OpenEBench 2023 update

better support and deeper interactions with benchmarking scientific communities

SCAN HERE
openebench.bsc.es

Carles Hernandez-Ferrer[1,2], Laura Portell-Silva[1,2], Anna Redondo Guitarte[1,2], Esther Vendrell[1,2], Laia Codó[1,2], José M. Fernández[1,2], Dmitry Repchevsky[1,2], Lidia López[1,2], Alfonso Valencia[1,2,3], Josep Ll. Gelpí[1,2,4], and Salvador Capella-Gutierrez[1,2]

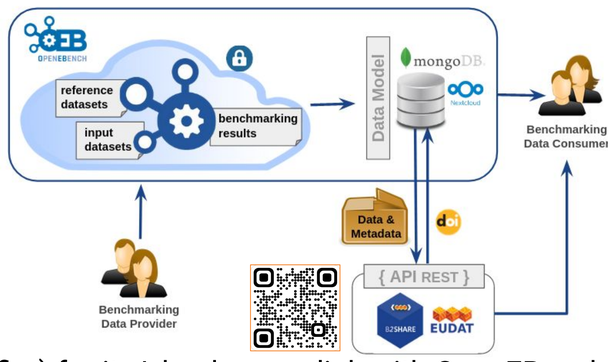
(1) Spanish National Bioinformatics Institute (INB/ELIXIR-ES), Barcelona, Spain; (2) Barcelona Supercomputing Center (BSC), Barcelona, Spain; (3) Catalan Institution for Research and Advanced Studies (ICREA), Barcelona, Spain; (4) Dept Biochemistry and Molecular Biomedicine, University of Barcelona (UB), Barcelona, Spain

OpenEBench (https://openebench.bsc.es/) is the ELIXIR platform for supporting community-led benchmarking activities and the technical monitoring of bioinformatics software. It has been part of the ELIXIR Tools Platform since its inception, contributing to its development and consolidation. OpenEBench collaborates with ***more than 20 different scientific communities within and beyond ELIXIR***. This interaction fosters the development of new components, both at the infrastructure level and extension of the visual interfaces. The Continuous Automated Model EvaluatiOn (CAMEO) runs weekly automated benchmarks on predictions of protein structures that are published in OpenEBench. The Intrinsically Disordered Proteins (IDP) has served to bring CAID into OpenEBench evaluating and improving the existing documentation. Our roadmap includes to integrate other relevant platforms, and we expect to work closely with APICURON to ***visibilize and credit the contributions*** made by the different members of the communities to the platform.

OpenEBench benchmarking engagement

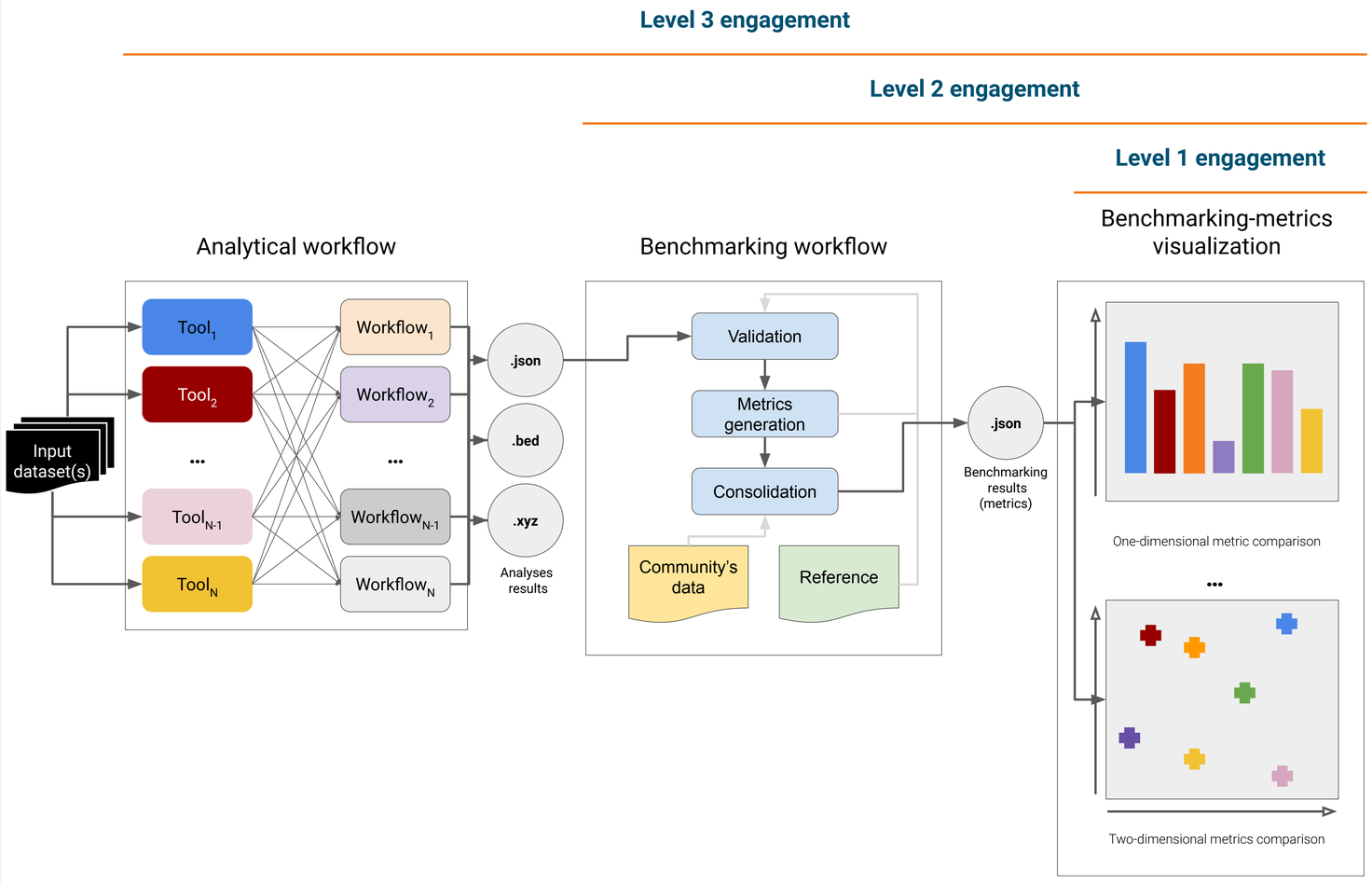
- ★ **Level 1**
  - Long-term storage of benchmarking results
  - Aiming at reproducibility and provenance
- ★ **Level 2**
  - The community defined a benchmarking workflows
  - Workflows compute one or more evaluation metrics given one or more reference datasets provided by the community
- ★ **Level 3**
  - Whole benchmarking process is performed at OpenEBench
  - Participants' tools that are run and evaluated by OpenEBench

Where is my data?



OpenEBench uses **EUDAT (B2SHARE)** for long-term storage of scientific benchmarking datasets. With one click, a user can export OpenEBench datasets to B2SHARE, which will mint a versioned DOI (digital object identifier) for it. A back cross-link with OpenEBench identifiers is kept so that published benchmarking data can be consumed from both B2SHARE and OpenBench platforms.

Benchmarking process and engagement



Analytical workflow

Along this step the communities provide all the input dataset(s), the tools to be benchmarked and the analytical workflows. OpenEBench runs them all on an agnostic environment for maximum reproducibility.

➔ This step is only perform when the communities have a *level 3 of benchmarking engagement*.

Benchmarking workflow

In this step, the community provide participants' results of the analytical workflows and OpenEBench runs the benchmarking workflow. The first sub-step, ***validation***, validates the prediction's file format. Secondly, during the ***metrics computation***, the predictions are compared with the reference datasets and the community's metrics are computed. Finally, during the ***consolidation***, the results of a particular tool are merged into the community's data.

➔ This step is perform when the communities have a *level 2 or level 3 of benchmarking engagement*.

Benchmarking-metrics visualization

The community submits the predictions' metrics and OpenEBench is in charge of its visualization.

➔ This step is perform when the communities have a *any level of benchmarking engagement*.

How to start?

For Level 1 of engagement

★ Adapt the data to the OpenEBench data model.

For Level 2 and 3 of engagement

★ **Scope of the benchmark:** the relevant scientific questions that the benchmarking event has to answer.
★ **Level of engagement:** what level of engagement would be useful for the community to adopt.
★ **Evaluation metrics:** what metrics are going to be used to evaluate the performance of the different tools/pipelines?
★ **Reference datasets:** which datasets can be useful as gold standards for the benchmarking evaluation?
★ **Tools to evaluate:** what pipelines are going to be part of the comparative evaluation?
★ **Benchmarking workflow:** once all the above is organized, the benchmarking workflow that will run in OpenEBench is developed.

Communities involved

ELIXIR communities: hCNV, Rare Diseases, Proteomics, 3D Bioinfo, IDP, SCO

Carles Hernandez-Ferrer
Senior Research Engineer
carles.hernandez2@bsc.es

Barcelona Supercomputing Center (BSC)
Spanish National Bioinformatics Institute (INB/ELIXIR-ES)

The EUCAIM project is co-funded by the European Union's Digital Europe Programme under Grant Agreement number 101100633.

The GDI project is co-funded by the European Union's Digital Europe Programme under grant agreement number 101081813.