



HAL
open science

EDAM, the life-science ontology for data analysis and management

Lucie Lamothe, Alban Gaignard, Mads Kierkegaard, Hager Eldakroury, Melissa Black, Bryan Brancotte, Jon Ison, Veit Schwämmle, Matúš Kalaš, Hervé Ménager

► To cite this version:

Lucie Lamothe, Alban Gaignard, Mads Kierkegaard, Hager Eldakroury, Melissa Black, et al.. EDAM, the life-science ontology for data analysis and management. JOBIM, Jul 2022, Rennes, France. 10.5281/zenodo.6769072 . hal-03702557

HAL Id: hal-03702557

<https://hal.science/hal-03702557v1>

Submitted on 23 Jun 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License



EDAM, life sciences ontology for data analysis and management

Lucie Lamothe¹, Alban Gaignard², Mads Kierkegaard³, Hager Eldakrouy⁴, Melissa Black⁵, Bryan Brancotte⁶, Jon Ison¹, Veit Schwämmle³, Matúš Kalaš⁷, Hervé Ménager^{1,6}, and the EDAM community

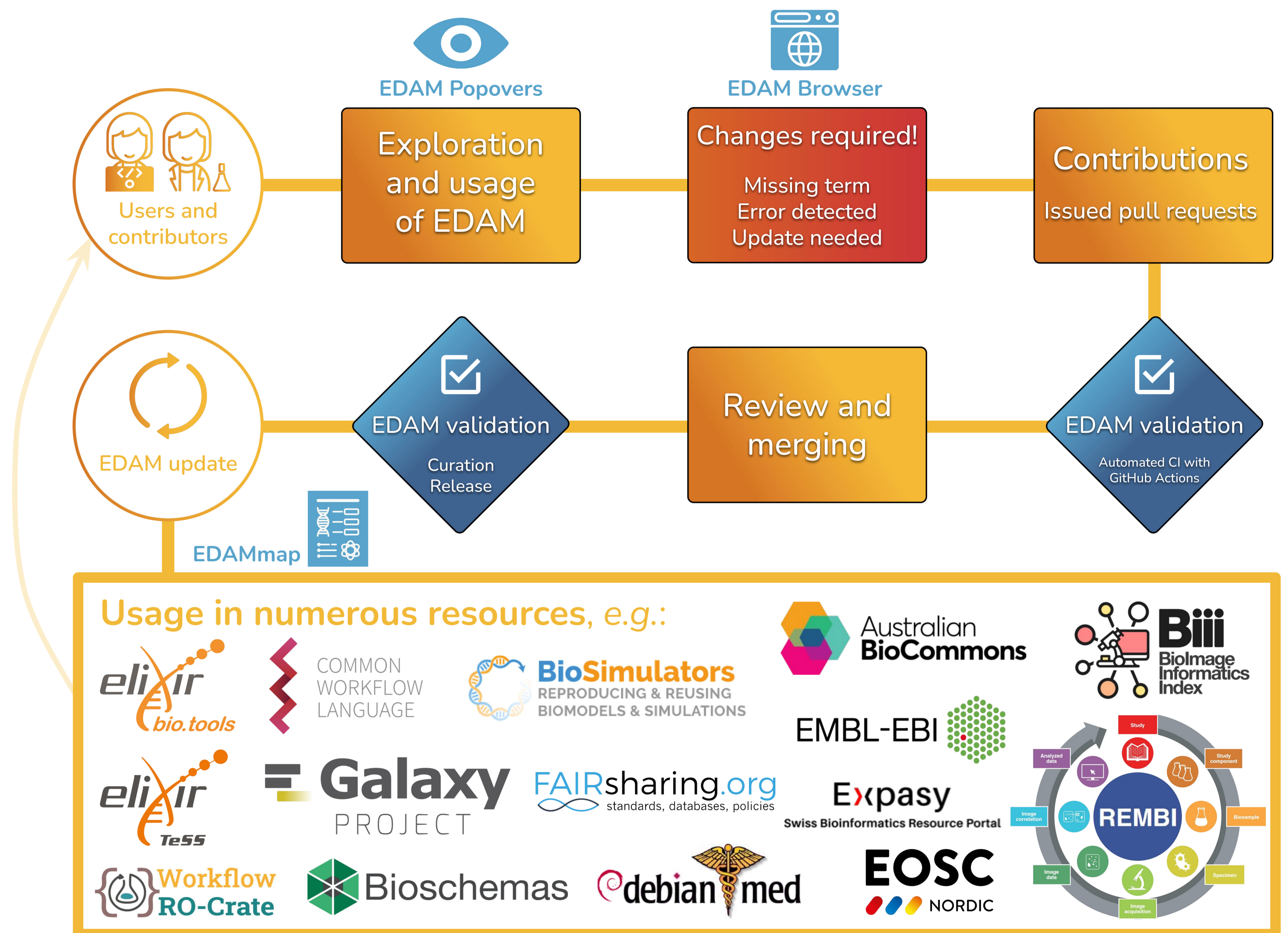
EDAM is an ontology of **data analysis** and **data management**, within and beyond life sciences. It comprises concepts related to data analysis, modelling, optimisation, and data life cycle. EDAM is a human and machine-actionable **controlled vocabulary**, used to annotate data, tools, workflows and training materials, so that they can be more **easily found, reused, and combined**.

EDAM is developed in a **participatory** and **transparent** fashion, within a broad and growing **community of contributors** from both inside and outside of ELIXIR.

Community extensions of EDAM are:

-  **EDAM Bioimaging** (including machine learning)
-  **EDAM Geo** for interdisciplinary applications, such as **global** and **planetary health**

The community-driven development comes with a set of **challenges**.

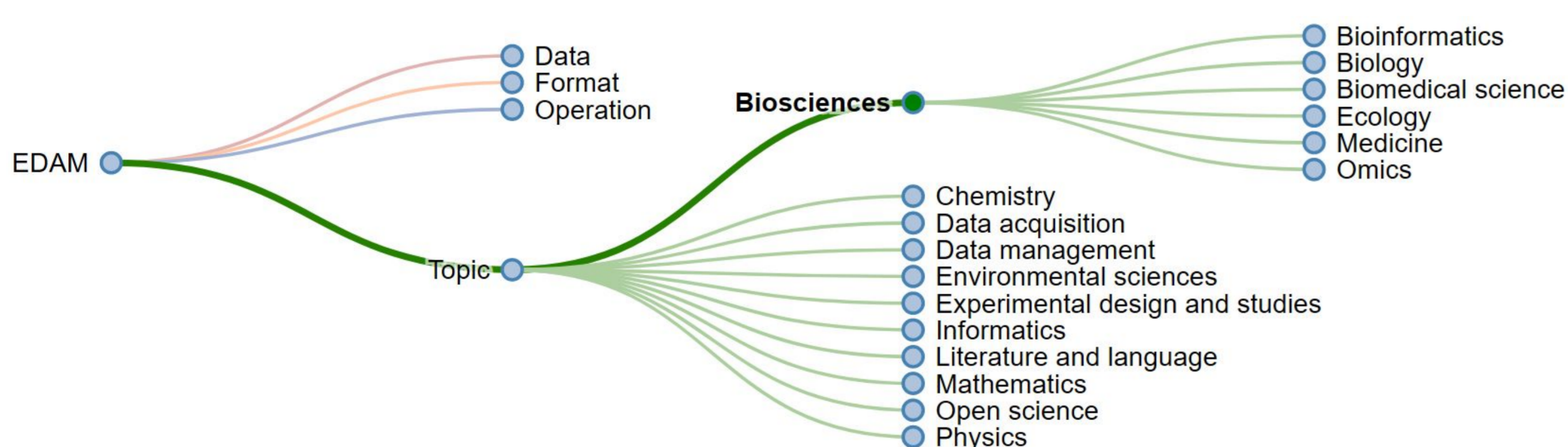


Design by Alban Sauvalle, IFB-core, ELIXIR France. Using icons from Flaticon.com

To help the community use and contribute to EDAM, a set of tools are provided:

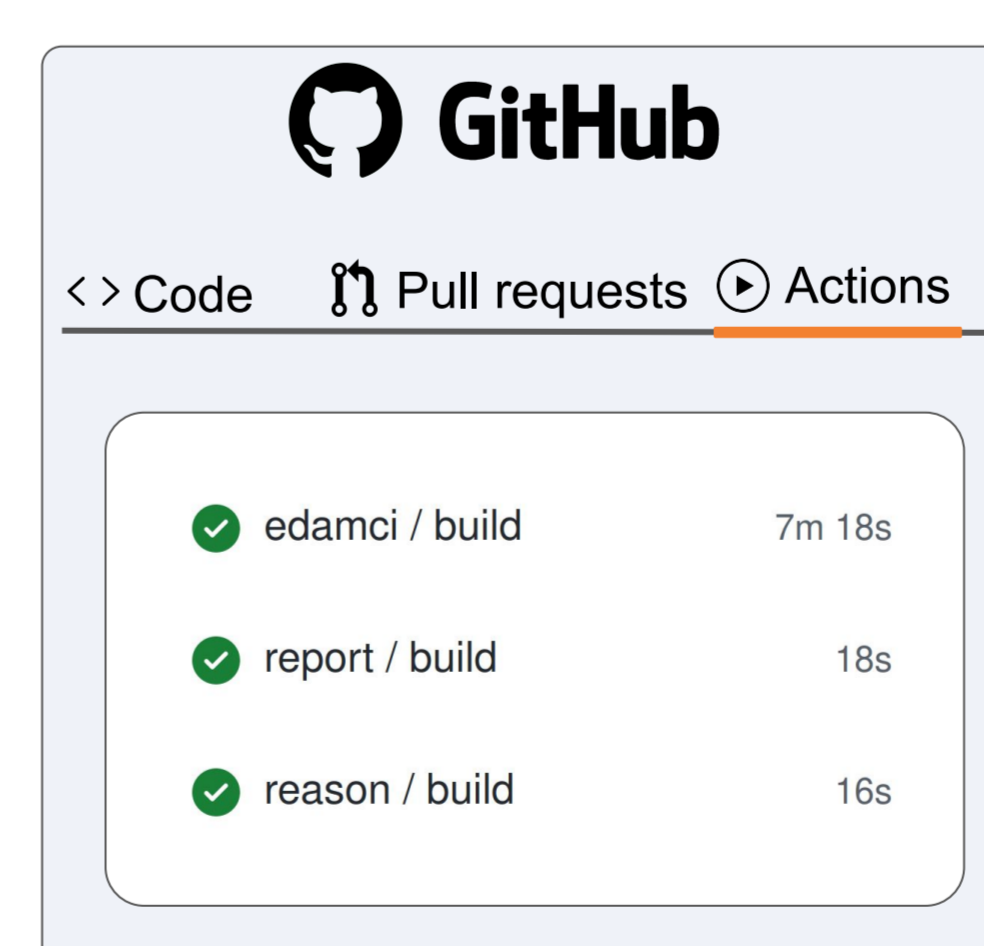
EDAM Browser

A web interface that lets users **explore EDAM** and its usage graphically. It also allows contributors to **submit a suggestion** to the GitHub repository using a graphical user interface.



EDAM validation

The quality-control tooling includes standard tools such as **ELK** and **ROBOT**, plus a custom validator working on both the syntactic and semantic levels. The validation tools are run by **continuous integration (CI)** workflows, using GitHub Actions. Additional tests are available for **curation**.



¹ CNRS, Institut Français de Bioinformatique, IFB-core, **France** (J.I. at the time of contribution)

² Nantes Université, CNRS, INSERM, L'institut du thorax, **France**

³ University of Southern Denmark, Ødense, **Denmark**

⁴ Outreachy intern (EDAM), Cairo, **Egypt** (at the time of contribution)

⁵ Outreachy intern (EDAM), São Paulo, **Brazil** (at the time of contribution)

⁶ Institut Pasteur, Université Paris Cité, Bioinformatics and Biostatistics Hub, **France**

⁷ University of Bergen, **Norway**

