

INSTITUT DATAIA

Science des données, Intelligence & Société

KICK-OFF

JEUDI 15 FÉVRIER 2018



dataia.eu

université PARIS-SACLAY *Inria*
inventeurs du monde numérique



Bienvenue,

Nous sommes heureux de vous accueillir à la journée de lancement de l'Institut DATAIA, Institut Convergence dédié aux sciences des données et à leurs enjeux socio-économiques en France.

Financé par l'Agence Nationale de la Recherche (ANR) dans le cadre des programmes d'« Investissements d'Avenir », l'Institut DATAIA a pour vocation de regrouper et de structurer des expertises pluridisciplinaires de grande ampleur et de forte visibilité pour mieux répondre à des enjeux majeurs des sciences des données et de l'intelligence artificielle, à la croisée des défis sociétaux et économiques et des questionnements de la communauté scientifique. L'objectif est de décloisonner les disciplines avec un couplage fort entre la recherche, la formation et l'innovation.

À l'issue de deux appels, dix instituts convergences ont été sélectionnés toutes disciplines confondues sur des sujets comme : Climat, Santé, Agriculture, Migration etc. Par l'ambition de son programme scientifique et l'ampleur de ses actions et de ses relations internationales, il a été identifié par le jury international comme le centre français des sciences des données et de l'intelligence artificielle avec ses interfaces à la fois avec les sciences humaines et sociales mais aussi avec les domaines applicatifs à fort impact.

L'Institut DATAIA se veut être au cœur de la co-construction de la transformation numérique par les sciences mathématiques et informatiques avec les sciences juridiques, économiques et sociales. Cette journée est l'occasion de vous présenter les défis à relever ainsi que les verrous scientifiques, technologiques, juridiques, économiques et éthiques totalement interdépendants dans les années à venir.



Nozha Boujemaa
Directrice de l'Institut DATAIA

DES DÉFIS INTERDISCIPLINAIRES

● Apprentissage automatique et Intelligence Artificielle

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 d'a priori dans l'apprentissage
Reproductibilité et apprentissage robuste
Inférence statistique et validation
Compositionnalité des architectures profondes

● 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

● Transparence, responsabilité et éthique

Responsabilité-by-design, explicabilité-by-design
Audit des systèmes algorithmiques : non-discrimination, loyauté, biais techniques, neutralité, équité
Mesure de la confiance et de l'appropriation du numérique
Transparence-by-design, équité-by-design
« Progressive user-centric-analytics » (monitoring interactif des systèmes de décisions : dataviz, dashboards, 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

● Protection, régulation et économie des données

« Privacy-by-design », GDPR
Apprentissage respectueux de la vie privée (« differential privacy »)
Développement de méthodologies éthiquement responsables, et de technologies pour réguler la collection, 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

- 10h00 - 10h10 ▶ **Mots de Bienvenue** : **Gilles Bloch**, (Université Paris-Saclay), **Hervé Biaußer** (CentraleSupélec), **Bertrand Braunschweig** (Inria),
- 10h10 - 10h20 ▶ **Intervention d'Yves Lecointe** (Agence Nationale de la Recherche)
- 10h20 - 10h45 ▶ **Présentation de l'Institut DATAIA et de sa feuille de route** : **Nozha Boujemaa** (Inria)
- 10h45 - 10h55 ▶ **Intervention de Patrick Garda** (DGRI - Ministère de l'enseignement supérieur, de la recherche et de l'innovation)
- 10h55 - 11h10 ▶ **Mission Villani**, **Cédric Villani** (Mathématicien et Député) et **Marc Schoenauer** (Inria)
- 11h10 - 11h20 ▶ **Intervention de Guillaume Boudy** (Secrétariat général pour l'investissement)
- 11h20 - 12h15 ▶ Table ronde des partenaires animée par **Philippe Watteau** (CEA) avec la participation de :
- Hervé Biaußer** (CentraleSupélec)
 - Michel Bidoit** (CNRS)
 - Frank Pacard** (Ecole Polytechnique)
 - Yves Poilane** (Institut Mines-Télécom)
 - Sylvie Retailleau** (Université Paris-Sud)
 - Bernard Salha** (EDF)
 - François Sillion** (Inria)
- 12h15 - 13h30 ▶ **Cocktail déjeunatoire**
- 13h30 - 14h15 ▶ Echange avec les membres du bureau sur les actions de l'Institut : actions recherche, formation & valorisation ; appels à projets (modalité, fréquence, financement) ; etc.
- 14h15 - 14h40 ▶ «La recherche et l'enseignement avec Codalab», **Isabelle Guyon** (Université Paris-Sud)
- 14h40 - 15h05 ▶ «Graphical models & Artificial Vision», **Nikos Paragios** (CentraleSupélec)
- 15h05 - 15h30 ▶ **Pause café**
- 15h30 - 16h00 ▶ «L'avènement d'une gestion individuelle de nos données personnelles. Regards croisés sur certains enjeux de vie privée.», **Célia Zolynski** (Université Versailles Saint-Quentin-en-Yvelines) et **Nicolas Anciaux** (Inria)
- 16h00 - 17h15 ▶ Table ronde sur le thème «Transparence et responsabilité des technologies de l'Intelligence Artificielle», animée par **Laurence Devillers** (Limsi-CNRS) avec la participation de :
- Pierre-Jean Benghozi** (Ecole Polytechnique - ARCEP)
 - Jean-Marc David** (Groupe Renault)
 - Nicolas Miailhe** (The Future Society - Harvard Kennedy School)
 - David Sadek** (Groupe Thales)
- 17h15 - 18h00 ▶ Q&A général sur les thèmes scientifiques
- 18h00 - 19h30 ▶ **Cocktail apéritif**

14h15 - Isabelle GUYON [Université Paris-Sud]

« La recherche et l'enseignement avec Codalab »

Les défis en science des données sont des concours étalés sur plusieurs semaines ou plusieurs mois pour résoudre des problèmes à l'aide de jeux de données fournis ou d'environnements simulés. Les défis peuvent être considérés comme des outils de crowdsourcing, d'analyse comparative et de communication. Ils ont été utilisés pendant des décennies pour tester et comparer des solutions concurrentes dans l'apprentissage automatique d'une manière juste et contrôlée, pour éliminer le biais «inventeur-évaluateur» et pour stimuler la communauté scientifique tout en favorisant la science reproductible. Cet exposé présentera la plate-forme Codalab, un projet "open-source" dont Paris-Saclay est "community lead". Cette plate-forme sur laquelle plus de 100 challenges ont déjà été organisés pourra vous aider vous aussi dans l'enseignement et la recherche en transformant vos problèmes en défis !

Isabelle Guyon est professeur d'informatique à l'Université Paris-Saclay (UPSud, Orsay), spécialisée dans l'analyse de données statistiques, la reconnaissance de formes et l'apprentissage automatique. Ses domaines d'expertise comprennent la vision par ordinateur, la bioinformatique et les systèmes de transmission d'électricité, et les applications de l'apprentissage automatique à la découverte de relations causales. Avant de rejoindre Paris-Saclay, elle a travaillé comme consultante indépendante et a été chercheuse à AT & T Bell Laboratories, où elle a été pionnière des applications de réseaux neuronaux (avec des collaborateurs comme Yann LeCun et Yoshua Bengio) et co-inventé avec Bernhard Boser et Vladimir Vapnik les Support Vector Machines (SVM). Elle a organisé de nombreux défis dans l'apprentissage automatique depuis 2003, soutenus par la Commission Européenne, NSF et DARPA, avec des prix sponsorisés par Microsoft, Google, Facebook, Amazon, Disney Research et Texas Instrument. Isabelle Guyon est titulaire d'un doctorat diplômé en Sciences physiques de l'Université Pierre et Marie Curie, Paris, France. Elle est présidente de Chalearn, une organisation à but non lucratif vouée à l'organisation de défis, rédactrice de la série Challenges in Machine Learning de Springer, rédactrice au Journal of Machine Learning Research, co-général chair de NIPS 2017.



14h40 - Nikos Paragios [CentraleSupélec]

« (Higher-order) Graphical models & Artificial Vision »

Computational vision and biomedical image have made tremendous progress of the past decade. This is mostly due the development of efficient learning and inference algorithms which allow better, faster and richer modeling of visual perception tasks. Graph-based representations are among the most prominent tools to address such perception through the casting of perception as a graph optimization problem. In this paper, we briefly introduce the interest of such representations, discuss their strength and limitations and present their application to address a variety of problems in computer vision and biomedical image analysis.

NIKOS PARAGIOS (45 ans) est Professeur des Universités de classe exceptionnelle et responsable de l'axe enseignement « Sciences des données » à CentraleSupélec, « IEEE Fellow » et membre senior d'Institut Universitaire de France en mathématiques. Il est une référence mondiale pour ses travaux de recherche en vision par ordinateur (éditeur en chef de journal « Computer Vision and Image Understanding » d'Elsevier), imagerie médicale et apprentissage statistique. Il a participé à la rédaction de 5 livres, a dirigé 30 doctorants, est l'auteur de plus de 250 publications et détenteur de 25 brevets. Il est lauréat d'une bourse «ERC Proof of Concept» (2017) et auparavant d'une «ERC consolidator» (2011) de l'«European research Council», et membre du conseil scientifique du groupe Safran, de l'Institut de Data Healthcare ainsi que de l'Université de Paris Descartes. Ses travaux de recherche sont à l'origine de la société TheraPanacea qui vise à révolutionner la prise en charge des malades de cancer par radiothérapie via l'intelligence artificielle.



15h30 - Célia Zolynski [UVSQ] & Nicolas Anciaux [Inria]

« L'avènement d'une gestion individuelle de nos données personnelles - Regards croisés sur certains enjeux de vie privée »

L'entrée en application en mai 2018 du RGPD et du droit à la portabilité des données personnelles d'une part, et l'émergence actuelle de solutions de clouds personnels d'autre part, laissent augurer l'avènement d'une gestion individuelle et sous contrôle de notre «patrimoine numérique» personnel. En cette période charnière, nous proposons une analyse multi-disciplinaire des enjeux liés à ces nouveaux principes juridiques et techniques.



Agrégée des Facultés de droit, Célia Zolynski est Professeur de droit privé à l'Université de Versailles Saint Quentin où elle dirige la mention de Master Droit du Numérique et co-dirige le Master 2 Propriété intellectuelle et droit des affaires numériques (PID@N). Elle est également membre du Comité de prospective de la CNIL et du Conseil supérieur de la propriété littéraire et artistique (CSPLA) et a été membre du Conseil national du numérique (CNNum) de février 2016 à décembre 2017.

Ses activités de recherche et d'enseignement portent sur le droit du numérique, le droit de la propriété intellectuelle, le droit du marché et les libertés fondamentales. Elle est l'auteur de différentes publications dans ces domaines, notamment sur les liens qu'entretiennent le droit interne et le droit de l'Union européenne. Elle anime plusieurs groupes de travail interdisciplinaires et projets de

recherches collectives sur la protection et la valorisation des données et la régulation des traitements algorithmiques.

Liste complète des publications : <http://www.uvsq.fr/zolynski-celia-339560.kjsp>
celiazolynski@uvsq.fr - @celiazolynski



Nicolas Anciaux est chercheur Inria et dirige l'équipe Petrus, commune à Inria et à l'Université de Versailles. Ses activités de recherche se concentrent actuellement sur les systèmes de gestion de données personnelles de confiance et le respect de la vie privée. Il s'intéresse notamment aux structures de données et algorithmes pour environnements sécurisés, embarqués ou distribués, et à leur intégration dans des architectures respectueuses de la vie privée. Nicolas conduit aussi des activités de recherche multidisciplinaires sur ces thèmes en coopération avec des juristes et des économistes. Il est co-concepteur de PlugDB, un serveur personnel et sécurisé, en cours d'industrialisation avec la société Hippocad en vue d'un déploiement sur le territoire des Yvelines pour le suivi à domicile de personnes dépendantes.

16h00 - « Transparence & confiance dans les technologies de l'IA »

Pierre-Jean Benghozi [Ecole polytechnique - ARCEP]

Ancien élève de l'Ecole polytechnique, titulaire d'un doctorat en gestion et d'une HDR en économie de l'Université Paris Dauphine, Pierre-Jean Benghozi est membre du Collège de l'Autorité des communications électroniques et des postes (ARCEP), Directeur de recherche CNRS et professeur à l'Ecole polytechnique, il y a dirigé, jusqu'en 2013, le Pôle de Recherche en Economie et Gestion et la Chaire « Innovation et Régulation des Services Numériques » qu'il avait fondé, en partenariat avec Telecom ParisTech. Pierre-Jean Benghozi est un des précurseurs des recherches sur l'économie de l'internet, la numérisation des entreprises et des modèles économiques, notamment dans les industries créatives. Au sein du Collège de l'Arcep, il suit notamment le marché entreprise, l'internet des objets et le déploiement des réseaux très haut débit. Auteur de plus de 150 publications académiques, Pierre-Jean Benghozi enseigne dans plusieurs grands établissements parisiens et étrangers et est régulièrement sollicité comme expert. Il est ainsi membre du comité de prospective de la CNIL.



Nicolas Miaillhe [The Future Society - Harvard]

Nicolas co-founded The Future Society at Harvard Kennedy School in 2014 which focuses on questions of impact and governance of emerging technologies including Artificial Intelligence through its AI Initiative. Nicolas is also the co-founder of Aletheion, an AI & cybersecurity start-up based in Paris. A recognized strategist and thought-leader, Nicolas advises multinationals, governments and international organizations. He has over fifteen years of professional experience working at the nexus of innovation, technology, government, industry and civil society across Europe, America and Asia. Nicolas teaches at the Paris School of International Affairs (Sciences Po), is a Senior Visiting Research Fellow with the Program on Science, Technology and Society at Harvard, and a Fellow with the Center for the Governance of Change at IE Business School in Madrid. Nicolas is also the co-founder of Yes Europe Lab: a pan-European civic entrepreneurship lab. An Arthur Sachs Scholar, he holds a Master in Public Administration from the Harvard Kennedy School, a Master in Defense, Geostrategy & Industrial Dynamics from Panthéon-Assas University, and a Bachelor of Arts in European Affairs and International Relations from Sciences Po Strasbourg.



Jean-Marc David [Renault]

Jean-Marc est Expert Leader en Intelligence Artificielle pour le groupe Renault. A ce titre il a la responsabilité d'animer l'utilisation des technologies de Big Data et d'Intelligence Artificielle dans l'ensemble des fonctions du groupe, de construire la vision moyen et long terme, d'identifier les compétences nécessaires dans ce domaine, et de contribuer à définir les partenariats stratégiques. Jean-Marc participe également à l'Equipe Transverse Innovation associée au CVC Renault-Nissan-Mitsubishi (un nouveau fond de Corporate Venture Capital lancé en janvier 2018, qui prévoit d'investir un milliard de \$ dans les 5 prochaines années pour supporter l'open innovation de l'Alliance). Dans son poste précédent, au sein de la Recherche de Renault, Jean-Marc avait la responsabilité du développement du véhicule autonome, des services connectés, ainsi que des nouveaux services de mobilité. A ce titre il a lancé le programme de recherche sur les robo-taxis qui est maintenant inscrit dans le Mid-Term Plan de l'entreprise. Par ailleurs, Jean-Marc est président du comité 'Big Data & Intelligence Artificielle' du pôle de compétitivité Systematic, dont l'objectif est de promouvoir ces technologies et de faire émerger des champions au sien de l'écosystème Systematic. Jean-Marc est ingénieur Telecom-Paris Tech, et docteur en Intelligence Artificielle.



Laurence Devillers [LIMSI - CNRS - Institut DATAIA]

Laurence Devillers est Professeur en informatique à Sorbonne-Université et chercheur au LIMSI-CNRS/Paris-Saclay sur les « Dimensions affectives et sociales dans les interactions parlées: enjeux technologiques et éthiques ». Elle est membre de la Commission de réflexion sur l'éthique de la recherche en sciences et technologies du numérique (CERNA) d'Allistène et est impliquée dans « The IEEE Global Initiative for Ethical Considerations in the Design of Autonomous Systems » depuis 2016 (<http://innovationatworkieee.org/courses/artificial-intelligence-and-ethics-in-design/>), elle anime notamment le groupe de réflexion P7008 sur le nudging. Elle est l'auteur de l'essai « Des robots et des hommes : mythes, fantasmes et réalité »,

(Plon, 2017). Elle est chargée de mission au sein du bureau de l'Institut convergence DATAIA et est fortement impliquée dans le déploiement de la plateforme TransAlgo



David Sadek [Thales]

David Sadek est Vice-Président Recherche, Technologie & Innovation à Thales. Il a été précédemment Directeur de la Recherche de l'Institut Mines-Télécom et auparavant Directeur délégué à la Recherche d'Orange. Docteur en Informatique et spécialiste en Intelligence Artificielle et technologies cognitives, il a créé et dirigé chez Orange pendant plus de quinze ans les activités puis le laboratoire de RGD sur les agents intelligents et le dialogue naturel humain-machine. Il est à l'origine des premières réalisations technologiques d'agents conversationnels, au niveau mondial, ainsi que du standard ACL de langage de communication inter-agent. Il a également piloté plusieurs programmes de transfert industriel et de déploiement de services. Il est ou a été membre de nombreux comités scientifiques et technologiques, nationaux et internationaux, et de pilotage et d'évaluation de la recherche et de l'innovation, ainsi que de groupes de réflexion, de prospective ou d'influence : Président du Comité d'évaluation du programme « Contenus numériques & Interactions » de l'ANR, Conseil Scientifique d'Inria et celui du département STIC du CNRS, groupe « Facteurs humains » du Conseil Scientifique Défense, Bureau Exécutif du RNRT, OISPG (Open Innovation Strategy Policy Group) de la Commission Européenne, Think Tank « Futur numérique », groupe Innovation du Medef, Comité de Coordination d'Allistène, Comité 21... Il a aussi été membre du groupe fondateur de la CERNA, Commission de réflexion sur l'éthique de la recherche en sciences et technologies du numérique d'Allistène. Il a été récemment membre du Conseil d'Orientation de l'initiative FranceIA de stratégie nationale en Intelligence Artificielle, et présidé dans ce cadre les groupes « Cartographie » et « Industrialisation et appropriation des résultats de la recherche par les entreprises ». Il a également été représentant de la France au « G7 des innovateurs ». David Sadek a eu pendant près de dix ans le titre d'Expert Emérite du groupe Orange. Il a reçu le Prix France Télécom de la meilleure innovation technologique, le Prix Orange de la meilleure innovation de service, le Trophée des Casques d'Or du Forum Européen de la Relation Client. Il est également lauréat de la Médaille Blondel.





INSTITUT **DATAIA**
Science des données, Intelligence & Société

dataia.eu



Institut Convergence 17-CONV-0003 INSTITUT DATAIA (I2-DRIVE)

1 rue Honoré d'Estienne d'Orves
Campus Polytechnique
91120 Palaiseau
com-dataia@inria.fr