

Lundi 1^{er} et mardi 2 décembre 2025

ANDOR

Assises
nationales
des
données
de la
recherche

DOSSIER de presse

université
PARIS-SACLAY

Sommaire

Éditos	3
Historique	5
Pourquoi partager les données de la recherche est un enjeu crucial pour la science ?	6
Intelligence artificielle, souveraineté informatique et enjeux environnementaux	6
La politique nationale des données de la recherche	7
Les données de la recherche à l'Université Paris-Saclay	7
Présentation des Assises nationales des données de la recherche, 1 ^{er} et 2 décembre 2025	8
Les intervenantes et intervenants	9
La semaine des données de la recherche : les événements satellites dans la continuité des ANDOR 2025	12

Éditos



Camille Galap, président de l'Université Paris-Saclay

Les données de la recherche constituent une matière première dont on extrait de la connaissance. Elles sont le fruit du travail de nos communautés scientifiques. Chercheurs et chercheuses, enseignants-chercheurs et enseignantes-chercheuses, doctorants et doctorantes, ingénieres et ingénieurs de recherche, personnels administratifs, bibliothécaires, archivistes, ... l'ensemble des corps de métier de la recherche contribuent à leur collecte, à leur sauvegarde ou à leur diffusion. Alors que le contexte international rend les enjeux de souveraineté numérique et scientifique de plus en plus prégnants, il nous incombe collectivement de préserver et de valoriser le potentiel de ce patrimoine scientifique immatériel.

En accueillant à l'Université Paris-Saclay, au sein de l'ENS Paris-Saclay, la deuxième édition des Assises nationales des données de la recherche, organisées conjointement avec le ministère de l'Enseignement supérieur, de la Recherche et de l'Espace, l'Université Paris-Saclay est heureuse de contribuer à placer au cœur des discussions cette problématique centrale pour la recherche.



Etienne Augé, vice-président délégué à la science ouverte de l'Université Paris-Saclay

Il n'y a pas de science véritablement ouverte sans bonne gestion des données de la recherche. La tenue des Assises nationales des données de la recherche, où seront remis les prix nationaux de la science ouverte, sera l'occasion de rappeler les liens étroits qui existent entre l'ouverture des résultats de la recherche expérimentale et les bonnes pratiques qui précèdent nécessairement le partage de jeux de données.

En la matière, l'Université Paris-Saclay s'inscrit pleinement dans la ligne des politiques nationales et européennes : il s'agit de publier nos données autant que possible, afin de rendre notre recherche à la fois plus cumulative, plus collaborative et plus efficiente. La transformation des pratiques vers laquelle nous allons est globale, et en même temps spécifique à chaque domaine de recherche. C'est cette réalité que s'attacheront à transcrire les ANDOR, à travers les enjeux liés aux données sensibles, aux évolutions de l'intelligence artificielle, ou encore aux questions de souveraineté sur nos infrastructures numériques.



Isabelle Blanc, Administratrice ministérielle des données, algorithmes et codes sources au ministère de l'Enseignement supérieur, de la Recherche et de l'Espace

Face à des défis planétaires comme le changement climatique, les pandémies ou les migrations, aucun pays ni aucune discipline ne peut prétendre disposer seul de toutes les réponses. Les données doivent circuler au-delà des frontières géographiques, linguistiques et disciplinaires, et s'intégrer dans des systèmes interopérables.

Les données de recherche, qu'elles proviennent de mesures quantitatives, d'entretiens, d'images, d'échantillons biologiques ou de corpus textuels, constituent le socle de la production scientifique. Elles ne se limitent pas à des chiffres ou des enregistrements : elles témoignent de phénomènes et d'observations qui, une fois analysées, nourrissent la construction de modèles, d'hypothèses et de prédictions. La fiabilité, la qualité et l'accessibilité de ces données conditionnent la robustesse des résultats et la reproductibilité des travaux, principe central de la démarche scientifique. Leur gestion rigoureuse, leur documentation, leur préservation, leur archivage et leur partage maîtrisé, est donc essentielle pour éviter les biais, renforcer la confiance et garantir que les savoirs produits puissent être vérifiés et réutilisés. À ce titre, les données sont désormais reconnues comme une production scientifique à part entière, et non comme un sous-produit des travaux de recherche.

Leur fiabilité, leur traçabilité, leur accessibilité et leur gouvernance conditionnent non seulement la robustesse des travaux de recherche mais également la capacité de notre pays à innover, coopérer et éclairer les décisions publiques.

Ces assises nationales des données de la recherche offrent un moment privilégié pour engager une réflexion collective sur les responsabilités qui nous incombent : garantir la pérennité des données, favoriser leur réutilisation, développer des infrastructures souveraines et ouvertes, et promouvoir une culture de la donnée qui traverse toutes les disciplines, de la recherche fondamentale aux sciences humaines et sociales.

Historique

2016

Adoption de la **loi pour une République numérique** qui a généralisé l'open data dans le secteur public

2019

Financement sur fonds publics des projets de recherche conditionné à la **mise en place d'un plan de gestion des données**

2022

Inauguration de Recherche Data Gouv, écosystème pour accompagner les équipes de recherche dans la gestion, la préservation et le partage des données de recherche

2025

2^e édition des Assises nationales des données de la recherche (ANDOR 2025) qui se tiennent à l'Université Paris-Saclay les 1^{er} et 2 décembre

2018

Premier **plan national pour la science ouverte**

2021

Le MESR se dote d'une **feuille de route sur la politique des données, des algorithmes et des codes sources et du 2^e plan national pour la science ouverte**

2024

1^{re} édition des Assises nationales des données de la recherche (ANDOR) qui se tiennent à Aix Marseille Université

En 2016, la Loi pour une République numérique a généralisé l'open data dans le secteur public, en posant le principe de la libre réutilisation des données de la recherche, lorsque les productions scientifiques sont financées au moins 50% sur fonds publics. Ces objectifs doivent se réaliser dans le respect du Règlement général sur la protection des données, adopté la même année par l'Union européenne, visant à protéger les droits des citoyennes et des citoyens européens, et trouvent un écho dans la gestion des données de santé. Le premier plan national pour la science ouverte publié en 2018 par le ministère de l'Enseignement supérieur et de la Recherche soutient et précise les actions pour accompagner la gestion maîtrisée et le partage des données de recherche.

Depuis 2019, les financeurs publics français et européens de la recherche demandent des Plans de Gestion des Données pour les projets financés, afin d'anticiper les moyens pour leur gestion et pour préparer les conditions de leur réutilisation dans un nouveau projet, une fois le programme de recherche initial achevé.

Dans ce contexte, le ministère de l'Enseignement supérieur et de la Recherche s'est doté en 2021 d'une feuille de route sur la politique des données, des algorithmes et des codes sources, ainsi que d'un deuxième plan national de la science ouverte. Crées en 2024, les Assises nationales des données de la recherche (ANDOR) sont un rendez-vous important pour les acteurs et actrices des données de recherche et de la science ouverte.

Pourquoi partager les données de la recherche est un enjeu crucial pour la science ?

La croissance rapide du volume de données produits par la recherche et l'importance des investissements humains et financiers pour les obtenir ont rendu la question de la gestion des données de la recherche incontournable pour les scientifiques et décideurs.

Ces données représentent un patrimoine de première importance.

Les principes de bonne gestion des données de la recherche passent notamment par le respect des principes FAIR (encadré) : les données doivent être faciles à trouver, accessibles, interopérables et réutilisables. Afin d'assurer la structuration et l'ouverture des données, dans le but d'en faciliter la réutilisation, les services d'appui à la recherche, et en particulier les bibliothèques universitaires, ont développé une expertise dédiée à ces questions.

La mise à disposition des données de recherche permet leur réutilisation et rend plus visibles les travaux de recherche ; une [étude récente](#) indique que les articles scientifiques avec des données ouvertes sont davantage cités que les autres. La diffusion et le partage des données constituent la base du principe de reproductibilité de la science, garantissant intégrité scientifique et éthique, suivant les critères établis par l'UNESCO pour la science ouverte.

Qu'est-ce que des données FAIR ?

FAIR (Facile à trouver, Accessible, Interopérable et Réutilisable) est une initiative lancée par l'organisation [Force 11](#) rassemblant des communautés d'académiques, bibliothécaires, archivistes, éditeurs et bailleurs de fonds de la recherche. Chaque principe FAIR se décline en un ensemble de caractéristiques que doivent présenter les données et les métadonnées pour faciliter leur découverte et leur utilisation par les humains mais aussi par les machines. Les données ne peuvent pas toujours être open, mais elles gagnent à minima à être FAIR.

Intelligence artificielle, souveraineté informatique et enjeux environnementaux

Les données de la recherche sont un enjeu majeur pour l'intelligence artificielle, en particulier l'intelligence artificielle générative qui a besoin de s'appuyer sur de grands volumes de données. Elle offre en retour de nouvelles perspectives pour la recherche, en particulier sur le traitement de très gros volumes de données peu structurés. Cependant, l'intelligence artificielle s'appuie sur des infrastructures (datacenters, supercalculateurs...) très gourmandes en énergie.

La recherche européenne dépend largement d'infrastructures américaines pour le traitement et le stockage de ses données. Alors que dans certaines disciplines telles que les sciences de la vie ou les sciences du climat les chercheuses et chercheurs accèdent de plus en plus difficilement aux données stockées aux États-Unis, la question de la souveraineté sur les données produites en France et en Europe se pose avec force.

La maîtrise de l'ensemble du « cycle de vie » des données (collecte, stockage et partage des données), sans dépendre de pays tiers extra-européens, représente un réel enjeu. L'Europe se trouve actuellement dans une période charnière, où elle doit consolider sa stratégie sur les données de la recherche, entre protection des données sensibles, concurrence industrielle et ouverture de la science.

La politique nationale des données de la recherche

Les actions du ministère et de ses opérateurs ont trois priorités :

- L'innovation, en partageant et exploitant le potentiel des données et des logiciels de l'ESR ;
- La confiance, en ouvrant pour la transparence de l'action publique ;
- La simplification, en facilitant la circulation des données entre administrations et établissements selon le principe du « Dites-le nous une fois »¹.

Les enjeux actuels des données s'organisent autour de quatre axes interdépendants :

- Promouvoir le partage et la valorisation des données et logiciels de recherche en favorisant leur potentiel d'innovation pour générer de nouvelles connaissances et services ;
- Renforcer la souveraineté scientifique française et européenne ;
- Développer une culture de la donnée qu'elle soit de recherche, de formation, de gestion ou de pilotage ;
- Exploiter des données institutionnelles comme instrument stratégique de pilotage des établissements.

Les données de la recherche à l'Université Paris-Saclay

L'Université Paris-Saclay s'est dotée en 2022 d'une stratégie pour la science ouverte, visant entre autres à accompagner les communautés de recherche dans la gestion et l'ouverture de leurs jeux de données scientifiques. Une stratégie sur les données codes et logiciels est en cours de formalisation, afin notamment de doter les jeunes chercheurs et chercheuses des bons réflexes en la matière. La réflexion sur la souveraineté des données de recherche est au cœur des discussions, avec le développement d'infrastructures et d'espaces de stockage qui permettent de gérer les données produites dans les laboratoires de l'université sur des serveurs hébergés et gérés par l'université.

DataSaclay

Depuis 2023, l'université a mis en place un Atelier de la donnée, DataSaclay, labellisé par le ministère et proposant un accompagnement complet avec des spécialisations disciplinaires, ainsi qu'un accompagnement juridique à la communauté scientifique et aux laboratoires affiliés. Les chercheurs et chercheuses sont accompagné·es pour comprendre comment gérer les données de leur recherche, rédiger leurs plans de gestion de données, déposer leurs données dans des espaces numériques, leur référencement et la curation de leurs métadonnées.

¹ Interview Isabelle Blanc Enssib 2025 : <https://www.enssib.fr/interview-biennale-du-numerique-isabelle-blanc>

Présentation des Assises nationales des données de la recherche, 1^{er} et 2 décembre 2025

Ces journées seront l'occasion d'un état des lieux de la situation et des bonnes pratiques en matière de données de la recherche autour de spécialistes ou d'exemple signifiants. La participation d'acteurs et actrices du monde économique permettra de faire le lien avec les enjeux industriels sur la réutilisation des données par les entreprises favorisant l'innovation.

Les ANDOR offrent une visibilité accrue des enjeux sur les données de la recherche auprès des communautés scientifiques en France afin de transformer les pratiques et permettre aux chercheurs et chercheuses de se saisir de la question dans l'exercice de leur métier.

Pendant ces deux jours, des acteurs et actrices académiques et du monde économique seront réunis pour aborder les enjeux liés à la production, la gestion, le traitement, la diffusion et la réutilisation des données de recherche, alors que la souveraineté européenne sur les données scientifiques est mise au premier plan par l'actualité internationale.

Les problématiques de traitement des données sensibles et de réutilisation des données, ainsi que d'intelligence artificielle sont notamment au programme.

Lors de cet événement se tiendra la cérémonie de remise des [Prix science ouverte](#), organisée par le ministère, le lundi 1^{er} décembre à 17h30.

Les discussions se prolongeront ensuite toute la semaine autour d'événements satellites consacrés aux données et plus particulièrement aux plateformes de recherche, aux sciences participatives et aux enjeux européens.

Les intervenantes et intervenants

Session 1 : L'accès aux données sensibles (lundi 1^{er} décembre, 14h30-17h30)



Bonnie Wolff-Boenisch

Titulaire d'un doctorat en géochimie isotopique de l'Institut Max-Planck de Mayence en Allemagne, Bonnie Wolff-Boenisch est directrice de CESSDA ERIC (Consortium of European Social Science Data Archives). Elle est membre de comités consultatifs scientifiques en Allemagne, en Italie, en France et aux États-Unis. CESSDA est une infrastructure européenne pour les centres nationaux d'archives en sciences sociales.



Kamel Gadouche

Directeur du Centre d'accès sécurisé distant aux données (CASD), dont le but est de fournir des environnements de calcul pour le traitement des données sensibles pour les chercheurs. Kamel Gadouche assure également la direction de Datastorm, organisme délivrant une expertise et des services data et IA aux entreprises et aux organismes publics.



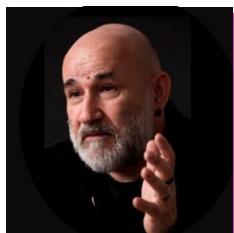
Hélène Hébert

Directrice de recherche au CEA, Hélène Hébert est spécialiste des tsunamis au Département analyse surveillance environnement (DASE) de la Direction des applications militaires du Commissariat à l'énergie atomique (CEA), où elle coordonne le Centre national d'alerte aux tsunamis (Cenalt).



Rémy Marquier

Diplômé statisticien-économiste de l'ENSAE Paris, Rémy Marquier s'est spécialisé en statistiques, probabilités et économie. Il a dernièrement occupé le poste de Délégué à la protection des données (DPO) et responsable data au Centre d'Accès Sécurisé aux Données, où il a œuvré pour la sécurisation et la conformité des traitements de données sensibles (notamment de santé), avant d'être nommé en 2025 directeur de projets Données de santé à la Drees-Ministères chargés de la Santé et des Solidarités.



Hugues Berry

Chercheur en bio-informatique et directeur de recherche à l'Inria/Inserm. Hugues Berry est adjoint au directeur scientifique, chargé de mettre en œuvre la stratégie d'Inria sur l'application des sciences du numérique à la biologie ou à la médecine.



Yves Vandenbrouck

Titulaire d'un doctorat en physiopathologie humaine et d'un Master en informatique, Yves Vandenbrouck est directeur de recherche au CEA. Ses travaux portent sur le développement d'approches numériques pour la découverte de biomarqueurs d'intérêt clinique et l'analyse des données massives multimodales (« omiques »). Il est également chargé de mission en santé numérique auprès de la Direction de Recherche fondamentale du CEA.



Marie Zins

Médecin épidémiologiste et professeure des universités en santé publique à l'Université Paris-Cité, Marie Zins dirige l'infrastructure nationale de recherche « Constances », la plus importante cohorte épidémiologique en population française.



Anne-Françoise Adam-Blondon

Directrice de recherche à INRAE, Anne-Françoise Adam-Blondon est directrice de l'Unité Ressources Génomique-Info (URGI), directrice adjointe de l'Institut Français de bioinformatique et Head of Node pour la France de l'infrastructure européenne ELIXIR.



Michael Alaux

Michaël Alaux est ingénieur de recherche à INRAE. Il est directeur adjoint de l'Unité ressources génomique-Info (URGI - INRAE / Univ. Paris-Saclay) et coordinateur scientifique de la plateforme bio-informatique. Au niveau international, il est président du comité d'experts sur le système d'information blé (WheatIS) de la Wheat Initiative.



Sarah Cohen-Boulakia

Professeure à l'Université Paris-Saclay, spécialiste en sciences des données en bio-informatique, directrice adjointe de DATAIA, Sarah Cohen-Boulakia coordonne le volet formation de l'IA cluster DATAIA et le projet CMA IA SacalAI-School. Elle a co-fondé et est membre du comité de pilotage du réseau national de recherche reproductible soutenu par le MESRE.



Hervé Roquet

Hervé Roquet est Ingénieur des ponts, des eaux et des forêts et actuellement directeur adjoint de l'enseignement supérieur et de la recherche à Météo-France. Il est également référent à l'intégrité scientifique auprès de l'Office français de l'intégrité scientifique (OFIS).



Stéphane Requena

Directeur de la technologie et de l'innovation chez GENCI - Grand équipement national de calcul intensif, Stéphane Requena a participé à plusieurs projets européens d'envergure comme PRACE (Partnership for Advanced Computing in Europe) et EXDCI (European eXtreme Data and Computing Initiative) et est actuellement chair de l'INFRAG un des deux comités d'experts de l'infrastructure européenne EuroHPC.



Valérie Masson-Delmotte

Valérie Masson-Delmotte est paléoclimatologue, chercheuse au Laboratoire des sciences du climat et de l'environnement (LSCE) du CEA. Elle a co-présidé le groupe n°1 du GIEC de 2015 à 2023, et est membre du Haut Conseil pour le climat, du Comité consultatif national d'éthique et de l'Académie des sciences.

Keynote : Droit des données de la recherche



Agnès Robin

Agnès Robin est maître de conférences HDR en Droit privé à l'Université de Montpellier. Elle est directrice adjointe du Laboratoire innovation, communication et marché (LICeM), Directrice du Master Droit de la propriété intellectuelle et du numérique, Responsable du DU Scientific data management (Gestion des données de la recherche). Elle pilote depuis 2017 le projet de recherche CommonData (« Les données de la recherche, des communs scientifiques ? ») et est autrice de plusieurs ouvrages dont l'ouvrage « Droit des données de la recherche » (Coll. Création, information, communication, Larcier, 2022). <https://univ-droit.fr/universitaires/4532-robin-agnes-robin-agnes>

Session 3 : Données et intelligence artificielle (mardi 2 décembre, 13h45-16h00)



Céline Hudelot

Céline Hudelot est professeur des universités en informatique à CentraleSupélec et directrice du laboratoire MICS (CentraleSupélec/Univ. Paris-Saclay) depuis octobre 2022. Elle a obtenu son doctorat en informatique de l'INRIA et de l'Université de Nice Sophia Antipolis en 2005 et l'Habilitation à Diriger des Recherches de l'Université Paris-Sud en 2014. Elle a encadré plus de 20 thèses en IA (soutenues ou en cours). Elle est responsable de l'axe transverse Intelligence Artificielle du laboratoire et co-directrice de l'option Intelligence Artificielle de la formation d'ingénieur de CentraleSupélec.

Ses travaux de recherche se situent à la frontière des domaines de l'apprentissage automatique et de la représentation et du raisonnement des connaissances. Ses recherches actuelles portent sur l'interprétabilité et l'explicabilité des algorithmes d'intelligence artificielle, l'apprentissage de représentations à partir de données non structurées (images, textes, documents...) avec différents paradigmes d'apprentissage (avec peu d'étiquettes, semi-supervisé, actif, incrémental, etc.), l'intelligence hybride et les modèles de fondations.



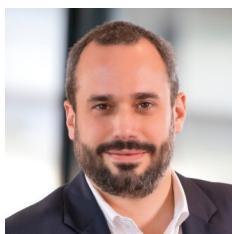
Renaud Vedel

Renaud Vedel est préfet, conseiller d'État, ancien coordinateur national pour l'intelligence artificielle. Il est spécialiste des enjeux technologiques et de sécurité, et met son expertise au service des politiques publiques du numérique.



Frédéric Pascal

Professeur des universités à CentraleSupélec, Université Paris-Saclay, spécialisé en Mathématiques appliquées pour l'Intelligence artificielle, au sein du Laboratoire des signaux et systèmes (L2S - CNRS / CentraleSupélec / Univ. Paris-Saclay). Il dirige l'Institut DataIA (cluster IA) depuis 2023. À la fin de 2024, il reçoit une mission du ministre de l'Enseignement supérieur et de la Recherche sur les questions d'IA et de pédagogie. Depuis avril 2025, Frédéric Pascal a été nommé vice-président IA de l'Université Paris-Saclay.



Olivier Biton

Ingénieur en informatique, diplômé de l'école Polytech Paris-Sud, Olivier Biton est directeur de la transformation technologique de Crédit Agricole S.A. et directeur général de Crédit Agricole Group Infrastructure Platform.



Marc De Falco

Marc de Falco est inspecteur général de l'éducation, du sport et de la recherche, ancien élève de l'École Normale Supérieure de Lyon en informatique et titulaire d'un doctorat de Mathématiques de l'Université d'Aix-Marseille. Il est, à l'IGÉSR, responsable du Pôle Données, IA et Numérique et membre des groupes mathématiques et informatique de l'IGÉSR. Il a participé à plusieurs missions sur l'usage de l'IA du point de vue pédagogique ou administratif dans les services, l'enseignement scolaire et l'enseignement supérieur.

La semaine des données de la recherche : les événements satellites dans la continuité des ANDOR 2025

L'événement satellite EUGLOH Research Data Days (3-4 décembre 2025)

Dans la continuité des ANDOR, ce rendez-vous propose de souligner l'importance de l'ouverture des données de recherche au sein de l'alliance EUGLOH (European University Alliance for Global Health).

Cet événement satellite mettra l'accent sur l'importance des enjeux et problématiques des « données ouvertes et de la transparence des résultats scientifiques » dans l'alliance EUGLOH (ainsi que dans le contexte international).

L'événement satellite de ces deux jours sera coorganisé au Lumen conjointement par l'Université Paris-Saclay et l'UiT – Université arctique de Norvège à Tromsø. Cet événement sera l'occasion de mobiliser et d'accueillir 16 étudiantes et étudiants internationaux en master et en doctorat d'EUGLOH venus par le programme Erasmus, et de visiter le campus, les laboratoires, les infrastructures de données et les expositions hébergées au Lumen (Data Art Work, posters des étudiantes et étudiants EUGLOH sur les données de recherche, etc.).

Dans ce cadre, trois collègues de la bibliothèque de l'Université de Tromsø viennent à l'Université Paris-Saclay pour travailler ensemble avec les collègues de la DiBISO pour concevoir un atelier d'une journée (le 4 décembre) au Lumen et intitulé « Data Stewardship » qui sera destiné à la communauté étudiante chercheuse sur les diverses questions de gestion de données de recherche et de la science ouverte.

Programme : <https://www.eugloh.eu/courses-trainings/activities/eugloh-research-data-days-physical-participation/>

L'événement satellite Plateformes de recherche

Cet événement réunit des plateformes de recherche lauréates d'un appel à manifestation d'intérêt interne à l'Université Paris-Saclay, ayant pour thème les data papers (articles scientifiques décrivant des jeux de données) pour un temps d'échange dédié à la valorisation et à la gestion des données de recherche.

Décliné autour de présentations plénières et d'une table ronde, cet événement met en débat les bonnes pratiques, les défis et les innovations en matière de données issues des infrastructures locales. Une initiative pour renforcer la collaboration entre acteurs et actrices, encourager la mutualisation des savoir-faire et renforcer la visibilité des plateformes de recherche de l'Université Paris-Saclay dans l'écosystème national.

Journée d'étude sur les Sciences participatives le 5 décembre

Cette journée d'étude sur les sciences participatives est à destination des chercheurs et chercheuses, enseignants-chercheurs et enseignantes-chercheuses, doctorant·es et étudiant·es en master.

La journée a pour objectif de rassembler des chercheurs et chercheuses souhaitant recueillir des pistes pour se lancer dans un projet de science participative, et pouvoir échanger autour de retours d'expérience et de bonnes pratiques.

Novembre 2025
universite-paris-saclay.fr

