About DataCite

Global non-profit membership organization working with 3000+ repositories in the world to provide DOIs for research outputs and resources.
A data citation is a reference to data, in the same way researchers routinely provide a bibliographic reference to other scholarly resources.
Joint Statement on Research Data: a role for everyone

Researchers

● When publishing their results, researchers deposit related research data and outputs in a trustworthy data repository that assigns persistent identifiers (DOIs where available). Researchers link to research data using persistent identifiers.
● When using research data created by others, researchers provide attribution by citing the datasets in the reference section using persistent identifiers.

Repositories

● Data repositories enable sharing of research outputs in a FAIR way, including support for metadata quality and completeness.
● Repositories and publishers connect articles and datasets through persistent identifier connections in the metadata and reference lists.
Repositories play an important role in data citation by linking data to related articles (or other outputs)

- Ask researchers to include associated publications in dataset metadata
- Include those links in DataCite metadata

<table>
<thead>
<tr>
<th>Related Identifiers</th>
<th>Identifiers of related resources.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>10.1111/jeb.13000</td>
</tr>
</tbody>
</table>

Must be a globally unique identifier. Visit our support website for the list of supported unique identifiers.

<table>
<thead>
<tr>
<th>* Related Identifier Type</th>
<th>DOI</th>
</tr>
</thead>
<tbody>
<tr>
<td>The type of the Related Identifier.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>* Relation Type</th>
<th>Is referenced by</th>
</tr>
</thead>
<tbody>
<tr>
<td>The type of the Relation.</td>
<td></td>
</tr>
</tbody>
</table>
The `RelatedIdentifier` field in the DataCite metadata schema can be populated in order to generate a link between the data and the publication that cites it.

Mandatory sub-properties:

- `relatedIdentifierType`: the type of identifier
- `relationType`: the type of relationship

https://support.datacite.org/docs/contributing-citations-and-references
How to create citations

The metadata of the dataset includes:
- IsReferencedBy
- IsCitedBy
- IsSupplementTo

The metadata of the article includes:
- References
- Cites
- IsSupplementedBy

Citation of a dataset

Citation of an article

https://support.datacite.org/docs/contributing-citations-and-references
DataCite Commons - surfacing citations

Corona Virus (COVID-19) Tweets Dataset (en) [Data set]. IEEE DataPort. https://doi.org/10.21227/781w-ef42

Description
Each database (*.db) contain three columns. First column: date and time of the tweet. Second column: tweet. Third column: sentiment score for the particular tweet, within the range [-1,1] with -1 being the most negative, 0 being the neutral and +1 being the most positive sentiment. The tweets have been collected by the LSTM model deployed here at sentiment-live. To make it easy for the NLP researchers to get access to the sentiment analysis of each collected tweet, the sentiment score out of TextBlob [2] has been appended as the last column. [1] https://sentiment-live/ [2] https://textblob.readthedocs.io/en/dev/

Open Access Dataset published 2020 in IEEE DataPort

17 Citations

https://commons.datacite.org /doi.org/10.21227/781w-ef42
The Data Citation Corpus

http://corpus.datacite.org/dashboard