



# Data papers and data journals

A **data paper** (data article, data descriptor) is a **publication that describes a scientific dataset** using structured information known as metadata. The data paper is a formalised way of sharing data and is unlike a traditional research article which tests hypotheses or presents new analyses. A data paper and a traditional research article are complementary.

## Why publish a data paper?

- To publicise the data's existence and make them easier to find;
- To credit authors (recognition, reliable reference) and enhance the data's value;
- To facilitate the data reuse by making them intelligible.

## Where to publish a data paper?

- In a data journal which is dedicated to this type of publication;
- In a traditional journal that publishes data papers alongside traditional articles.

## What is in a data paper?

The structure of data papers can vary a great deal from one journal to another but they do have some [common component parts](#).



### A descriptive section

- Elements common to classical articles;
- Specific elements of data.

### Example

- Title, summary, keywords, etc.
- Metadata, reuses, etc.



### Access to data

They can be cited in the article after having been deposited in a repository. If this is the case, the data identifier (DOI) [links from the data paper to the data](#).

## How do you write a data paper?

### Instructions and template

- Some journals provide templates;
- Others provide writing tools;
- Others permit data submissions from external platforms.

### Examples

- The '[Data in Brief](#)' template
- The [arpha Writing Tool](#)
- [Scratchpads](#), [GBIF](#), etc.

# Focus on data papers

The content and length of the data paper can vary considerably depending on the journal and its requirements, ranging from a very brief summary to a fully comprehensive article.

Here is an example of a data paper structure for the data journal [Data In Brief](#):



**Data in Brief**  
Volume 2, March 2015, Pages 42-47



**Title:** the title should focus on the specific data shared and should not be the research article title.

**Author(s):** name, affiliations, email, etc.

**DOI :** <https://doi.org/10.1016/j.dib.2014.12.001>

**Licence type:** CC-BY...

**Article date:** article submission date, publication date, validation date, revision date

**Summary:** presentation of the context in which the data were obtained (research question, etc.)

Metadata:	
Field of research	Physics, chemistry, psychology, etc.
A more specific field	e.g. nuclear physics
Type of data	Table, image (X-ray, microscopy, etc.), text, graph, figure, etc.
Data acquisition methods	Microscopy, survey (general view), SEM, NMR, mass spectroscopy, etc.
Data format	Raw, filtered, analysed, etc.
Experimental factors	e.g. A brief description of the preparation of samples
Experimental characteristics	e.g. A very brief experimental description
Data source location	City, country, GPS coordinates for samples or data
Access to data	Repository name, data DOI or direct URL
The related research article	If the data accompanies a research article, this should be cited

**Value of the dataset:** describe the scientific value of these data

**Data description:** brief description of shared data

**Materials and methods:** full description of how the data were obtained including any figures/tables that help understand the data.

**Acknowledgements**

**References:** References of data cited