

# Searching for Research Data

A workshop for researchers

Sacha Zurcher (sazu@kb.dk) – Special Advisor Research Data  
Roskilde University Library

Nordic Love Data Week

February 11, 2026



**DET KGL.  
BIBLIOTEK**  
Royal Danish Library

# Agenda

- Who are we/you
- Searching for research data – why this workshop
- Workshop structure and learning objectives
- Where to search – five strategies
- How to assess data
- Enhancing reuse – Open Educational Resources (OER)
- Future developments
- Q & A



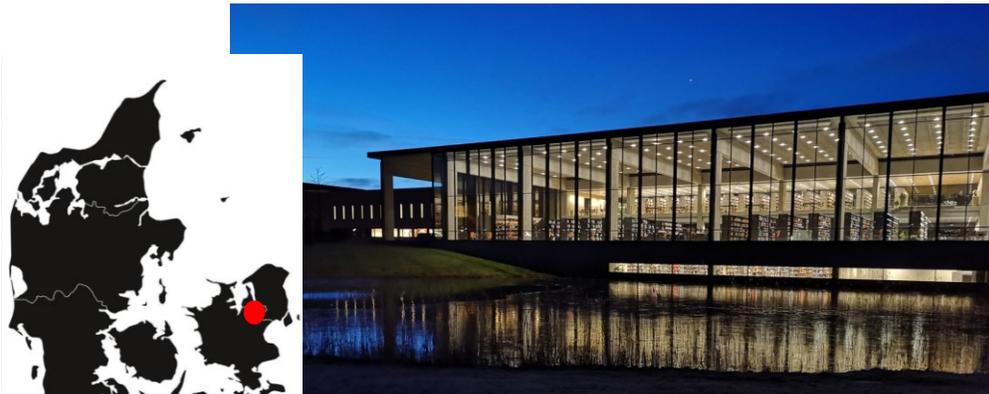
# FAIR data network – Royal Danish Library



[Aarhus University Library](#)



[Copenhagen University Library](#)



[Roskilde University Library](#)



# Who are you?

What is your main role in data management?



# Searching for research data – new service

Why a workshop for researchers "Searching for research data"

- Encourages reuse of existing data - as aimed for in the FAIR principles and Open Science.
- Horizon Europe funding requires that beneficiaries check whether it is possible to reuse existing data.
- Supplements information literacy – now in relation to research data.
- Fills a "gap" in educational activities on FAIR data.
- May result in more FAIR datasets.



# Workshop – Searching for research data

- Duration: 4 hours
- Target group: Researchers (including Ph.D. students)
- Structure: Combination of traditional classroom teaching, hands-on exercises, a workshop and a concluding reflexive exercise. Modular structure to enhance different combinations depending on available time and/or prior knowledge of participants.
- Materials: PowerPoint presentation and three handouts:
- Logbook;
  - List of resources where to search;
  - Selection of pre-designed cases.



# Structure and content of the workshop

The PowerPoint presentation is divided into 6 main sections:

1. Introduction to research data, the FAIR principles and searching for data.
2. What to expect when searching for data and five data search strategies.
3. Introduction to good search practice, incl. assessment and documentation.
4. Presentation of the five data search strategies incl. short exercises.
5. Workshop.
6. Final discussion.



# Learning objectives

Upon completion of this (4 hours) workshop, the participant has acquired:

- Basic knowledge of research data – definitions, data types and formats, etc.
- Knowledge of different search strategies and methods for searching for data.
- Knowledge of how to assess research data.
- Knowledge of conditions for reuse (licenses).
- Skills to find and navigate relevant research data repositories.
- Skills to apply search engines and bibliographic databases in data search.
- Knowledge of how to document the search process and findings.



# No standardized search strategies

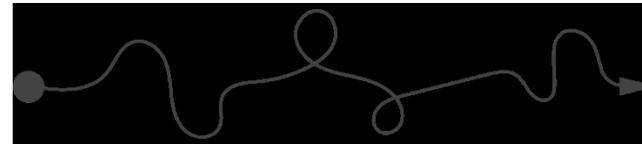
Relevant data



(Research) data is not collected and indexed in the same way as literature, and there are no standardized search strategies.

This means:

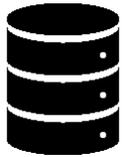
- Generally a **less structured** and systematic process compared to a literature search.
- An element of **serendipity**: You are likely to find interesting or relevant data by chance and often in places you do not expect.
- Expect **dead ends**.
- Be prepared to be **flexible** in your approach - use a combination of strategies.
- The search process often **extends over time** and usually involves a **social element** (knowledge sharing with colleagues, etc.).



Relevant literature



# Search strategies - (Research)data can be found in ...



**Strategy 1**  
Repositories



**Strategy 2**  
Bibliographic  
databases



**Strategy 3**  
Connected to  
monographs,  
articles, etc.



**Strategy 4**  
Outside academia  
(organisations,  
governments, etc.)



**Strategy 5**  
Search engines  
and AI



# Strategy 1 - Research data in a repository

Example: [Dataverse Norway](#)

## Replication Data for: Climate Nags: Affect and the Convergence of Global Risk in Online Networks

Version 1.1



Robinson, Jessica Yarin, 2023, "Replication Data for: Climate Nags: Affect and the Convergence of Global Risk in Online Networks", <https://doi.org/10.18710/G1CIXA>, DataverseNO, V1

Cite Dataset -

Learn about [Data Citation Standards](#).

Access Dataset -

Contact Owner

Share

Dataset Metrics ?

20 Downloads ?

Abstract/summary

Metadata, terms of use and versions of the dataset

Description ?

This data set contains the IDs of the 1,186,322 tweets used in "Climate Nags: Affect and the Convergence of Global Risk in Online Networks" (published in Continuum, 2023). The data was collected from Twitter's Streaming API using the DMI-TCAT during the first four months of the Coronavirus pandemic; 2020 U.S. presidential race; and the early stages of the 2022 Russia-Ukraine War. These collections were then filtered based on key words related to climate change (see README file for more details). (2023-07-12)

Subject ?

Arts and Humanities; Earth and Environmental Sciences; Social Sciences

Keyword ?

climate, Twitter, social media, global warming, sentiment, tweets

Related Publication ?

"Climate Nags: Affect and the Convergence of Global Risk in Online Networks," Continuum, 2023 [issn: 1469-3666](#)

License/Data Use Agreement



[CC0 1.0](#)

General information about the dataset

Readme file with information about the dataset

Files

Metadata

Terms

Versions

Search this dataset..



Filter by

File Type: All

Access: All

Sort -

Download

1 to 2 of 2 Files



00\_README.txt

Plain Text - 2.8 KB

Published Jul 12, 2023

9 Downloads

MD5: c69...b7e

Readme file with information about dataset



ClimateNags\_AllData\_afterRemovingBots.csv

Comma Separated Values - 45.3 MB

Files for download



DET KGL.  
BIBLIOTEK

Royal Danish Library

January 2026

# Strategy 2 - Research data in a bibliographic database

Example: [Web of Science \(Data citation Index\)](#)

Searching in Data Citation Index via Web of Science

Click on the title to read the metadata and view the source

Number of citations: click to get access to the list of citations

The screenshot displays the Web of Science search results page. At the top, the search bar contains the query "pregnancy AND 'physical activity' (Topic)". Below the search bar, it indicates "278 results from Data Citation Index for:". The results are listed in a table-like format. The first result is titled "Description of physical activity energy expenditure in MET-h/week during pregnancy, according to intensity and type of activity based on PPAQ." and has 1 citation. The second result is titled "Physical activity distribution during pregnancy from Actigraphs GT1M recording." and also has 1 citation. The interface includes various filters on the left, such as "Refine results" and "Filter by Marked List".

Item	Title	Citations
1	Description of physical activity energy expenditure in MET-h/week during pregnancy, according to intensity and type of activity based on PPAQ.	1 Citation
2	Physical activity distribution during pregnancy from Actigraphs GT1M recording.	1 Citation

# Strategy 3 - Indirect search for research data

Web of Science™ Smart Search **Advanced Search** Research Assistant

DOCUMENTS RESEARCHERS

Search in: Web of Science Core Collection Editions: All

FIELD SEARCH QUERY BUILDER CITED REFERENCES STRUCTURE

Title Example: water consum\*

+ Add row + Add date range

Clear Search



ORIGINAL ARTICLE | Open Access | CC BY-NC-ND

## Anticipated administrative burdens: How proximity to upcoming compulsory meetings affect welfare recipients' experiences of administrative burden

Martin Baekgaard, Jonas Krogh Madsen

First published: 21 April 2023 | <https://doi.org/10.1111/padm.12928> | VIEW METRICS

SECTIONS

PDF TOOLS SHARE

### OPEN RESEARCH BADGES



This article has earned an Open Data badge for making publicly available the digitally-shareable data necessary to reproduce the reported results. The data are available at <https://doi.org/10.5281/zenodo.7742956>.

### DATA AVAILABILITY STATEMENT

The data that support the findings of this study will be made openly available at Zenodo (<https://zenodo.org>) upon acceptance for publication.

<https://doi.org/10.1111/padm.12928> and <https://zenodo.org/records/7742956>

EN DA that target group members are likely to sychological costs when interacting with



# Strategy 4 - Data from outside the world of research

Example: [WorldBank Data](#)

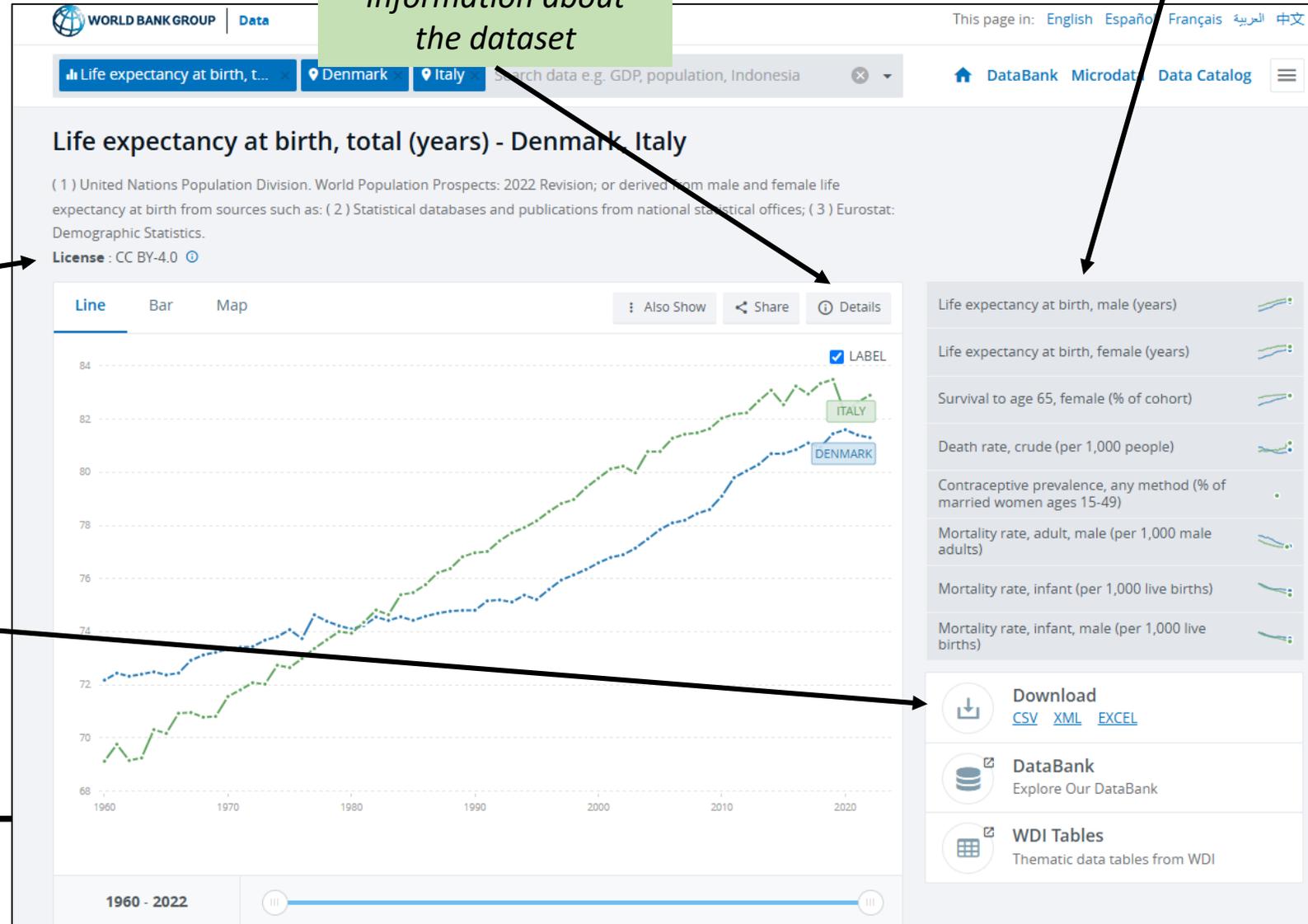
A database with free and open access to data on global development

License

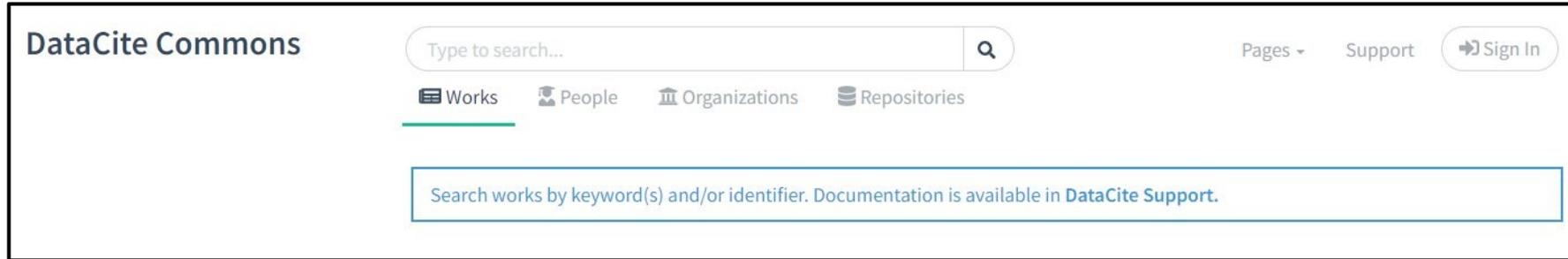
Files for download

Information about the dataset

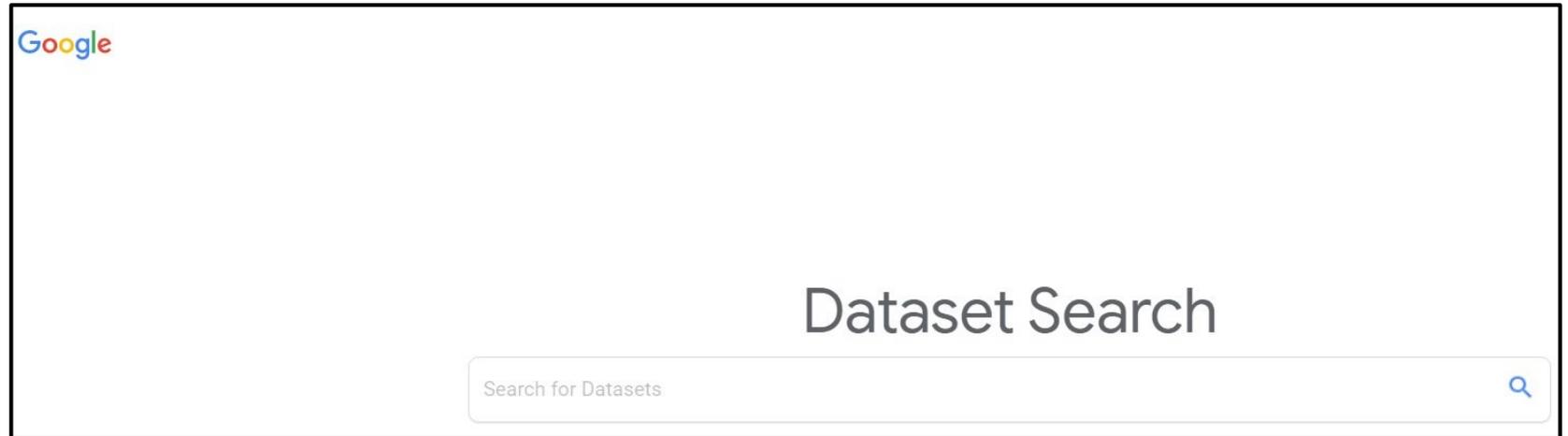
Similar datasets



# Strategy 5 - Search engines



[commons.datacite.org](https://commons.datacite.org)



[datasetsearch.research.google.com](https://datasetsearch.research.google.com)



# Strategy 5 - Search engines

Example: [Google Dataset Search](#)

The screenshot shows a Google Dataset Search interface. At the top, the Google logo is on the left, and a search bar contains the query 'youth participation site:zenodo.org'. Below the search bar are several filter buttons: 'Last updated', 'Download format', 'Usage rights', 'Topic', 'Provider', and 'Free'. The main content area displays '15 datasets found'. The first result is highlighted in a light blue box. It is a Zenodo dataset titled 'CATCH-EyoU: Meanings and Practices of Youth Participatio...' with the URL 'zenodo.org' and an update date of 'Aug 2, 2024'. To the right of the search results, a detailed view of the selected dataset is shown. It includes the Zenodo logo, the full title 'CATCH-EyoU: Meanings and Practices of Youth Participation and Youth Civic and Political Participation Initiatives', an 'Explore at' link to 'zenodo.org', a 'Unique identifier' (DOI: https://doi.org/10.5281/zenodo.1313558), the 'Dataset updated' date (Aug 2, 2024), the 'Dataset provided by' (Zenodo), a list of 'Authors' (Shakuntala Banaji, Isabel Menezes, Veronika Kalmus, Frosso Motti-Stefanidi, Erik Amnå, Elvira Cicognani, Peter Noack, Petr Macek), the 'License' (Attribution-NonCommercial 4.0 (CC BY-NC 4.0)), and a 'Description' section. The description states that the dataset integrates documentary evidence of national youth civic and political participation initiatives, including pamphlets, brochures, leaflets, newsletters, social media, and websites, and that it consists of a spreadsheet and a textual report.

TIP: To narrow down search and search a specific website, add

- site: URL

E.g.: youth participation site: zenodo.org



# Assessment of (research) data - Once you have found (meta)data

To determine if you can (re)use data, consider the following:

- What is **the source** of the data?
- Are the data properly **documented**?
- Consider potential criteria for data evaluation, e.g. **credibility, quality, and relevance**.
- Are the data **accessible**?
- Are the data **interoperable**?
- Check the conditions for **reuse** (license).



# Recap

## Databases

E.g. Web of Science:

- Associated Data
- Data Citation Index

## Repositories

Zenodo

Figshare

Re3data.org

## Search engines

DataCite Commons

Google Dataset Search

OpenAIRE Explore

## When finding data

Be critical of your sources:

- Assess the data and metadata
- Consider the context: does it make sense to reuse in your project?
- Remember to look for the license



# Handouts

## Where to search

### Repositories

- Harvard Dataverse: <https://dataverse.harvard.edu/>
- Zenodo: <https://zenodo.org/>
- Figshare: <https://figshare.com/>
- Dryad: <https://datadryad.org/>
- Mendeley Data: <http://www.mendeley.com/data/>

### Websites with data from outside the world of research

- ESA Climate Office: <https://climate.esa.int/en/data/#/dashboard>
- Statistics Denmark: <https://www.dst.dk/da/> or <https://www.statistikbanken.dk/stathank5a/default.asp?w=1133>

## Data search exercise

### Bibliographic databas

- Web of Science: <https://www.webofscience.com/>
- Dimensions: <https://www.dimensions.org/>
- OpenAlex: <https://openalex.org/>

### Bibliographic databas

- Start your search in [Scopus](#), [Web of Science](#)

**Case 1:  
Political leadership and interactive governance**

You are interested in research data on political leadership, collaboration policies, and interactive governance, where community actors - especially young people - are involved. The datasets should preferably be based on Scandinavian municipalities and be published alongside or in support of peer-reviewed journal articles.

**Case 2:  
HIV patients' attitudes towards treatment**

You are interested in finding research data regarding HIV patients' attitudes towards treatment, either with a single pill a day or with multiple pills a day. It is preferred that the datasets are published alongside or in support of peer-reviewed journal articles.

**Case 3:  
Participatory education and community planning**

You are interested in data relevant for research where active citizen involvement is part of the method. Datasets that include interviews and/or surveys are preferred, and the study can take place in Denmark or Scandinavia, but other locations may also be relevant.

## Logbook

### Part 1: Searches

Please record your searches below.

Repository/search engine	Keywords/search strings	No. of hits	Profitable/useful search?	Notes

Notes (e.g., filters chosen, source of reference)

Notes (e.g., filters chosen, source of reference)	Link or DOI to found datasets	Date

# Open Educational Resources

## Searching for research data (teaching material for PhDs and researchers)

datasegning DET-KGL-BIBLIOTEK engelsk evet svar

datasegning

Research data is not search strategies. The process compared

On this page, you will find teaching materials

Searching for Research Data is for researchers. It is also for researchers who are new to the field

### Teacher's guide

#### Searching for Research Data – Teacher's Guide

##### Contents

Introduction.....	1
Learning objectives.....	2
Structure and content .....	3
Step-by-step activity instructions.....	4
Search strategy exercises .....	4
Workshop .....	5
Final discussion.....	6
Mix & Match / Combination possibilities .....	7
Scenario 1.....	7
Scenario 2.....	8

SIDE 1 AF 16

► Download the teacher's guide



SLIDE 1 UD AF 47

## Searching for Research Data: Introduction

A libguide on how to find existing data to reuse for research.

Search this Guide

Search

- Introduction
- Where and how to search
- How to document your search
- Research data assessment
- Open Data Sources
- Training & other resources
- Glossary

### Introduction to Research Data Search

This guide is for researchers, students and other scholars who want to learn how to find existing data to reuse for research. The guide is a result of a collaboration between Aarhus University Library (AUL), Copenhagen University Library (KUB) and Roskilde University Library (RUB), all part of the Royal Danish Library (KB). Note that access to certain sources that are mentioned throughout this guide may depend on your affiliation to a university or university library (e.g. Web of Science Data Citation Index).

Data are the foundation of most research – the raw material, if you will. Many researchers create their own data through e.g. measurements, observations or simulations, but a good deal of research is done by re-using data from other researchers or data that have been collected in another context. This includes data from citizen science projects, previous research projects, public registries or commercial databases, for example.

Data are being shared in many different ways and can be found in many different places – including places that are not indexed in the same way as e.g. peer reviewed articles are. This means that a different approach to searching data is needed - compared to searching articles. In fact, a combination of strategies is more suitable.

This guide introduces two distinct, yet complementary approaches to searching for data: a direct and an indirect search strategy.

The **direct strategy** involves targeted searches, focusing on specific keywords, databases, and repositories to locate relevant data. This method is ideal for researchers seeking precision and efficiency in their searches, and who know what they are looking for.

- Introduction
- Where and how to search
- How to document your search
- Research data assessment
- Open Data Sources
- Training & other resources
- Glossary

### New Open Data sources?

If you know of any open data resources that would fit nicely in the collection, please let us know. Send us a message: [forskertsupport@kb.dk](mailto:forskertsupport@kb.dk).

Please check licence terms of the original provider, as some of these resources use third-party data.

#### General

- The Danish National Archives**  
The Danish National Archives preserves Denmark's history. You can find material from the 12th century to the present day from both authorities and private actors. Danish.

#### Social Sciences - Denmark

- Censuses (Folketællinger)**  
Censuses are available when they are 75 years old. The first is from 1787, and the newest on Arkivalier online is from 1940. Danish/English.
- Danish Demographic Database**

#### Humanities

- CLARIN-DK**  
CLARIN-DK is an infrastructure where researchers can deposit, share and download language-based materials, such as texts, transcriptions, encyclopedias, glossaries, audio and video files. Danish/English.

#### Natural Sciences

- Climate Trace**  
Climate TRACE emissions data are free and publicly available for download and via API. English.
- Danish Meteorological Institute Open Data**  
Open Data API's provide free and open

[learninglib.org](http://learninglib.org)

[libguides.ruc.dk](http://libguides.ruc.dk)

February 2026

► Download the teaching material

► Download a Danish version of the teaching material

# Future development

- Searching for research data – workshop for students
- Improving the FAIRness of the teaching material
- Considering different formats (e.g. online/in-person workshops)

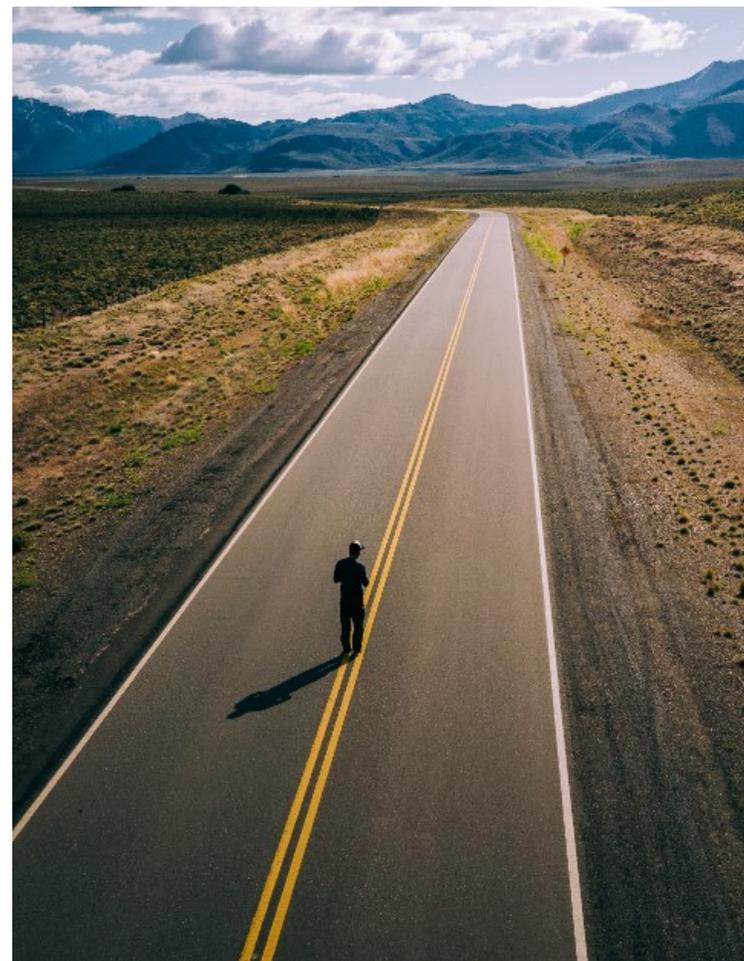


Photo by [Martin Sanchez](#) on [Unsplash](#)



# Q & A



**DET KGL.  
BIBLIOTEK**

Royal Danish Library