



# ML BIOMARKERS

### DATAIA PARIS-SACLAY INSTITUTE

Located within the **Paris-Saclay University** (16<sup>th</sup> in the Shanghai ranking, 1<sup>st</sup> in mathematics), it is the **first French** ecosystem in Data Sciences, AI and their societal impacts.

#### MISSION

To bring together **multidisciplinary expertise and boost the collective strength of its partners** in the Paris-Saclay cluster with the aim of combining big data and AI technologies with social sciences and humanities for an AI at the service of humans.

#### **IN FIGURES**





The Industrial Affiliation Plan (PAI) aims to boost the collective strength of the Institute's academic ecosystem and its industrial members. The services offered in response to the respective needs expressed include:

- Joint actions to support research;
- Sharing of experiences and collective needs;
- Facilitated access to recruitment;
- Access to training, seminars, workshops, etc.;
- Implementation of dedicated events (hackathons, challenges, etc.);
- Access to working places to increase exchanges.



The D2C system aims **upstream**, to present the priority research issues and to match them with the problems of industry. **Downstream**, to monitor contacts and opportunities for collaboration identified until they are set up and launched. It is part of the ambition to facilitate the establishment of several levels of collaboration and create a constructive dynamic:

- 1. Expertise / Student projects / Internships
- 2. Research collaborations / CIFRE theses
- 3. Joint laboratories / Joint teams
- 4. Multi-partner chairs

### **OBJECTIVES & PROGRAM**



The main objectives of this D2C are focusing on :

- Identification of biomarker signatures on treatment and disease progression;
- Optimization of biomarker selection by unsupervised analysis on multi-omics data;
- Integration of biomedical database to study gene function.

2pm - 2:10pm	Introduction
2:10pm - 3pm	DATAIA researchers pitches on prospective research topics followed by industrialists on related issues
3pm - 4pm	Individual meetings with a view to setting up new collaborations
4pm - 4:10pm	Conclusion and action plan



#### Identification of genes responsible for phenotypes

Christophe Ambroise (UEVE, LaMME) Supervised and unsupervised learning based on probabilistic models



#### Querying and ranking biological and biomedical data

Sarah Cohen-Boulakia (Paris-Saclay University, LISN) Integration, querying and ranking in the context of biomedical database



#### Identification of new markers and classification based on molecular profiles

Paul-Henry Cournède (CentraleSupélec, MICS) Biomathematics, mathematical modelling of biological systems



#### Specific and dedicated statistical approaches

**Bertrand Thirion** (Inria, Parietal) Statistical modeling and machine learning applied to brain imaging data

### **DATAIA RESEARCHERS**



#### Response to treatment with genomic data

Farida Zerhaoui (UEVE, IBISC) Interpretation of learning models, multi-source, multiobjective classification, clustering



Blaise Hanczar (UEVE, IBISC) Deep learning, supervised learning, predictive systems, performance evaluation



#### New high-dimensional statistical methods for multiomics data

Julien Chiquet (INRAE, MIA) Statistical learning applied to the analysis of data from life sciences



# Designing ML models to discover brain imaging signatures of mental disorders

Edouard Duchesnay (CEA, Neurospin) Transfer learning algorithms to bridge the gap between heterogeneous and homogeneous datasets



# Imaging data correlated with patient treatment responses

Nicolas Gogin - Senior Manager deep learning and image analytics

# sanofi

**GE Healthcare** 

Integration of biological knowledge: gene function, protein networks Development of ML models for biomarker signature discovery in medical imaging

Caroline Paccard - Biomarker statistics head Franck Augé - Translational sciences, bioinformatics group head Elton Rexhepaj - Senior data scientist, bioimaging and deep learning

Gene linkage modeling to optimize marker selection, mixing of different types of data within the same ML model

Laura Xuereb - Biostatistics Manager, biomarkers Perrine Soret - Tranlational analytics and statistics



### **GUEST COMPANIES**

# TRIBVN

#### Learning data to challenge algorithms in biopsy sorting

Jean-François Pomerol - CEO Salma Ben Hadj - IA Director Rutger Fick - Senior Data Scientist



Optimization of biomedical data in R&D and pharmacovigilance

Romain Clément - CEO



Integration of multi-omics data for the characterization of patients resistant to anti-cancer treatments

Sébastien Vacheric - OncoSNIPE Program Director

## **INSTITUTIONAL PARTNERS**























#### **INSTITUT DATAIA**

Centre de Recherche Inria Saclay - Île-de-France Campus de l'Ecole Polytechnique - Bâtiment Alan Turing 1 rue Honoré d'Estienne d'Orves - 91120 Palaiseau

SERVICE COMMUNICATION com-dataia@inria.fr



in

www.dataia.eu

@institut\_dataia

@institut-dataia