

FAIRification of biological data from literature

Aravind Venkatesan

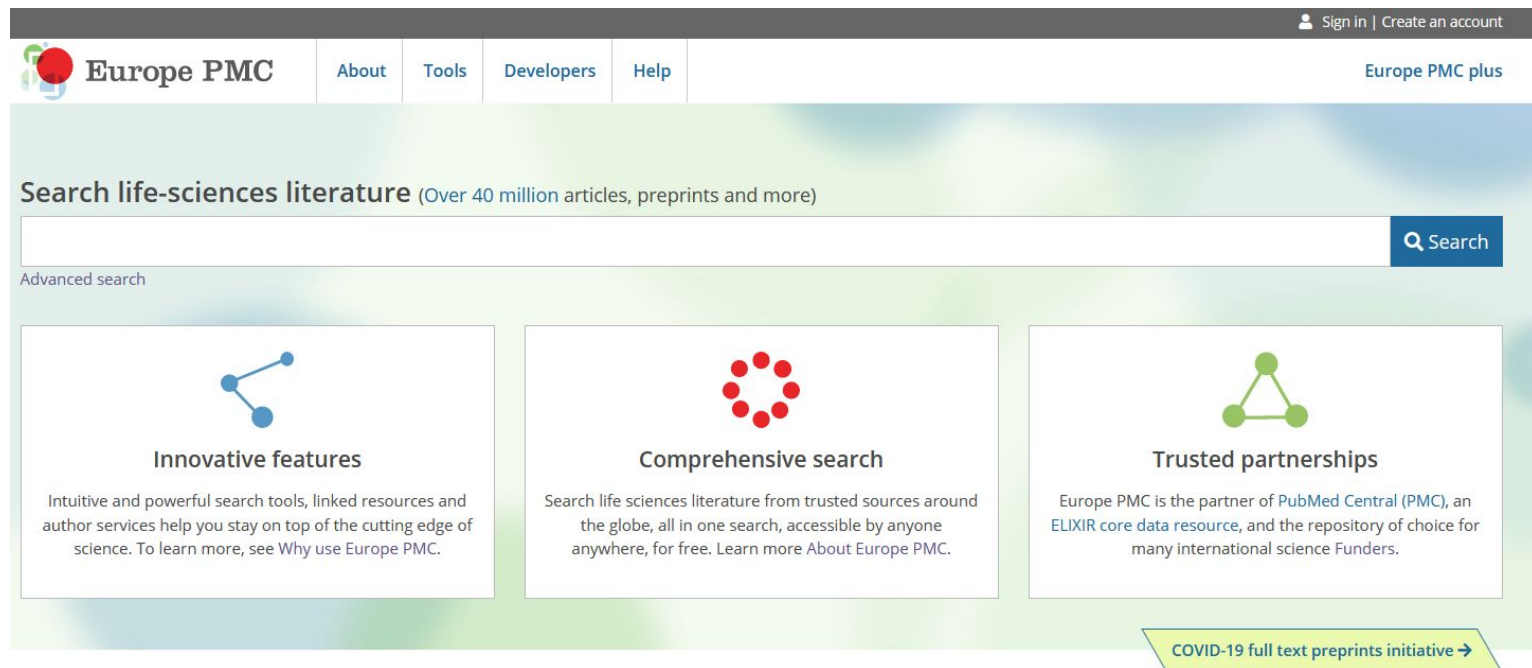
Senior Data Scientist
Literature services, EMBL-EBI



What is Europe PMC

A free digital archive of
biomedical and life sciences research publications

www.europepmc.org



The screenshot shows the Europe PMC website interface. At the top, there is a dark navigation bar with the Europe PMC logo on the left and links for 'Sign in' and 'Create an account' on the right. Below this is a white navigation bar with the Europe PMC logo and links for 'About', 'Tools', 'Developers', and 'Help'. The main content area has a light blue background with a search bar and a 'Search' button. Below the search bar, there is a section titled 'Search life-sciences literature' with a subtext '(Over 40 million articles, preprints and more)'. Underneath this, there is a link to 'Advanced search'. The main content area is divided into three columns, each with a different icon and title: 'Innovative features' (blue network icon), 'Comprehensive search' (red circular icon), and 'Trusted partnerships' (green triangle icon). Each column contains a brief description of the feature or partnership. At the bottom right, there is a yellow banner with the text 'COVID-19 full text preprints initiative' and a right-pointing arrow.

Sign in | Create an account

Europe PMC

About Tools Developers Help

Europe PMC plus

Search life-sciences literature (Over 40 million articles, preprints and more)

Search

Advanced search

Innovative features

Intuitive and powerful search tools, linked resources and author services help you stay on top of the cutting edge of science. To learn more, see Why use Europe PMC.

Comprehensive search

Search life sciences literature from trusted sources around the globe, all in one search, accessible by anyone anywhere, for free. Learn more About Europe PMC.

Trusted partnerships

Europe PMC is the partner of PubMed Central (PMC), an ELIXIR core data resource, and the repository of choice for many international science Funders.

COVID-19 full text preprints initiative →

The literature in Europe PMC

Abstracts and metadata

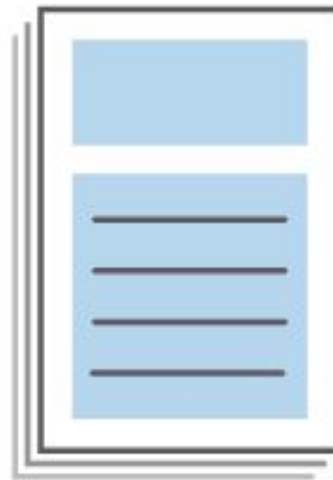
from PubMed, Agricola



41.7 million

Full text publications

from PubMed Central



8.6 million

Preprint abstracts
and metadata

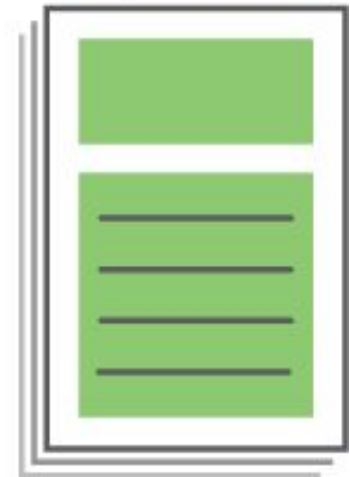
from various platforms



1/2 million

Preprint full text

COVID-19 and funder
preprints subset



34 thousand



Europe PMC

Supported by:



Biotechnology and
Biological Sciences
Research Council



National Centre
for the Replacement
Refinement & Reduction
of Animals in Research

**VERSUS
ARTHRITIS**



Medical
Research
Council

mnda
motor neurone disease
association



ZonMw



The **Academy of
Medical Sciences**

IndiaAlliance
DBT wellcome

DIABETES UK
KNOW DIABETES. FIGHT DIABETES.



NIHR

National Institute
for Health Research



myrovlytis **XXXX** trust



**CHIEF
SCIENTIST
OFFICE**



**Medical
Research
Foundation**



Remarkable
research for
healthy ageing
THE DUNHILL MEDICAL TRUST



**CANCER
RESEARCH
UK**



Ymchwil Iechyd
a Gofal **Cymru**
Health and Care
Research **Wales**



PARKINSON'S^{UK}
CHANGE ATTITUDES.
FIND A CURE.
JOIN US.



**BREAST
CANCER
NOW**
The research
& care charity

**worldwide
cancer
research**

**Pancreatic
Cancer
UK**

**RN
I:D**

**Blood
cancer
UK**

EMBL



**World Health
Organization**



**Swiss National
Science Foundation**

FWF

Der Wissenschaftsfonds.



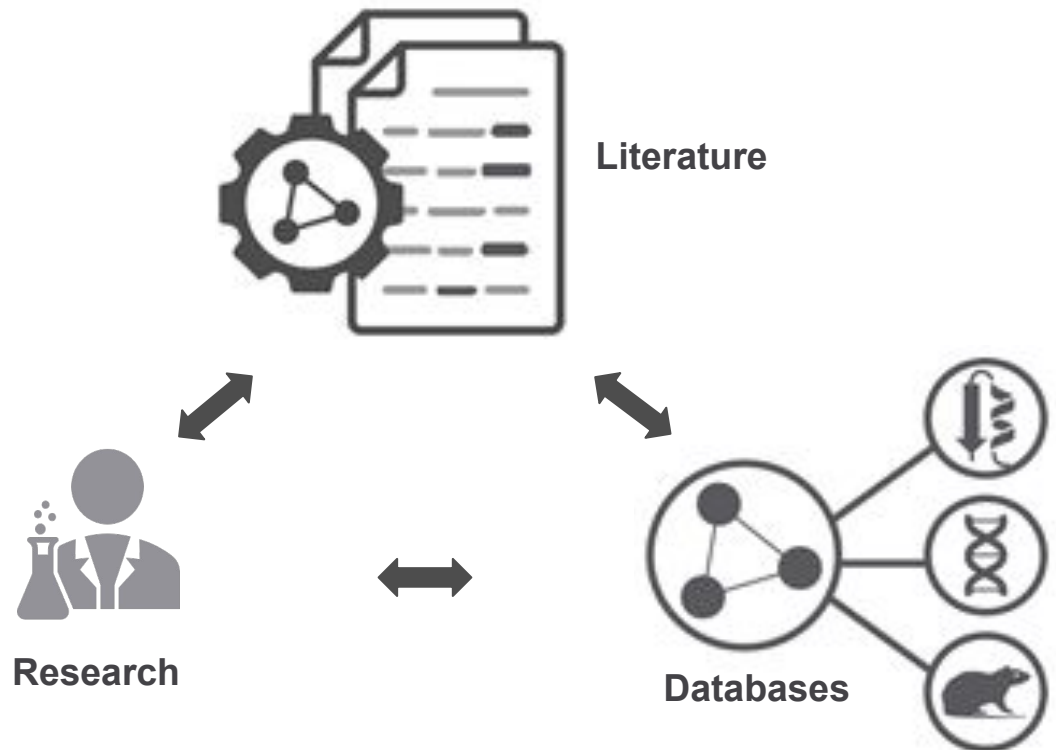
Europe PMC

Data and services in Europe PMC



Literature-Data Integration

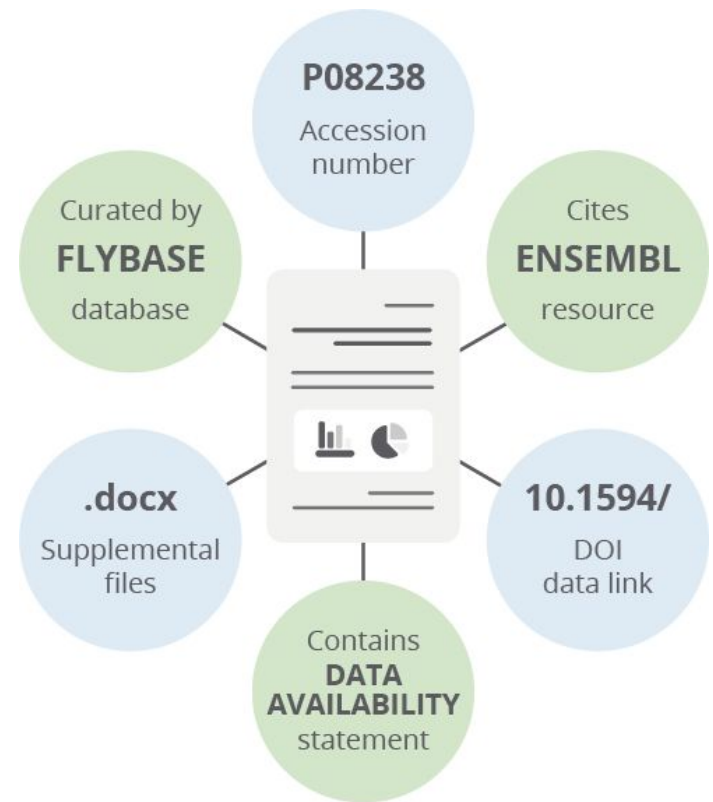
- Creating a inter-connected data ecosystem
- Provide Text/Data Mining and other engineering solutions
- Support scalable curation - triage/searching biological facts/assertions



Literature-Data Integration

Different ways of linking Literature - data

- Explicit data links
- Implicit data mentions



Literature-Data linking in Action



Europe PMC

[About](#)

[Tools](#)

[Developers](#)

[Help](#)

Europe PMC plus

[Abstract](#)

[Figures \(6\)](#)

[Free full text ▶](#)

[Citations & impact](#)

[Data](#)

[Similar Articles](#)

[Funding](#)

Data that cites the article

This data has been provided by curated databases and other sources that have cited the article.

Genes & Proteins

[Spastin](#)
(UniProt - A0A0B4LHJ5)

Protein Structures

[spastin hexamer in complex with substrate](#)
(PDB - 6p07)
[View structure ▶](#)

FlyBase

http://flybase.org/cgi-bin/uniq.html?db=fbrf&field=pubmed_id&context...

QuickGO Annotations

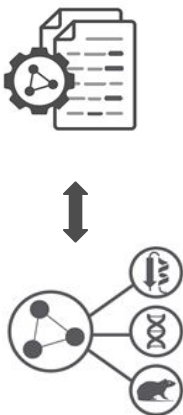
<https://www.ebi.ac.uk/QuickGO/annotations?reference=31285604>

[Annotations \(574\)](#)

[Get citation](#)

[Open PDF](#)

[Claim to ORCID](#)



PMCID:
PMC6761829



Europe PMC

Literature-Data linking in Action



Europe PMC

About

Tools

Developers

Help

Europe PMC plus

Abstract

Figures (6)

Free full text ▶

Citations & impact

Data

Similar Articles

Funding

Data

Data behind the article

This data has been text mined from the article, or deposited into data resources.

BioStudies: supplemental material and supporting data

<http://www.ebi.ac.uk/biostudies/studies/S-EPMC6761829?xr=true>

Genes & Proteins

UniProt - Q9UBP0

(1 citation)

Protein Structures (2)

PDBe - 3B9P

(3 citations)

[View structure ▶](#)

PDBe - 6P07

(2 citations)

[View structure ▶](#)

Annotations (574)

Get citation

Open PDF

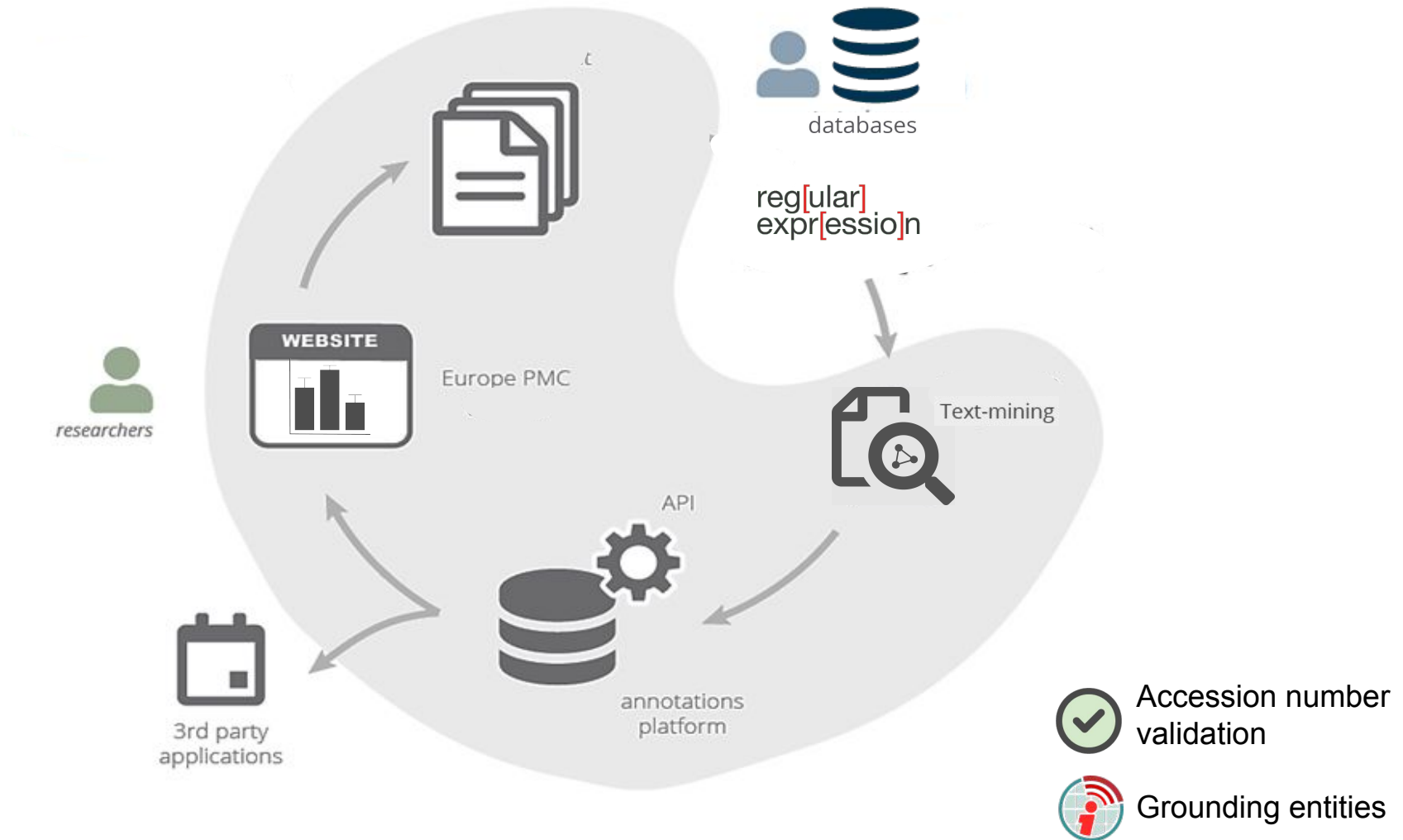
Claim to ORCID

PMCID:
PMC6761829

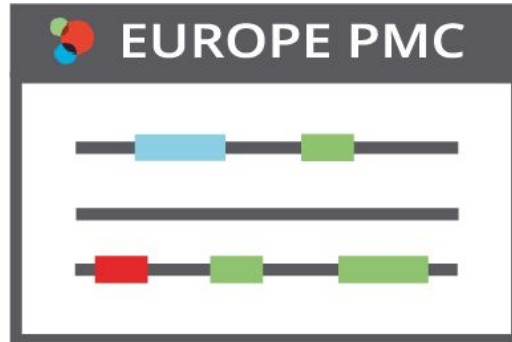


Europe PMC

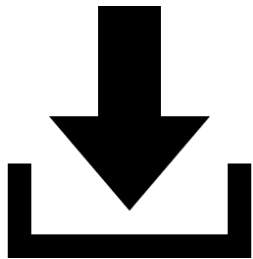
Data citation workflow



Data access



www.europepmc.org



www.europepmc.org/downloads



www.europepmc.org/AnnotationsApi

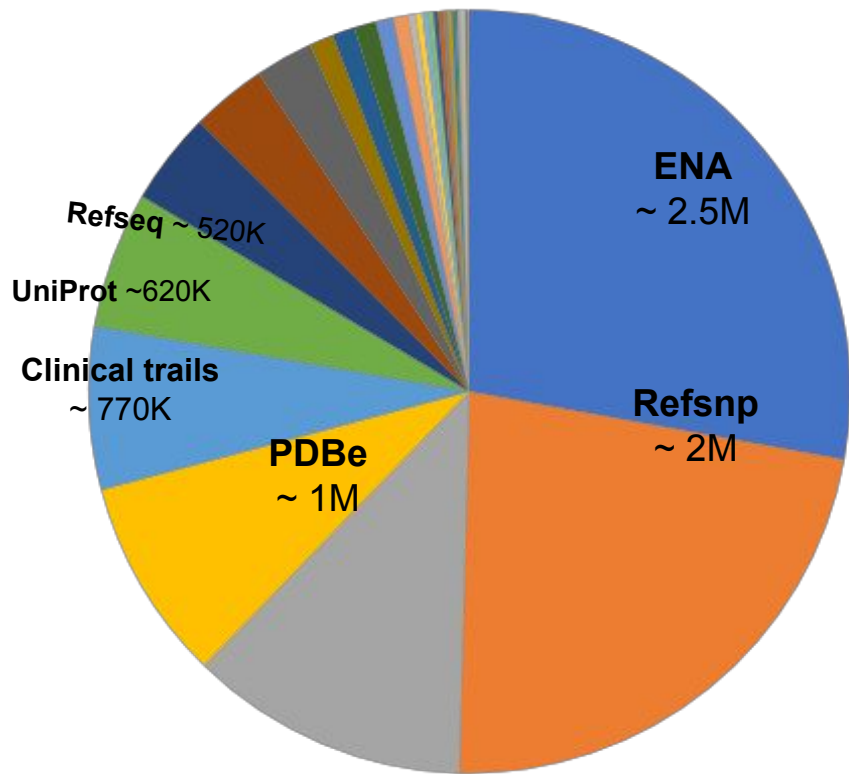
User Stories

Metrics for monitoring resource usage

- Accession number mentions
- Data citation patterns
- Aids in sustained funding



Database mentions in full text articles



~ 9 million database
accession numbers



User Stories



Acta Crystallogr F Struct Biol Commun. 2019 Nov 1; 75(Pt 11): 665–672.

Published online 2019 Nov 5.

doi: [10.1107/S2053230X1901210X](https://doi.org/10.1107/S2053230X1901210X)

PMCID: PMC6839820

PMID: 31702580

Automatic annotation of protein residues in published papers

Robert Firth,^{a,*} Francesco Talo,^b Aravind Venkatesan,^b Abhik Mukhopadhyay,^b Johanna McEntyre,^b Sameer Velankar,^b and Chris Morris^a



<https://about.west-life.eu>



Europe PMC

Acknowledgements

Team members



Melissa Harisson



Zunaira Shafique



Aravind Venkatesan



Santosh Tirunagari



More questions? Contact us

My contact details

avenkat@ebi.ac.uk

Contact us

annotations@europemc.org



@EuropePMC_News
@aravindvenkates

Thank you!



Europe PMC