

Data librarianship

–

Preparing librarians for their future tasks


Prof. Dr. Konrad U. Förstner


ZB MED – Information Centre for Life Sciences, Cologne, Germany &
TH Köln – University of Applied Sciences, Cologne, Germany

25th July, 2019, LibPMC, Aberystwyth, Wales




<https://doi.org/10.5281/zenodo.3351357>

- 
- I am here to bring a new perspective not old wisdom
 - I am a scientist and not a librarian
 - My very opinionated views / wishes
 - Mostly focussing on research libraries

- 
- I am here to bring a new perspective not old wisdom
 - I am a scientist and not a librarian
 - My very opinionated views / wishes
 - Mostly focussing on research libraries

And, yes, I might be wrong! Let's have a discussion.

#63 BE HONEST.
EVEN IF IT HURTS

A close-up photograph of a blue-painted metal surface, possibly a door or a cabinet. The paint is chipped and peeling in several places, revealing a rusty metal underneath. A brass padlock is attached to the surface, and a keyhole is visible. The padlock has the word "STERLING" and a logo of an eagle with spread wings. A key is partially inserted into the keyhole.

The good news:

Open Access is on its way to become the default publishing mode.



Plan S

Making full and immediate Open Access a reality

Plan S

Making full and
immediate Open
Access a reality

Why Plan S

Principles and Implementation

Who's involved

About


Contact

About Plan S

Plan S is an initiative for Open Access publishing that was launched in September 2018. The plan is supported by cOAlition S, an international consortium of research funders. Plan S requires that, from 2021, scientific publications that result from research funded by public grants must be published in compliant Open Access journals or platforms.

<https://www.coalition-s.org/>



A close-up photograph of a blue-painted metal surface, possibly a door or a cabinet. The paint is chipped and peeling in several places, revealing a rusty metal underneath. A brass padlock is attached to the surface, and a keyhole is visible. The padlock has the word "STERLING" and a logo of a bird with spread wings. A key is partially inserted into the keyhole.

Even more good news:

Basically all scientific publications are already now as full-text online available and easy to get at a single location.



the first website in the world to provide mass & public access to research papers




SCI-HUB

...to remove all barriers in the way of science

enter URL, PMID / DOI or search string




<https://sci-hub.tw>


A close-up photograph of a blue-painted metal surface, possibly a door or a cabinet. The paint is chipped and peeling in several places, revealing a rusty, brown metal underneath. A brass padlock is attached to the surface, and a key is inserted into the keyhole. The padlock has the word "STERLING" and a logo of an eagle with spread wings embossed on it. The background is a solid blue color with some white speckles and scratches.

Even more good news:

We have excellent discovery services for scientific articles ...



 My profile

 My library

SIGN IN


Google Scholar



Articles Case law

New! 2019 Scholar Metrics Released

Stand on the shoulders of giants



Even more good news:
We have excellent discovery services for books ...




Sign in



Search the world's most comprehensive index of full-text books.

[My library](#)



Still not done - even more good news:
We have excellent discovery services for datasets!



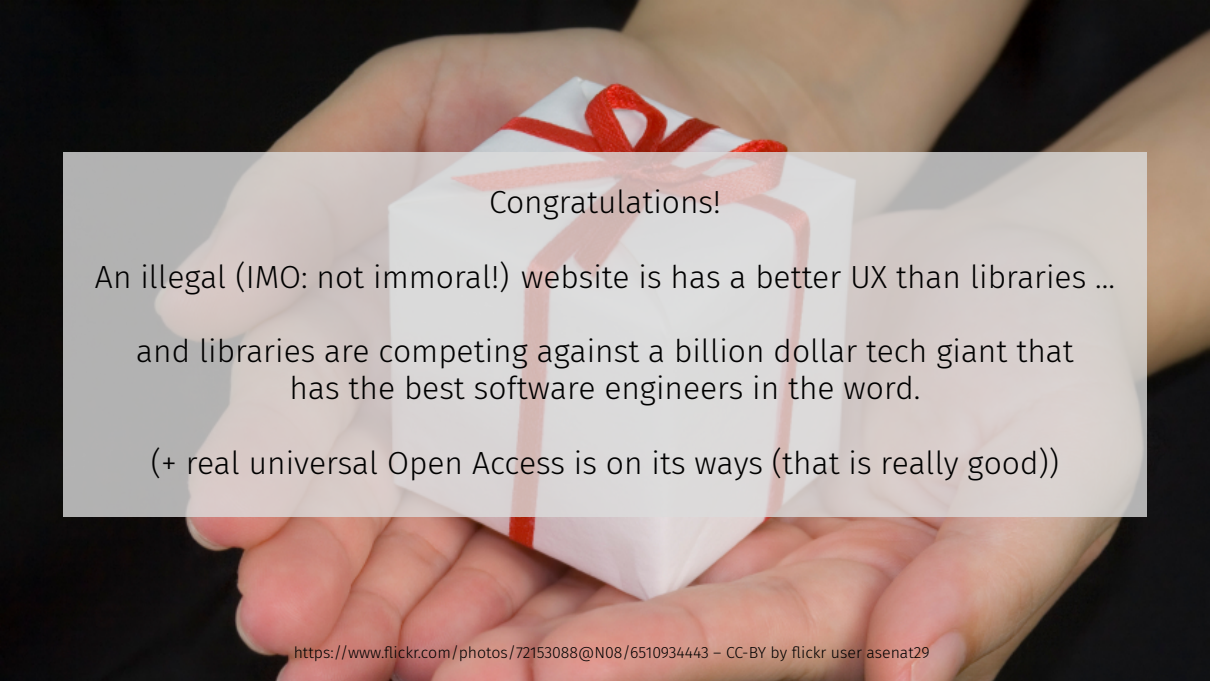
Google Dataset Search Beta

Search for Datasets



Try [boston education data](#) or [weather site:noaa.gov](#)


[Learn more](#) about including your datasets in Dataset Search.

A pair of hands is shown from the front, gently holding a small, white, cube-shaped gift box. The box is wrapped in white paper and secured with a vibrant red ribbon that forms a bow on top. The hands are positioned as if presenting the gift. The background is dark, making the hands and the gift stand out.


Congratulations!

An illegal (IMO: not immoral!) website is has a better UX than libraries ...
and libraries are competing against a billion dollar tech giant that
has the best software engineers in the word.
(+ real universal Open Access is on its ways (that is really good))



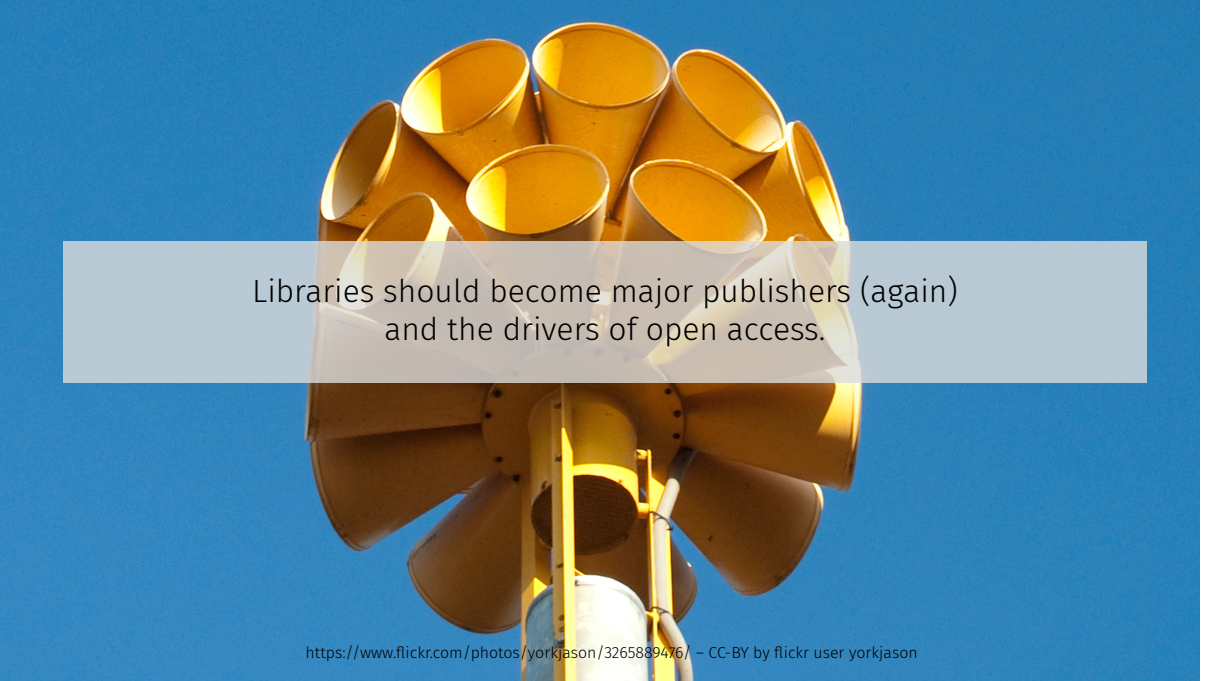


Dear librarians,
don't take the existence of libraries for granted.

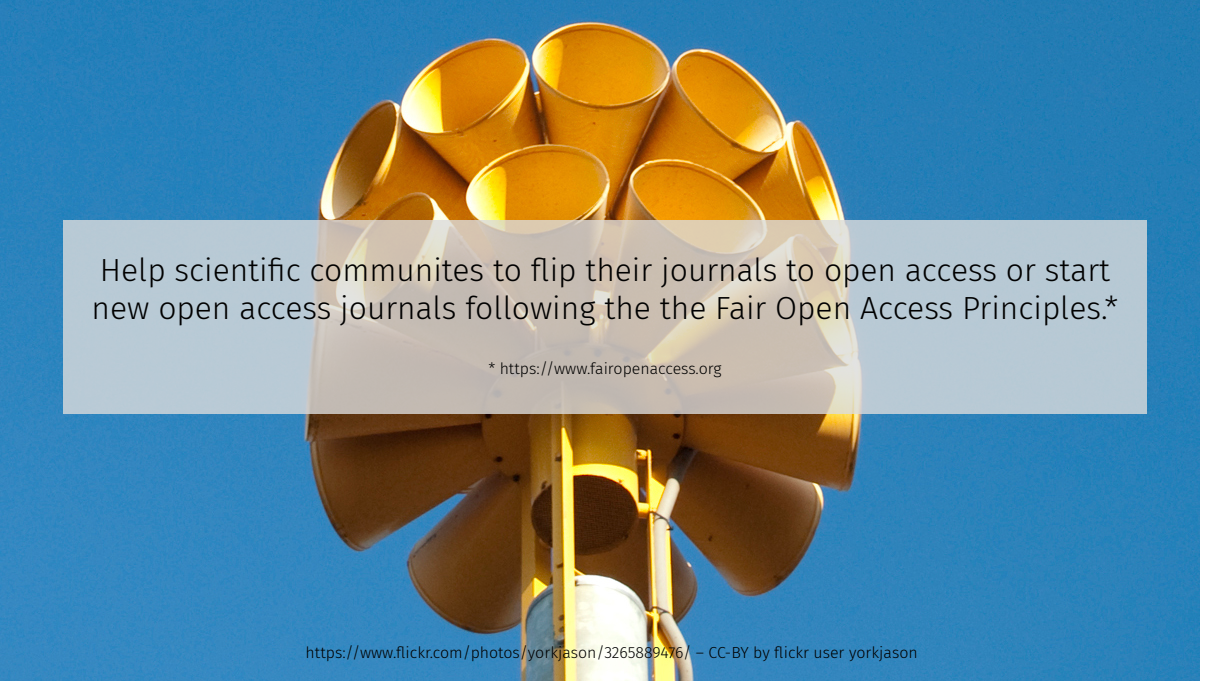
An abstract painting featuring concentric circles. The outermost ring is red, followed by a blue ring, then a yellow ring, and finally a central blue circle. The texture of the paint is visible, giving it a rough, layered appearance.

The mission remains the same:

Research libraries must preserve research output
and provide access to it.



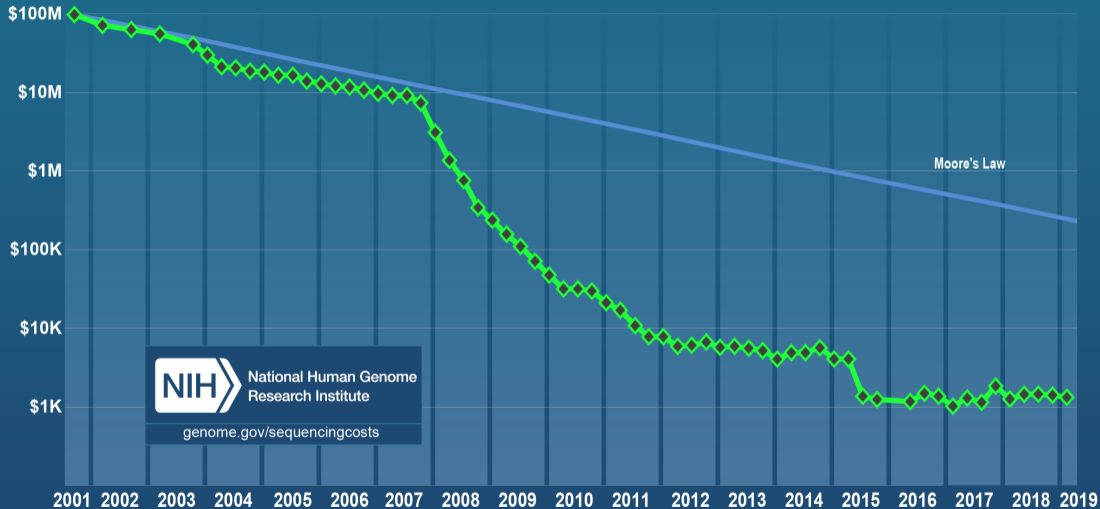
Libraries should become major publishers (again)
and the drivers of open access.

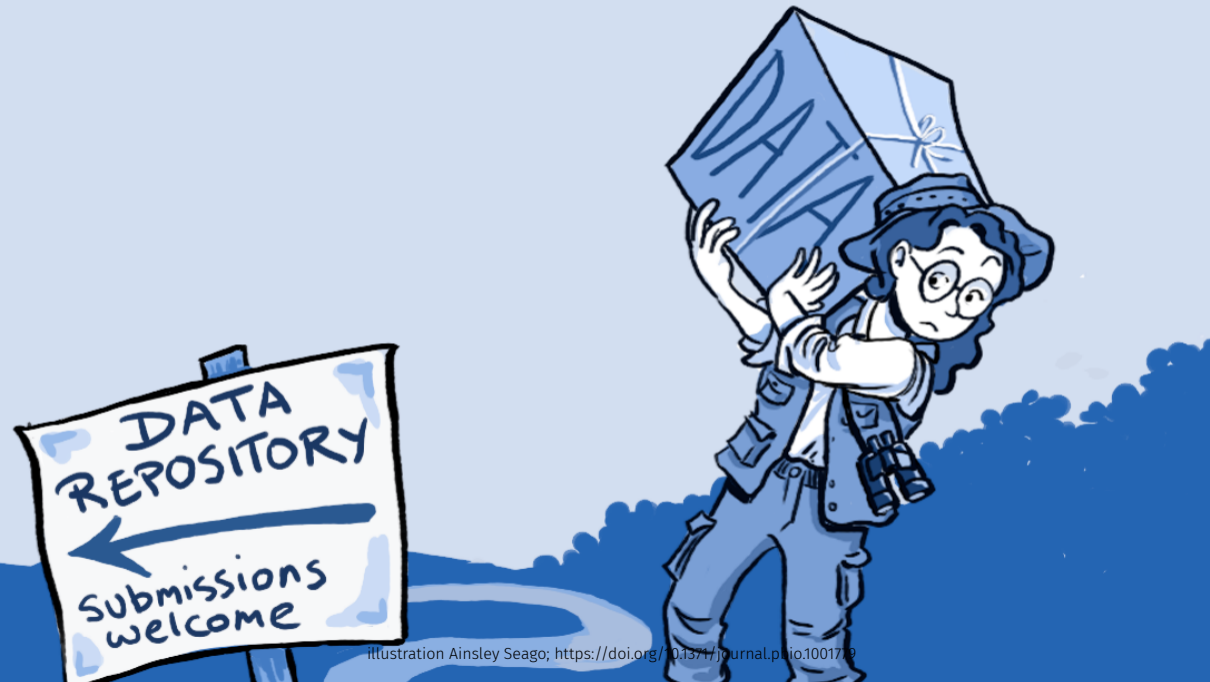


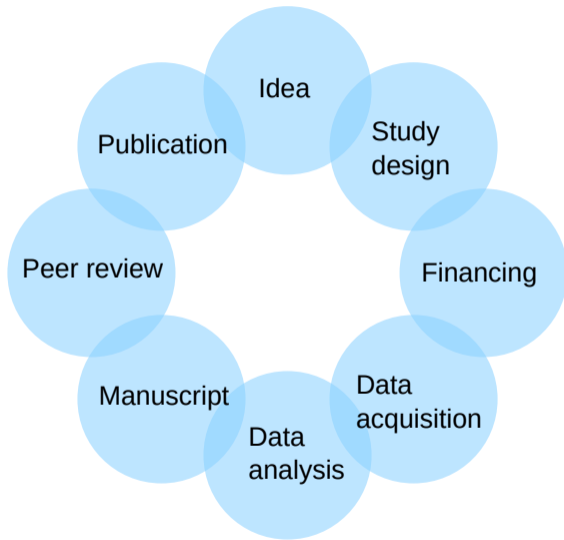
Help scientific communities to flip their journals to open access or start new open access journals following the the Fair Open Access Principles.*

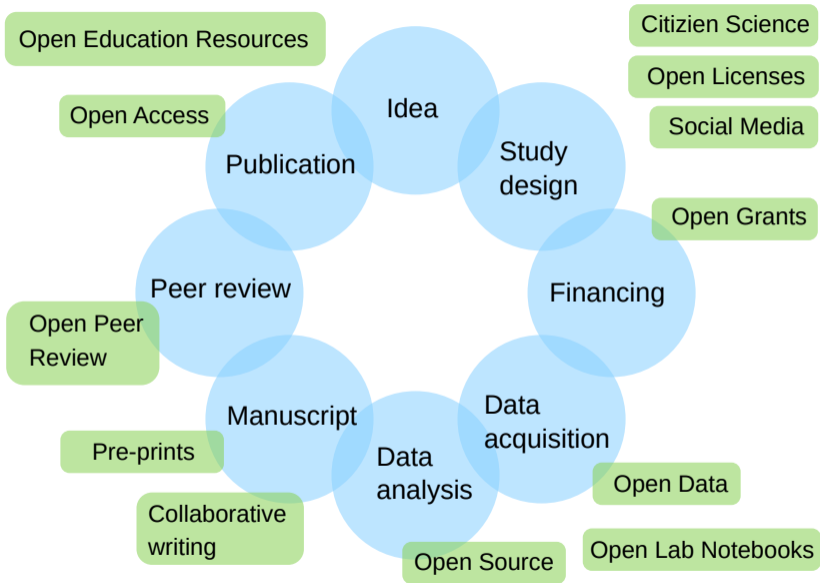
* <https://www.fairoopenaccess.org>

Cost per Genome











Not trivial!

Numerous different file formats, licences, privacy issues, legal requirements, provenance and interaction between different items ...



Not trivial!

Numerous different file formats, licences, privacy issues, legal requirements, provenance and interaction between different items ...

Librarians can become consultant for all these topics.


Currently my favorite example ...

Science \Leftrightarrow Technology

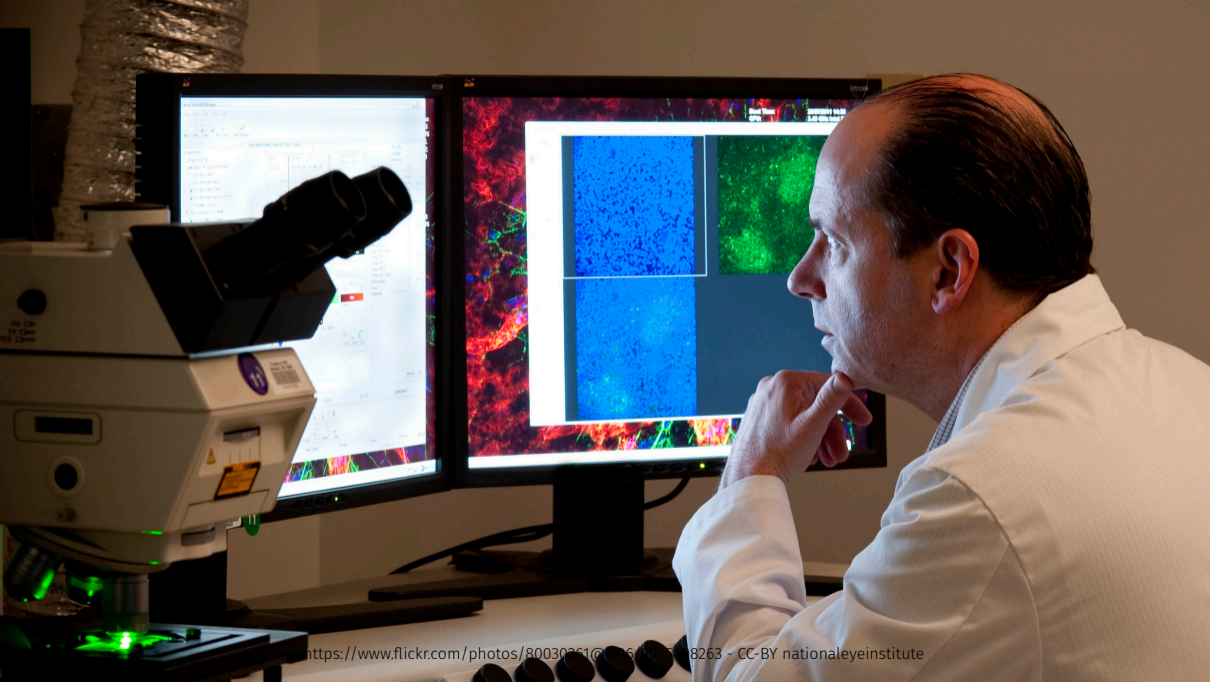




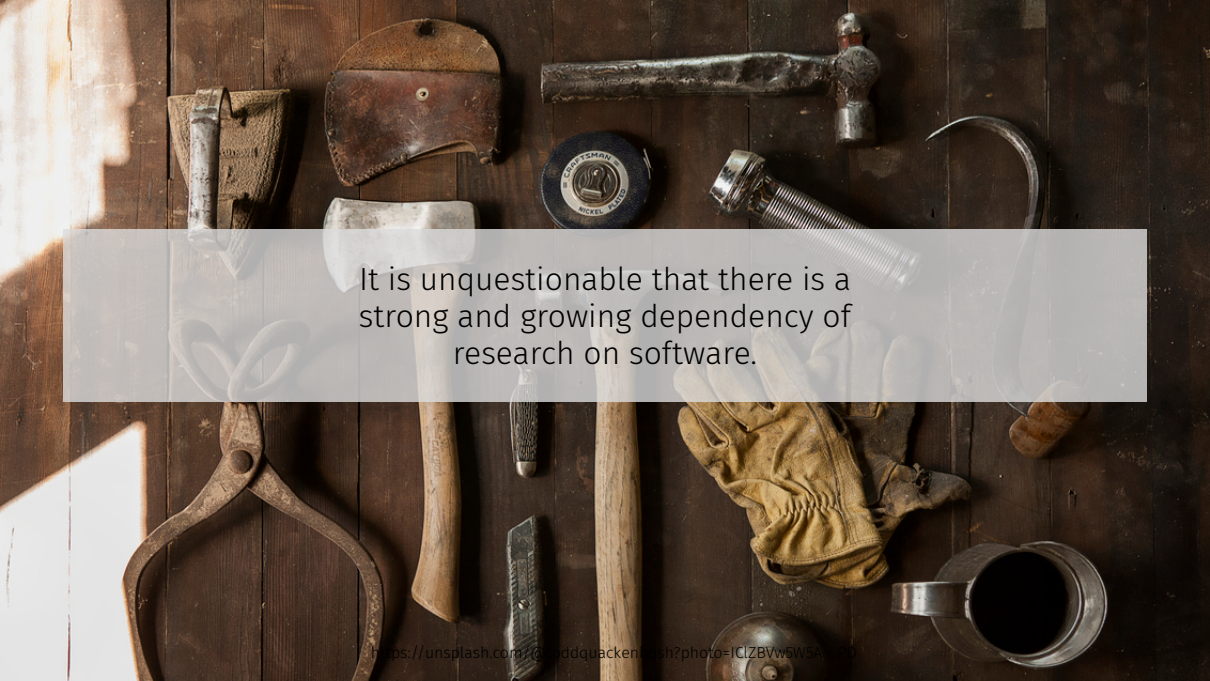


A top-down photograph of various vintage tools and objects arranged on a dark wooden surface. The tools include a large axe with a wooden handle, a hammer with a metal head and wooden handle, a curved saw, a pair of large metal tongs, a pair of worn leather gloves, a utility knife, a folding knife, a circular object with a logo, and a metal cup. The scene is lit from the left, creating strong shadows.


Software
- an ubiquitous research tool






A top-down photograph of various tools and objects arranged on a dark wooden surface. The items include a hammer, a tape measure with a 'CRAFTSMAN' logo, a pair of work gloves, a mug of coffee, a pair of pliers, a utility knife, a saw, and several other hand tools. The scene is lit from the left, creating strong shadows.

It is unquestionable that there is a strong and growing dependency of research on software.



Software is also a result of
scientific work.

Quality, accessibility, citability, etc.
have to be ensured.

A top-down photograph of various vintage tools and objects arranged on a dark, weathered wooden surface. The tools include a large axe with a curved blade, a hammer with a wooden handle, a pair of large pliers, a hand saw with a curved blade, a utility knife, a tape measure with a blue case labeled 'CRAFTSMAN', a pair of worn tan work gloves, and a metal mug. The scene is lit from the left, creating strong shadows and highlights on the wood and tools.

The importance of software for research
is widely ignored.




Recommendations on the


Development, Use and Provision of Research Software

Research Software Working Group
in the Priority Initiative Digital Information
of the Alliance of German Science Organisations


<https://doi.org/10.5281/zenodo.1172988>



Ultimate access creates new problems:
Increased effort to find and process information.




Train skills to find research items beyond articles efficiently.




And yes - compete with Google and build better, open discovery services.

Together we can be strong.

See Peter Kraker's #DontLeaveltToGoogle initiative.

A red Swiss Army knife is shown with several tools extended, including a large blade, a pair of pliers, a corkscrew, and a small saw. The knife is resting on a wooden surface. A semi-transparent text box is overlaid on the center of the image.

Libraries must train the skills and build tools to find, connect and integrate, literature, information and data to create new knowledge.

A photograph of the Earth from space, showing the Americas and the Atlantic Ocean. The Earth is curved, with the top part showing North America and the bottom part showing South America and the Atlantic Ocean. The text is overlaid on a semi-transparent grey rectangle in the center.

Think big, think global!

Work on the knowledge commons.

Imagine a world in which every single human being can freely share in the sum of all knowledge.



[← About](#)

Wikimedia vision

Imagine a world in which every single human being can freely share in the sum of all knowledge. That's our commitment.



[← About](#)

Wikimedia vision

This should be the vision of every library!

Imagine a world in which every single human being can freely share in the sum of all knowledge. That's our commitment.

The Wikipedia Library/1Lib1Ref

< The Wikipedia Library

[Translate this page](#); This page contains [changes](#) which are not marked for translation.

[1] LIBRARIAN REFERENCE

1lib1ref.org



#1Lib1Ref runs every January 15th to February 5th and every May 15th to June 5th. Join the Campaign! [Translate the banner](#)



Home



Participate



Host a coffee hour


















Resources



Connect with us

Other languages:

Bahasa Indonesia  • Bahasa Melayu  • Deutsch  • English  • Esperanto  • Nederlands  • Tiếng Việt  • català  • dansk  • español  • euskara  • français  • galego  • italiano  • latviešu  • magyar  • norsk bokmål • polski • portuguese • português do Brasil • română • suomi • svenska • čeština • Ελληνικά • македонски • русский • українська • עברית • اروو • العربية • سنڌي • فارسی • हिन्दी • বাংলা • தமிழ் • മലയാളം • မြန်မာဘာသာ • 中文 • 日本語

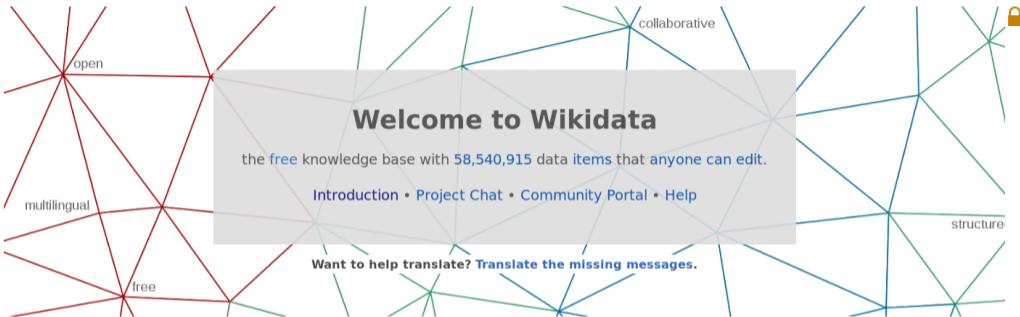
- Main page
- Wikimedia News
- Translations
- Recent changes
- Random page
- Help
- Babel

- Community
- Wikimedia Resource Center
- Wikimedia Forum
- Mailing lists
- Requests
- Babylon
- Reports
- Research
- Planet Wikimedia

- Beyond the Web
- Meet Wikimedians
- Events
- Movement affiliates
- Donate

- Print/export
- Create a book
- Download as PDF
- Printable version

- Tools
- What links here



Welcome to Wikidata

the [free](#) knowledge base with **58,540,915** data items that anyone can edit.

[Introduction](#) • [Project Chat](#) • [Community Portal](#) • [Help](#)

Want to help translate? [Translate the missing messages.](#)

- Main page
- Community portal
- Project chat
- Create a new item
- Create a new Lexeme
- Recent changes
- Random item
- Query Service
- Nearby
- Help
- Donate

In other projects

- Wikimedia Commons
- MediaWiki
- Meta-Wiki
- Wikispecies
- Wikibooks
- Wikimania
- Wikinews
- Wikipedia
- Wikiquote
- Wikisource
- Wikiversity
- Wikivoyage
- Wiktionary
- Print/export
- Create a book
- Download as PDF



Welcome!

Wikidata is a free and open knowledge base that can be read and edited by both humans and machines.

Wikidata acts as central storage for the **structured data** of its Wikimedia sister projects including Wikipedia, Wikivoyage, Wiktionary, Wikisource, and others.

Wikidata also provides support to many other sites and services beyond just Wikimedia projects! The content of Wikidata is [available under a free license](#), exported using standard formats, and can be interlinked to other open data sets on the linked data web.



Learn about data

New to the wonderful world of data? [Develop and improve your data literacy through content](#) designed to get you up to speed and feeling comfortable with the fundamentals in no time.



Page





WikiCite

[Home](#)

[Instructions](#)

[Media and events](#)

[Administration](#)

[Development](#)

[Discuss](#)

[Main page](#)

[Wikimedia News](#)

[Translations](#)

[Recent changes](#)

[Random page](#)

[Help](#)

[Babel](#)

Community

[Wikimedia Resource Center](#)

[Wikimedia Forum](#)

[Mailing lists](#)

[Requests](#)

[Babylon](#)

[Reports](#)

[Research](#)

[Planet Wikimedia](#)

Beyond the Web

[Meet Wikimedians](#)

[Events](#)

[Movement affiliates](#)

[Donate](#)

In other projects

[Wikimedia Commons](#)

[Wikidata](#)

[Print/export](#)

[Create a book](#)

{ } wikicite

WikiCite is a Wikimedia initiative to develop a database of open citations and linked bibliographic data to serve free knowledge.

The idea is curate the collection of citations in Wikidata, or perhaps in local instances of the Wikibase platform at particular institutions. To the extent that Wikidata has capacity and data is compatible, then information is in Wikidata. When the data seems too large, or not formatted for Wikidata, or when there is another constraint, anyone can host it in their own Wikibase instance and still make it compatible with Wikicite for possible later integration.

WikiCite is among the most popular projects in Wikidata. This popularity is a legacy of an old desire in the Wikipedia community to better manage citations across Wikimedia projects and languages. While WikiCite cannot quickly fulfill all dreams and ambitions, some community wishes for WikiCite is that it leads to better research discovery for Wikipedia and for all research in general. Potential applications include ease of discovering publications on a given topic, profiling of authors and institutions, and visualizing knowledge sources in new ways. WikiCite promises to open citation data, which many people are surprised to hear is a closed dataset in many fundamental ways.

Wikidata is the database which anyone can edit within the Wikipedia and Wikimedia platform. Wikicite is the community and project within Wikidata which focuses on source metadata.

Contents [\[hide\]](#)

[1 About](#)

[2 Events](#)

[3 News](#)

<https://meta.wikimedia.org/wiki/WikiCite>





LibPMC

International Conference on Performance Measurement in Libraries

23rd July – 25th July 2019. Aberystwyth, Wales

REGISTRATION & TRAVEL ▼

PROGRAMME ▼

ABOUT ▼

SPONSORS & AWARDS ▼



<https://libraryperformance.org/>



LibPMC

International Conference on Performance Measurement in Libraries

Question for you: How do we measure all that?

23rd July – 25th July 2019. Aberystwyth, Wales

REGISTRATION & TRAVEL ▼

PROGRAMME ▼

ABOUT ▼

SPONSORS & AWARDS ▼



<https://libraryperformance.org/>

We are facing a huge skill gap.

Caution
Gap



Shifting to Data Savvy: The Future of Data Science In Libraries

Burton, Matt and Lyon, Liz and Erdmann, Chris and Tijerina, Bonnie (2018)



Library Carpentry

Software and data skills for people working in library- and information-related roles



What we do

Library Carpentry develops lessons and teaches workshops for and with people working in library- and information-related roles. Our goal is to create an on-ramp to empower this community to use software and data in their own work as well as be advocates for and train others in efficient, effective and reproducible data and software practices. Our workshops are based on **our lessons**. Workshop hosts, Instructors, and learners must be prepared to follow The Carpentries **Code of Conduct**.

[More >](#)



Who we are

We are a diverse, global community of **volunteers**. Our community includes Instructors, helpers, Trainers, **Maintainers**, **Mentors**, community champions, **member organisations**, supporters, workshop organisers, and **staff**. Library Carpentry is guided by a **Advisory Group** and a **Curriculum Advisory Committee**. Our **audience** are primarily people working in library- and information-related roles.

[More >](#)

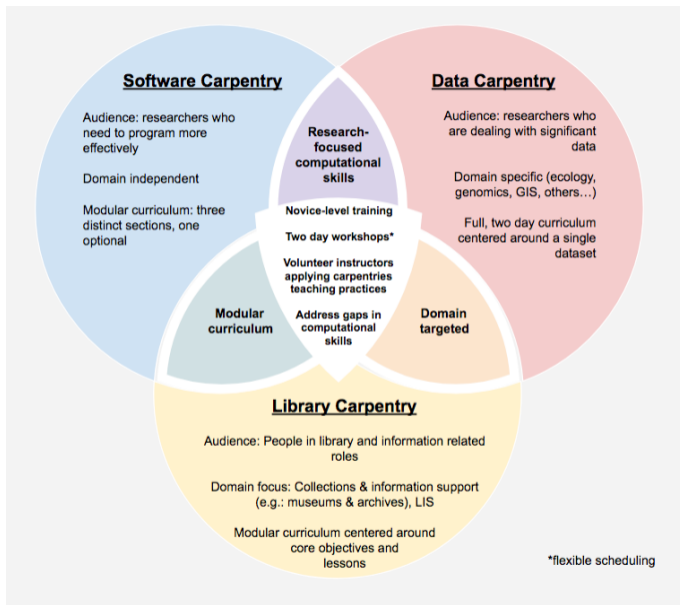


Get involved

See all the **ways you can engage and get involved** with Library Carpentry. Follow us on **Twitter**.

[More >](#)

<https://librarycarpentry.org/>

















Our Core Curriculum

Our Core Curriculum consists of the lessons in the table below. These have been taught many times, and have been further refined after instructor and learner feedback. For more information regarding core lessons and workshops, see [Our Workshops](#).

The lessons introduce terms, phrases, and concepts in software development and data science, how to best work with data structures, and use regular expressions in searches. We introduce the Unix-style command line interface, and teach basic shell navigation, as well as the use of `loops` and pipes for linking shell commands. We also introduce `grep` for searching and subsetting data across files. Exercises cover the counting and mining of data. In addition, we cover working with OpenRefine to transform and clean data, and the benefits of working collaboratively via Git/GitHub and using version control to track your work.





















Lessons

Lesson	Site	Repository	Reference	Instructor Notes	Status	Maintainer(s)
Introduction to Data					Stable	Carmi Cronje, Paul R. Pival, Shari Laster*, Anton Angelo (Past Maintainer: James Baker)
The UNIX Shell					Stable	Danielle Kane*, Nilani Ganeshwaran, John Wright, Anna Oates, Belinda Weaver, Tim Dennis
OpenRefine					Stable	Owen Stephens, Juliane Schneider, Paul R. Pival, Kristin Lee, Erin Carrillo* (Past Maintainer: Carmi Cronje)
Introduction to Git					Beta	Silvia di Giorgio, Chris Erdmann, Thea Atwood, Eric Lopatin, Drew Heles, Eva Seidlmayer (Past Maintainers: Katrin Leinweber, Belinda Weaver, Jez Cope)

Extended Curriculum

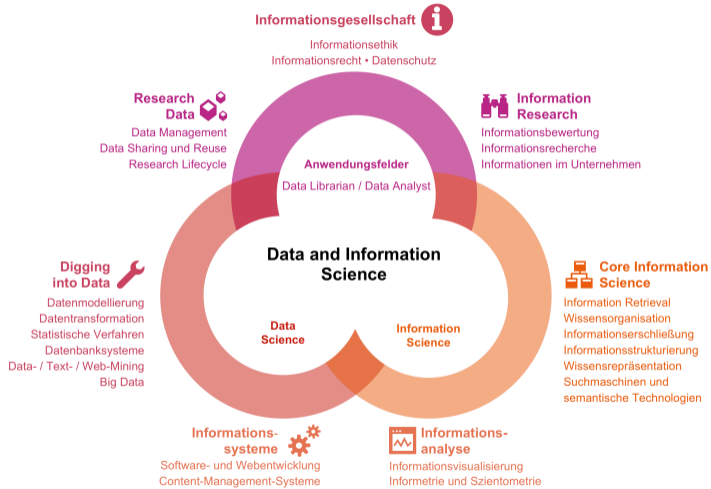
The following Library Carpentry lessons can also be taught in addition to our core curriculum. These lessons have been taught infrequently and still need further work. We would value any feedback on these lessons.


Lessons

Lesson	Site	Repository	Reference	Instructor Notes	Status	Maintainer(s)
Tidy Data					Beta	Tim Dennis, Thea Atwood, Sherry Lake*, Erika Mias (Past Maintainer: Jez Cope)
SQL					Beta	Jordan Pedersen* , Kristin Lee, Chris Erdmann, Lise Doucette(Past Maintainers: Elaine Wong, Janice Chan)
Webscraping					Alpha	Joshua Dull*, Thomas Guignard (Past Maintainer: Belinda Weaver)
Introduction to Python					Alpha	Konrad Foerstner, Drew Heles, Elizabeth Wickes, Laura Wrubel* (Past Maintainers: Carlos Martinez, Richard Vankoningsveld)
Introduction to Data for Archivists					Alpha	Jeanine Finn, Katherine Koziar*, and Scott Peterson (Past Maintainers: Jenny Bunn, Noah Geraci, and James Baker)

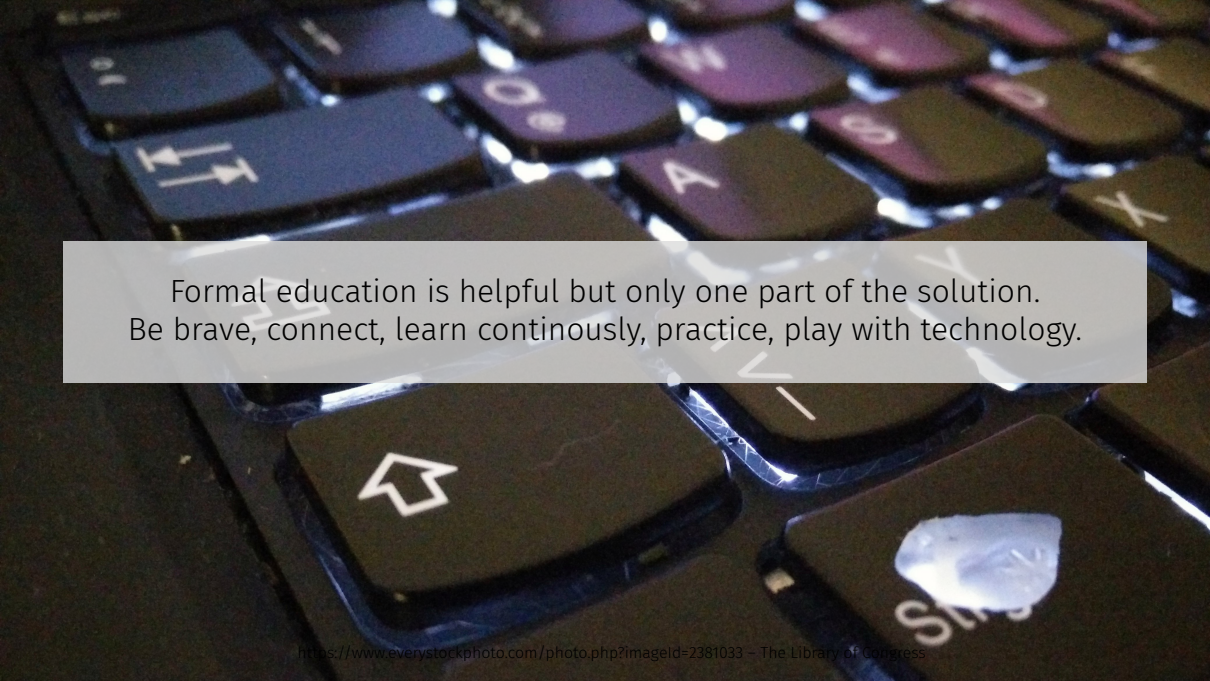
* Indicates Lead Maintainer

New Bachelor at TH Köln: *Data and Information Science*



A welder wearing a dark, worn leather protective suit and a helmet with the number '253' on it is working on a large, cylindrical metal pipe. The welder is holding a welding torch and a rod, and a bright light is visible from the welding process. The background is dark and industrial.

Currently under construction at the ZBIW of the TH Köln:
Certificate course *Data Librarian*



Formal education is helpful but only one part of the solution.
Be brave, connect, learn continuously, practice, play with technology.

./ HackyHour Cologne

Code|Tools|Discussion|Community

Download as .zip

Download as .tar.gz

 View on GitHub

```
22 open("HackyHour", "r")
23 for everybody in library_and_life_sciences:
24     if everybody.courious:
25         print("I'm coming!")
26     print("HackyHour :-) ")
```

About

Come to talk about code, tools and science, in a social environment. If you have a problem in one of those areas: bring it along and get help. If you want to know how computational tools and code can improve your science or you just want to drink a coffee: join us. The HackyHour is open to all intrested librarians and scientists of the Cologne area and beyond.

When and Where

Our meeting takes place on Thursdays from 2-3pm in the Meeting Room at [ZB_MED](#). In order to secure a peer to peer interaction frontal teaching longer than 15min is dispensed for HackyHour. Discussions are preferred to be held in English or English and German enable international researchers to join the meeting.

You can come without registration but feel free to join the discussion in the linked pad (see table below) if you have any questions beforehand.

Topics

In addition to the open discussions we also aim to talk about specific topics at each HackyHour. However prior knowledge or interest in these topics is not a prerequisite. Online discussions for these topics are accessible via the following links.

<https://hackyhour.github.io/Cologne/>

What are your questions?

konrad.foerstner.org / @konradfoerstner

zbmed.de / @ZB_MED

th-koeln.de / @th_koeln



Technology
Arts Sciences
TH Köln