

CASE STUDY



FAIRplus

FAIRplus use case IMI COMBINE: The importance of data standards

Authors: Erwin Boutsma, Ilse Custers, Jan-Willem Boiten,
Philip Gribbon, Rake! Arrazuria, Yojana Gadiya, Jon Ulf Hansen



Challenge

To be able to combat pathogenic drug-resistant microbes, the world needs new antimicrobial tools and treatments. Better use of existing data is essential to stimulate this. However, many of the current AMR data is hardly reusable due to lack of interoperability.

Solution

The so-called FAIRplus squad team identified that a good ontology is lacking for AMR *in vivo* models. After thorough investigation, the team decided to develop a new application ontology that combined several existing ontologies.

Overview

A growing number of microbial infections that used to be treatable with antibiotics, now again pose a threat to those infected. The reason for this is antimicrobial resistance (AMR): some pathogenic microbes have acquired defense mechanisms against existing antimicrobial drugs.

The European [antimicrobial resistance \(AMR\) Accelerator program](#) comprises several projects with the shared goal of progressing the development of new medicines to treat or even prevent resistant bacterial infections in Europe and worldwide. The COMBINE project was created to coordinate the AMR Accelerator projects and provide them with the resources needed to achieve their goals. These include data management guidelines and an IT infrastructure to enable the collection, aggregation, storage, sharing and analysis of datasets generated by AMR Accelerator projects.

The IMI FAIRplus project aims to develop tools and guidelines for making life science data FAIR (Findable, Accessible, Interoperable, Reusable). In the past year, the so-called 'squad teams' from FAIRplus, consisting of experts working in universities and pharmaceutical companies, have been actively working to FAIRify data sets from large IMI projects, such as APPROACH, eTOX, RESOLUTE, and COMBINE. The developed tools and methods are subsequently added as 'recipes' to the FAIR cookbook, enabling projects and companies with similar FAIR data challenges to apply this consolidated know-how to increase the FAIRness of their data.

Aims and impact

The IMI projects to be supported by FAIRplus are selected in an [unbiased process](#), based on a recorded set of criteria, among others societal impact. The aim of FAIRifying project data is to facilitate sharing data outside the consortium to conduct further research.

COMBINE was selected as one of those projects to be FAIRified by the FAIRplus squad teams. COMBINE is a centralized hub which supports the AMR community and is important for scientific capacity building of the nine AMR Accelerator projects. FAIRplus work package leader **Philip Gribbon** explains: '*Initially, we were looking for a protocol that would describe precisely how the in vivo experiments were performed. We tested over ten different ontologies that were available for AMR, but they were all focused on different domains, and captured different aspects of the protocol. We wanted a single ontology that could do all of this. Also, we couldn't get the terms of those protocols in a machine-readable format.*'

The squad team decided to develop a new ontology to improve the consistency of the experimental metadata. At the moment, the ontology is being developed. A template is already in place functioning as a 'data dictionary', as Gribbon puts it. The new ontology is expected to be ready by the end of 2022. It will then be published and made available to other AMR projects dealing with *in vivo* data.

As it turned out, there is also an important additional effect of the new ontology. *'If you have good quality studies with consistent metadata, you are likely to need less animals'*, COMBINE researcher **Rakel Arrazuria** states. *'Reproducibility is a huge problem in preclinical studies. The new ontology aims to improve consistency within and between data sets and improve the interoperability of preclinical study metadata. This will undoubtedly increase reproducibility, thus decreasing the number of animals needed.'*

The FAIRification of the AMR Accelerator *in vivo* experiments is an inspiring example of how two different scientific needs were met, **Gribbon** concludes. *'There was a gap on the COMBINE side in a sense that we needed a new ontology, and there was a gap on the FAIRplus side that we did not have an example for the [FAIRplus Cookbook](#) about creating an ontology.'* He adds that coming up with a new ontology could only work if you bring technical experts and scientific domain experts together. *'You need to establish whether the gap between what you have and what you want is wide enough to justify the substantial amount of time needed to develop a new ontology.'*

The benefit



The new ontology for *in vivo* studies developed for the IMI COMBINE project will result in a more consistent approach for applying metadata in *in vivo* experiments within COMBINE, greatly enhancing usability of the data. As an added benefit, the new ontology thus supports the goal to reduce animal use in biomedical research.

Also, the new ontology can be applied to other *in vivo* studies within the AMR Accelerator projects. It is expected that broad adoption of the new standardized ontology will contribute to new solutions for AMR, and a more efficient translation of AMR research to clinical practice.

