EDAM The ontology of bioinformatics operations, topics, data, and formats (update 2020)

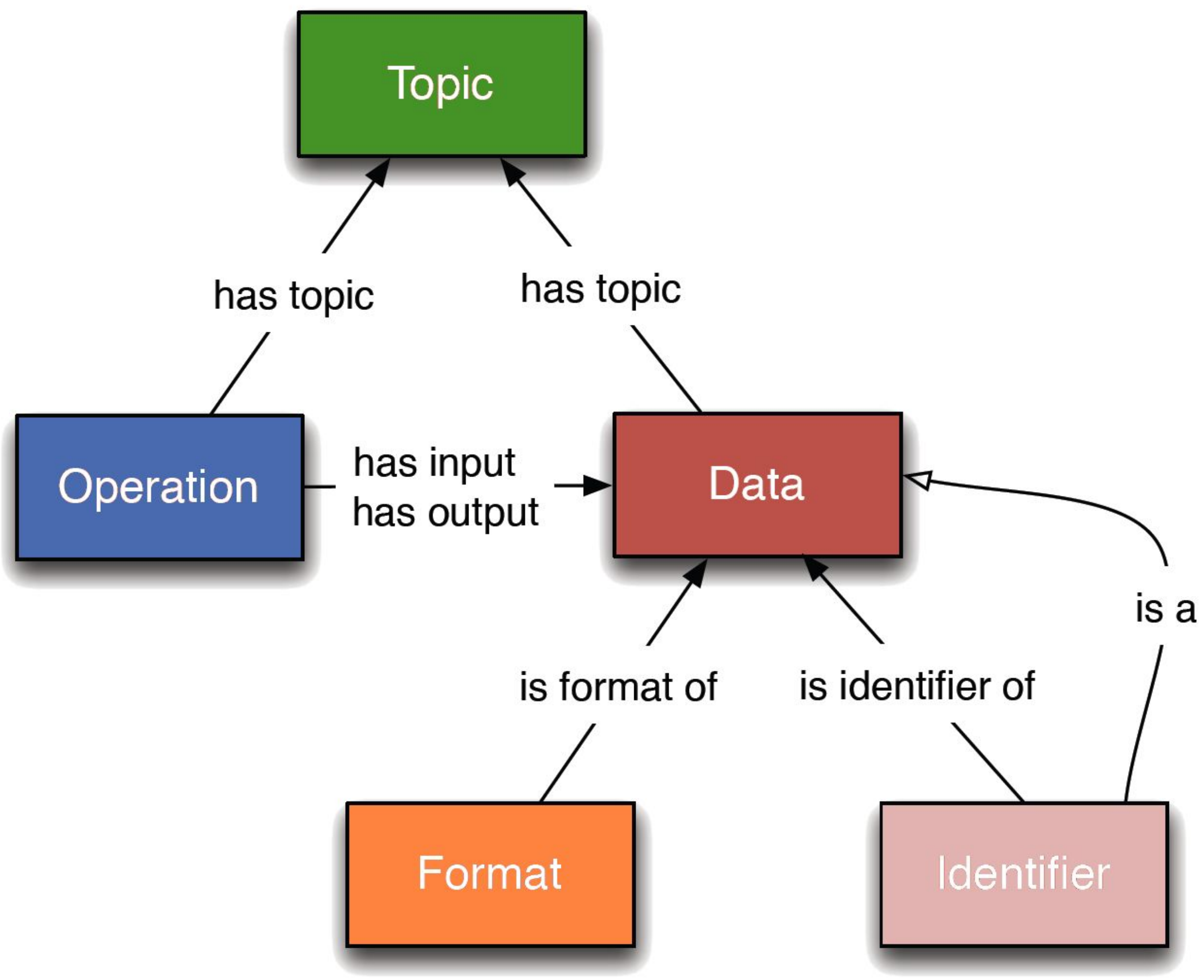
Matúš Kalaš[1], Hervé Ménager[2], Alban Gaignard[3], Veit Schwämmle[4], Jon Ison[5], and the EDAM contributors and advisors

All logoi are with links, clickable

http://edamontology.org

EDAM is an ontology of well-established, familiar concepts that are prevalent within bioinformatics, and bioscientific data analysis in general. The scope of EDAM includes types of data and data identifiers, data formats, operations, and topics. EDAM has a relatively simple structure, and comprises a set of concepts with terms, synonyms, definitions, relations, links, and some additional information (especially for data formats).

Architecture



Open source and community-driven continuous development

DOI 10.5281/zenodo.822690

@edamontology

Used in diverse applications and projects

… and many more

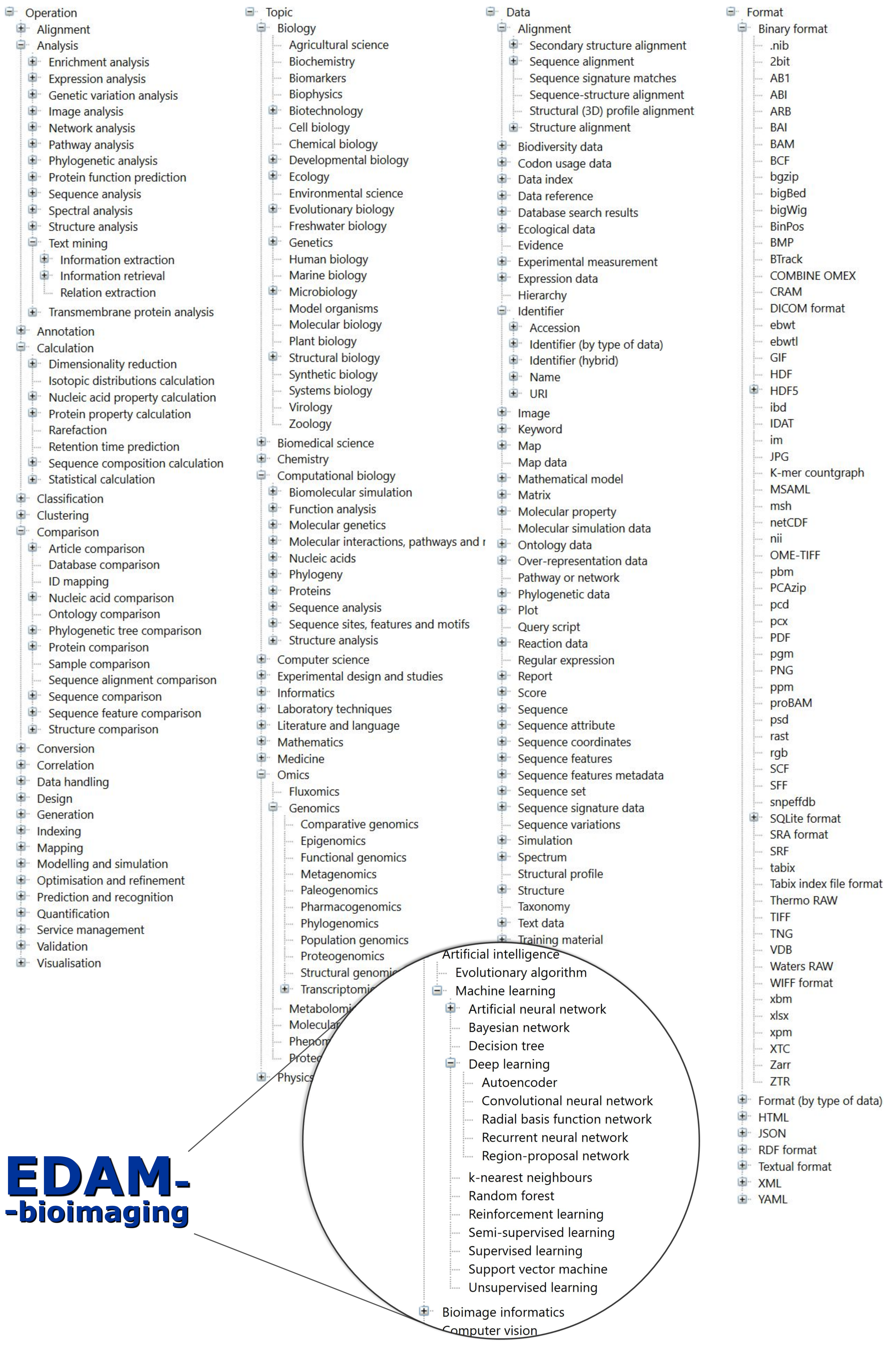
Available in various ontology browsers and download formats

OWL TSV CSV

Overview

EDAM version 1.24 DOI 10.5281/zenodo.3608238

Browse interactively at
https://bioportal.bioontology.org/ontologies/EDAM



- Operation
  - Alignment
  - Analysis
    - Enrichment analysis
    - Expression analysis
    - Genetic variation analysis
    - Image analysis
    - Network analysis
    - Pathway analysis
    - Phylogenetic analysis
    - Protein function prediction
    - Sequence analysis
    - Spectral analysis
    - Structure analysis
    - Text mining
      - Information extraction
      - Information retrieval
      - Relation extraction
    - Transmembrane protein analysis
  - Annotation
  - Calculation
    - Dimensionality reduction
    - Isotopic distributions calculation
    - Nucleic acid property calculation
    - Protein property calculation
    - Rarefaction
    - Retention time prediction
    - Sequence composition calculation
    - Statistical calculation
  - Classification
  - Clustering
  - Comparison
    - Article comparison
    - Database comparison
    - ID mapping
    - Nucleic acid comparison
    - Ontology comparison
    - Phylogenetic tree comparison
    - Protein comparison
    - Sample comparison
    - Sequence alignment comparison
    - Sequence comparison
    - Sequence feature comparison
    - Structure comparison
  - Conversion
  - Correlation
  - Data handling
  - Design
  - Generation
  - Indexing
  - Mapping
  - Modelling and simulation
  - Optimisation and refinement
  - Prediction and recognition
  - Quantification
  - Service management
  - Validation
  - Visualisation

- Topic
  - Biology
    - Agricultural science
    - Biochemistry
    - Biomarkers
    - Biophysics
    - Biotechnology
    - Cell biology
    - Chemical biology
    - Developmental biology
    - Ecology
    - Environmental science
    - Evolutionary biology
    - Freshwater biology
    - Genetics
    - Human biology
    - Marine biology
    - Microbiology
    - Model organisms
    - Molecular biology
    - Plant biology
    - Structural biology
    - Synthetic biology
    - Systems biology
    - Virology
    - Zoology
  - Biomedical science
  - Chemistry
  - Computational biology
    - Biomolecular simulation
    - Function analysis
    - Molecular genetics
    - Molecular interactions, pathways and r
    - Nucleic acids
    - Phylogeny
    - Proteins
    - Sequence analysis
    - Sequence sites, features and motifs
    - Structure analysis
  - Computer science
  - Experimental design and studies
  - Informatics
  - Laboratory techniques
  - Literature and language
  - Mathematics
  - Medicine
  - Omics
    - Fluxomics
    - Genomics
      - Comparative genomics
      - Epigenomics
      - Functional genomics
      - Metagenomics
      - Paleogenomics
      - Pharmacogenomics
      - Phylogenomics
      - Population genomics
      - Proteogenomics
      - Structural genomi
      - Transcriptomi
    - Metabolomi
    - Molecular
    - Phenom
    - Prote
  - Physics

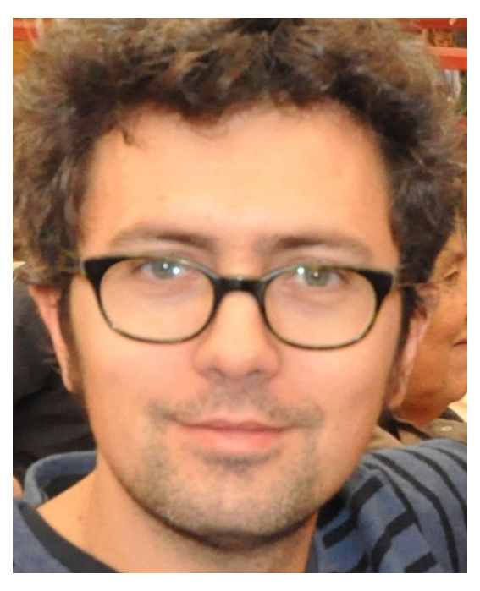
- Data
  - Alignment
    - Secondary structure alignment
    - Sequence alignment
    - Sequence signature matches
    - Sequence-structure alignment
    - Structural (3D) profile alignment
    - Structure alignment
  - Biodiversity data
  - Codon usage data
  - Data index
  - Data reference
  - Database search results
  - Ecological data
  - Evidence
  - Experimental measurement
  - Expression data
  - Hierarchy
  - Identifier
    - Accession
    - Identifier (by type of data)
    - Identifier (hybrid)
    - Name
    - URI
  - Image
  - Keyword
  - Map
  - Map data
  - Mathematical model
  - Matrix
  - Molecular property
  - Molecular simulation data
  - Ontology data
  - Over-representation data
  - Pathway or network
  - Phylogenetic data
  - Plot
  - Query script
  - Reaction data
  - Regular expression
  - Report
  - Score
  - Sequence
  - Sequence attribute
  - Sequence coordinates
  - Sequence features
  - Sequence features metadata
  - Sequence set
  - Sequence signature data
  - Sequence variations
  - Simulation
  - Spectrum
  - Structural profile
  - Structure
  - Taxonomy
  - Text data
  - Training material

- Format
  - Binary format
    - .nib
    - 2bit
    - AB1
    - ABI
    - ARB
    - BAI
    - BAM
    - BCF
    - bgzip
    - bigBed
    - bigWig
    - BinPos
    - BMP
    - BTrack
    - COMBINE OMEX
    - CRAM
    - DICOM format
    - ebwt
    - ebwtl
    - GIF
    - HDF
    - HDF5
    - ibd
    - IDAT
    - im
    - JPG
    - K-mer countgraph
    - MSAML
    - msh
    - netCDF
    - nii
    - OME-TIFF
    - pbm
    - PCAzip
    - pcd
    - pcx
    - PDF
    - pgm
    - PNG
    - ppm
    - proBAM
    - psd
    - rast
    - rgb
    - SCF
    - SFF
    - snpeffdb
    - SQLite format
    - SRA format
    - SRF
    - tabix
    - Tabix index file format
    - Thermo RAW
    - TIFF
    - TNG
    - VDB
    - Waters RAW
    - WIFF format
    - xbm
    - xlsx
    - xpm
    - XTC
    - Zarr
    - ZTR
  - Format (by type of data)
  - HTML
  - JSON
  - RDF format
  - Textual format
  - XML
  - YAML

EDAM-bioimaging

- Artificial intelligence
  - Evolutionary algorithm
  - Machine learning
    - Artificial neural network
    - Bayesian network
    - Decision tree
    - Deep learning
      - Autoencoder
      - Convolutional neural network
      - Radial basis function network
      - Recurrent neural network
      - Region-proposal network
    - k-nearest neighbours
    - Random forest
    - Reinforcement learning
    - Semi-supervised learning
    - Supervised learning
    - Support vector machine
    - Unsupervised learning
- Bioimage informatics
- Computer vision


Matúš Kalaš[1]


Hervé Ménager[2]


Alban Gaignard[3]


Veit Schwämmle[4]


Jon Ison[5]

[1] Computational Biology Unit, Department of Informatics, University of Bergen, Norway
[2] Hub de Bioinformatique et Biostatistique, Institut Pasteur, Paris, France
[3] University of Nantes, France
[4] University of Southern Denmark, Ødense, Denmark
[5] Institut Français de Bioinformatique (ELIXIR France), France