

# FICHE

## Données de la recherche et science ouverte : le principe d'un accès « aussi ouvert que possible, aussi fermé que nécessaire »

### PARTIE 1 : SAVOIR

- Le contexte
- Aussi ouvert que possible, aussi fermé que nécessaire ?
- Quelques définitions

### PARTIE 2 : EN PRATIQUE

- Quels repères ? Les zones verte, orange, rouge
- Quel appui ? Vos référents

## PARTIE 1 : SAVOIR

### Le contexte

- Le plan national pour la science ouverte, lancé le 4 juillet 2018 par le Ministère chargé de l'enseignement supérieur et de la recherche, consolide la place de la recherche publique dans le dispositif juridique d'ouverture des données publiques défini par la Loi pour une République numérique du 7 octobre 2016.
- L'accès aux données publiques et les modalités de leur libre réutilisation sont actuellement principalement organisés par les dispositions de la Loi Valter (2015) et de la Loi pour une République numérique (2016).
- Ces lois marquent un tournant majeur car elles indiquent, en substance, que **sauf exception, les données produites par une administration dans le cadre de sa mission de service public sont par défaut ouvertes et réutilisables gratuitement.**
- Conformément à la mission de service public qui lui est confiée<sup>1</sup> et suivant les cadres juridiques et politiques précités, INRAE organise l'accès aux données qu'il détient selon le principe directeur « **aussi ouvert que possible, aussi fermé que nécessaire** ».
- Cette fiche s'attache à préciser, en complément du logigramme illustrant ce principe, les aspects juridiques qui conditionnent l'ouverture, le partage ou, au contraire, la fermeture de l'accès aux données.
- D'autres éléments, liés à la stratégie scientifique ou partenariale, ou aux conséquences négatives potentielles d'une ouverture d'accès sont également à prendre en compte.



### Aussi ouvert que possible, aussi fermé que nécessaire ?

- Le principe « aussi ouvert que possible, aussi fermé que nécessaire » est la traduction de l'expression "As open as possible, as closed as necessary" utilisée par la Commission européenne<sup>2</sup>. Il signifie qu'un accès aux données sans restrictions, libre et gratuit, ne peut

<sup>1</sup> Selon les dispositions du code rural et de la pêche maritime (art. R831-1), la mission de service public d'INRAE comprend « l'organisation du libre accès aux données scientifiques et aux publications ».

<sup>2</sup> Pour plus de précisions, voir le [volet du guide H2020](#) de la Commission européenne consacré à la gestion des données. Ce guide sera modifié lors du passage au programme-cadre Horizon Europe à partir de 2021.

# FICHE

## Données de la recherche et science ouverte : le principe d'un accès « aussi ouvert que possible, aussi fermé que nécessaire »

s'entend qu'à la condition préalable d'avoir vérifié le caractère public de ces données et l'absence d'exceptions prévues par la loi (existence de secrets légaux, notamment).

- Ce principe illustre également le fait que le régime de l'accès aux données publiques est un régime parmi d'autres, qui coexiste avec d'autres règles d'accès, spécifiques à certaines données (par exemple, les données de santé), ou se combine avec des situations partenariales.

### Quelques définitions<sup>3</sup>

- **Mission de service public :**  
Il s'agit d'une activité d'intérêt général exercée par ou sous le contrôle d'une personne publique. La mission de service public confiée à INRAE est celle indiquée à l'article R. 831-1 du code rural et de la pêche maritime, soit, principalement, de réaliser, d'organiser et de coordonner, à son initiative ou à la demande de l'Etat, tous travaux de recherche scientifique et technologique dans les domaines de l'agriculture, de l'alimentation, de la forêt, de l'environnement, de l'eau, de la biodiversité, de la bioéconomie, de l'économie circulaire, de la gestion durable des territoires et des risques dans les champs de compétence précités. Elle comprend notamment l'organisation du libre accès aux données scientifiques et aux publications.
- **Donnée publique :**  
Il s'agit d'une donnée communicable contenue dans un document administratif c'est-à-dire dans tout document détenu par une personne morale publique ou privée produit ou reçu dans le cadre de l'exercice d'une mission de service public.
- **Données de la recherche :**  
D'après le plan national pour la Science ouverte<sup>4</sup> dévoilé en 2018 par le Ministère chargé de la Recherche, il s'agit de « données issues de recherches financées sur projet ». Ces données comprennent les enregistrements factuels (chiffres, textes, images, son, vidéo, ...) utilisés comme sources primaires pour la recherche et habituellement acceptés par la communauté scientifique comme étant nécessaires pour valider les résultats de la recherche. Si les échantillons ne sont pas considérés comme des données, mais comme des matériels (biologiques, par exemple), ils obéissent également à des règles de gestion.
- **« Open Data » des données publiques :**  
Ce terme générique renvoie à la fois aux politiques d'ouverture des données mais également aux données ouvertes elles-mêmes c'est-à-dire les données du secteur public, produites dans le cadre de l'exécution d'une mission de service public, librement accessibles en ligne dans un format compréhensible, non propriétaire permettant une réutilisation libre. L'Open data est *un* mode d'accès et de réutilisation de données : les données publiques. **Voir la fiche « FAIR<sup>5</sup> ».**
- **Donnée « achevée »**  
Il s'agit d'une caractéristique indispensable à vérifier avant toute décision d'ouverture et d'une condition sine qua non de la mise en « Open data ». Elle est gage de fiabilité et de qualité de la donnée : le traitement de la donnée est finalisé et l'on s'est assuré de la complétude des métadonnées et des documents liés à sa traçabilité. **Voir la fiche « FAIR ».**
- **Droit d'accès :**  
Dans le cadre juridique posé par la loi pour une République numérique, c'est la concrétisation de la décision d'ouverture des données prise par l'institut. Ce droit se manifeste par l'accès de toute personne aux données publiques détenues par l'institut. Cet accès aux données peut avoir lieu après vérification préalable du caractère public de la donnée, de l'absence de facteurs limitants ou bloquants dûs à l'existence d'un partenariat (combinaison de plusieurs natures de données) ou d'une réglementation spécifique.

<sup>3</sup> Ces définitions sont adaptées du Guide CNIL-CADA-ETALAB 2019.

<sup>4</sup> [Plan national pour la science ouverte](#) (PNSO) du 4 juillet 2018.

<sup>5</sup> FAIR : Acronyme signifiant « findable, accessible, interoperable, reusable ». Voir Wilkinson, M. D. et al. The FAIR Guiding Principles for scientific data management and stewardship. *Sci. Data* 3:160018 <http://dx.doi.org/10.1038/sdata.2016.18> (2016)

# FICHE

## Données de la recherche et science ouverte : le principe d'un accès « aussi ouvert que possible, aussi fermé que nécessaire »

- **Droit de réutilisation**

Dans le cadre juridique posé par la loi pour une République numérique, la réutilisation d'un document administratif ou d'une donnée publique est le corollaire du droit d'accès. Il désigne l'utilisation de la donnée publique par un tiers à d'autres fins que celles de la mission de service public pour laquelle ledit document ou la donnée a été produit ou reçu. Ce droit permet au tiers réutilisateur d'utiliser les données à son propre bénéfice, gratuitement et librement, dans des conditions fixées dans une licence de réutilisation « homologuée ». **Voir la fiche « réutilisation ».**

- **CRPA :**

Il s'agit du « Code des relations entre le public et l'administration », entré en vigueur le 1<sup>er</sup> janvier 2016. Son objectif est d'unifier l'ensemble des sources juridiques tenant aux relations entre l'administration et le public. Il est structuré en cinq livres : 1) les échanges du public et de l'administration, 2) les actes unilatéraux pris par l'administration, 3) l'accès aux documents administratifs, 4) le règlement des différends avec l'administration, 5) les dispositions relatives à l'outre-mer.

# FICHE

## Données de la recherche et science ouverte : le principe d'un accès « aussi ouvert que possible, aussi fermé que nécessaire »

### PARTIE 2 : EN PRATIQUE

#### Quels repères ? Les zones verte, orange, rouge du logigramme

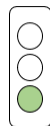
- Au fil d'une activité de recherche, la **combinaison de plusieurs types, scientifiques ou juridiques, de données** (publiques, privées, faisant l'objet de réglementations spécifiques) invite à un examen du contexte, au cas par cas, avant toute décision d'ouverture.
- Dans les faits, il convient de détecter l'existence d'éventuelles clauses partenariales limitantes et/ou de données répondant à une réglementation spécifique.
- Cette combinaison des règles juridiques applicables aux données permet de dégager un schéma de diagnostic en trois zones, illustrant le principe que l'accès aux données devrait être « **aussi ouvert que possible, aussi fermé que nécessaire** » :

**Zone verte** : Les conditions d'un accès libre et gratuit des données par tous sont réunies.

**Zone orange** : Un contrat limite l'accès des données à certains acteurs ou la nature des données implique de se mettre en conformité avec une réglementation spécifique, qui peut en limiter l'accès.

**Zone rouge** : Des facteurs bloquants, soit issus des clauses d'un partenariat, soit liés à la nature des données ou à des questions de sécurité, interdisent un accès libre aux données.

- **Ce schéma diagnostic en trois zones est fondé sur des arguments juridiques.** En complément de ces derniers, et sans les mettre en contradiction les uns avec les autres, d'autres éléments sont à prendre en compte pour déterminer les options les plus appropriées en matière de gestion des données<sup>6</sup> :
  - Les facteurs de conduite liés à la stratégie scientifique ou partenariale (exemple montage de consortium international pour le séquençage d'un génome d'intérêt) .
  - Les modulations possibles et les modalités d'ouverture liées au caractère « non achevé » des données, ou les délais possibles pour terminer l'exploitation scientifique des données et la publication des résultats.
  - La possibilité de valorisation via la propriété intellectuelle, en particulier pour les données intégrées dans les résultats scientifiques que sont les bases de données, logiciels, procédés, produits et publications (**voir fiche « Données, partenariat et innovation dans un contexte de science ouverte »**) .
  - L'impact négatif possible sur l'environnement et/ou dans le domaine socio-économique.
  - Le respect des valeurs éthiques et d'intégrité de la recherche.
- **Sur le plan technique**, les processus de FAIR-isation des données doivent être suivis à l'égard de toute donnée rendue accessible.



<sup>6</sup> Pour plus de précisions, se reporter à la note « Principes de gouvernance des données » dont est extrait ce passage.

# FICHE

## Données de la recherche et science ouverte : le principe d'un accès « aussi ouvert que possible, aussi fermé que nécessaire »

### ZONE VERTE : OUVERTURE

La zone verte correspond aux données concernées par le régime de l'**open data des données publiques**. Il s'agit :


- des documents qualifiés d'administratifs, de « données publiques », c'est-à-dire issus d'une mission de service public **et** présentant un caractère « achevé » (art L 311-2 CRPA). Ces données peuvent être récentes ou anciennes, peu importent leur forme ou leur support (cf. art L300-2 CRPA).
- des mises à jour régulières des données précédentes et qui présentent un intérêt économique, social, environnemental, ou ou sanitaire (art. L312-1-1 CRPA). A noter que l'Etat a créé un Service public de la donnée qui publie certaines de ces données, dites « [données de référence](#) ».

### BON A SAVOIR

- **Le régime de l'open data des données publiques est d'ordre public : il ne peut être écarté que par des modérations ou des exceptions prévues par la loi (cf. les rubriques « orange » et « rouge »).**
- **La typologie des données est variée et il existe d'autres régimes favorables à un accès ouvert des données**, soit en dehors des règles de l'open data, soit en interaction avec ces dernières.

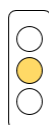
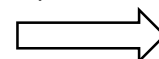
C'est la raison pour laquelle le logigramme signale que le fait qu'une donnée soit soumise à une réglementation n'équivaut pas nécessairement à un facteur limitant en termes d'accès.

C'est par exemple le cas de la réglementation INSPIRE issue de la directive européenne établissant une infrastructure d'information géographique dans la Communauté européenne, ainsi que plusieurs autres dispositions du code de l'environnement français.

- Pour en savoir plus sur la directive européenne INSPIRE  , nous vous conseillons la visite du [site de la Commission européenne](#) et du [site du ministère de la transition écologique et solidaire](#).
- Pour en savoir plus sur le droit d'accès aux informations relatives à l'environnement, nous vous conseillons la lecture de la [circulaire sur la mise en œuvre des dispositions régissant le droit d'accès à l'information relative à l'environnement](#) publiée le 11 mai 2020 par le ministère de la transition écologique et solidaire. Cette circulaire a été diffusée pour exécution aux établissements publics tels qu'INRAE.
- 

### LA BOUCLE DE VERIFICATIONS

- Le logigramme est construit sur une « boucle de vérifications » permettant la mise en œuvre du principe de gouvernance selon lequel l'accès aux données est « aussi ouvert que possible, aussi fermé que nécessaire ».
- Cela signifie qu'ouvrir l'accès aux données suppose d'avoir vérifié au préalable, et au cas par cas, que l'on ne se trouve pas dans l'un des cas signalés en zones orange ou rouge ci-après. Suivez le guide !



## FICHE

### Données de la recherche et science ouverte : le principe d'un accès « aussi ouvert que possible, aussi fermé que nécessaire »

#### ZONE ORANGE : PARTAGE ET CONTROLE D'ACCES

#### (CONTRATS, REGLEMENTATIONS OU DROITS SPECIFIQUES, EXCEPTIONS)

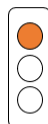
La zone orange correspond à des situations de **partage des données avec certains acteurs ou de contrôle de l'accès aux données, compte tenu de clauses de partenariat ou de contraintes liées à la nature des données** (par exemple : la possibilité d'accéder à une donnée est accordée sous réserve de se mettre en conformité avec une réglementation spécifique).

Le sort des jeux de données se situant dans cette zone peut parfois consister, après examen au cas par cas, en l'occultation de certaines données ou en la mise en accès des données selon des principes différents de ceux de l'open data.

#### BON A SAVOIR

La zone orange correspond donc à des situations diverses :

- **Une accessibilité aux données sous conditions de vérification des clauses contractuelles de partenariat** : Dès lors que les recherches ayant généré les données sont conduites en partenariat, la nature des partenaires et la combinaison possible avec le statut et la nature des données (publiques, privées) peuvent conduire à un partage négocié des données entre les partenaires. En complément, certains partenariats avec des acteurs privés peuvent définir, à travers les clauses contractuelles, les règles de d'accès à ces données au public.
- **Une accessibilité aux données limitée, sur la base de réglementations spécifiques** : De nombreuses réglementations portent des conditions d'accès et d'usage de données différentes, certaines allant dans le sens d'une ouverture des données (exemple des données relevant du régime d'INSPIRE), d'autres constituant des facteurs limitants voire très limitants (exemple des données à caractère personnel). Ces réglementations spécifiques peuvent également concerner des données issues de matériel tangible (ex : les données d'informations environnementales et de biodiversité, les données de recherches impliquant la personne humaine, ...).
- **Une accessibilité aux données sous conditions de vérification que le traitement de données a été réalisé conformément à la réglementation**. Exemple : il convient de satisfaire les exigences relatives à l'accès aux données à caractère personnel avant d'ouvrir les données (anonymiser ou brouter les données individuelles, cf. L322-2 CRPA).
- **Modération de l'open data : l'accès aux données peut être réservé à la personne concernée par ces données (c'est-à-dire « à l'intéressé », cf. article L. 311-6 CRPA) ou « tronquée » pour tenir compte de certains secrets : secret des affaires, secret médical**. Exemple : l'ouverture des données peut être freinée par l'obligation de satisfaire les exigences de la réglementation sur les recherches impliquant la personne humaine.
- **Modération de l'open data** : L'article L311-4 du CRPA indique quant à lui que « les documents administratifs sont communiqués ou publiés sous réserve des droits de propriété littéraire et artistique ».



#### ZONE ROUGE : FERMETURE

## FICHE

### Données de la recherche et science ouverte : le principe d'un accès « aussi ouvert que possible, aussi fermé que nécessaire »

La zone rouge correspond à la présence de **facteurs bloquants, soit issus des clauses d'un partenariat, soit liés à la nature des données** ou à des questions de sécurité, qui en interdisent l'accès.

#### BON A SAVOIR

Deux cas principaux sont concernés :

- **Lorsque les données ne sont pas considérées comme des informations publiques (art L321-2 CRPA) et sont donc hors champ de l'obligation d'ouverture :**

Il s'agit de données non issues d'une mission de service public, données communiquées par des tiers, données non achevées. En outre, **les partenariats, publics ou privés, peuvent comporter des clauses de confidentialité** qui font obstacle à toute publication des données.

- **Lorsque d'une réglementation proscrit l'accès aux données :**

**Exceptions à l'open data :** L'article L311-5 CRPA liste les informations qui ne sont pas communicables, même si elles sont issues d'une administration et ont été acquises ou générées dans le cadre de l'exercice d'une mission de service public. A titre d'exemple :

- les documents réalisés en exécution d'un contrat exécuté pour le compte d'une ou de plusieurs personnes déterminées (les contrats de partenariat peuvent être concernés à ce titre) ;
- les documents dont la diffusion pourrait porter atteinte au secret de la défense nationale, à la conduite de la politique extérieure de la France, à la sûreté de l'Etat, à la sécurité publique, à la sécurité des personnes ou à la sécurité des systèmes d'information des administrations ;
- les documents dont la diffusion pourrait porter atteinte au déroulement des procédures engagées devant les juridictions ou d'opérations préliminaires à de telles procédures (sauf autorisation donnée par l'autorité compétente), à la recherche et à la prévention d'infractions de toute nature ;
- les documents dont la diffusion pourrait porter atteinte aux autres secrets protégés par la loi.

**Quel appui ? Vos référents :** DipSo, DPTI, DAJ, DPO. **+cellule nationale**