BIOINFORMATICS

FAIR your data

A toolkit facilitating FAIRness assessment.

Digital scholarly resources such as protein structure data and bioinformatics software are invaluable resources for the community. The FAIR Guiding Principles — standing for findable, accessible, interoperable and reusable — aim to serve as a guideline for data management and stewardship. Although widely acknowledged, the implementation of FAIR in the real world is often not straightforward. On top of that, “there is currently no real incentive for biomedical data producers to spend a lot of effort in making the data they produce FAIR,” says Avi Ma’ayan, Icahn School of Medicine at Mount Sinai in New York.

A common task when evaluating FAIRness of a digital resource is finding a suitable set of criteria. Given the immense and evolving diversity of biomedical data types, this may not be easy, and sometimes there is little existing consensus in a field. Ma’ayan and colleagues decided to take a democratic approach. In FAIRshake, a tool they have developed for FAIRness evaluation, a user has the flexibility to associate selected FAIR metrics (questions that assess whether a certain aspect of FAIR is followed) and rubrics (a set of FAIR metrics) with the object being evaluated. Existing metrics collected by FAIRshake are provided as a first choice for FAIRness evaluation, in the hope that “reusing FAIR metrics will increase community interoperability and convergence on what it means to be FAIR,” comments Ma’ayan. By this design, it aims to encourage democratization while also nurturing community consensus. In his words, “while there are some people who aim to become the authority on what it means to be FAIR, the FAIRshake platform aims to give the community the power to decide. It also enables different communities to establish different FAIR standards, highlighting what is most important to improve for their community.”

A second key strength of FAIRshake lies in its usability and automation. It is a toolkit consisting of different components, such as a search engine, API (application programming interface), YouTube or

Jupyter tutorials, bookmarklet and browser extension. Collectively they build a succinct and well-defined workflow for FAIRness evaluation for a single or a collection of digital objects under the umbrella of a project. After registering digital objects and associating them with FAIR metrics or rubrics, assessment can be performed in manual or automatic manner. “Currently, the FAIRshake framework/toolkit is the most robust and flexible implementation for assessing FAIRness,” says Ma’ayan. “It is an open-source project that encourages the community to use it as well as contribute to it.”

Along with the evaluation results, FAIRshake also generates an insignia for an evaluated object that represents the FAIR score in the form of a grid of red, blue and purple squares. This FAIR insignia can further be embedded within the websites hosting those digital resources. As case studies, Ma’ayan’s team used FAIRshake to perform FAIRness assessment to digital objects belonging to a number of high-profile projects such as the Alliance of Genome Resources (AGR) and dbGAP. While some aspects of the FAIR principles were overall well met, there are other areas that warrant improvement.

While FAIRshake was developed to facilitate FAIRness evaluation, its real purpose is simpler. “The goal of FAIRshake is to make people aware of things that they may overlook. So the most important part is to see if FAIRshake helps a person to add or implement improvements to the digital objects they serve or produce in a way that would make it easier to reuse, find, integrate with other resources, and understand these digital objects by others,” says Ma’ayan. Thus by making the evaluation process robust and easy, it reveals ways to improve FAIRness.

Lin Tang

Published online: 4 February 2020

<https://doi.org/10.1038/s41592-020-0742-y>

Research paper

Clarke, D. J. B. et al. FAIRshake: toolkit to evaluate the fairness of research digital resources. *Cell Syst.* 27, 417–421 (2019).

**Working together
we open the potential
of scientific innovation**

Our enzymes have been discovering facts about life for decades.

That’s why Worthington, the primary enzyme producer, has more citations in respected journals like *Nature Methods* and rank higher in search engines like *Bioz*®.

COLLAGENASES → 99% Bioz Rating

DNASE I → 99% Bioz Rating

PAPAIN → 99% Bioz Rating

Unlock the power of Worthington with our new **Introduction to Enzymes** guide. Go to: Worthington-Biochem.com