EOSC-Life

Cloud deployment of FAIR resources: Success stories from the EOSC-LIFE RI community

OmicsDI Cloud

Henning Hermjakob, EMBL-EBI
hhe@ebi.ac.uk

Project description

Develop a cloud-deployed production version of the Omics Discovery Index (OmicsDI), to support efficient integration with cloud-based workflows. In addition, we propose to enable local satellite deployments, to demonstrate OmicsDI use as a community coordination tool for the European and international systems biology community, and to provide API access to OmicsDI impact metrics.

- D 1.1: Production deployment of OmicsDI on EOSC cloud infrastructure.
- D1.2: Open source provision of OmicsDI for local satellite deployment, with optional synchronisation with the OmicsDI public instance.
- D 2.1: Integration of FAIRDOMhub and Virtual Parts Repository into the OmcisDI public cloud deployable instance (D1.1).
- D2.2: BioHackathon, COMBINE.org (or similar event 2021) project for inclusion of additional systems and synthetic biology resources into OmicsDI.
- D 2.3: Infrastructure (API and Web) to enable grant agencies and journals to retrieve the impact metrics of each dataset deposited in OmicsDI. Similar to Altmetrics, this will allow to display a version of the current OmicsDI “rosette” on external websites, or retrieve the data for further external processing.

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 824087

OmicsDI.org

Omics Discovery Index

Consistent dataset discovery across 23 repositories from multiple omics types

D1.1 still in progress: Kubernetes deployment currently in refactoring

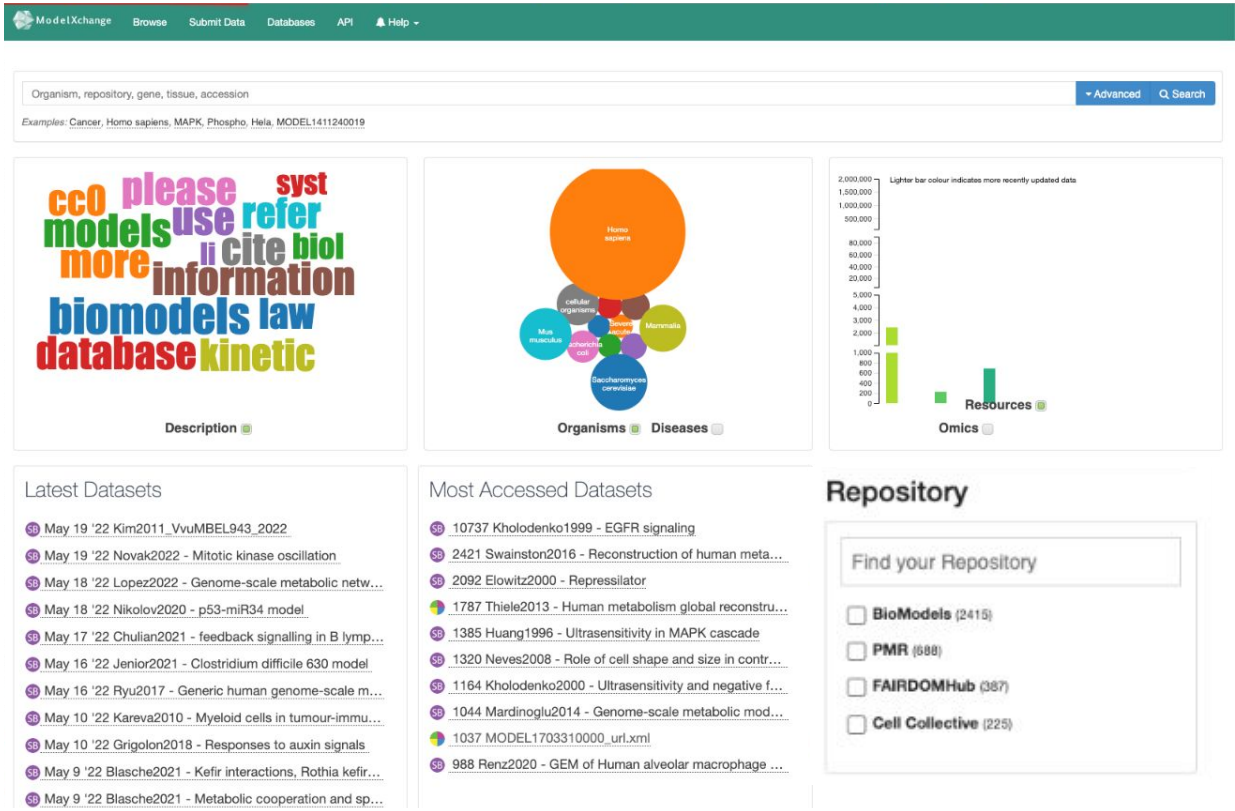
This project has receiv
2020 research and inn

Satellite Deployment: ModelXchange

Created satellite instance for systems biology (D1.2).

Included three new repositories (PMR, FAIRDOMHub, Cell Collective) (D2.1)(D2.2).

Supports cross-repository search in fragmented systems biology domain, increasing FAIRness.



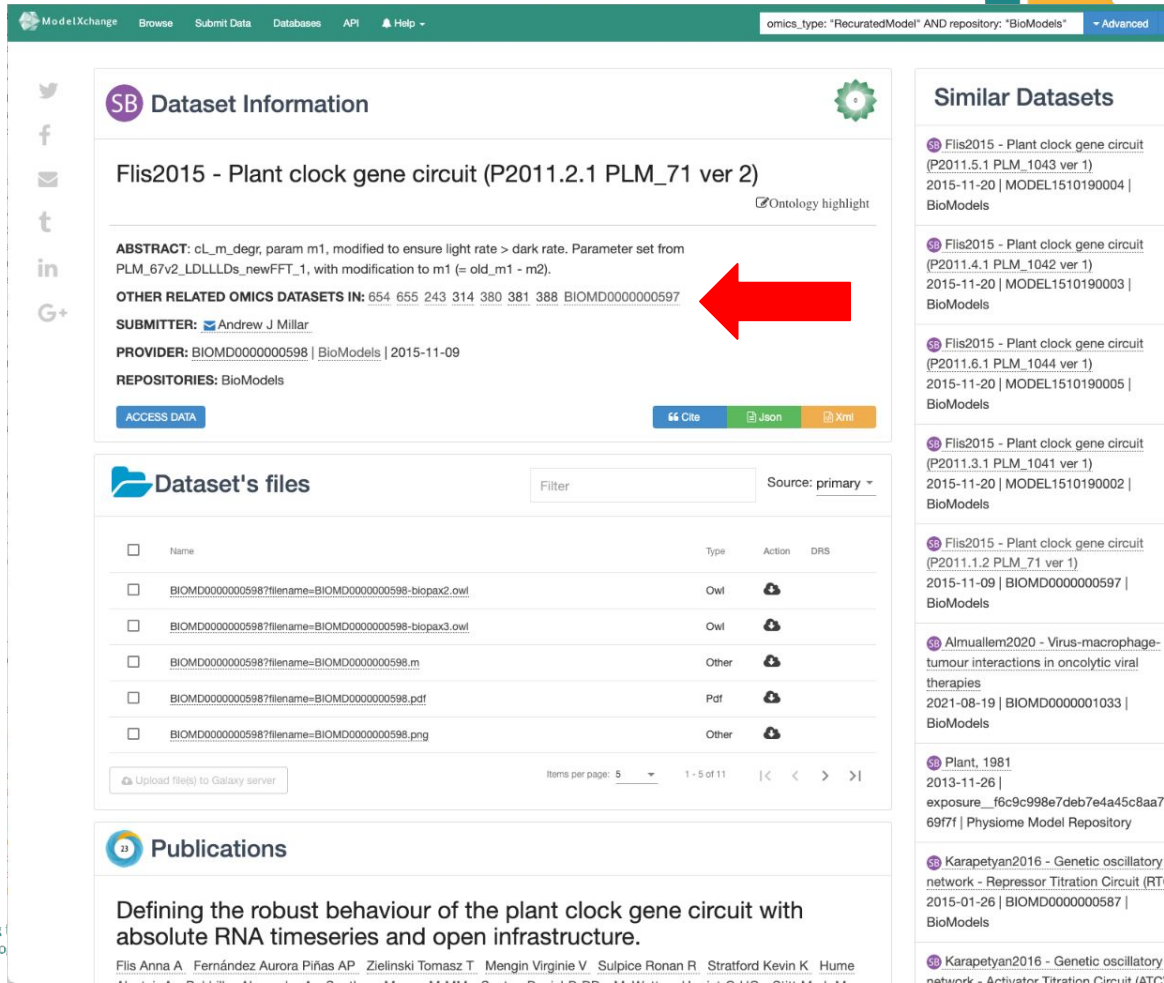
This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824087

http://wwwdev.ebi.ac.uk/Tools/omicsdi/

Redundancy detection

Based on shared publication reference, we can identify systems biology models closely related or duplicated across multiple repositories, and reference them.

ModelXchange consortium aims to avoid redundant curation.



This project has received funding
2020 research and innovation pro

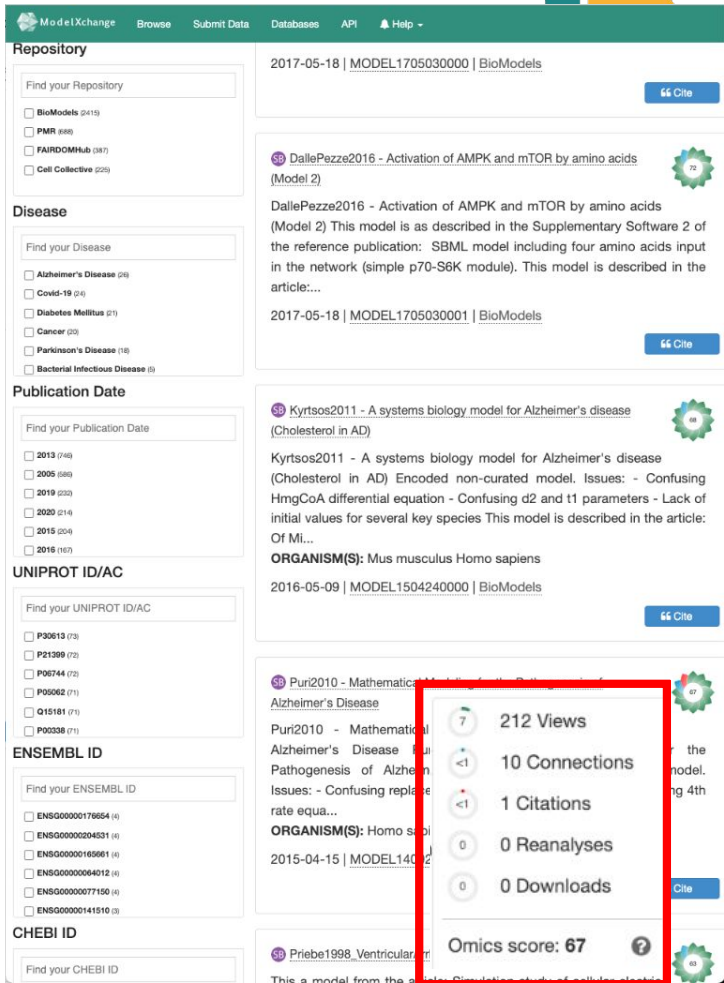
Impact Metrics

OmicsDI provides prototype impact metrics for data based on views, downloads, citations, and references among data objects.

Developed API to allow third parties to use these metrics in their sites (D2.3).

Deployed by BioModels and PRIDE resources.

Potential for use by authors, organisations, funders, etc.



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824087

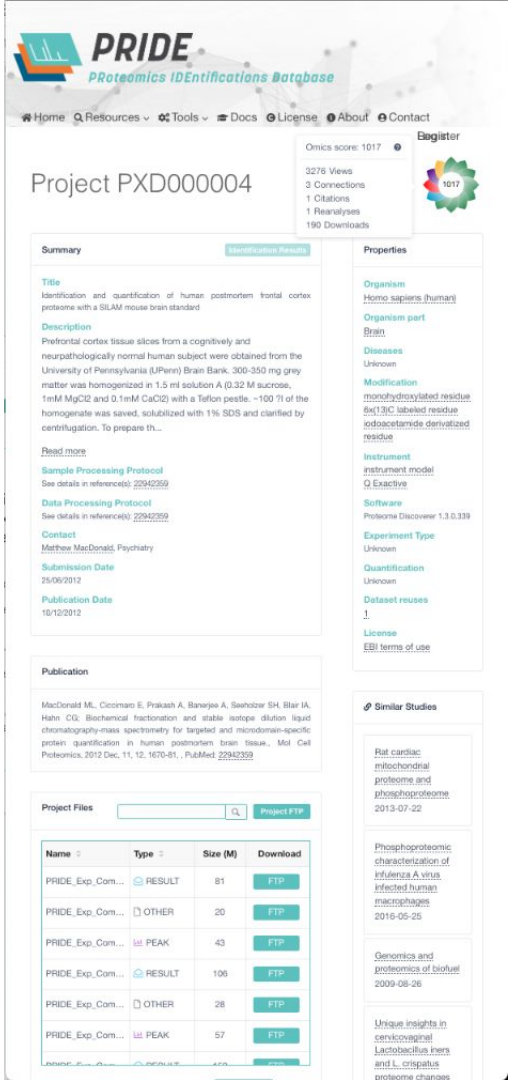
Impact Metrics

OmicsDI provides prototype impact metrics for data based on views, downloads, citations, and references among data objects.

Developed API to allow third parties to use these metrics in their sites (D2.3).

Deployed by BioModels and PRIDE resources.

Potential for use by authors, organisations, funders, etc.



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824087

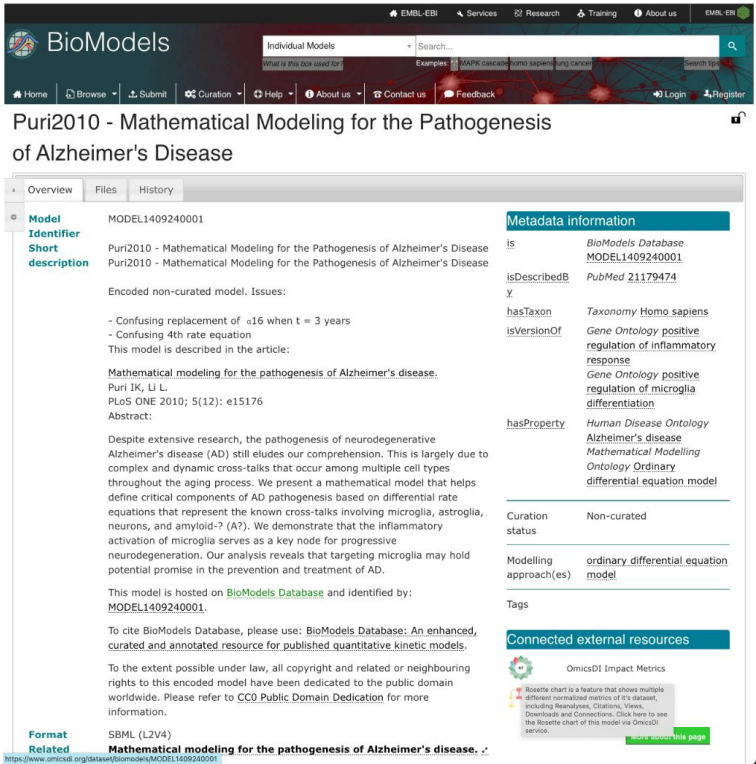
Impact Metrics

OmicsDI provides prototype impact metrics for data based on views, downloads, citations, and references among data objects.

Developed API to allow third parties to use these metrics in their sites (D2.3).

Deployed by BioModels and PRIDE resources.

Potential for use by authors, organisations, funders, etc.



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824087

Impact of the Project on the Research Infrastructure(s)

- Limited technical impact:
  - Implemented “standard” Kubernetes deployment on EMBL-EBI embassy cloud
- Significant community impact:
  - Specialisation of OmicsDI to ModelXchange may have pilot impact for other fields.
  - ModelXchange is expected to have a significant impact on the systems biology community, enabling cross-repository metadata-based discovery.
    - A previous, related effort (ProteomeXchange) contributed to changing data culture in proteomics from “closed” to “open”.
  - Impact metrics for data are a logic follow-up to the FAIR data efforts.
    - Long term indirect impact.

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 824087

Impact of the Project on EOSC

- Multi-omics, multi-repository metadata discovery and metrics may have pilot function for EOSC.
- Specialisation of multi-omics infrastructure to valuable tool for specific omics field is a “simplified case study” for the EOSC challenge of providing general infrastructures which are useful for specific domains without too much adaptation overhead.

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 824087

Experience of working in EOSC-LIFE and technical teams

- Less interaction with main project than we could have had.
  - Approached as technical project
  - Regular EOSC-LIFE meetings were “just another meeting”, and so often skipped (by me)
  - To look beyond my own immediate project was challenging as a result.
- Focus on existing EBI infrastructure, rather than “broadening out”
  - Easy, but perhaps not best choice

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 824087

Future work/Sustainability of the project outcome

- Kubernetes deployment of OmicsDI still ongoing, but will be the long term deployment of choice in a hybrid cloud environment which many organisations use/develop.
- ModelXchange is an informal, but active collaboration of systems biology data resources. The presented work is a prototype that will be sustained and further developed in a collaborative context, and support grant applications.
- Dataset impact metrics are a prototype, supporting efforts to value scientific output as more than just publications. Long term perspectives: Exploratory.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824087

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824087

Acknowledgements

- EMBL-EBI:
  - Gaurhari Dass
  - Robert Petryszak
  - Yasset Perez-Riverol
  - Tung Nguyen
  - Helen Parkinson
- Phil Gribbon, ITMP, Germany
- Wolfgang Müller, HITS gGmbH
- Tomas Helikar, University of Nebraska
- David Nickerson, University of Auckland