D2C

DATAIA CLUB CONNECTION

ML BIOMARKERS

January 28, 2021

université PARIS-SACLAY

INSTITUT DATAIA
Science des données, Intelligence & Société

The DATAIA Paris-Saclay Institute

Located within the Paris Saclay University (16th in the Shanghai ranking, 1st 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

12
academic members

42
partner laboratories

1200
professors-researchers

10
IA Chairs out of 40 national

18
research projects launched

450
PhD Students

Industrial Affiliation Plan (PAI)

The Industrial Affiliation Plan (IAP) 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.

DATAIA Club Connection (D2C)

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 and program

The main objectives of the D2C « *ML Biomarkers* » to address are:

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

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

DATAIA researchers

Querying and ranking biological & biomedical data

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

Specific and dedicated statistical approaches

Bertrand Thirion (Inria, Parietal)
Research: statistical modeling and machine learning applied to brain imaging data

Identification of genes responsible for phenotypes

Christophe Ambroise (Université d'Evry, LaMME)
Research: supervised and unsupervised learning based on probabilistic models

Identification of new markers and classification based on molecular profiles

Paul-Henry Cournède (CentraleSupélec, MICS)
Research: biomathematics, mathematical modelling of biological systems

DATAIA researchers

Response to treatment with genomic data

Farida Zerhaoui (Université d'Evry, IBISC)
Research: interpretation of learning models, multi-source, multi-objective classification / clustering

Blaise Hanczar (Université d'Evry, IBISC)
Research: deep learning, supervised learning, predictive systems, performance evaluation

New high-dimensional statistical methods for multi-omics data

Julien Chiquet (INRAE, MIA)
Research: 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)
Research: transfer learning algorithms to bridge the gap between heterogeneous) and homogeneous datasets

DATAIA Club PAI Companies

Imaging data correlated with patient treatment responses

GE Healthcare

Nicolas Gogin - *Senior manager deep learning and image analytics*

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

SANOFI

Caroline Paccard - *Biomarker statistics head*
Franck Auge - *Translational sciences, bioinformatics group head*
Elton Rexhepaj - *Senior data scientist - Bioimaging and deep dearning*

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

SERVIER

Laura Xuereb - *Biostatistics manager - biomarkers*
Perrine Soret - *Translational analytics and statistics*

Invited companies

Learning data to challenge algorithms in biopsy sorting

TRIBVN HEALTHCARE

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

Optimization of biomedical data in R&D and pharmacovigilance

AS

Romain Clement - *CEO*

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

Onco design BIOTECHNOLOGY

Sebastien Vachenc - *OncoSNIPE Program Director*

Institutional partners

université PARIS-SACLAY

INSTITUT DATAIA
Science des données, Intelligence & Société

Institut Convergence 17-CONV-0003 INSTITUT DATAIA (I2DRIVE)

Institut DATAIA Paris-Saclay
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

www.dataia.eu
@institut_dataia