Scientific communities management in OpenEBench

Fernández JM[1,2], Codó L[1,2], Sundesha V[1,2], Garrayo J[1,2], Valencia A[1,2,3], Gelpí JL[1,2,4], Capella-Gutierrez S[1,2]

[1] Barcelona Supercomputing Center (BSC), Barcelona Spain. [2] Spanish National Bioinformatics Institute (INB), Barcelona Spain. [3] Catalan Institution for Research and Advanced Studies (ICREA), Barcelona Spain. [4] Universidad de Barcelona (UB), Barcelona Spain. Molecular Modeling and Bioinformatics group, Instituto de Investigación Biomédica (IRB Barcelona), Barcelona Spain.
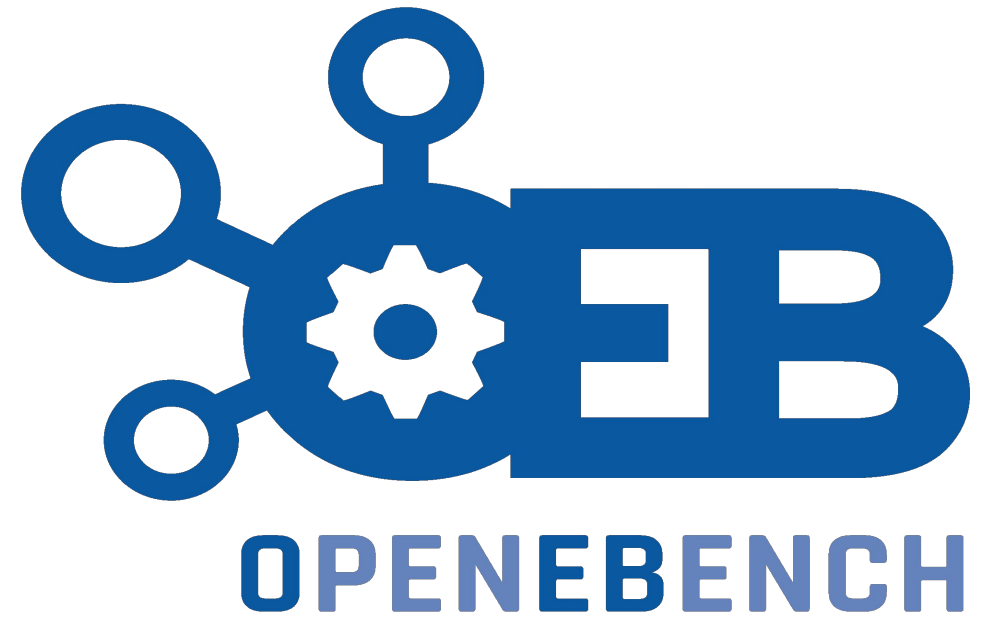
OpenEBench (https://openebench.bsc.es) is the ELIXIR benchmarking and technical monitoring platform for bioinformatics tools, web servers and workflows. The infrastructure provides the resources to enable an objective, reproducible and unbiased comparison of bioinformatic software through metrics-based critical measurements on technical and scientific performance

OpenEBench scientific benchmarking gravitates around three user roles:

Scientific community managers administer and operate benchmarking events for a specific research domain

Researchers objectively compare bioinformatic methods through the community defined metrics

Software developers evaluate their tools against community defined datasets and scientific challenge events

Community managers are an essential figure for incorporating scientific communities into OpenEBench and maintaining its activities up-to-date once the community is integrated into the platform. Thus, OpenEBench strives to provide easy-to-use tools and mechanisms to these privileged users as part of the interactions with its communities.
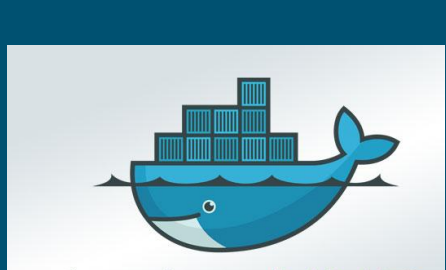
Responsibilities

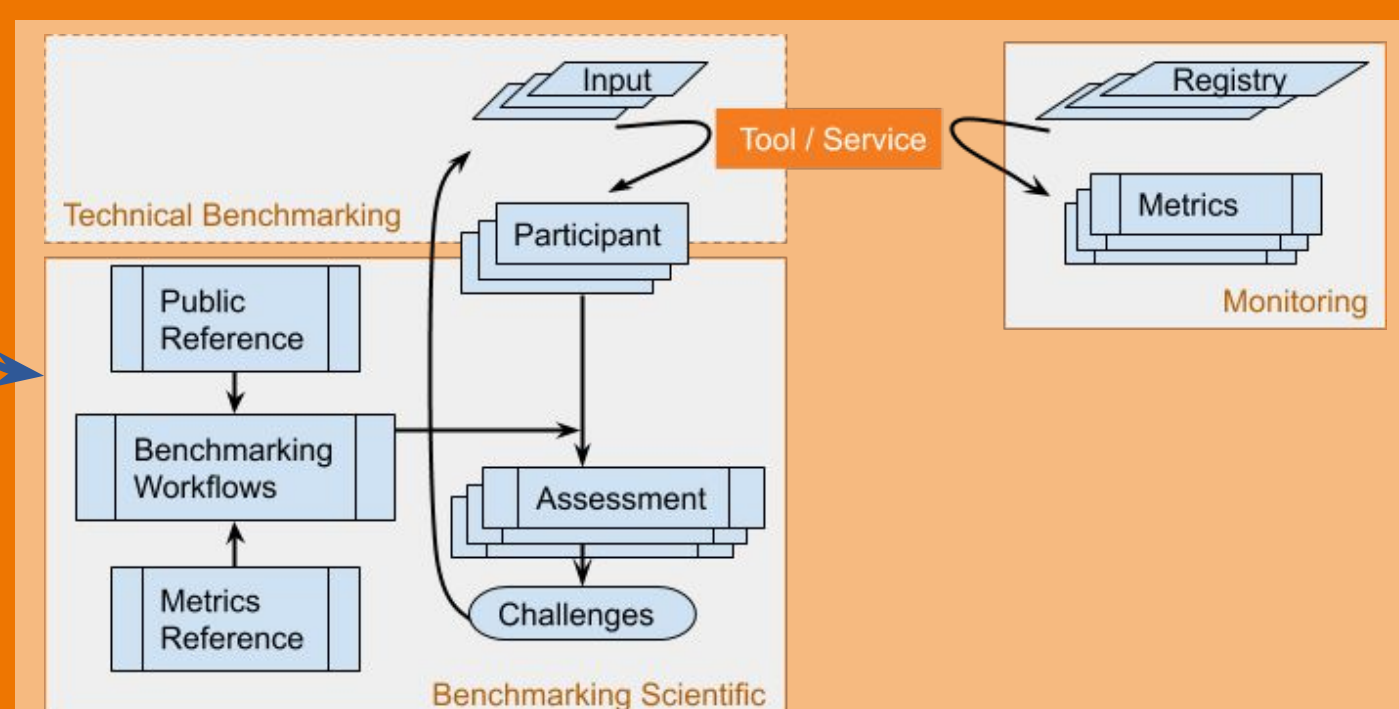1. Design and develop evaluation workflows

Prepare reference datasets
B2SHARE
RESEARCH DATA REPOSITORY

Define & implement scientific metrics
docker HUB

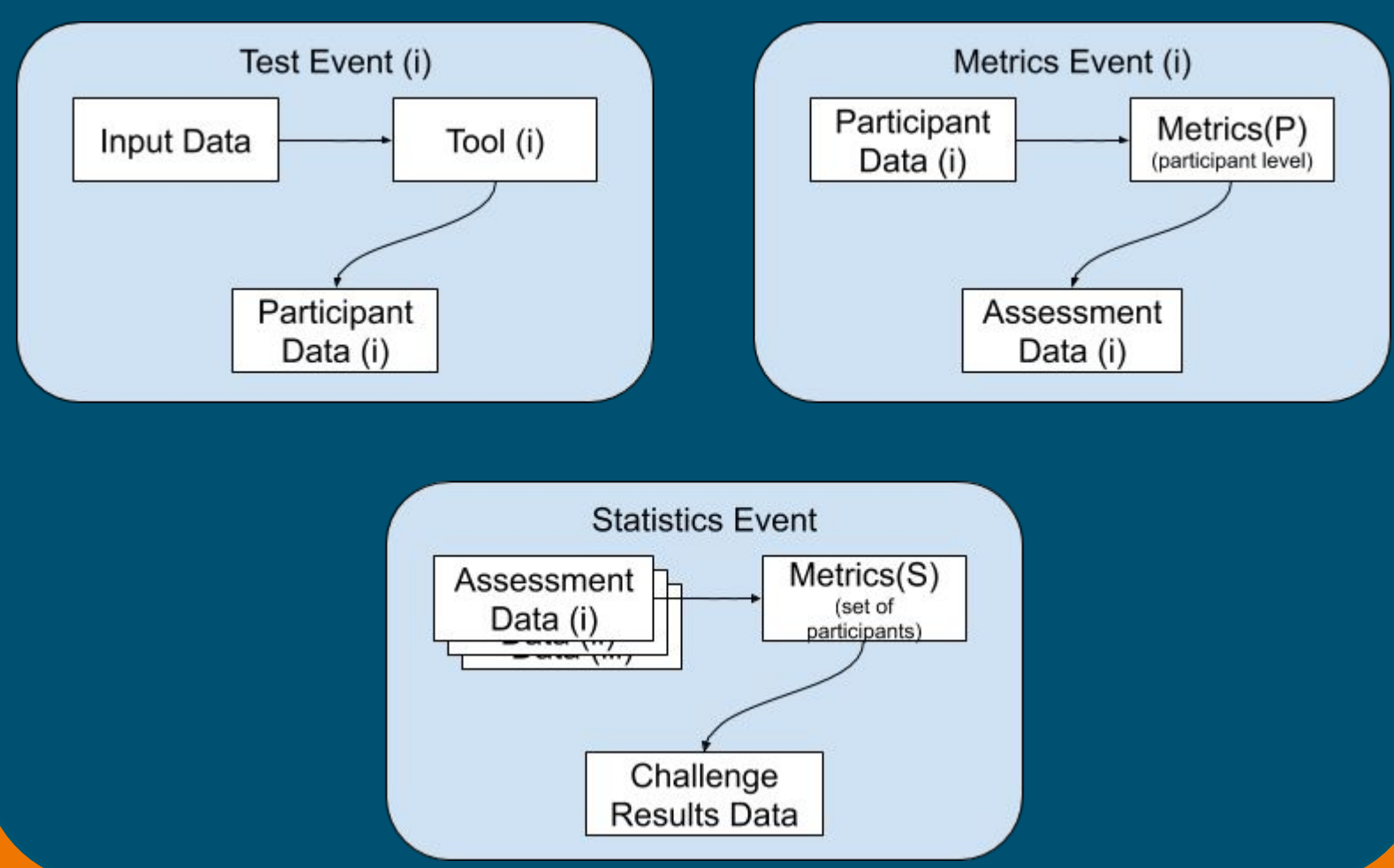Consolidate Benchmarking Workflow
OEB DATA REPOSITORY
nextflow

2. Frame the evaluation workflow into scientific benchmarking challenges

Schedule benchmarking events
OEB Virtual Research Environment

Test Event (i)
Input Data → Tool (i)
Participant Data (i)

Metrics Event (i)
Participant Data (i) → Metrics(P) (participant level)
Assessment Data (i)

Statistics Event
Assessment Data (i) → Metrics(S) (set of participants)
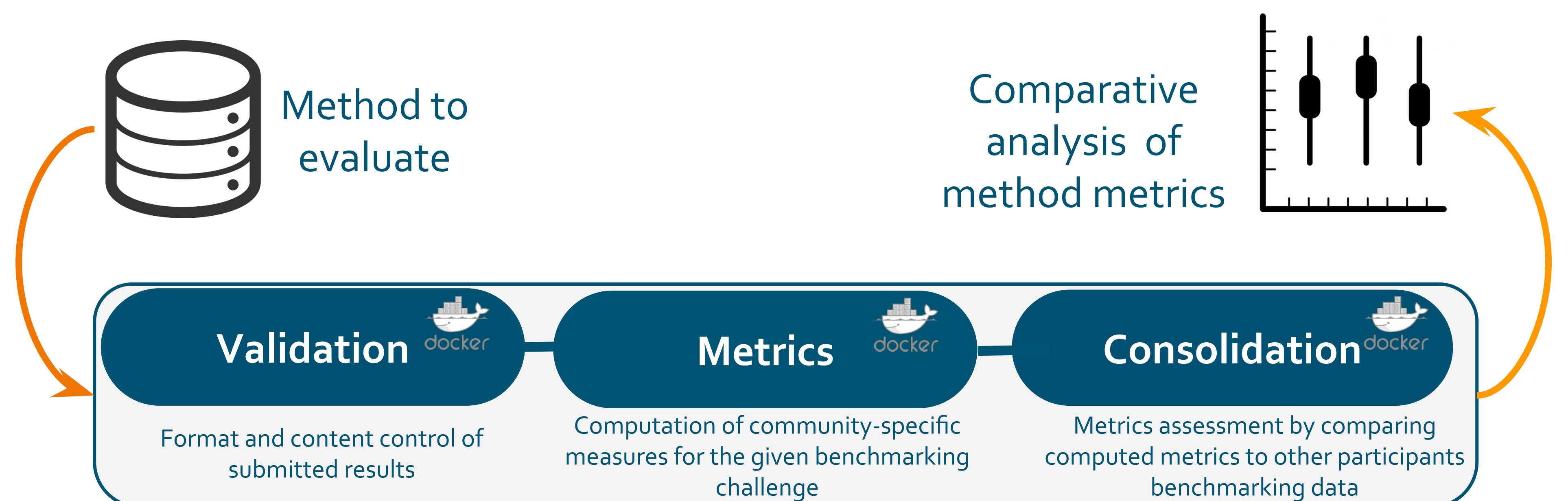Challenge Results Data

Scientific Benchmarking Data Model
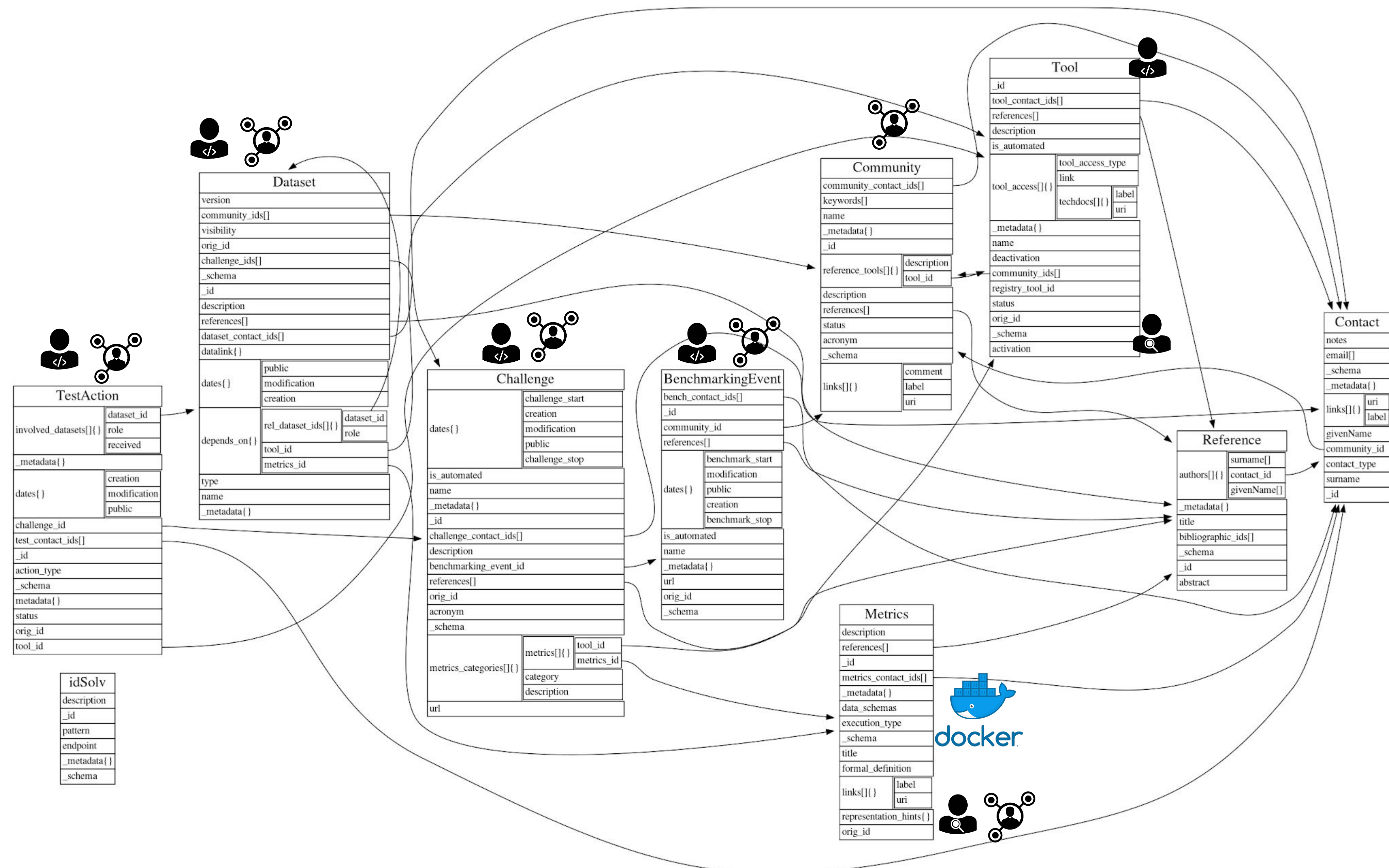
Scientific communities involvement is essential for having real benchmarking activities, as they promote collaboration, transparency and harmonised standards. Importantly, communities control the publication life cycle of the **benchmarking challenges** by which evaluation workflows are made available on the platform. In similar terms, **community managers** also provide different reference datasets, under public or private terms, associated with each challenge. **Reference datasets** are used as input for participants to generate their predictions and as golden reference when evaluating participants input data.

Benchmarking workflow

Method to evaluate

Comparative analysis of method metrics

Validation docker
Format and content control of submitted results

Metrics docker
Computation of community-specific measures for the given benchmarking challenge

Consolidation docker
Metrics assessment by comparing computed metrics to other participants benchmarking data