

# Appel à projets

## Institut Convergence DATAIA

### Contexte général de l'appel

L'institut Convergence DATAIA en Sciences des Données, Intelligence Artificielle et Société, est l'Institut Convergence en France dédié aux sciences des données et à leurs interfaces disciplinaires et applicatives. Il rassemble des Universités, des organismes de recherche nationaux et des Grandes Ecoles et regroupe 130 enseignants-chercheurs. Financé par l'Agence Nationale de la Recherche (ANR) dans le cadre des "Programmes d'Investissements d'Avenir", l'Institut DATAIA vise à :

- 1/ rassembler des compétences de recherche multiples visant à produire des savoirs nouveaux par la mobilisation conjointe de différentes compétences disciplinaires ;
- 2/ développer, en lien avec ces recherches interdisciplinaires, des formations d'excellence innovantes aux niveaux master et doctorat, en formation initiale comme en formation continue ;
- 3/ déployer dans la durée une dynamique de structuration autour de l'enjeu principal de l'institut, appuyée sur une ambition scientifique de haut niveau international,

### Périmètre scientifique de l'appel et objectifs :

Les réponses à cet appel devront montrer en quoi elles concourent aux objectifs de l'institut DATAIA et notamment à la structuration du domaine de la science des données au sein du Campus Paris-Saclay et à la création de synergies nouvelles entre les différents acteurs scientifiques en interface avec des utilisateurs ou des producteurs de données.

En conséquence, toutes les propositions doivent s'inscrire dans au moins un des 4 défis principaux de DATAIA et peuvent impacter un domaine applicatif (voir la liste des « Axes de travail » en annexe 1). Elles devront s'appuyer sur la collaboration d'au moins deux personnes de deux établissements fondateurs de l'institut DATAIA n'appartenant pas au même laboratoire ni au même établissement hébergeur.

Cet appel 2018 visera différents types de soutien aux actions de DATAIA :

- AAP recherche (5 projets maximum) : chaque projet de recherche inclut le financement d'une thèse et d'un CDD de 2 ans (post-doc ou ingénieur) ou de deux thèses et éventuellement des frais de fonctionnement. La demande budgétaire par projet ne pourra pas excéder 220k€
- Organisation de workshop (à hauteur de 5k€ maximum - 10 slots disponibles)

- Invitation de professeurs pour des séjours thématiques (durée du séjour entre 1 mois et 6 mois - 12 mois d'invitation disponibles)

## Périmètre partenarial de l'appel

Cet appel est réservé aux équipes relevant du périmètre de l'institut de convergence DATAIA. Seuls les laboratoires des établissements fondateurs figurant dans la convention de financement pourront être porteurs et gestionnaires des crédits alloués (cf liste en annexe 2). Les consortiums pourront compter des partenaires hors des membres fondateurs de DATAIA mais ceux-ci ne seront pas financés.

## Modalités de soumission des propositions :

Chaque dossier comportera une partie de description scientifique du projet et une annexe financière. Les dossiers soumis hors délai ou incomplets ne seront pas examinés.

Format du dossier de soumission :

Le(s) porteur(s) de projet soumet(tent) un document comportant les informations suivantes :

- La nature de la demande : projet de recherche, organisation de workshop ou invitation de professeur
- ACRONYME et titre du projet
- Nom, laboratoire et coordonnées des porteurs
- Les laboratoires (ou équipes) de l'institut de convergence qui bénéficieront des financements, le contact scientifique et le contact administratif au sein de chaque laboratoire (équipe) du partenaire bénéficiaire
- Durée et dates envisagées du projet

Projet scientifique (5 pages max) :

Le projet décrira le contexte du projet et l'état de l'art, les objectifs du travail de recherche, son caractère innovant et les verrous, sa valeur ajoutée, les critères de succès, les pistes de travail envisagées, sa portée applicative, ainsi que l'adéquation avec les objectifs et axes de travail de l'institut de convergence.

Pour l'évaluation des projets, les critères suivants seront considérés : l'excellence scientifique, la synergie entre les partenaires, l'interdisciplinarité et l'impact applicatif potentiel lié au sujet traité.

Tous les projets doivent être soumis sur un site dont l'adresse sera communiquée ultérieurement. Tous les projets sont à soumettre avant le 9 février 2018 à 12h (heure de Paris).

Cet appel à projets est disponible sur le site <http://dataia.eu>

Vous pouvez poser vos questions sur cet appel à l'adresse [viviane.hoang@inria.fr](mailto:viviane.hoang@inria.fr)

Résultats de l'appel : 30 mars 2018

Mise à disposition des crédits alloués à partir d'avril 2018 et début de contractualisation entre l'Université Paris-Saclay avec les bénéficiaires.

## Modalité de sélection des projets :

L'éligibilité des projets sera traitée par les membres du bureau. Une pré-sélection des projets éligibles sera effectuée par le comité des programmes, qui pourra avoir recours à des rapporteurs extérieurs. Les porteurs des propositions présélectionnées passeront éventuellement une audition devant le comité des programmes qui délibérera pour décider des projets retenus. Les décisions finales seront publiées le 30 mars 2018.

## Validité des financements DATAIA :

Les projets de recherche sont financés pour une durée de 3 ans, et doivent être menés à terme le 31/12/2021. Chaque thèse financée devra impérativement démarrer en 2018.

## Engagement en cas d'acceptation :

En cas de financement du projet, il conviendra de communiquer à l'institut de convergence le plus rapidement possible le CV du/des candidat(s) retenu(s) (doctorant, post-doctorant ou ingénieur). La validation des candidats en thèse est conditionnée à l'approbation de l'école doctorale.

Le porteur de projet s'engage à utiliser les crédits conformément à la demande validée par le comité des programmes de DATAIA, à rendre compte annuellement de l'activité scientifique et de l'utilisation des crédits. En fin de projet, il remettra un rapport décrivant les réalisations et listant les publications effectuées dans le cadre du projet.

Les bénéficiaires du soutien de l'institut de convergence doivent faire apparaître dans toutes les communications scientifiques (ou non) associées aux activités financées la mention de DATAIA et des programmes d'investissements d'Avenir (PIA) en insérant la mention : « This research was supported by DATAIA convergence institute as part of the « Programme d'Investissement d'Avenir », (ANR-17-CONV-0003) operated by [Partner] XXX.

Si l'institut de convergence juge que les crédits n'ont pas été utilisés conformément aux objectifs annoncés, il pourra être demandé un remboursement. De même, les crédits non utilisés devront être remboursés par l'établissement partenaire.

## Annexe 1

### Axes de travail

Le programme scientifique de chaque proposition devra aborder explicitement **au moins un des quatre** axes de recherche esquissés ci-dessous, qui sont les quatre défis interdisciplinaires de DATAIA.

- **Des données à la connaissance, des données à la décision**
  - Données hétérogènes, complexes, incomplètes, semi-structurées et/ou incertaines ;
  - Fast big data : structuration de la donnée pour pouvoir l'exploiter ;
  - Apprentissage en ligne, méthodologie pour les données massives, méthodes efficaces ;
  - Amélioration du stockage, calcul et estimation pour la science des données ;
  - Modélisation des interactions entre agents (humains ou artificiels) par théorie des jeux ;
  - Représentation et algorithmes multi-échelle et multimodaux ;
  - Analyse théorique de méthodes heuristiques (théorie de la complexité, géométrie de l'information, théorie des chaînes de Markov) ;
  - Coévolution Humain-Machine dans les systèmes autonomes : agents conversationnels, voitures, robots sociaux.
  
- **De l'Apprentissage Statistique Profond à l'Intelligence Artificielle**
  - Nouvelles approches en apprentissage statistique et IA : sens commun, adaptabilité, généralisation ;
  - Apprentissage profond et apprentissage adversarial ;
  - Apprentissage automatique et hyper-optimisation ;
  - Optimisation pour l'apprentissage, e.g. améliorations des méthodes de gradient stochastique, optimisation bayésienne, optimisation combinatoire ;
  - Lien apprentissage-modélisation, intégration d'a priori dans l'apprentissage
  - Reproductibilité et apprentissage robuste ;
  - Inférence statistique et validation ;
  - Compositionnalité des architectures profondes.
  
- **Transparence et Confiance numérique**
  - Audit des systèmes algorithmiques : non-discrimination, loyauté, biais techniques, neutralité, équité ;
  - Mesure de la confiance et de l'appropriation du numérique ;
  - Explicabilité par construction ;
  - Transparence par construction, équité par construction ;

- Analyse progressive centrée utilisateur (supervision interactive des systèmes de décision : visualisation, tableaux de bord, IHM) ;
  - Responsabilité du traitement de l'information et de la prise de décision : contrôle d'usage des données et fact-checking ;
  - Découverte causale, traçabilité des inférences à partir des données sources, interprétabilité des architectures profondes.
  - Méthodologies éthiquement responsables à travers la transparence et la redevabilité des systèmes algorithmiques ;
- **Economie, régulation et protection de la donnée**
    - Responsabilité by-design ;
    - « Privacy-by-design », régulation générale de la protection des données (RGPD), incluant « differential privacy » ;
    - Apprentissage distribué respectueux de la vie privée;
    - Développement de méthodologies éthiquement responsables, et de technologies pour réguler la collecte, l'usage et le traitement des données personnelles, et l'exploitation des connaissances tirées de ces données ; abordant des problèmes de concentration et de monopole
    - Sécurité informatique des chaînes de traitement de données ;
    - Sécurité/cryptographie : blockchain et tiers de confiance.

Les projets pourront aborder des sujets plus spécifiques à des domaines d'application. Une liste de ces domaines est donnée ci-dessous à titre indicatif. Toutefois, il est demandé que les projets abordent des questions de science des données d'intérêt général, c'est-à-dire non restreintes spécifiquement à un domaine d'application.

- Energie : optimisation de la gestion (production et distribution), et régulation.
- Santé, nutrition personnalisée et bien-être.
- Urbanisation & Mobilité (véhicule connecté et autonome, villes intelligentes)
- « Business Analytics » (finance et assurance)
- Internet des personnes et des objets
- e-Sciences.

## **Annexe2**

### **Liste des membres fondateurs de DATAIA**

CentraleSupélec

CEA

CNRS

Ecole Polytechnique

ENSAE

HEC

IFP-EN

Inria

Institut Mines-Telecom (Telecom Paris Tech, Telecom Sud Paris, Telecom Ecole de Management)

Inra

Univ. d'Evry Val d'Ess.

Univ. Paris-Saclay

Univ. Paris-Sud

Univ. Versailles St-Quentin