

# D2C

## DATAIA CLUB CONNECTION

### ML Biomarqueurs

28 janvier 2021

université  
PARIS-SACLAY

INSTITUT DATAIA  
Science des données, Intelligence & Société

# Sommaire

---

<b>L'Institut DATAIA Paris-Saclay</b>	<b>3</b>
<b>Plan d’Affiliation Industrielle (PAI)</b>	<b>4</b>
<b>DATAIA Club Connection (D2C)</b>	<b>4</b>
<b>Objectifs et programme</b>	<b>5</b>
<b>Les chercheurs DATAIA</b>	<b>6</b>
<b>Les entreprises Club PAI DATAIA</b>	<b>8</b>
<b>Les entreprises invitées</b>	<b>9</b>
<b>Les partenaires institutionnels</b>	<b>10</b>
<b>Les D2C à venir</b>	<b>11</b>

# L'Institut DATAIA Paris-Saclay

Situé au sein de l'Université Paris Saclay (14e au classement de Shanghai, 1ère en mathématique), il est le premier écosystème français en sciences des données, IA et leurs impacts sociétaux.

## SA MISSION

Regrouper les expertises pluridisciplinaires et dynamiser la force collective de ses partenaires du cluster Paris-Saclay avec l'objectif de croiser les technologies big data et IA avec les SHS pour une IA au service de l'humain.

## EN CHIFFRES



12

membres académiques



42

laboratoires partenaires



1000

chercheurs



16

Chaires IA sur 40 nationales



13

projets de recherche lancés



300

étudiants M2/an

# Plan d'Affiliation Industrielle (PAI)

---

Le Plan d'Affiliation Industrielle (PAI) a pour objectif de dynamiser la force collective de l'écosystème académique de l'Institut et des acteurs industriels adhérents.

Les services proposés en réponse aux besoins respectifs exprimés sont notamment :

- Actions conjointes de soutien à la recherche ;
- Partage d'expériences et de besoins collectifs ;
- Accès facilité pour le recrutement ;
- Accès aux formations, séminaires, workshops, ... ;
- Mise en oeuvre d'événements dédiés (hackathon, challenges...) ;
- Accès à des working places pour augmenter les échanges.

## DATAIA Club Connection (D2C)

---

Le dispositif D2C vise :

- En amont, à présenter les enjeux de recherche prioritaires et à les faire correspondre avec les problématiques des industriels ;
- En aval, à suivre au plus près les contacts et les opportunités de collaborations identifiées jusqu'à leur montage et lancement.

Il s'inscrit dans l'ambition de faciliter le montage de plusieurs niveaux de collaboration et créer une dynamique constructive :

1. Expertises / projets étudiants / stages
2. Collaborations de recherche / thèses CIFRE
3. Laboratoires communs / équipes communes
4. Chaires multi-partenaires

# Objectifs et programme

---

Le principaux objectifs du D2C « ML Biomarqueurs » à adresser sont :

- Identification des signatures de bio-marqueurs sur le traitement et la progression de la maladie ;
- Optimisation de la sélection de bio-marqueurs par l'analyse non supervisée sur données multi-omiques ;
- Intégration de données de bases biomédicales pour étudier le fonctionnement de gènes

**14h00 - 14h10**

Introduction

**14h10 - 15h00**

Pitchs chercheurs DATAIA autour de sujets de recherche prospectifs suivis des industriels autour de problématiques associées

**15h00 - 16h00**

Rendez-vous individuels en perspective de monter de nouvelles collaborations

**16h00 - 16h10**

Conclusion et plan d'action

# Les chercheurs DATAIA

---

## Querying and ranking biological & biomedical data



**Sarah Cohen-Boulakia** (Université Paris-Saclay, LISN)  
Recherche : intégration, interrogation et classement dans le contexte de données de bases biomédicales

## Approches statistiques particulières et dédiées



**Bertrand Thirion** (Inria, Parietal)  
Recherche : modélisation statistique et apprentissage automatique appliqué aux données d'imagerie cérébrale

## Identification des gènes responsables de phénotypes,



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

## Identification de nouveaux marqueurs et classification à partir des profils moléculaires



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

# Les chercheurs DATAIA

---

## Réponse à un traitement avec des données génomiques



**Farida Zerhaoui** (Université d'Evry, IBISC)  
Recherche : interprétation modèles d'apprentissage, classification / clustering multi-sources, multi-objectifs



**Blaise Hanczar** (Université d'Evry, IBISC)  
Recherche : *deep learning*, apprentissage supervisé, systèmes prédictifs, évaluation des performances

## Nouvelles méthodes statistiques à grande dimension autour des données multi-omiques



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

# Les entreprises Club PAI DATAIA

---

Données d'imagerie corrélées à des réponses aux traitements des patients



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

Intégration connaissances biologiques : fonctionnement des gènes, ensemble réseaux de protéines

Développement des modèles ML pour la découverte des signature de type biomarqueurs dans l'imagerie médicale



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

Modélisation de liens entre les gènes pour optimiser la sélection de marqueurs

Mélange de données de différents types au sein d'un même modèle ML



**Laura Xuereb** - *responsable biostatistiques - biomarqueurs*  
**Perrine Soret** - *Translational analytics and statistics*



# Les entreprises invitées

---

Données d'apprentissage pour challenger des algorithmes dans le tri des biopsies



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

Intégration de données multi-omiques pour la caractérisation des patients résistants aux traitements anticancéreux



**Sebastien Vachenc** - Directeur Programme  
OncoSNiPE

Optimisation des données biomédicales en R&D et en pharmacovigilance



**Romain Clement** - CEO

# Les partenaires institutionnels

---



# Les D2C à venir

---

## D2C SERVICES MOBILITÉ

*Février 2021*

Analyse comportementale des utilisateurs et prédiction.

## D2C MACHINE LEARNING POUR L'IMAGERIE

*Mars 2021*

Solutions de traitement, reconstruction, recalage des images médicales.

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'École Polytechnique - Bâtiment Alan Turing  
1 rue Honoré d'Estienne d'Orves  
91120 Palaiseau

**Service communication**  
com-dataia@inria.fr

 [www.dataia.eu](http://www.dataia.eu)

 [@institut\\_dataia](https://twitter.com/institut_dataia)