GENERATIVE AI

DATAIA PARIS-SACLAY INSTITUTE

Located within the **Paris-Saclay University** (16th in the Shanghai ranking, 1st in mathematics), it is the f**irst 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
DATAIA members

42
laboratories partners

1,2K
full-time researchers

10
IA chairs out of 40 national

30
IA theses

450
PhD students per year

PAI
Programme d'Affiliation Industrielle

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.);
- Participation to DATAIA governance through Industrial External Commitee.

D2C
DATAIA CLUB CONNECTION

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 :

- **Evaluation of generative AI**: qualification of the performance and operation of an algorithm using generative AI on specific use cases;

- **Generative AI for text/image generation**: radiology, medical reports;

- **Generative AI for code generation**: new functionalities, ultra-customization, alpha code, bug detection, prompt engineering to generate code in Python, societal impact for service companies;

- **Generative AI for design**: generate drawings for product design, assistance.

DATAIA RESEARCHERS

Instrumental for the responsible adoption of promising NLP generative methods

Pierre Colombo (CentraleSupélec, MICS)
Robustness of transformers architecture, reliable models evaluation, fariness to measure bias and correct them

Generative AI for text generation (terminology, natural language, ontology)

Arnaud Ferre (INRAE, MaIAge)
Acquisition and formalization of knowledge from text

Increase the learning base, simulate patient trajectories, identification of biomarkers, improve the confidentiality of personal data

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

DATAIA CLUB PAI COMPANIES

GE Healthcare

Generative AI in the post-processing of medical images
Etienne Perot - *Senior Data Scientist*

SERVIER
moved by you

Integration of bio-physical constraints
Model improvements by experimental validations
Thierry Dorval - *Head of Data Science Lab*
Jérémy Grignard - *Data Scientist*

Quantmetry

LLM audit and conversational search engine
Performance metrics and gain in performance
Data anonymization without loss of performance
Code generation
Nicolas Brunet - *Scientific Director*
Julien Roussel - *Data Scientist*
Aymen Mejri - *Data Scientist*

GUEST COMPANIES

Support for the production of low-code applications
Bruno Berstel Da Silva - *CTO*
Kajetan Wojtacki - *AI Engineer*

Generative AI for the multi-document scientific summary
Fréjus Laleye - *CTO*

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

www.dataia.eu

@institut_dataia

@institut-dataia