

Cellule “Données de la recherche” d’AgroParisTech

Missions et services

I. Présentation générale

La cellule « Données de la recherche » d'AgroParisTech, prévue par la [politique de l'établissement sur les données de la recherche](#), rassemble les compétences de plusieurs services internes. Cette cellule vise à instruire et prendre en charge, sous les conditions définies ci-dessous, les demandes et besoins des agents et doctorants d'AgroParisTech sur la gestion, l'ouverture et la valorisation des données de la recherche.

Cet appui s'organise selon les modalités suivantes :

- Réponse aux demandes d'informations en s'appuyant sur la politique de l'établissement, la documentation d'appui et le développement progressif des outils d'accompagnement. Un réseau de correspondants de proximité sera progressivement mis en place au niveau des unités de recherche. La cellule agira à terme en suppléance de ce réseau.
- Prise en charge progressive de services spécifiques tels que définis dans la suite du document.
- Instruction de tout besoin non défini dans le présent document, et qui ne pourrait pas être géré à un niveau unité, en lien avec les différents interlocuteurs concernés (problématique appelant une réponse établissement, expression de besoins spécifiques, etc.).

La cellule a vocation à évoluer et s'adapter autant que possible aux demandes exprimées par la communauté de recherche d'AgroParisTech.

II. Position de la cellule

Plus particulièrement, la cellule « Données de la recherche » intervient en appui à :

- L'instruction du sujet des données de la recherche dans la phase de montage des projets de recherche ;
- La formalisation de la gestion des données au niveau des unités de recherche.

III. Composition de la cellule

Elle se compose actuellement des personnes suivantes :

- Les chargées de mission "données de la recherche" de l'établissement : Gaëlle Jaouen (gaelle.jaouen@ecofog.gf) et Eva Legras (eva.legras@agroparistech.fr)
- Le directeur de la Direction des Systèmes d'Information : Dominique Cagnon (dominique.cagnon@agroparistech.fr)
- Le directeur de la Direction des Documentations, du Musée du vivant et du CIRE : David Gasparotto (david.gasparotto@agroparistech.fr)
- Le chargé de mission juridique d'AgroParisTech Innovation : Kévin Angers (kevin.angers@agroparistech.fr)
- Le Délégué à la Protection des Données (DPD) d'AgroParisTech : Tahar Zouzou (tahar.zouzou@agroparistech.fr)

La cellule agit dès que nécessaire en étroite interaction avec les services et acteurs concernés, notamment : AgroParisTech Innovation, la Direction de la recherche et de la valorisation, etc.

IV. Conditions de sollicitation de la cellule

Toute demande doit être de préférence adressée par mail à l'alias : donnees-recherche@agroparistech.fr

Il demeure possible de contacter individuellement un ou des membres de la cellule pour des besoins spécifiques (l'information est alors transmise en interne par les collègues sollicités).

Afin de garantir un accompagnement du meilleur niveau, il est expressément demandé de solliciter la cellule **le plus en amont possible**, afin de laisser un temps d'instruction raisonnable selon la question posée. Toute demande réalisée dans l'urgence ne pourra être traitée qu'au mieux des disponibilités et capacités de la cellule.

Les sollicitations ne correspondant pas aux actions définies ci-dessous seront susceptibles d'être instruites selon des échéances plus longues.

V. Périmètre des actions prises en charge par la cellule

1) Au niveau des unités de recherche

La cellule intervient sur demande en appui aux réflexions menées par les directions d'unité pour formaliser les modalités de gestion courante des données de la recherche au sein de l'unité. Cet appui peut prendre les formes suivantes :

- mise à disposition de ressources d'information (documentation d'appui, plan de gestion de structure, etc.) ;
- aide à l'identification des besoins (informatiques, etc.) et accompagnement dans l'instruction de ces besoins ;
- formation des agents sur les principes généraux relatifs à la gestion et l'ouverture des données de la recherche ;
- instruction des enjeux juridiques autour de la gestion et de l'ouverture des données au cours des phases de montage de projets de recherche.

Par ailleurs, un travail sera progressivement mené, en lien avec les partenaires d'AgroParisTech, afin de développer autant que possible un réseau de correspondants de proximité au niveau des unités. Le profil et les actions de ce.tte correspondant.e pourront varier en fonction des possibilités et des besoins. Ce réseau, autant que possible partagé avec les différentes tutelles de l'unité, aura vocation à terme à traiter les questions génériques ou pouvant obtenir une réponse immédiate, tandis que la cellule continuera à être sollicitable en cas de besoin pour les questions plus complexes ou bien nécessitant une réponse de l'établissement.

2) Au niveau des projets de recherche

a) Appui à la rédaction de Plans de Gestion de Données (PGD)

Dans le cadre du montage d'un projet de recherche, les participants en charge de la rédaction du PGD peuvent solliciter la cellule pour être orientés vers des ressources, internes ou externes, d'aide au choix d'un modèle de PGD et d'aide à la rédaction du PGD. Notamment,

la cellule fournira très prochainement deux modèles de PGD propres à l'établissement (en cours de finalisation). L'un des modèles est destiné aux projets de recherche, l'autre concerne les structures (unité de recherche, plateforme, équipe, laboratoire...). Selon le type de projet de recherche, un modèle de PGD peut être fourni par le financeur.

Il est possible de contacter les membres de la cellule si des éléments du PGD sont complexes à remplir, notamment si ces éléments concernent l'établissement (infrastructures informatiques ou de documentation partagées par exemple).

Dans tous les cas, il est nécessaire d'anticiper le remplissage de ce PGD le plus en amont possible du dépôt du projet afin que l'aide apportée par la cellule et son réseau soit la plus pertinente possible.

Toute demande formulée à moins d'un mois d'une échéance définie (date limite pour un dossier de financement, etc.) ne pourra être garantie d'une prise en charge.

b) Appui à l'instruction des besoins d'infrastructure numérique

L'analyse de la demande doit permettre d'identifier la nature des besoins, notamment :

- la volumétrie
- les contraintes de localisation du stockage (local, réseau interne, réseau externe)
- les modalités d'accès aux données
- les besoins de sauvegarde
- le niveau de sécurisation attendu et réglementaire (en lien avec l'appui juridique)

Les échanges avec une personne de la Direction des Systèmes d'Information (DSI) permettront d'identifier dans un deuxième temps une éventuelle infrastructure existante pouvant héberger les données ou l'anticipation de l'achat d'une solution lorsque c'est possible.

L'accompagnement et le conseil proposés par la DSI le plus en amont possible permettront de mieux anticiper et identifier les solutions et, le cas échéant, d'identifier les budgets et les délais nécessaires ; en fonction du besoin, une solution pourra être identifiée dans le cadre du budget d'un projet donné.

Ces demandes permettront de constituer une cartographie des besoins afin de mieux anticiper les demandes.

Le travail de préfiguration en cours d'AgroParisTech et de l'université Paris-Saclay fera sans doute émerger des offres mutualisées de stockage. Une réflexion propre à AgroParisTech est également en cours afin de préconiser des solutions de stockage locales mais sécurisées et moins chères et qui conviendraient aux grands volumes de données (achat à la charge des unités).

Point de vigilance : le stockage, notamment réseau, a un impact budgétaire important qui n'est à ce stade pas encore complètement anticipé faute d'avoir une visibilité suffisante des besoins. Ces besoins peuvent également largement dépasser les capacités de financement de notre établissement. Il conviendra de rester réaliste sur les volumes de stockage demandés.

c) Appui juridique

La cellule apporte sur demande un soutien juridique sur les points suivants :

- **Données de la recherche couvertes par le Règlement Général sur la Protection des Données (RGPD)** : toutes les données personnelles sont soumises au RGPD et les données de la recherche ne font pas exception. Il en résulte qu'en cas de doute sur ses données, et sur la nature des actions à mettre en œuvre, le DPD est à contacter.
- **Droit de la propriété intellectuelle** : expertise quant à la protection des inventions issues de la recherche propre des UMR et des résultats issus de la collaboration Public/Privé ou Public/Public ; une expertise complémentaire pourra être mobilisée en cas de besoin pour les questions les plus complexes relatives aux données dans ce contexte.
- **Droit des données publiques et *open data*** : capacité d'instruction sur le cadre juridique global relatif à l'ouverture des données publiques.

d) Appui documentaire, terminologique et textuel

La cellule apporte sur demande un accompagnement sur le traitement documentaire de jeux de données, notamment :

- un appui sur la définition et la mise en œuvre de vocabulaires structurés, thésaurus, ontologies et outils équivalents, choix et définition des métadonnées ;
- un appui sur l'identification et la constitution de corpus de jeux de données ouvertes (identification d'entrepôts, recherches transversales de jeux de données via des répertoires appropriés en ligne) ;
- un appui sur la méthodologie et l'orientation vers des services et des ressources appropriées sur les techniques de fouille de texte et de données

Pour chacune de ces actions, suite à la sollicitation de la cellule, un personnel de documentation, désigné en fonction de ses compétences et du domaine considéré, interviendra auprès du demandeur afin de clarifier et répondre aux questions au plus près des enjeux exprimés. Si la demande fait état d'une expertise particulière pour laquelle des compétences ou des outils qui ne sont pas encore disponibles sont requis, cette personne donnera des recommandations et des pistes à envisager au demandeur en vue de faciliter la rapidité de traitement.

e) Aide à la stratégie d'ouverture des données

La cellule « Données de la recherche » intervient sur demande pour accompagner les agents et doctorants dans la définition de leur stratégie d'ouverture des données de la recherche. Sur la base et en complément des documents d'appui produits par la cellule, cet accompagnement peut prendre les formes suivantes :

- Aide à l'identification des points d'attention sur le cadre juridique et contractuel relatif à l'ouverture des données ;
- Aide à l'identification d'un entrepôt adapté pour le dépôt et l'ouverture de jeux de données ;
- Aide à la définition d'une stratégie de publication

Comme pour les points précédents, cet accompagnement doit être anticipé afin de pouvoir être assuré dans de bonnes conditions. Toute demande réalisée à moins d'un mois d'une échéance définie ne pourra être garantie de prise en charge.

VI. Mode de fonctionnement de la cellule

La cellule « données de la recherche » se réunira *a minima* 3 à 4 fois par an pour :

- Produire un bilan des actions réalisées (avec une synthèse annuelle) ;
- Capitaliser les actions réalisées afin de développer les outils et ressources d'accompagnement proposées à la communauté de recherche ;
- Instruire les éventuelles questions de fond soumises à la cellule allant au-delà d'une aide ponctuelle.

Compte tenu des évolutions rapides du domaine, les membres de la cellule seront attentifs à monter en compétence en cas de besoin sur leurs domaines respectifs.