Discover Data Citation Index and accelerate your research with open data

User guide

March 2023
Open Science to accelerate discovery
An increasing number of funders is promoting Open Science

“Horizon Europe will set a new standard for dissemination of knowledge and new skills across European societies. With clear and immediate open access requirements for beneficiaries, the Open Europe Research publishing platform and a strengthened European Open Science Cloud, we are well underway in making truly open science a reality.”

Mariya Gabriel, Commissioner for Innovation, Research, Culture, Education and Youth

Today, the White House Office of Science and Technology Policy (OSTP) updated U.S. policy guidance to make the results of taxpayer-supported research immediately available to the American public at no cost. In a memorandum to federal departments and agencies, Dr. Alondra Nelson, the head of OSTP, delivered guidance for agencies to update their public access policies as soon as possible to make publications and research funded by taxpayers publicly accessible, without an embargo or cost. All agencies will fully implement updated policies, including ending the optional 12-month embargo, no later than December 31, 2025.
Promoting Open Science

One of the key blocks in the Open Science model is Open Data

Open Data must be:
- Findable
- Accessible
- Interoperable
- Re-usable
Facts and Figures for open research data

Figures and case studies related to accessing and reusing the data produced in the course of scientific production.

Researchers' attitude towards data sharing

This indicator shows how researchers share data and their underlying attitudes, based on the survey for the 2017 report "Open data: The researcher perspective" and the results of a second online survey carried out between October/November 2018 (see the methodological note for additional information).

Benefits of sharing research data: % of researchers per benefit

Reference dates: 2016 and 2018

- More possibilities for collaboration (66% in 2018, 55% in 2016)
- Reproducibility of research (57% in 2018, 51% in 2016)
- Encourage other researchers to make their data publicly available (55% in 2018, 50% in 2016)
- Article more likely to be cited (51% in 2018, 49% in 2016)
- Data reuse (50% in 2018, 49% in 2016)
- Article more likely to be accepted for publication (42% in 2018, 41% in 2016)
- Research aggregation (e.g., meta-analysis) (39% in 2018, 39% in 2016)
- Compliance with journal or publisher requirements (28% in 2018, 21% in 2016)
- Compliance with funding body mandates (25% in 2018, 21% in 2016)
- Other (15% in 2018, 15% in 2016)
Why Data Citation Index?
Why Data Citation Index?

Research data is scattered across many subject specific and generalist repositories and can be hard to find. Moreover, data from a single institution can be scattered across multiple data repositories. So, there is the need for a standard search tool.
Web of Science platform content

Gain a comprehensive view of worldwide research across the sciences, social sciences, and arts & humanities

34,000+
Journals across the platform

109 million
Patents for over 56 million inventions

21,900+
Total journals in the Core Collection

14.5 million+
Data Sets and Data Studies

2.2 billion+
Cited references

Backfiles to 1900
With cover-to-cover indexing

196 million+
Records

300,000+
Conference proceedings

20 million +
Records with funding data

137,000+
Books

DCI was created in 2012

Statistics as of October 2022
Why Data Citation Index?
Discover and download data sets faster

>95%
Over 95% of data meeting Data Citation Index’s selection criteria is open.

Repositories are selected by Clarivate editorial team based on qualitative and quantitative criteria

Track citations to the data and software you have shared to demonstrate the impact and influence of your work.

Save time searching for relevant data across hundreds of repositories and easily navigate between published papers and data sets.

• Quickly find data to validate studies or to reuse in your own work.
• Access a trusted reference of resources for depositing data.
• Promote data sharing and citation.
• Over 13 million datasets
• Over 450 repositories

Updated weekly
Finding research data
Existing associated data is indicated in Web of Science search results

When searching Web of Science databases, you can identify papers with associated data indexed in Data Citation Index.

- Exposing research data in search results puts millions of data sets and data studies directly into your discovery workflow.
- Related data supports Open Science by bringing more transparency to the research process.
- Available only for Data Citation Index subscribers

The Associated Data filter is available in All Databases and each database (only exception is Derwent Innovations Index).
Access associated data from a document record

The Association Between Men’s Mental Health During COVID-19 and Deterioration in Economic Status

By: Hadar Shoval, D (Hadar-Shoval, Dorit) [1]; Alon-Tirosh, M (Alon-Tirosh, Michal) [2]; Asraf, K (Asraf, Klim) [1]; Tannous-Haddad, L (Tannous-Haddad, Lubna) [2]; Tzochinsky, O (Tzochinsky, Oron) [3]

View Web of Science ResearcherID and ORCID (provided by Clarivate)

1 View all associated records in Data Citation Index

2 Click on the title to open the DCI record

3 Direct access to the data at the top

Data
Generalized anxiety and depressive symptoms in various age groups during the COVID-19 lockdown in Poland.

Gambin, Malgorzata

2020 | Harvard Dataverse | Data study | 1

Dataset includes data from the following sections:
(i) Generalized anxiety symptoms (GAD-7) in men and women
(ii) Depression (CES-D)
(iii) Differences in perceived mental health

Abstract

Generalized anxiety and depressive symptoms in various age groups during the COVID-19 lockdown in Poland.

By: Gambin, Malgorzata[1]

Harvard Dataverse
Volume: 1
DOI: 10.7910/DVN/ONP102
Published: 2020
Indexed: 2020-11-17
Document Type: Data study

Dataset includes data from the study that investigated:
(i) Differences in levels of depressive and generalized anxiety symptoms in four age groups in the Polish population during the COVID-19 lockdown;
(ii) Differences in perceived difficulties related to the pandemic in these groups; and
(iii) Which factors and difficulties related to the pandemic are the best predictors of generalized anxiety and depressive symptoms in various age groups during the COVID-19 lockdown. A total of 1,115 participants (aged 18-80) took part in the study. The sample was representative of the Polish population in terms of sex, age, and place of residence. Participants completed the following online: the Patient Health Questionnaire-9, the Generalized Anxiety Disorder-7, a Scale of Perceived Health and Life Risk of COVID-19, a Social Support Scale, and a Scale of Pandemic-Related Difficulties.
Searching Data Citation Index

Data Citation Index presents a different indexation and thus different search options.

For a research field, you can use any of the following: topic, title, abstract, subject descriptors (subject descriptors are all the indexing fields outside of title, abstract or author keywords).

Note that DCI uses a combined taxonomy from Zoological Record and BIOSIS.

Search taxonomic descriptors in the thesaurus and add them to the query.

Example: Geospatial OR Biophysics OR Rodent

Search:

Subject Descriptors

- Taxonomic Data (includes Taxa Notes and Super Taxa terms)
- Web of Science Category

Examples:
- Fungi
- Cultural studies
- Chordata

Search in: Data Citation Index ▼ Editions: All ▼

Topic:
- Enter Topic terms to search the following fields and tables within a record:
  - Title
  - Abstract
  - Author keywords

Examples:
- Cardiovascular disease
- ecology
Searching Data Citation Index

Data Citation Index presents the powerful Web of Science options for exploring a body of information.

Several filters are specific to Data Citation Index.

Whereas data types are assigned by the repository or the depositor of the data and usually vary a lot from one place to another.
Searching Data Citation Index

Data Citation Index present the powerful Web of Science options for exploring a body of information

Web of Science categories are assigned manually for each repository

The advanced search allows you to write a search with a series of field tags

Where the author's affiliation is available from the repository it is included in DCI records as provided by the repository

Research areas or Subject areas constitute a classification scheme that is shared by all Web of Science databases.
Searching Data Citation Index

Reading a data record in Data Citation Index

University Student Health and Well-being Survey 2021

Group Author: Finnish Institute for Health and Welfare
Finnish Social Science Data Archive

Source URL: http://urn.fi/urn:nbn:fi-fe:fs03016

Published: 2022
Indexed: 2023-02-14
Content Type: Data study
Data Type: Quantitative

Abstract
This national survey examined the health and health-related behaviour of university and university of applied sciences (polytechnic) students in Finland. Questions covered, among other topics, physical, mental and dental health, health-related behaviour, use of health services, and satisfaction with health services. Factors related to the ability to study and well-being, such as social relationships and income, were also examined. The survey also included questions on the COVID-19 pandemic, ongoing at the time. The study was funded by the Social Insurance Institution, the Ministry of Education and Culture, and the Finnish Institute for Health and Welfare. General state of health was charted with a number of questions relating to long-term illnesses, disabilities or disorders, and current well-being in general. The respondents were asked if they had suffered from a variety of symptoms in the previous month. Birth control methods used by the respondent and her/his partner within the previous month were surveyed. Some questions covered the respondents' weight, height, and attitude toward food. Psychological or social symptoms causing problems were investigated. Psychosocial health was further studied using items of the General Health Questionnaire (GHQ12). Use of healthcare services by different types of service providers and satisfaction with the services of the Finnish Student Health Service (FSHS) were surveyed, along with reasons for using the services of other healthcare service providers. Time spent sitting during the day as well as engagement in sporting activities and less vigorous physical activities were surveyed. Eating habits were investigated with a number of questions, for instance, asking whether the respondents ate at student restaurants, what diet they followed, on how many days per week they ate certain meals or snacks (e.g., breakfast, lunch, afternoon snack, dinner), and the consumption of fruit and vegetables. The respondents were also asked how often they brushed or flossed their teeth and used toothpaste or xylitol products, and whether they had problems with teeth grinding, facial pain or jaw locking. Health-related behaviour was further studied by asking the respondents about their
Research data citation
Track citations to the data and software you have shared

Demonstrate the impact and influence of your work

<table>
<thead>
<tr>
<th>View data</th>
<th>Cite this Resource</th>
<th>Export ▼</th>
<th>Add To Marked List ▼</th>
<th>&lt; 2 of 7,662 &gt;</th>
</tr>
</thead>
</table>

**University Student Health and Well-being Survey 2021**

**Group Author:** Finnish Institute for Health and Welfare  
Finnish Social Science Data Archive

**Source URL:** http://urn.fi/urn:nbn:fi:fssd:2-FSD3616

**Published:** 2022  
**Indexed:** 2023-02-14  
**Content Type:** Data study  
**Data Type:** Quantitative

**Abstract**
This national survey examined the health and health-related behaviour of university and university of applied sciences (polytechnic) students in Finland. Questions covered, among other topics, physical, mental and dental health, health-related behaviour, use of health services, and satisfaction with health services. Factors related to the ability to study and well-being, such as social relationships and income, were also examined. The survey also included questions on the COVID-19 pandemic, ongoing at the time. The study was funded by the Social Insurance Institution, the Ministry of Education and Culture, and the Finnish Institute for Health and Welfare. General state of health was charted with a number of questions relating to long-term illnesses, disabilities or disorders, and current well-being in general. The respondents were asked if they had suffered from a variety of symptoms in the previous month. Birth control methods used by the respondent and her/his partner within the previous month were surveyed. Some questions covered the respondents' weight, height, and attitude toward food. Psychological or social symptoms causing problems were investigated. Psychosocial health was further studied using items of the General Health Questionnaire (GHQ-12). Use of healthcare services by different types of service providers and satisfaction with the services of the Finnish Student Health Service (FSSH) were surveyed, along with reasons for using the services of other healthcare service providers. Time spent sitting during the day as well as engagement in sporting activities and less vigorous physical activities were surveyed. Eating habits were investigated with a number of questions, for instance, asking whether the respondents ate at student restaurants, what diet they followed, on how many days per week they ate certain meals or snacks (e.g. breakfast, lunch, afternoon snack, dinner), and the consumption of fruit and vegetables. The respondents were also asked how often they brushed or flossed their teeth and used toothpaste or xylitol products, and whether they had problems with teeth grinding, facial pain or jaw locking. Health-related behaviour was further studied by asking the respondents about their...
How do we count citations to data?

The citation information is gathered from three main sources:

- from repositories where available
- from formal cited references lists in publications
- from informal citations in the literature full text (captured from text of articles from life science databank accession numbers)
How to cite research data?

DCI provides a citation format based on the DataCite’s citation guidelines (https://www.datacite.org)

Clarivate recommends citing this resource as:

# Elements of a Data Citation

Metadata elements for the DataCite citation guideline

<table>
<thead>
<tr>
<th>Elements</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Author/Creator</td>
<td>Individuals or organizations that created or contributed to the data set; this metadata element is vital to guarantee attribution and credit for data contributor, and to provide metrics for their nontraditional scholarly output.</td>
</tr>
<tr>
<td>Year</td>
<td>The year of “publication” of the data; when it is made publicly available, such as through deposition in a repository.</td>
</tr>
<tr>
<td>Title</td>
<td>The title of the data object, which may differ from the title of the parent research paper/project.</td>
</tr>
<tr>
<td>Publisher</td>
<td>The data repository that houses the data and/or the governing organization responsible for publishing, (i.e., making available) the data.</td>
</tr>
<tr>
<td>Version</td>
<td>Dynamic data sets or those where new editions may be issued (such as with error corrections or new values) must employ proper version control to guarantee accuracy and uniqueness in data citation.</td>
</tr>
<tr>
<td>Permanent Identifier</td>
<td>A unique and persistent identifier should be assigned; for example, a Digital Object Identifier (DOI); in Data Citation Index citations, this bibliographic element may take the form of a unique URL, databank accession number, or other permanent identifier such as Handle (hdl) (<a href="http://www.handle.net">http://www.handle.net</a>).</td>
</tr>
</tbody>
</table>
Recommendations
to cite data

Recommended practices
to promote scholarly data
citation and tracking

The role of the Data Citation Index

Web of Science
Trust the difference

Data citation
A guide to best practice

Everyday millions of people use data. They search, publish, reuse, analyse! But very often, a major difficulty for the users is finding the right data. And how to help them? It is simple: Cite the data! The publication ‘Data citation: a guide to best practice’ is a practical guide that includes recommendations and best practices on how to cite data. It is divided in 3 parts: part 1 – issues in data citation, general recommendations; part 2 – recommendations on specific formats and other elements for data citation and part 3 – other information (e.g., specific differences between recommendations). At the end the reader will find annexes with checklists, diagrams and examples.
Why relying on Data Citation Index?
The repositories

Started in 2012, with 50 repositories, DCI is now indexing +450 repositories and we are aiming to expand coverage.

This shows the top 25 repositories by number of records in DCI

Over 95% of the repositories in the DCI provide free access to content.
Repositories are selected by Clarivate editorial team, like journals are selected for the Web of Science Core Collection. The editorial team considers many factors, both qualitative and quantitative.

The repository must also show a long-term data preservation strategy and be actively updating content.

The repositories are globally distributed and provide quality control and editorial policies for managing the data. The repository must also show a long-term data preservation strategy and be actively updating content.
The geographical coverage by repository location

Researchers do not necessarily deposit their data in a local repository

Over 1700 repositories identified; 50% acceptance
The geographical coverage by author address

The records in the Data Citation Index span > 240 countries

This shows the top 25 countries by number of records in the DCI.
Organizations represented

Note that affiliation names are not unified in DCI (only in the Core Collection)

This shows the top 25 organisations by number of records in the DCI
This shows the top 25 categories by the number of records in the DCI.

- Genetics Heredity
- Crystallography
- Microbiology
- Spectroscopy
- Oceanography
- Chemistry Multidisciplinary
- Multidisciplinary Sciences
- Geosciences Multidisciplinary
- Social Sciences Interdisciplinary
- Physics Particles Fields
- Biodiversity Conservation
- Meteorology Atmospheric Sciences
- Geosciences Multidisciplinary
- Engineering Multidisciplinary
- Humanities Multidisciplinary
- Cultural heritage, language corpus, image collections, recordings...
- Geography
- Environmental Sciences
- Geochimistry Geophysics
- Materials Science Multidisciplinary
- Marine Geohazards
- Geochemistry
- Geology
- Geophysics
- Geotechnical Engineering
- Geophysics
Data coverage by publication year

Although DCI indicates coverage from 1900, you will find records older than 1900. In fact, there is no lower limit to the year in terms of the data we gather; we index the whole content of the data repository irrespective of year.
How do we curate data?

Our editorial team work on all the data from the repositories.

- Pick up descriptive metadata feed from repository
- Analyse raw metadata
- Add additional content and citation data

As a result, improved records for Repositories, Data Studies, Data Sets and Software are added to the Data Citation Index.
- The DCI spans from 1900 and earlier, and includes all available citation data
How do we organize data?

We evaluate and select the best data repositories for content, persistence and stability, and searchability. Data that we index is organized into three content types to enhance searchability and discoverability:

- **Repository**: This resource consists of data studies and data sets and provides access to the data.
- **Data Study**: This is a description of the study or experiment with the associated data used in the study. Includes serial or longitudinal studies over time.
  - **Data Set**: A single or coherent set of data, or a data file, provided by the repository as part of a collection, data study, or experiment.
  - **Software**: A computer program or package in source code or compiled form, which can be installed on another machine and used to support & analyze research.
How will Data Citation Index help you?
Data Citation Index can help you

**RESEARCHERS**
- Save time looking for data in one place rather than searching across repositories
- Avoid duplication of efforts
- Reach more people, have greater impact
- Track citations and get recognition for the data and software you are sharing

**LIBRARIANS**
- Promote the value and the importance of sharing data
- Provide a trusted reference for researchers of where to deposit their data
- Preserve data for future researchers
- Validate attribution of single datasets by researchers

**RESEARCH OFFICES**
- Improve your research visibility and impact
- Align your strategy with Open Science
Data Citation Index

Connecting data to the research it informs

Go to product  Contact us

For questions, contact:
WoSG.support@clarivate.com

© 2022 Clarivate. All rights reserved. Republication or redistribution of Clarivate content, including by framing or similar means, is prohibited without the prior written consent of Clarivate. Clarivate and its logo, as well as all other trademarks used herein are trademarks of their respective owners and used under license.