

# Managing food and microbiome studies data using Fairspace, a flexible and FAIR data management platform

Ewelina Grudzien<sup>1</sup>, Eelke van der Horst<sup>1</sup>, Frank van den Bergh<sup>1</sup>, Maria H. Traka<sup>2</sup>, Duncan Ng<sup>2</sup>, Falk Hildebrand<sup>2</sup>, Chris T. Evelo<sup>3</sup>, Susan L. M. Coort<sup>3</sup>, Duygu Dede Şener<sup>3</sup>, Elisa Cirillo<sup>1</sup>

<sup>1</sup> *The Hyve B.V., Arthur van Schendelstraat 650, Utrecht, The Netherlands*

<sup>2</sup> *Quadram Institute Bioscience, Norwich Research Park, United Kingdom*

<sup>3</sup> *Department Bioinformatics – BiGCaT NUTRIM, Maastricht University, Maastricht, The Netherlands*

## Abstract

Fairspace is an open-source research data management platform that adheres to the FAIR principles. The Hyve created Fairspace in 2016 and has developed it ever since, customizing it for several organizations' use cases. We present the implementation of this tool within the FNS-Cloud consortium, in which Fairspace became the user browser that allows microbiome and food data exploration within public resources mapped to a common (meta)data model and vocabularies/ontologies.

## Keywords

FAIR, Research Data Management, Microbiome, Food & Nutrition Security, Fairspace

## 1. Introduction

[Fairspace](#) is an open-source research data management platform that adheres to the FAIR principles. It offers a collaborative environment for managing research data (files, collections) and a metadata catalog with flexible data model and search interface. It is built on semantic web technologies, such as RDF and SHACL, and it is FAIR-by-design. [The Hyve](#) created Fairspace in 2016 and has developed it ever since, customizing it for several organizations' use cases, such as for [Institut Curie](#), a cancer research hospital in Paris, and the Food Nutrition Security Cloud (FNS-Cloud) microbiome use case, which is presented here.

## 2. The FNS-Cloud challenges

One of the ambitions of the FNS-Cloud consortium is to reduce the fragmentation of food & nutrition security (FNS) research resources such as datasets and tools. Most of these resources have been developed independently, with different interfaces or data formats. Therefore, the FNS-Cloud project aims to set up a first-generation food & nutrition security cloud that integrates already existing research resources and shares these in a uniform manner in the cloud. Several demonstrators are planned in the project in order to provide the use cases for guidance to properly develop and integrate the FNS-Cloud tools. The focus for Fairspace development concerns a Microbiome Demonstrator, in which a researcher wants to find studies regarding diet interventions and wants to combine them with gut microbiome data for further analysis. In addition, the consortium partners want to leverage the collaboration with the [ELIXIR Infrastructure](#), which already includes several European food and nutrition resources. However, ELIXIR does not have a solution for food and microbiome data yet that allows users to find those data across multiple public data sources for analysis. With this focus the Hyve designed a customized solution using Fairspace as the main component that could potentially fill this gap.

---

14th International SWAT4HCLS Conference, February 13–16, 2023, Basel, Switzerland

✉ ewelina@thehyve.nl (A. 1); eelke@thehyve.nl (A. 2); frank@thehyve.nl (A. 3); maria.traka@quadram.ac.uk (A. 4); Duncan.Ng@quadram.ac.uk (A. 5); Falk.Hildebrand@quadram.ac.uk (A. 6); chrisevelo@maastrichtuniversity.nl (A. 7); susan.coort@maastrichtuniversity.nl (A. 8); d.dedesener@maastrichtuniversity.nl (A. 9); elisa@thehyve.nl (A. 10)



© 2023 Copyright for this paper by its authors.  
Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

CEUR Workshop Proceedings (CEUR-WS.org)

### 3. Solution Specification

There are several aspects of FairSPACE that make it the right solution for the above challenges. Firstly, the tool uses a (semantic) metadata model which was developed by the consortium to facilitate semantic data integration: FNS-Harmony. In particular, FNS-Harmony was used to map data source-specific fields and entities to classes and attributes in the ontology. In addition, The Hyve developed a SHACL model based also on specific ontologies and user requirements. FairSPACE uses this model as its content data model, for validation (data integrity) and user interface generation. With the above aspects the (meta)data search can be performed both by users and machines.

Secondly, The Hyve implemented in FairSPACE a set of ETL (extract, transform, load) processes working together in order to fetch data from several public sources such as: [ENA](#), [MGnify](#), [MetaboLights](#) and [dbNP](#). The data is mapped to the common data model defined in FairSPACE, using selected ontologies and vocabulary. Lastly, a JupyterHub environment integrated with FairSPACE allows users to analyse metadata and data using R, Python or Julia. It includes direct access to a shared FairSPACE storage of predefined scripts for data pre-processing and analysis, as well as metadata selected in FairSPACE search interface. Users can access the harmonized metadata through sparql and custom API and (file) data.

This work is currently still in progress since the FNS-Cloud project will run until September 2023, and the ETL customization is not open source.

### 4. Conclusions

FairSPACE has proven to be an easy-to-use open-source research data management tool that is able to resolve different types of data management challenges faced by different organizations, and it does it in a way that follows the FAIR principles.

In particular, for the FNS-Cloud consortium, FairSPACE became the user browser that allows microbiome and food data exploration within public resources mapped to a common (meta)data model and vocabularies/ontologies. Using the built-in faceted search interface the researcher can export or analyze selected data in a secure analysis environment such as Jupyter Notebook. Finally, in line with the open source philosophy The Hyve takes care to develop code which is well readable, easy to maintain and with a high test coverage.

### 5. Acknowledgements

Food Nutrition Security Cloud (FNS-Cloud) has received funding from the European Union's Horizon 2020 Research and Innovation programme (H2020-EU.3.2.2.3. – A sustainable and competitive agri-food industry) under Grant Agreement No. 863059 – [www.fns-cloud.eu](http://www.fns-cloud.eu)

We thank all the other partners of FNS-Cloud that have contributed to discussions and collateral work related to the Microbiome Demonstrator usecase: Institut Jozef Stefan, University of Florence, Premotec, and ScaleFocus.

