

OpenEBench opts Bioschemas

Dmitry Repchevsky^{1,2}, Eva Martin Del Pico^{1,2}, Salvador Capella-Gutierrez^{1,2}, Josep Lluís Gelpí^{1,2,3}

¹ Barcelona Supercomputing Center (BSC), Barcelona Spain.
² Spanish National Bioinformatics Institute - Instituto Nacional de Bioinformatica (INB-ISCIII), Barcelona Spain.
³ Universitat de Barcelona (UB), Barcelona Spain.



ELIXIR All Hands , 5-8 June 2023, Dublin, Ireland

Since its inception, the [OpenEBench](#) platform has been providing technical monitoring for the bioinformatics tools collected from a set of sources such as [Biotoools](#), [Bioconda](#), [Galaxy](#) and [DebianMed](#). Being a part of the Research Software Ecosystem initiative, it provides a set of periodically updated FAIR metrics for Research Software that are further committed to the Research Software Ecosystem GitHub [repository](#). Historically, the platform used its own custom model for the tool descriptions which was greatly influenced by the [biotooolsSchema](#) format. Interoperability challenges faced during Research Software Ecosystem participation led to the adoption of the Bioschemas Computational Tool [profile](#) as a format for the research tools description. Here we present a REST API to store Research Tools description using Bioschemas semantics. OpenEBench Research Tools Registry is based on the [Jakarta EE](#) (Enterprise Edition) and uses MongoDB database.

```
{
  "name": "trimAl",
  "description": "Tool for the automated removal of spurious sequences or poorly aligned regions from a multiple sequence alignment."
}
```

{JSON}

```
>curl -k -X PUT -u redmitry:xxx -H "Content-Type: application/json"
http://openebench.bsc.es/bioschemas/tools/biotoools:trimal:1.4/cmd -d @trimal.json
```

PUT application/json

GET application/ld+json

```
{
  "@context": {
    "@import": "https://openebench.bsc.es/bioschemas/oebtools.jsonld"
  },
  "@id": "https://openebench.bsc.es/bioschemas/tools/biotoools:trimal:1.4/cmd",
  "@type": "https://openebench.bsc.es/bioschemas/oebtools#cmd",
  "rdfs:label": "trimal",
  "sc:name": "trimAl",
  "sc:softwareVersion": "1.4",
  "sc:applicationCategory": "Command-line tool",
  "sc:url": "http://trimal.cgenomics.org",
  "sc:description": "Tool for the automated removal of spurious sequences or poorly aligned regions from a multiple sequence alignment.",
  "sc:operatingSystem": ["Linux", "Windows", "Mac"],
  "sc:programmingLanguage": ["C++"],
  "sc:dateModified": "2023-05-23T12:28:32.561Z"
}
```

JSON-LD



The OpenEBench Research Tools Monitoring stores Research Tools descriptions in Bioschemas-friendly JSON format which allows on-the-fly generation of the Bioschemas JSON-LD tools descriptions compliant with Bioschemas Computational Tool profile. To achieve this, the API analyses Bioschemas ontology and dynamically extends stored JSON descriptions with JSON-LD context.

It is also possible to download complete dataset either as a JSON Array of objects or JSON-LD graph.

One of the compelling feature of the Repository is tracking property changes over the time. All modifications are logged in a form of RFC 6902 Patch object together with update timestamp and data provenance (a user performed the data update). The REST API provides a way to get the history of any property.

```
https://openebench.bsc.es/bioschemas/tools/log/biotoools:trimal:1.4/cmd/description
```

```
[
  {
    "date": "2023-05-25T14:24:27.057Z",
    "value": "Tool for the automated removal of spurious sequences or poorly aligned regions from a multiple sequence alignment."
  }
]
```

REST response with a list of 'description' property updates

```
{
  "@context": {
    "@import": "https://openebench.bsc.es/bioschemas/oebtools.jsonld"
  },
  "@graph": [
    {
      "@id": "https://openebench.bsc.es/bioschemas/tools/biotoools:trimal:1.4/cmd",
      "@type": "https://openebench.bsc.es/bioschemas/oebtools#cmd",
      "rdfs:label": "trimal"
    }, ...
  ]
}
```

A complete dataset as JSON-LD graph

```
{
  "_id": {
    "id": "biotoools:trimal:1.4/cmd",
    "date": "2023-05-25T14:24:27.057Z"
  },
  "@provenance": "redmitry",
  "patch": [
    {
      "op": "add",
      "path": "/description",
      "value": "Tool for the automated removal of spurious sequences..."
    }
  ]
}
```

tools.log

MongoDB Document corresponding to the 'trimAl' record data update

<https://gitlab.bsc.es/inb/elixir/openebench/tools-monitoring>

Contact

Dmitry Repchevsky
Research Software Engineer
dmitry.repchevski@bsc.es



This project has received funding from the European Union's Horizon Europe Programme under GA 101058427 - EOSC4Cancer - HORIZON-INFRA-2021-EOSC-01

