# Omics Challenges in Health and Biology

Justine Engel[1], Vincent Ruzand[1], Guillaume Guérard[1], Sonia Djebali[1]

Léonard de Vinci Pôle Universitaire, Research Center, 92 916 Paris La Défense, France

{first_name, last_name}@devinci.fr

**Keywords** : *omics, deep learning*

## 1 Classification of the Omics

The term "omics" encompasses various biological fields that analyze comprehensive data representing complete biological systems at specific levels. When combined with data from medical imaging, clinical records, and other medical technologies, omics has significantly advanced precision medicine, aligning with patient biology and pathology.

Omics can be broadly categorized into two blocks [1, 4] (refer to Figure 1): the Big Four, which include genomics, transcriptomics, proteomics, and metabolomics, each with its own subsections such as interactomics or epiomics; and the expansive realm of techno-applied omics, denoting the application of omics data to existing technologies. It is important to note that the integration of multi-omics, representing diverse omics data, should not be treated as a specific omics but rather viewed as a method for integrating independent observations.

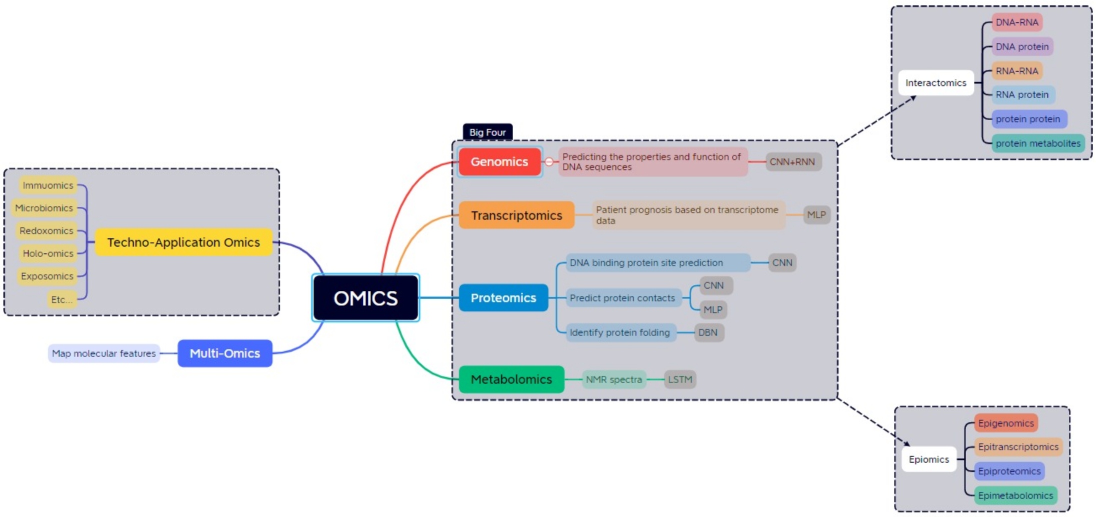


FIG. 1: Omics MindMap

## 2 Needs in Decision Making

Due to the exponential growth of omics data over the past decade, scientists have recognized the impracticality of manually processing such vast amounts of information. Consequently, they have sought new and promising approaches for decision-making [2, 3]. Omics data find applications in diverse use cases, ranging from genomics to protein-metabolite interactomics, employing various technologies. Examples include RNA and drug classification, protein interaction studies, and protein classification. Different omics fields serve specific purposes. For instance: **Genomics** – Involves DNA sequence structure, gene expression regulation, genotype, and drug-related studies ; **Transcriptomics** – Encompasses RNA sequence structure, RNA-based disease prediction, and RNA-drug classification ; **Proteomics** – Focuses on protein classification, drug design, and the study of protein interactions ; **Metabolomics** – Monitors changes in metabolite levels in various processes.

The application of Deep Learning to these use cases has introduced novel approaches. The versatility of sub-techniques, such as using Convolutional Neural Networks (CNN) for understanding DNA sequence structures or Recurrent Neural Networks (RNN) for protein classification and pattern recognition, allows for flexibility across a broad range of applications. However, it is crucial to note that while this method excels at processing massive amounts of data quickly, its main drawback is the inherent black box effect, making it challenging to comprehend the decision-making process.

## 3 Explainable Models for Precision Medicine

The rapid evolution of omics has significantly expanded our ability to comprehend biological systems at a holistic level. Notably, the perception of multi-omics as a distinct category is challenged, and it is positioned as a methodological approach. This approach enhances our understanding of complex biological phenomena, providing a more comprehensive view compared to studying individual omics in isolation.

The fusion of omics and deep learning holds promise for unveiling novel insights into the intricate mechanisms of complex pathologies. While deep learning shows great potential in analyzing omics data for decision-making, it is crucial to acknowledge associated drawbacks and challenges in this domain, including data requirements (overfitting, computational complexity, noise), bias (integration of prior biological knowledge is limited), computational complexity, and interpretability. Given that the primary objective is decision-making for precision medicine, the decisions made must be understandable and explainable to humans.

To address this, the field is open to explainable artificial intelligence and combinatorial methods aimed at extracting knowledge from the substantial amount of untapped data. This signifies a commitment to ensuring transparency and interpretability in decision-making processes within the realm of precision medicine.

## References


[1] Xiaofeng Dai and Li Shen. Advances and trends in omics technology development. *Frontiers in Medicine*, 9:911861, 2022.

[2] Guillermo de Anda-Jáuregui and Enrique Hernández-Lemus. Computational oncology in the multi-omics era: state of the art. *Frontiers in oncology*, 10:423, 2020.

[3] Katy Vandereyken, Alejandro Sifrim, Bernard Thienpont, and Thierry Voet. Methods and applications for single-cell and spatial multi-omics. *Nature Reviews Genetics*, pages 1–22, 2023.

[4] Timothy D Veenstra. Omics in systems biology: current progress and future outlook. *Proteomics*, 21(3-4):2000235, 2021.