New framework for making data findable, accessible, interoperable and reusable

June 29 2023, by Oana Stroe

Credit: Karen Arnott/EMBL-EBI

The FAIRplus consortium, including EMBL-EBI and pharmaceutical companies, has launched a new framework to help organizations make their data FAIR.

The COVID-19 pandemic highlighted the weak points in molecular data

sharing on a massive scale. These include service incompatibilities, data access restrictions, missing data or metadata. Such issues delayed research, resulted in duplication of efforts, and hindered the pandemic response.

Despite FAIR data management principles (findable, accessible, interoperable, reusable) being a priority for many research funding organizations, including the European Commission's Horizon Europe program, and despite the known benefits of FAIR data sharing, detailed technical guidance for organizations is still lacking.

To address the gap, the FAIRplus consortium—a unique initiative that brought together partners from academia, including EMBL-EBI, and major pharmaceutical companies—launched a new FAIR Cookbook and FAIRification framework, supporting organizations and projects undertaking a FAIR transformation.

The framework is a major collaborative project, developed with over 17 data-producing research projects from the Innovative Medicines Initiative (IMI), which also funds the FAIRplus consortium. These collaborations enabled the authors to apply the framework to clinical interventional study datasets and data generated in the lab to elucidate molecular interactions, as well as to real-world and clinical observational data.

"There is a lot of goodwill and energy in the life science community to make research data easier to access and use," explained Tony Burdett, Team Leader at EMBL-EBI. "But the journey can be daunting. We hope that our FAIRification framework will guide organizations in academia and the private sector looking to start a FAIR transformation and improve their capabilities for data sharing, access, and reproducibility. These efforts are truly worthwhile because they benefit the community, and society as a whole."

The power of the framework lies in providing a template that is:

- easy to apply, establishing processes that are applicable to any dataset, as well as disciplines beyond the life sciences
- easy to adapt depending on an organization's specific processes and resources
- reproducible, making it a sustainable solution for data producers.

"Data that can't be found is data that can't be reused," said Melanie Courtot, Director of Genome Informatics at the Ontario Institute for Cancer Research. "The FAIRification framework provides practical, actionable steps for data owners to increase the value of their work and make their results available to the community. Standardized interoperable data benefits the global research community, who can link different projects and outcomes, and ultimately benefits human health."

"Making research data findable, accessible, interoperable, and reusable is essential for speeding up the pace of discovery. Most private companies are actively working on this, and still have much to gain from this approach," explained Herman Van Vlijmen, Head of Computer-Aided Drug Design at Janssen. "This FAIRification framework is a great starting point to guide our efforts."

The paper is published in the journal *Scientific Data*.