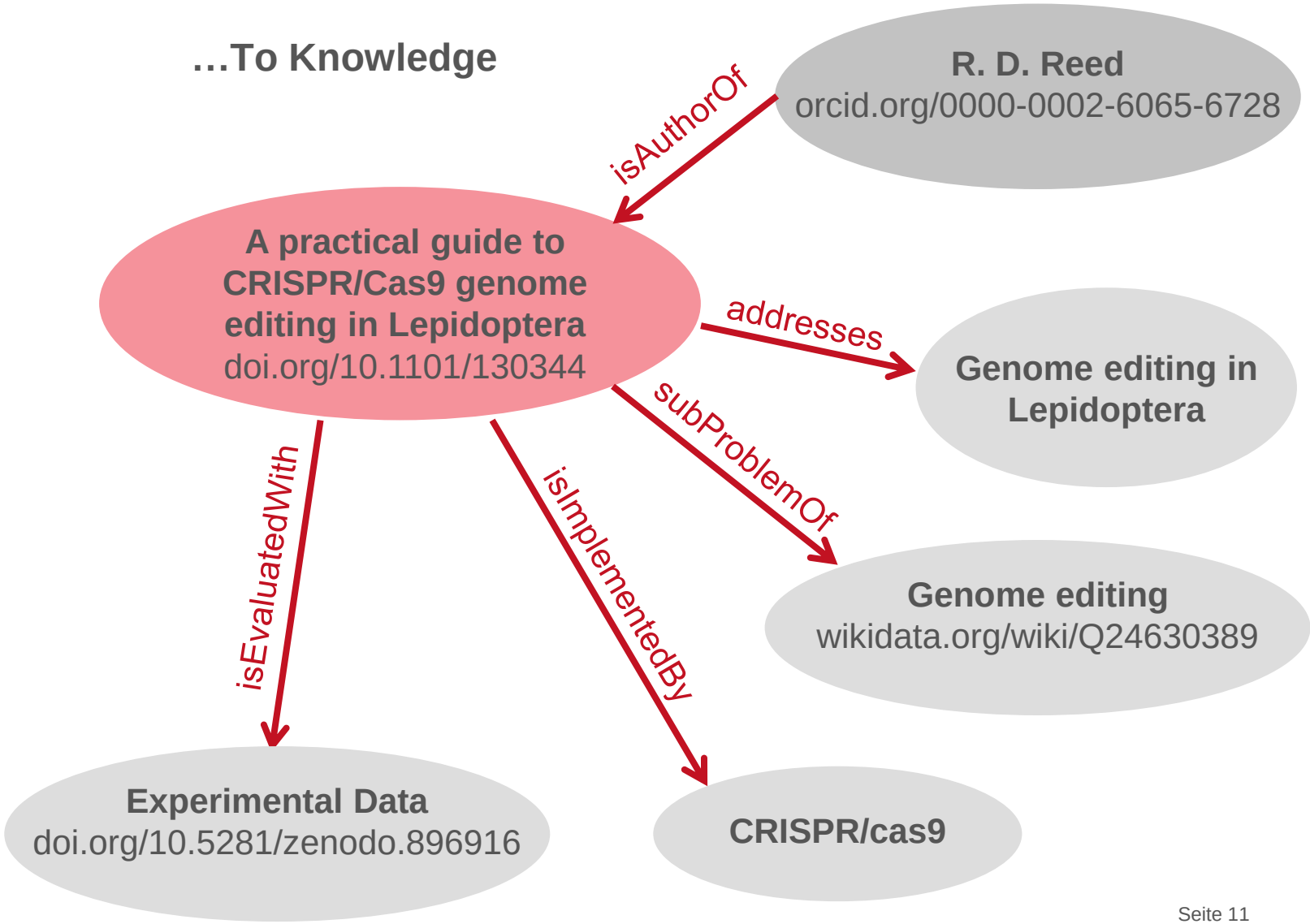
https://www.slideshare.net/Frank.van.Harmelen/adoption-of-knowledge-graphs-late-2019

Knowledge Representation in Graphs

From papers…

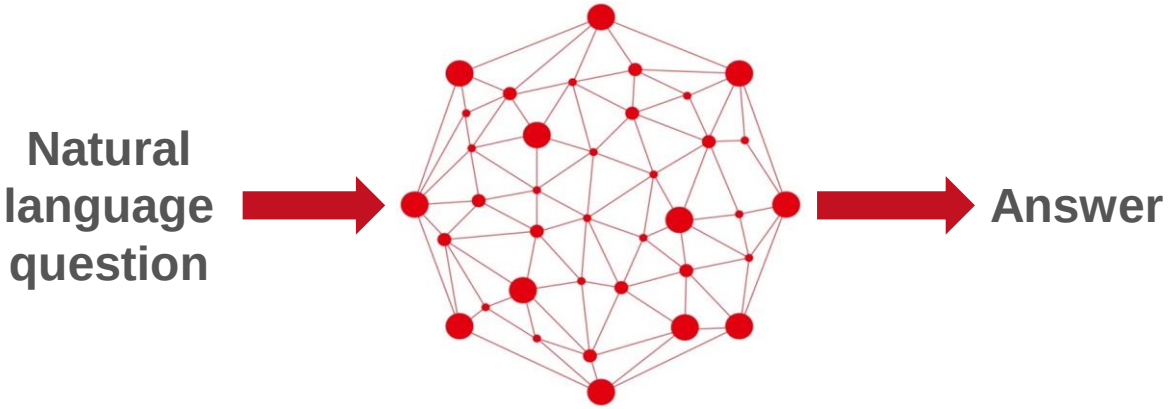


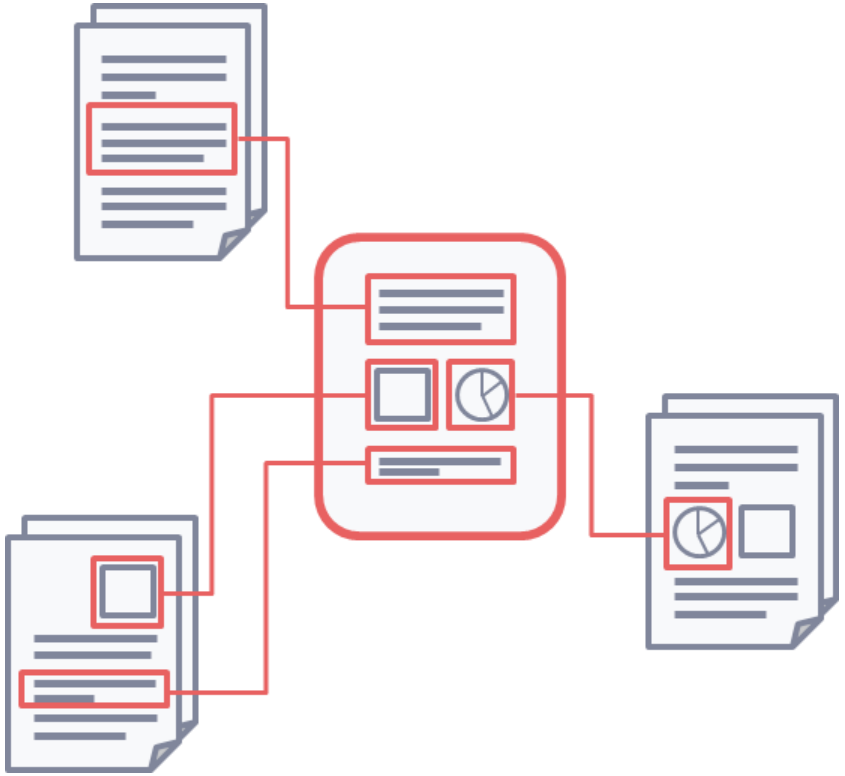
…To Knowledge



Advantages of a Graph-Based Approach

- Machine-actionable
- Automated finding and linking of research contributions towards a specific problem
- Natural language question answering possible
  e.g. „How do different genome editing techniques compare?“



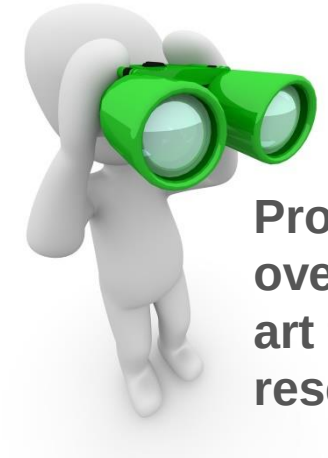


- Explore knowledge in entirely new ways

An Example: COVID-19 Basic Reproduction Number

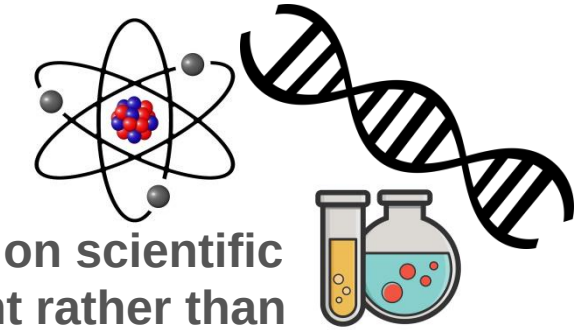
Properties	The early phase of the COVID-19 outbreak in Lombardy, Italy *2020 - Contribution 1*	Transmission potential of COVID-19 in Iran *2020 - Contribution 1*	Transmission potential of COVID-19 in Iran *2020 - Contribution 2*	Estimating the generation interval for COVID-19 based on symptom onset data *2020 - Contribution 1*
location	Lombardy, Italy	Iran	Iran	Singapore
Time period	Time interval	Time interval	Time interval	Time interval
has beginning	2020-01-14	2020-02-19	2020-02-19	2020-01-21
has end	2020-03-08	2020-02-29	2020-02-29	2020-02-26
Basic reproduction number	Basic reproduction number estimate value specification	Basic reproduction number estimate value specification	Basic reproduction number estimate value specification	Basic reproduction number estimate value specification
Has value	3.1	3.6	3.58	1.27
Confidence interval (95%)	Confidence interval (95%)	Confidence interval (95%)	Confidence interval (95%)	Confidence interval (95%)
Lower confidence limit	2.9	3.4	1.29	1.19
Upper confidence limit	3.2	4.2	8.46	1.36
Method*		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

ORKG‘s Objectives

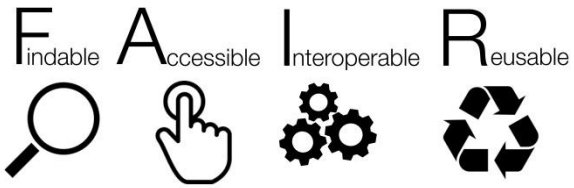


Provide overview over the state-of-the-art for specific research problems

Foster collaboration



Focus on scientific content rather than document



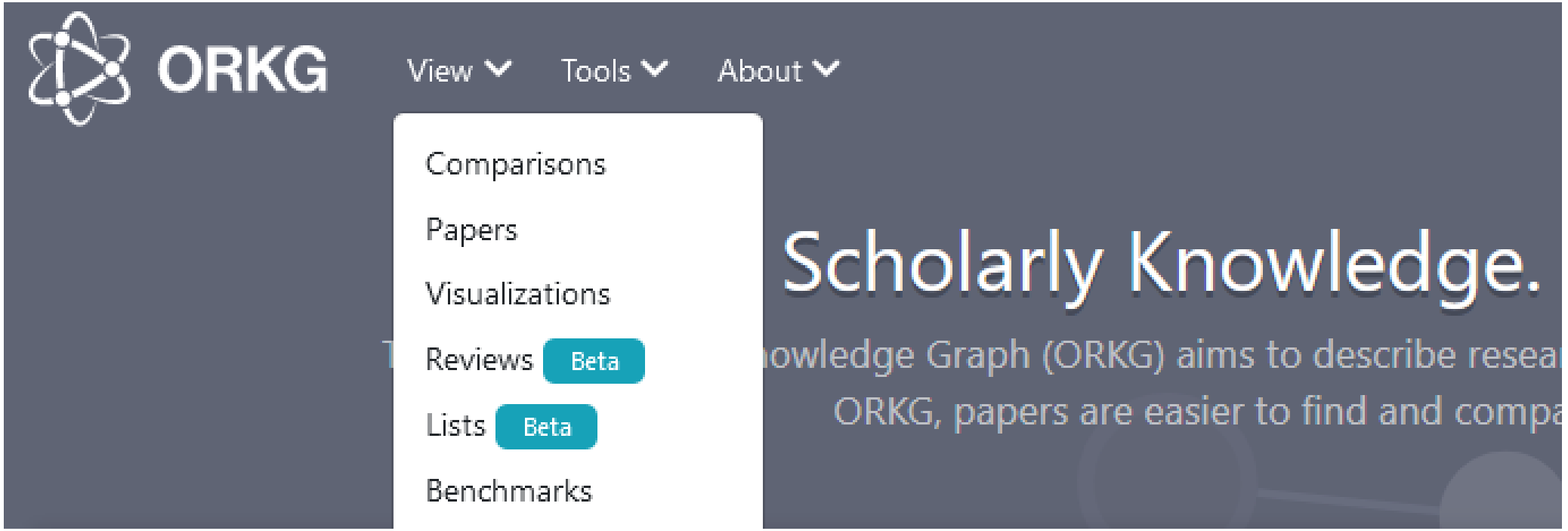
Make research FAIR

Tackle interdisciplinary challenges such as climate change research, disease prevention, etc.

Finally bring scholarly communication to the 21st century!

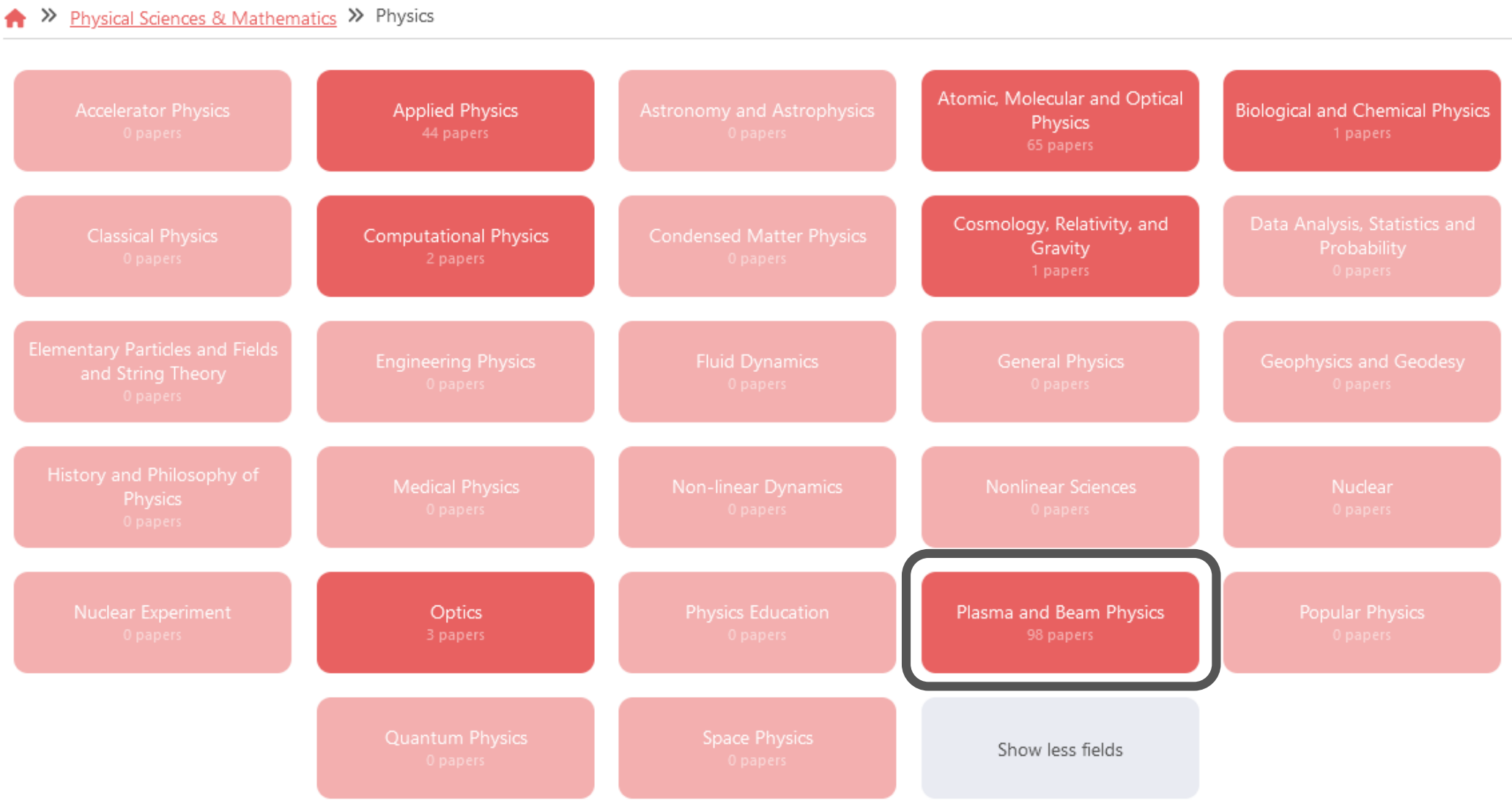
ORKG: Lighthouse in the Publication Flood

ORKG‘s Current Features



We are constantly improving and developing new features.
Also interesting: Describe datasets & software

So, what about Plasma Physics?



So, what about Plasma Physics?

- Already some content: 98 described papers

- Last year: Curation Grants
  - 1 grantee from laser & plasma physics
  - 14 Comparisons, 1 Review
  - This year: Same grantee → Content will increase

- Could be more…

Who creates ORKG content?

The answer is YOU!

We follow a **crowd based approach** for the curation process

Everyone can create, edit, add, complement, reuse, etc.

Domain experts

Field specific knowledge

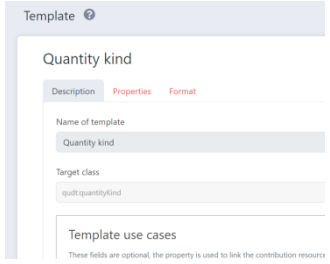
Data Curators

Knowledge on data modelling

Observatories: Taking the Lead in Content Curation

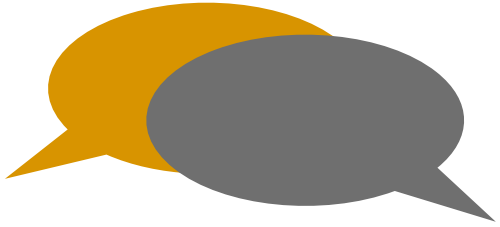
Organize research in your field

Ensure high quality standard



Create templates and simplify using ORKG for beginners

Promote ORKG



Stay in contact with development team: Issues & Requests will be prioritized

ORKG Observatory

How to get an Observatory started

- **Collect typical research questions from your field: What type of questions must the system be able to answer to be beneficial for your discipline?**
  - E.g. „By how much do the theoretical values of transition energies deviate from measurements?“
- **What is needed to answer these questions? And where can it be found in literature?**
  - E.g. system characteristics (Element, ionization state, transition, which experiment etc.), results, numerical values

Data Curators

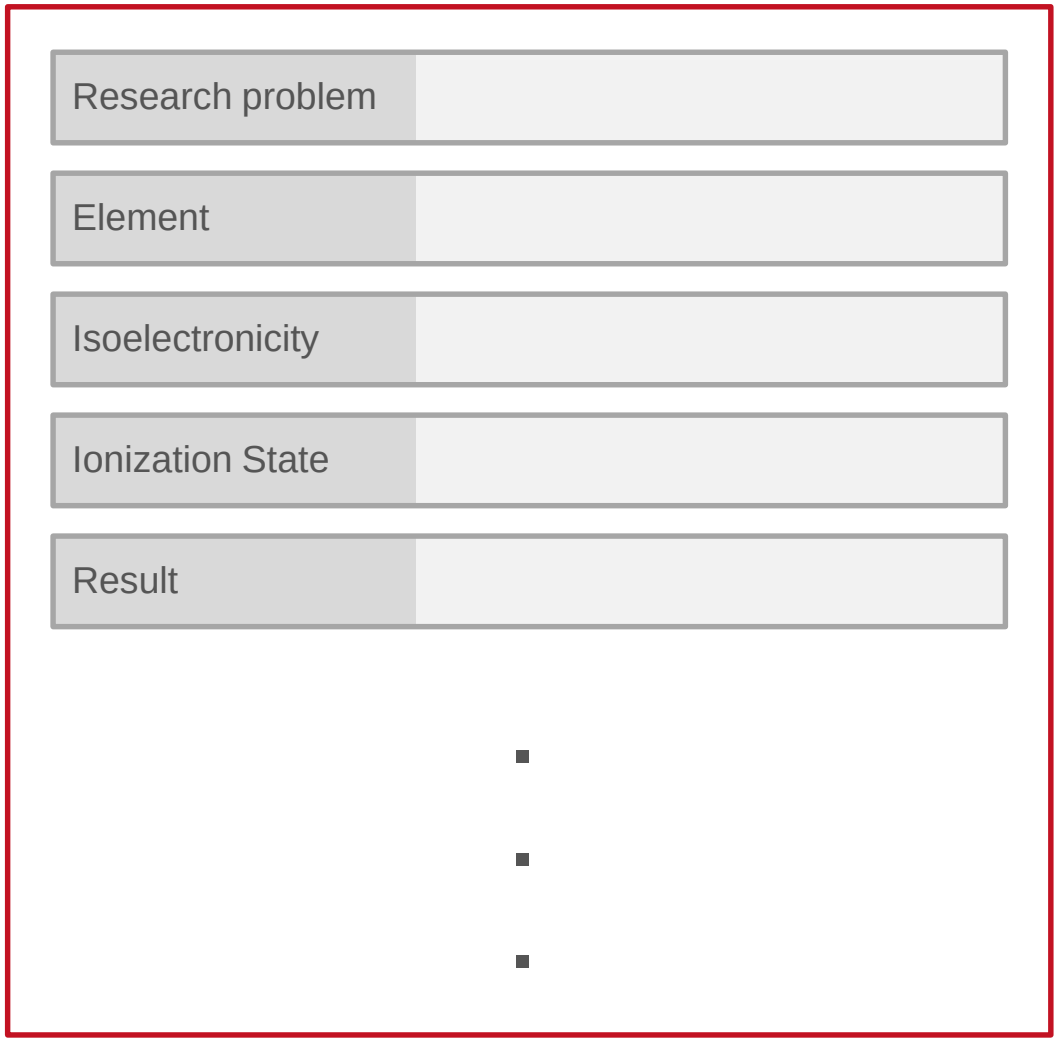
- Create a data model
- Template

Templates

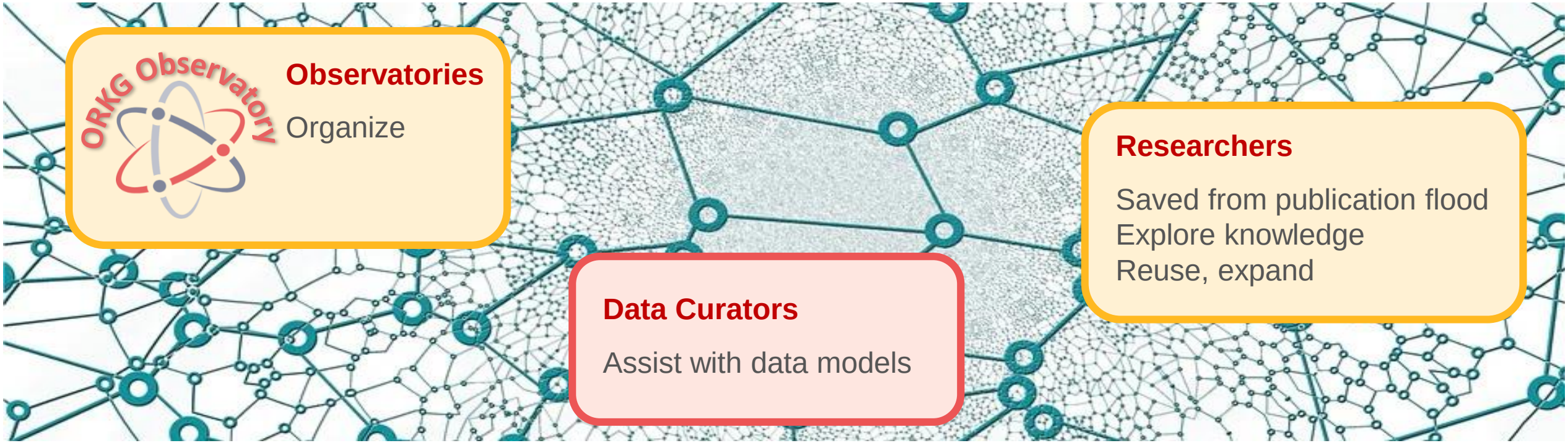
- Describe papers, datasets, software
- Forms with predefined properties
- Easy to fill out
- Increase interoperability

Researchers

- Enter data: focus on content, no worries about data model
- Can reuse & expand



Create a Community Knowledge Graph for your discipline!



Observatories

Organize

Researchers

Saved from publication flood
Explore knowledge
Reuse, expand

Data Curators

Assist with data models

Profit from organized knowledge & machine assistance
Increase FAIRness

Team

Lead

Prof. Dr. Sören Auer

Co-Lead

Dr. Markus Stocker

Post-Docs

Dr. Jennifer D‘Souza

Dr. Lars Vogt

Dr. Oliver Karras

Dr. Vinodh Ilangovan

PhD Students

Allard Oelen

Golsa Heidari

Hassan Hussein

Muhammad Haris

Yaser Jaradeh

Salomon Kabenamualu

Developers

Kheir Eddine Farfar

Manuel Prinz

Omar Arab Oghli

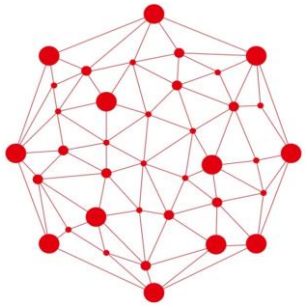
Community Building

Dr. Anna-Lena Lorenz

Summary

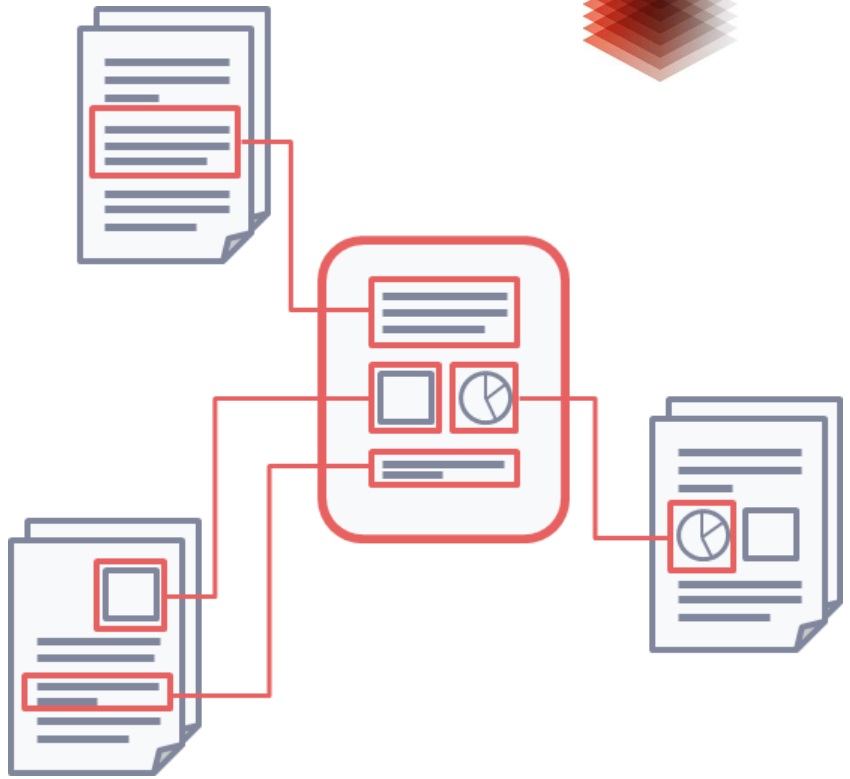
Rethink scholarly communication

Machine-actionable knowledge representation for FAIR research



Crowd-based approach

Get involved with an Observatory and build a community knowledge graph for plasma physics



Learn more: orkg.org
Contact us: info@orkg.org
Follow us: @orkg_org

LEIBNIZ INFORMATION CENTRE
FOR SCIENCE AND TECHNOLOGY
UNIVERSITY LIBRARY

TIB

MORE INFORMATION

www.orkg.org

Follow us: @orkg_org

Contact
info@orkg.org

Leibniz Association