# Biography L. Gatto

Laurent Gatto is the head of the CBIO group at UCLouvain and teaches bioinformatics in the Facutly of Pharmacy and Biomedical Sciences (FASB).

# Biography S. Jodogne

Sébastien Jodogne is Professor of Computer Science applied to Life Sciences at UCLouvain. An expert in e-Health, he is particularly active in the fields of medical imaging, image analysis and patient education.

# e-OMIX homepage description (short)

The e-OMIX project aims to develop and deploy a cutting-edge bioinformatic hub that meets the demand for integrated multiomics data solutions.

With the recent development of different type of “omics” methods (technologies able to characterize and quantify entire spectra of biomolecular content, such as DNA- and RNA-sequencing, mass spectrometry applied to protein, etc.), biologists have been able to generate large datasets accounting for a wide range of cellular processes. Integrating these different streams of data however, revealed to be an arduous task, both for technical and theoretical reasons.

e-OMIX strives to overcome these obstacles by combining the integration, management, storage, sharing and analysis of different types of multiomics data in a single platform. The end product will be a set of open-source tools tailored for local academic institutions, healthcare networks and R&D departments conducting either fundamental or applied research, with the ambition to drive the development of new therapeutics, help identifying novel biomarkers, and accelerate the comprehension of complex cellular processes.

# e-OMIX project description (long)

## State of the art

Over the past two decades, the life science community has witnessed a significant rise in "omics" research, specialized branches of biology dedicated to characterizing and quantifying an extensive spectrum of biologically significant molecules. Genomics, which offers an insight into an individual's complete DNA sequence, emerged as a pioneering molecular tool, was probably the first of these disciplines to have been developed and gained widespread recognition following the landmark achievement of the Human Genome Project in 2003. Comparable techniques were swiftly developed to quantify other molecular contents, such as gene variants, RNA transcripts, epigenomic markers, proteins, and metabolic by-products.

The field is still in constant mutation and is a source of continuous innovation for biological, clinical, and pharmaceutical research. Initially, omics methods were constrained to the analysis of entire tissue samples, known as "bulk omics." However, recent technological advancements have enabled a more refined approach, permitting the analysis of specific cell types, a domain known as "single-cell omics."

Concurrently with these technical advancements, the field of computational biology has undergone significant expansion. It has established a repertoire of statistical tools tailored to both generate and rigorously test hypotheses, utilizing the vast reservoir of data generated by omics analyses. Together, omics-based quantification methods and bioinformatics algorithms have paved the way to great progresses in the understanding of biological processes, and, in fine, the improvement of human health.

## Objectives

While a multitude of single omics pipelines are readily available and are used to answer experimental hypotheses through specific molecular lenses, combining multiple type of omics would considerably broaden the scope of investigations. Such an endeavor however comes with significant hurdles. Therefore, the deployment of bioinformatic platforms where omics-specific pipeline converges to build a centralized, standardized, and accessible datasets would be a tremendous improvement.

An example of these obstacles is the heterogeneity of molecular biology data. Each instrument is often associated with proprietary formats, hindering the sharing of analysis between research groups. Additional specifications might result from experimental workflow, the type of samples or other constraints. One of the objectives of our data treatment pipelines would then be to standardize and transform proprietary data into open-source formats, as well as resolve disparities among data types. In that perspective, e-OMIX will adhere to HL7’s FHIR standards, ensuring interoperability of biomedical data, and make its own resources available for the scientific community, facilitating exchanges and collaboration.

In subsequent development steps, we will implement scalable semi-automated statistical analyses to identify pertinent networks, hubs, and links between cellular constituents. Finally, e-OMIX will be able to incorporate clinical data as well, providing enhanced insights into both normal and pathological phenotypes. As an integrated tool, e-OMIX could then be leveraged to advance our understanding of cell biology, identify new biomarkers, and facilitate therapeutic developments.