

Appel à Professeurs invités

Institut DATAIA

Contexte général de l'appel

DATAIA est l'institut d'intelligence artificielle de l'Université Paris-Saclay dédié aux sciences des données, l'IA et à leurs interfaces disciplinaires et applicatives. Il rassemble plus de 800 enseignants-chercheurs provenant d'Universités, d'organismes de recherche nationaux et de Grandes Écoles. Financé en tant qu'institut convergence par l'Agence Nationale de la Recherche (ANR) dans le cadre des « Programmes d'Investissements d'Avenir », DATAIA vise à :

- Mobiliser des compétences pluridisciplinaires afin de produire des savoirs nouveaux ;
- Développer, en lien avec ces recherches interdisciplinaires, des formations d'excellence innovantes aux niveaux master et doctorat, en formation initiale comme en formation continue ;
- Structurer dans la durée des collaborations internationales et industrielles de haut niveau scientifique autour des principaux défis relevés par l'Institut (<https://www.dataia.eu/linstitut-dataia-paris-saclay>)

Périmètres scientifiques 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. En conséquence, toutes les propositions doivent s'inscrire dans au moins un des 4 défis principaux de DATAIA et peuvent éventuellement concerner un domaine applicatif (voir la liste des « Axes de travail » en Annexe 1). Les invités devront être rattachés à au moins deux équipes de recherche appartenant aux établissements partenaires de l'institut DATAIA (voir la liste en annexe 2).

Cet appel financera des invitations, au fil de l'eau, pour 3 à 12 mois, sans contrainte de date de début. La personnalité invitée doit occuper une position de chercheur ou d'enseignant-chercheur dans une université ou un organisme de recherche étranger. L'institut DATAIA finance la visite à hauteur de 4 000€ par mois sous forme d'allocation qui seront utilisés pour la rémunération du visiteur et/ou pour ses frais de séjour. DATAIA attribue une indemnité complémentaire pour les frais de voyage :

- Jusqu'à **1 000 €** pour un invité en provenance d'un pays d'Europe ;
- Jusqu'à **1 500€** pour un invité en provenance d'un pays hors d'Europe

Attendus/Engagements

Le visiteur s'engage à proposer un cycle de cours ou de séminaires aux niveaux master ou doctorat dans les laboratoires partenaires. Pour les séjours de longue durée, le jury appréciera l'engagement du visiteur dans les activités de co-encadrement d'étudiants (e.g., stages de M2, PhD, post-doctorants...) et de montage de collaboration.

Une synthèse d'une à deux pages des activités du visiteur sera établie par l'un des laboratoires d'accueil dans un délai de 1 mois après la visite. Ce rapport détaillera les activités menées durant le séjour, les faits marquants et les perspectives de collaborations futures. L'invitant s'engage par ailleurs à rendre compte à l'institut DATAIA des faits permettant de mesurer l'impact de la visite (projets ou publications communes en particulier).

Les bénéficiaires du soutien de l'Institut DATAIA doivent faire apparaître dans toutes les communications scientifiques associées aux activités financées la mention de DATAIA et des Programmes d'Investissement 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”

Périmètre partenarial de l'appel

Cet appel est réservé aux équipes relevant du périmètre de l'Institut DATAIA. Seuls les laboratoires des établissements partenaires figurant dans la convention de financement pourront être porteurs et gestionnaires des crédits alloués (cf. liste en Annexe 2).

Modalités de soumission des propositions

Le dossier sera élaboré conjointement par les laboratoires d'accueil et la personnalité pressentie pour une invitation. Le dossier sera soumis par un des laboratoires d'accueil par email à : submission-dataia@inria.fr. **Vous pouvez poser vos questions sur cet appel à : contact-dataia@inria.fr.**

Le dossier devra comporter d'une part, les informations administratives et financières, la description du dossier scientifique avec les éléments suivants :

- CV du/de la candidat(e) (3-5 pages indiquant ses coordonnées, son statut et son établissement d'origine ainsi que les principales contributions scientifiques) ;

- Axe scientifique de l'institut DATAIA (cf. Annexe 1) ;
- Programme de travail durant la visite (1 page max) ;
- Proposition de cours / séminaires (1-2 page(s) max) ;
- Retombées attendues pour l'Institut DATAIA.

Modalité de sélection des projets

Les projets seront étudiés et sélectionnés par le Bureau (comité exécutive) de DATAIA ainsi que son comité des programmes et une réponse sera apportée dans le mois suivant la soumission.

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 AUX CONNAISSANCES, DES DONNÉES À LA DÉCISION

La disponibilité croissante de données massives repousse les frontières techniques dans de nombreux champs. D'une part, la nature hétérogène, semi-structurée, incomplète ou incertaine des données remet en cause les modèles statistiques usuels ainsi que les algorithmes dédiés à la décision. D'autre part, la gestion des données soulève de nouvelles contraintes d'opérabilité telles que la sécurité, l'intégrité et la traçabilité.

En outre, produire de la connaissance nécessite de construire des modèles qui fournissent des décisions explicables, statistiquement valides et calculables. L'acceptation des résultats requiert aussi que la confidentialité et la loyauté soient renforcées. En parallèle, de nouveaux développements en optimisation doivent permettre d'améliorer les procédures d'estimation.

Défis :

- Données hétérogènes, complexes, incomplètes, semi-structurées et/ou incertaines
- Données massives : algorithmique et structuration des données
- 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-échelles 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

APPRENTISSAGE ET INTELLIGENCE ARTIFICIELLE

Récemment, la recherche en apprentissage profond a réalisé des avancées spectaculaires en vision par ordinateur et en traitement du langage naturel. Au-delà de l'arrivée des données massives, de l'accroissement de la puissance de calcul et des efforts de conception, les causes de ces progrès, encore mal connues, soulèvent au moins trois questions. Quelle théorie de l'apprentissage permettra d'analyser les architectures profondes ? Comment gérer la compositionnalité de ces architectures et leur capacité à appréhender des objets plus complexes ? Comment ouvrir la boîte noire pour mettre à jour les représentations apprises ?

Défis :

- Apprentissage automatique innovant 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 stochastiques, optimisation Bayésienne), optimisation combinatoire
- Lien apprentissage-modélisation, intégration à priori dans l'apprentissage
- Reproductibilité et apprentissage robuste
- Inférence statistique et validation
- Compositionnalité des architectures profondes.

TRANSPARENCE, IA RESPONSABLE ET ÉTHIQUE

La confiance numérique s'instaure à partir de l'implantation de méthodologies éthiquement responsables à travers la transparence et la redevabilité des systèmes algorithmiques ; de la régulation de la collecte, de l'usage et du traitement de données personnelles ; du renforcement de la régulation au travers de procédures numériques appropriées. La confidentialité par conception est une forme de régulation qui inclut la protection des données personnelles dans toutes les étapes de la collecte et du traitement. Le traçage des outils appliqués aux données doit également être développé de manière à faciliter l'explication du modèle pour les experts comme pour les utilisateurs rendant les systèmes algorithmiques auditables. Les principes de confidentialité, bien que faciles à formuler, nécessitent de modifier les infrastructures de stockage et de traitement, avec d'importants impacts législatifs, sociologiques et économiques. Les techniques de transparence des systèmes algorithmiques seront développées en focalisant sur : l'équité, la loyauté et la non-discrimination et la redevabilité-par-construction.

Défis :

- Responsabilité et explicabilité par conception
- Transparence et équité par conception
- Audit des systèmes algorithmiques : non-discrimination, loyauté, biais techniques, neutralité, équité
- Mesure de la confiance et de l'appropriation du numérique
- Progressive user- analyse de donnée progressive (monitoring interactif des systèmes de décisions)
- Responsabilité du traitement de l'information et de la prise de décision : contrôle d'usage des données et vérification des faits
- Découverte causale, traçabilité des inférences à partir des données sources, interprétabilité des architectures profondes

PROTECTION, RÉGULATION ET ÉCONOMIE DE LA DONNÉE

Les entreprises impliquées dans l'économie de la donnée ont continuellement besoin de repenser leur structuration : elles doivent adopter une organisation orientée-projet

avec des changements rapides dans l'allocation de ressources. L'économie de la donnée soulève aussi des problèmes de concentration et de monopole. Un petit nombre d'entreprises (GAFAM) détiennent la plupart des données. Cette concentration de marché peut conduire à une compétition déloyale dont l'innovation dans les petites et moyennes entreprises risque de pâtir. Les citoyens attendent des gouvernements qu'ils interviennent dans l'économie digitale pour éviter trop de concentration et de monopole. Les gouvernements doivent empêcher la fuite des informations pour préserver la souveraineté des états et le respect des régulations.

Défis :

- Respect de la vie privée par conception GDPR
- Apprentissage respectueux de la vie privée (differential privacy)
- Développement de méthodologies éthiquement responsables, et technologies pour réguler la collecte, l'usage et le traitement des données personnelles, et l'exploitation des connaissances tirées de ces données
- Sécurité informatique des chaînes de traitement de données
- Sécurité/crypto : block-chain 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...)*
- *Analyse pour la finance les assurances*
- *Internet des personnes et des objets*
- *E-Sciences*

Annexe 2 : Liste des partenaires de DATAIA éligibles à l'appel

<ul style="list-style-type: none">• Agro ParisTech• CentraleSupélec• CEA• CNRS• ENS Paris-Saclay• IFP- Énergies Nouvelles• Inria	<ul style="list-style-type: none">• Institut Mines-Télécom Business School• Inrae• ONERA• Université d'Evry Val d'Essonne• Université Paris-Saclay• Université Versailles St-Quentin-en-Yvelines
--	---