Call for National and International 2022

Introduction

DATAIA is the AI Institute of the University of Paris-Saclay, dedicated to data science, AI and the impact of their applications on society. It brings together more than 1,200 professors and researchers from universities, national research organizations and Grandes Ecoles. Funded as a convergence institute by the French National Research Agency (ANR) as part of the "Programmes d'Investissements d'Avenir", DATAIA aims to :

1. Mobilize multi-disciplinary skills to produce new knowledge ;
2. Develop innovative training programs of excellence at the master's and doctoral levels, for both initial and continuing education, based on this interdisciplinary research;
3. To structure long-term international and industrial collaborations at a high scientific level around the main challenges addressed by the Institute (https://www.dataia.eu/linstitut/les-priorites-scientifiques)

1. Objectives and scientific scope of the call

This 2022 call will finance on an ongoing basis, for a period of 3 to 6 months :

- Stays of PhD students from laboratories outside the Institute's partner institutions. These laboratories may be located in France or abroad.
- Visits abroad by doctoral students from laboratories of DATAIA partner institutions.

The DATAIA Institute will fund these types of stays in the form of stipends, which will be used to cover the doctoral student's living expenses :

- 1,000 €/ month for doctoral students from a laboratory located in France ;
- 1 000 € to 2 000 € / month for PhD students coming from a foreign laboratory ;
- 1,000 to 2,000 €/month for outgoing mobility abroad.

The DATAIA Institute may provide an additional allowance for travel expenses :

- up to 1 000 € for a mobility from and/or to a European country ;
- up to 1,500€ for a mobility from and/or to a country outside of Europe.

Responses to this call for proposals must demonstrate how they contribute to the objectives of the DATAIA Institute. Accordingly, all proposals must address at least one of DATAIA's four main challenges, and may address an application area (see list of "Workstreams" in Appendix 1).

Expectations/Commitments

The doctoral student commits to participate in seminars at the master's or doctoral level of the host laboratory. In the case of incoming mobility, the PhD student will also be expected to participate in the scientific activities of the DATAIA Institute (seminars, workshops, etc.).

A one-page summary of the student's activities will be prepared by the host laboratory within one month of the visit. This report will detail the activities carried out during the visit, the highlights and the perspectives of future collaborations. The host laboratory will also report to the DATAIA Institute on the impact of the visit (in particular, joint projects or publications). Recipients of DATAIA convergence institute support must acknowledge DATAIA and the PIA program in all scientific communications associated with the funded activities by inserting the following statement: This research was supported by DATAIA convergence institute as part of the " Programme d'Investissement d'Avenir ", (ANR17-CONV-0003) operated by [Partner] XXX."

Partnership scope of the call

This call for proposals is restricted to teams within the scope of DATAIA. Only laboratories of partner institutions listed in the funding agreement will be eligible to apply for and manage funding (see list in Appendix 2).

How to submit a proposal

Applications must be submitted by a laboratory or research team from a DATAIA partner institution by email to : submission-dataia@inria.fr

Questions about this call can be sent to : contact-dataia@inria.fr

The application must include the following information :

- Name, laboratory and contact information of the applicant ;
- The DATAIA Institute laboratory (or team) that will receive funding, the scientific contact and the administrative contact in the laboratory (team) of the beneficiary partner ;
- Scientific focus of the DATAIA Institute (see Appendix 1)
- Expected duration and dates of the stay ;
- Expected benefits for the DATAIA Institute.

The scientific file will include the following elements :

- CV of the applicant (2/3 pages including contact information, status and institution of origin, and major scientific contributions) ;
- Work program, courses and seminars during the visit (max. 1 page)

Method of project selection

Projects will be reviewed and selected by the DATAIA Office and a response will be provided within one month of submission.

Appendix 1 : Areas of work

The scientific program of each proposal should explicitly address at least one of the four research areas listed below, which are the four interdisciplinary challenges of DATAIA.

FROM DATA TO KNOWLEDGE, FROM DATA TO DECISION

The increasing availability of massive data is pushing back the technical boundaries in many fields. On the one hand, the heterogeneous, semi-structured, incomplete or uncertain nature of data calls into question the usual statistical models as well as the algorithms dedicated to decision making. On the other hand, data management raises new operability constraints such as security, integrity and traceability.

In addition, producing knowledge requires building models that provide explainable, statistically valid and computable decisions. Acceptance of the results also requires that confidentiality and trustworthiness be strengthened.

In parallel, new developments in optimization should allow to improve estimation procedures.

Challenges :

- Heterogeneous, complex, incomplete, semi-structured and/or uncertain data
- Massive data: algorithmic and data structuring
- E-learning, methodology for massive data, efficient methods
- Improved storage, computation and estimation for data science
- Modeling interactions between agents (human or artificial) using game theory
- Multiscale and multimodal representation and algorithms
- Theoretical analysis of heuristic methods (complexity theory, information geometry, Markov chain theory)
- Human-machine coevolution in autonomous systems: conversational agents, cars, social robots.

LEARNING AND ARTIFICIAL INTELLIGENCE

Recently, deep learning research has made spectacular advances in computer vision and natural language processing. Beyond the arrival of massive data, increased computing power and design efforts, the causes of these advances, which are still not well known, raise at least three questions.

Which learning theory will allow us to analyze deep architectures?

How to manage the compositionality of these architectures and their ability to understand more complex objects?

How to open the black box to update the learned representations?

Challenges :

- Innovative machine learning and AI: common sense, adaptability, generalization
- Deep learning and adversarial learning
- Machine learning and hyper-optimization
- Optimization for learning, e.g. improvements in stochastic gradient methods, Bayesian optimization), combinatorial optimization
- Learning-modeling link, a priori integration in learning
- Reproducibility and robust learning
- Statistical inference and validation
- Compositionality of deep architectures.

TRANSPARENCY, RESPONSIBLE IA AND ETHICS

Digital trust is built on the implementation of ethically responsible methodologies through the transparency and accountability of algorithmic systems; the regulation of the collection, use and processing of personal data; and the strengthening of regulation through appropriate digital procedures.

Privacy by design is a form of regulation that includes the protection of personal data in all stages of collection and processing. The tracing of tools applied to data must also be developed in a way that facilitates the explanation of the model for experts and users alike, making algorithmic systems auditable. Privacy principles, while easy to formulate, require changes in storage and processing infrastructures, with important legislative, sociological and economic impacts. Transparency techniques for algorithmic systems can be addressed with a focus on fairness, equity, non-discrimination, and accountability-by-construction.

Challenges :

- Accountability and explicability by design
- Transparency and fairness by design
- Audit of algorithmic systems: non-discrimination, fairness, technical bias, neutrality, equity
- Measuring trust and digital ownership
- "Progressive user-analysis of progressive data" (interactive monitoring of decision systems)
- Accountability for information processing and decision making : data usage control and fact checking
- Causal discovery, traceability of inferences from source data, interpretability of deep architectures.

PROTECTION, REGULATION AND ECONOMICS OF DATA

Companies involved in the data economy continually need to rethink their structure: they must adopt a project-oriented organization with rapid changes in resource allocation. The data economy also raises issues of concentration and market power. A small number of companies (GAFAM) hold most of the data. This market concentration can lead to unfair competition, which can have a negative impact on innovation in small and medium-sized enterprises. Citizens expect governments to intervene in the digital economy to avoid too much concentration and abuse of dominant position. Governments must also prevent the leakage of information in order to preserve the sovereignty of States and the respect of regulations.

Challenges :

- Privacy by design RGPD
- Differential privacy learning
- Development of ethically responsible methodologies and technologies to regulate the collection, use and processing of personal data and the exploitation of knowledge derived from such data
- Computer security of data processing chains
- Security/crypto: blockchain and trusted third parties

Projects may address more specific topics in application areas. A list of these domains is given below as an indication. However, it is requested that projects address data science issues of general interest, i.e. not specifically restricted to an application domain.

- Energy : optimization of management (production and distribution), and regulation
- Health, personalized nutrition and well-being
- Urbanization & Mobility (connected and autonomous vehicles, smart cities...)
- Analysis for finance, insurance - Internet of people and things
- E-Sciences.

Appendix 2 : List of DATAIA Partner Institutions

Agro ParisTech
CentraleSupélec
CEA
CNRS
ENS Paris-Saclay
IFP-Energies Nouvelles
Inria
Institut Mines-Télécom Business School
INRAE
Université d'Evry-Val-d'Essonne
Université Paris-Saclay
Université Versailles St-Quentin-en-Yvelines