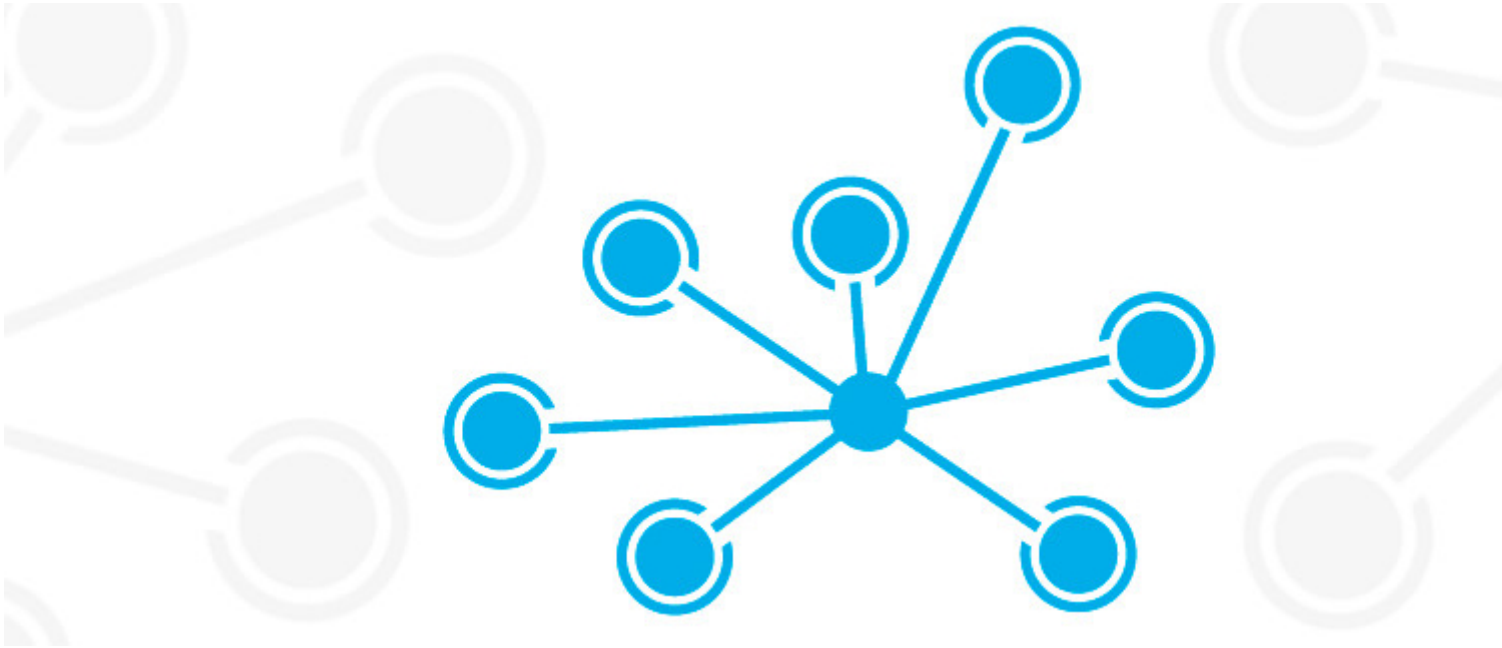


Biological data integration and management



Present bioinformatics faces an exponential growth of data. Genomics, clinical records, or simulation data accumulate terabytes of data that require finding new ways of storage or data transmission, the adaptation of analytical tools, and efficient ways to present results.

Summary

Data management is a key step in present biological projects. The amount of data generated in genomics studies, especially in the biomedical context, and also in the biomolecular simulations field, requires initiatives focussed in data: transmission, security, storage, and efficient distribution. Our group is developing some actions in these directions.

As part of the Spanish node of the ELIXIR European infrastructure is collaborating in the definition of the guidelines, and to generate tools to ease biological data management. We provide support for data management to a number of large scale genomics projects: ImidKit, The study of the genomics of Chronic Lymphocytic Leukemia, integrates in the International Cancer Genome Consortium (ICGC), the PanCancer Analysis of Whole Genomes (PCAWG), Blueprint.

Objectives

- Develop strategies for the efficient use of biological ontologies in the generation of bioinformatics

tools

- Develop new models of web-service registries
- Develop strategies for automated verification and benchmarking of bioinformatics tools
- Interface bioinformatics pipelines to HPC and cloud infrastructures

Barcelona Supercomputing Center - Centro Nacional de Supercomputación

Source URL (retrieved on 10 Mar 2025 - 13:21): <https://www.bsc.es/ca/research-development/research-areas/bioinformatics/biological-data-integration-and-management>