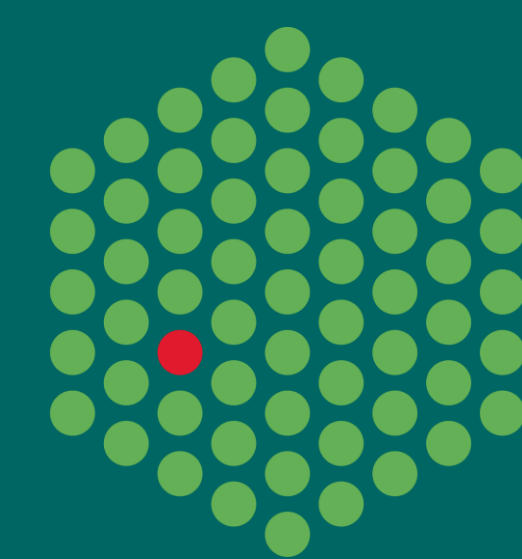


# BioStudies database: aggregating all outputs of a life sciences study

EMBL-EBI



BioStudies is a database at EMBL-EBI that facilitates transparency and reproducibility of research, by aggregating all the outputs of a study (a 'data package') in a single place.

Time

## Data generation

- Linking data in community archives
- Support for data without a community archive
- Support for project data sharing and transition to publication

## Article publication

- Direct deposition
- Integration with journals (aspiration)

## Post-publication

- Import from Europe PMC
- Makes data citable and findable independent of the article

## Data rescue

- Home for large datasets from retired projects
- Cost-efficient data sustainability

Better for: reproducibility, open science, FAIR data

## Data submissions

1 # This section is used to describe details of the study  
2 Submission  
3 Title  
4 ReleaseDate  
5  
6  
7  
8 Study  
9 Title  
10 Description  
11 Study type  
12 Key words  
13 Study Organism  
14 # This section is used to describe contact details of the people who have worked up  
15 # Copy and paste the section downwards if more than one contributor to the study  
16  
17 Name  
18 Email  
19 Role  
20 ORCID  
21 Affiliation  
22  
23 Organism  
24 Name  
25 # The  
26 # This  
27 # Copy  
28 Link  
29 Type  
30  
31 # How

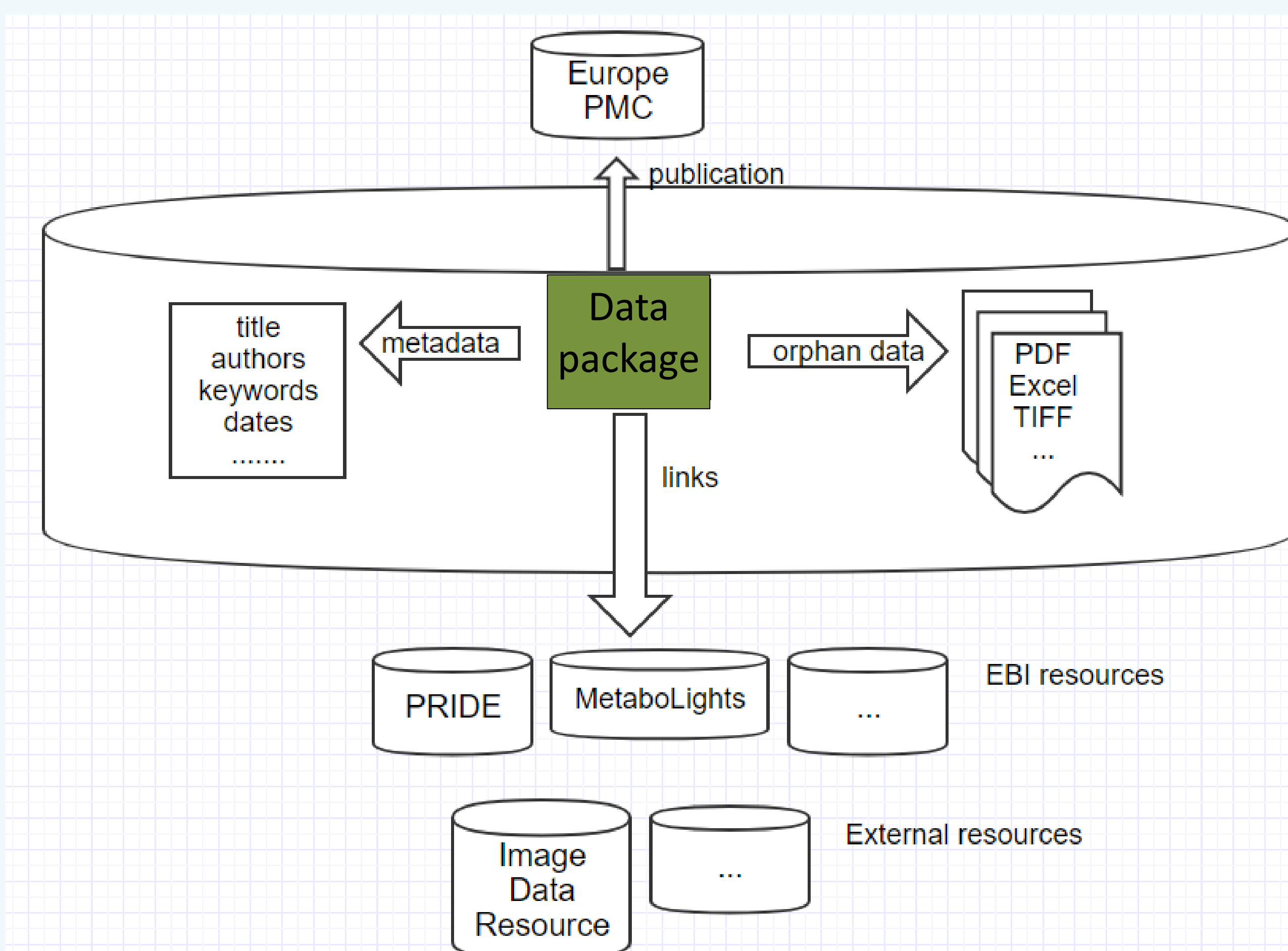
- Automatic ingest pipelines

Europe PMC SOURCE DATA Making data discoverable

- Programmatic submission of JSON, XML and TSV  
A simple data format for flexibility and fast prototyping

- Online web submission tool

## Generic data model



## Data access

BioStudies. Search BioStudies. Text search

Unique accessions

Downloads available in a number of formats

Ability to share studies privately

Data files and links

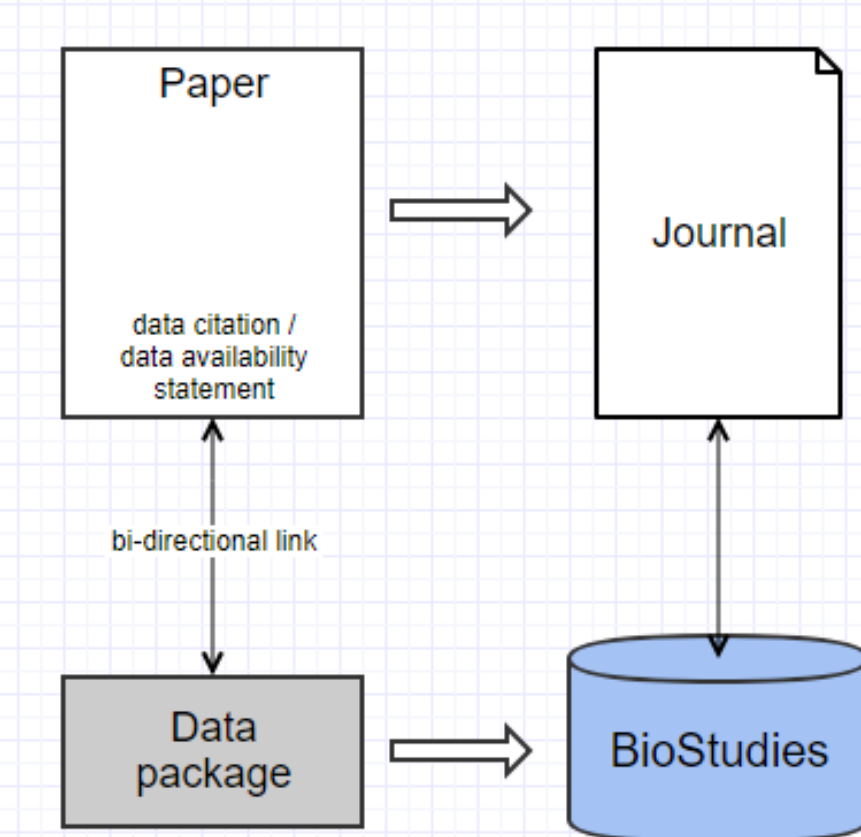
Metadata

Data searchable via facets

## Some future developments

More data packages created pre-publication:

- Links from paper to data can be created
- Mechanisms to ensure better annotation for supplementary data, specific for journal/community/project



BioStudies as a pilot infrastructure for light microscopy data management, providing capabilities of visual representations for images and easy means to explore and download large datasets