

Aspects juridiques de la gestion et du partage des données

Romain Boistel, Frédérique Bordignon, Lionel Maurel

► **To cite this version:**

Romain Boistel, Frédérique Bordignon, Lionel Maurel. Aspects juridiques de la gestion et du partage des données. Journées Nationales de la Science Ouverte 2019, Nov 2019, Paris, France. hal-02372271

HAL Id: hal-02372271

<https://hal-enpc.archives-ouvertes.fr/hal-02372271>

Submitted on 20 Nov 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Aspects juridiques de la gestion et du partage des données

Romain Boistel¹ - Frédérique Bordignon¹ - Lionel Maurel²

¹ École des Ponts ParisTech, Direction de la Documentation, Pôle IST

² InSHS, CNRS

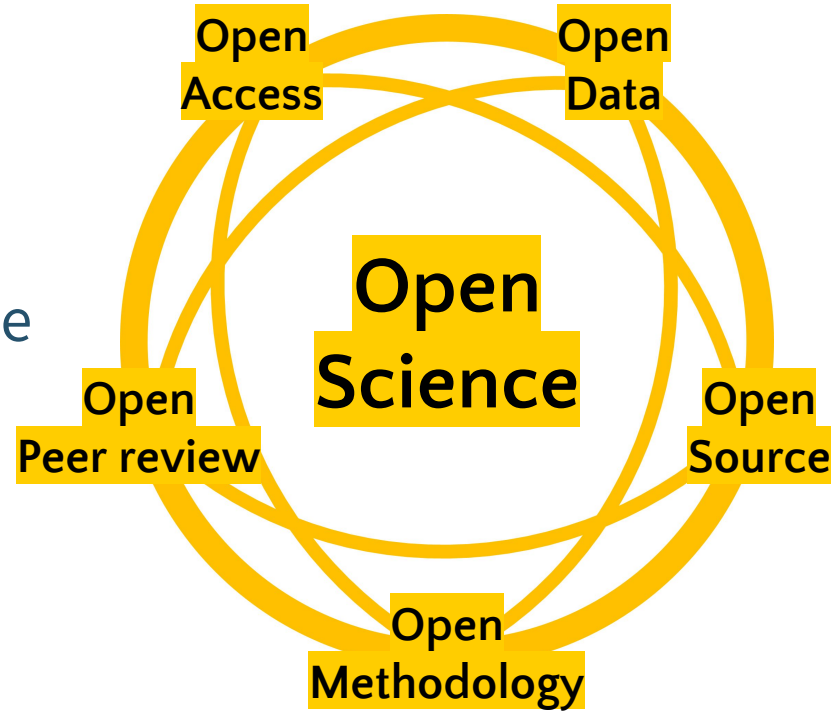


Contexte

L'Open Science et ses bénéfices

Ouvrir la science...

- » Validation
- » Reproductibilité
- » Visibilité
- » Réutilisation inattendue
- » Partage citoyen



Définition informelle des données

- » Données d'observation
- » Données expérimentales
- » Données de simulation numérique
- » Données dérivées ou compilées
- » Données de référence
- » **Inclut le code**

Définitions formelles des données

OCDE, 2017

« Enregistrements factuels (chiffres, textes, images, sons) utilisés comme source principale pour la recherche scientifique et généralement reconnus par la communauté scientifique comme nécessaires pour valider les résultats de la recherche. Un ensemble de données de recherche constitue une représentation systématique et partielle du sujet faisant l'objet de la recherche. »

Directive Open Data PSI, 2019

« Documents se présentant sous forme numérique, autres que des publications scientifiques, qui sont recueillis ou produits au cours d'activités de recherche scientifique et utilisés comme éléments probants dans le processus de recherche, ou dont la communauté scientifique admet communément qu'ils sont nécessaires pour valider des conclusions et résultats de la recherche. »

OCDE : <http://www.oecd.org/fr/sti/sci-tech/38500823.pdf>

Directive PSI : <https://eur-lex.europa.eu/legal-content/FR/TXT/HTML/?uri=CELEX:32019L1024&from=EN#d1e1721-56-1>

Définition des données en droit français

Pas de définition légale.

Les données et le code sont considérés :

- comme des documents administratifs
- comme des informations publiques
- » on regarde dans la loi ce qui se rapporte à ces concepts

On ne considère que les données achevées.
Le code est communicable même inachevé.

Document administratif : [Art. L300-2 CRPA](#)
Information publique : [Art. L321-1 CRPA](#)

Les textes

Lois ou codes ?

En France la plupart des lois sont regroupées dans des codes

- » Code de la Propriété Intellectuelle (CPI)
- » Code des Relations entre le Public et l'Administration (CRPA)
- » Code de l'Éducation
- » Code de la Recherche
- » Code Civil
- » Code de l'Environnement

Loi pour une République Numérique

Un texte majeur adopté en 2016, à la fois pour le libre accès aux publications, et la diffusion des données de la recherche.

Ses articles sont transposés dans plusieurs codes.

Devoir, pouvoir, ne pas avoir le droit...

Déterminer les titulaires des droits

Chercheurs

Établissements

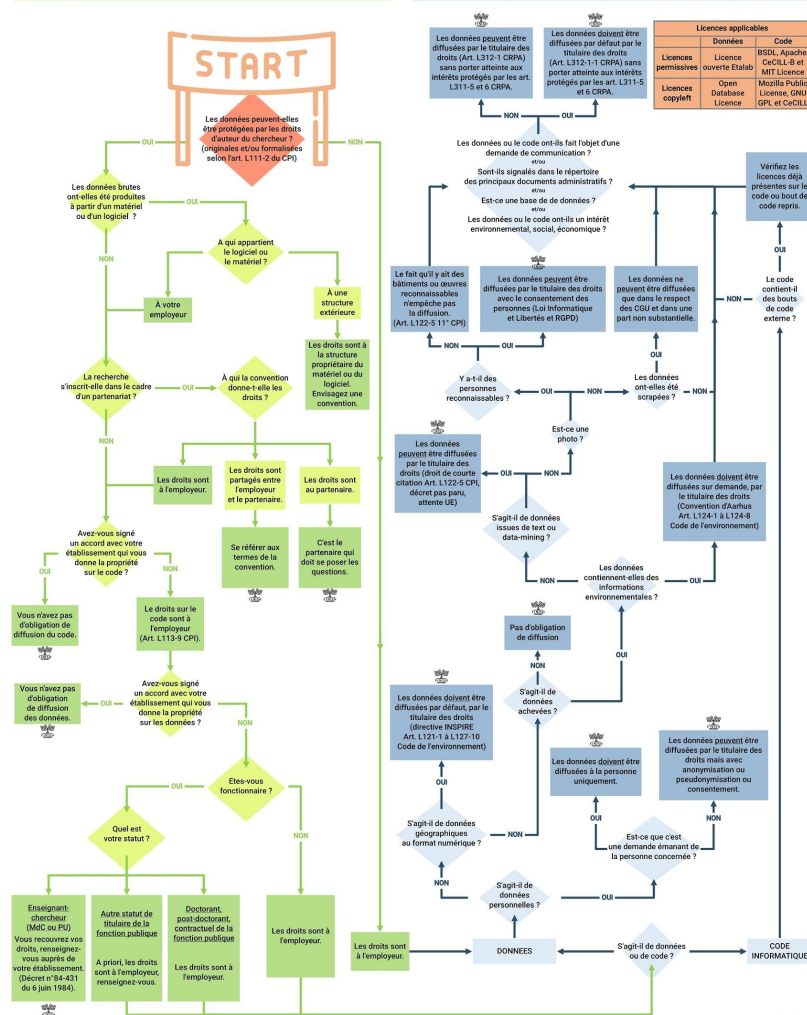
Déterminer les conditions de partage

Devoir

Pouvoir

Ne pas avoir le droit

Titularité des droits



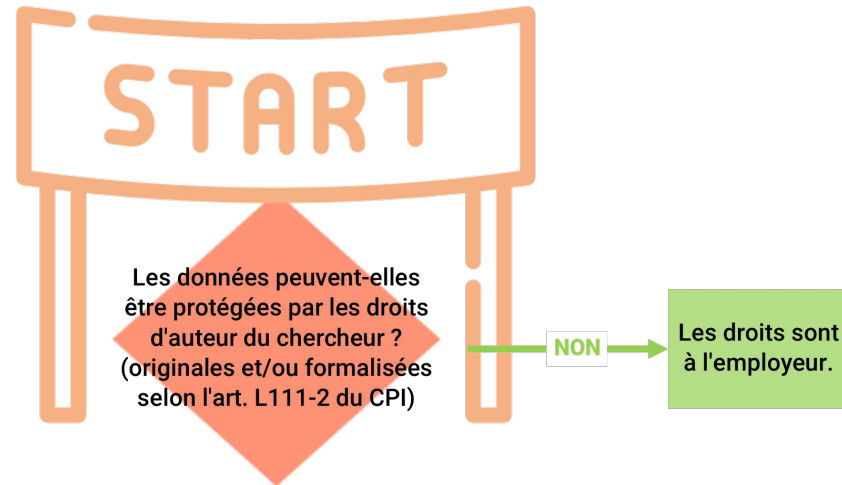
Titularité des droits

Les données peuvent-elles être protégées par les droits d'auteur du chercheur ?

Les données, en tant qu'œuvres de l'esprit, sont protégées par le droit d'auteur du chercheur à deux conditions :

1. qu'elles soient **formalisées** (donc pas les idées, les algorithmes...)
2. qu'elles soient **originales**

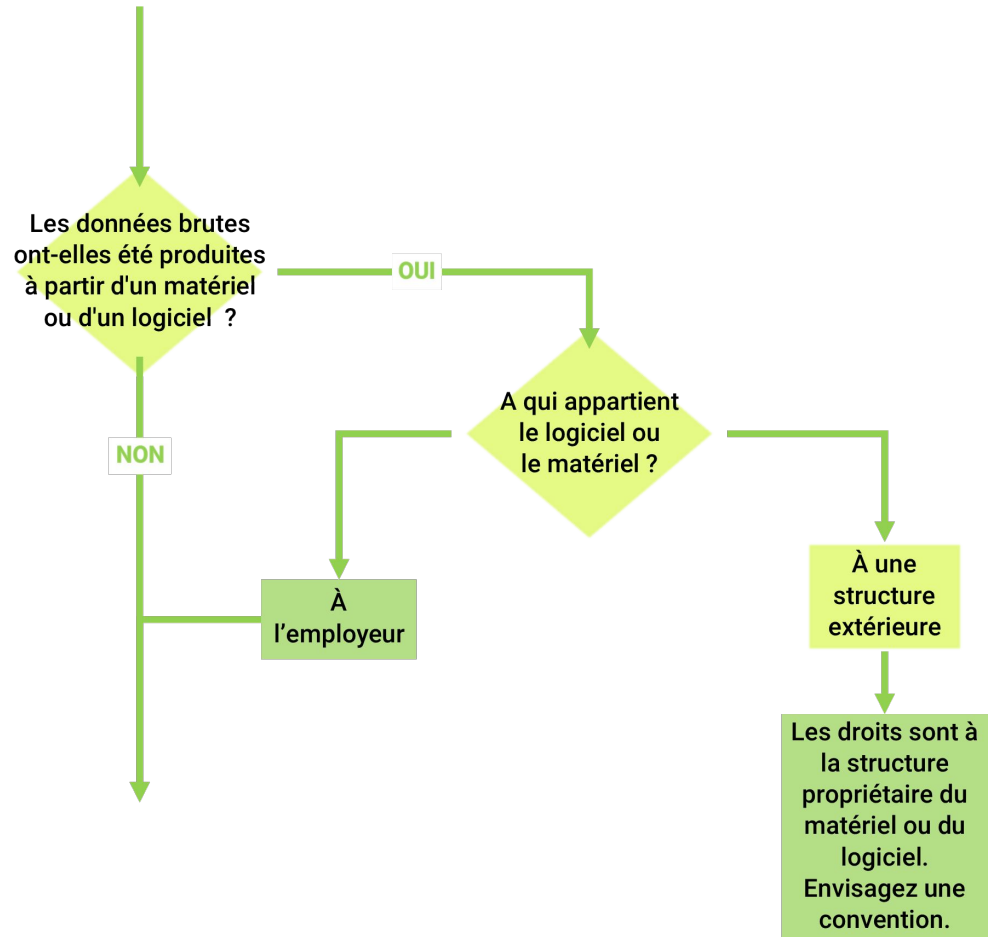
→ Difficile de justifier que les données soient des œuvres originales
→ L'établissement en est le producteur au sens juridique



[Art. L111-2 du CPI](#)
+ jurisprudence

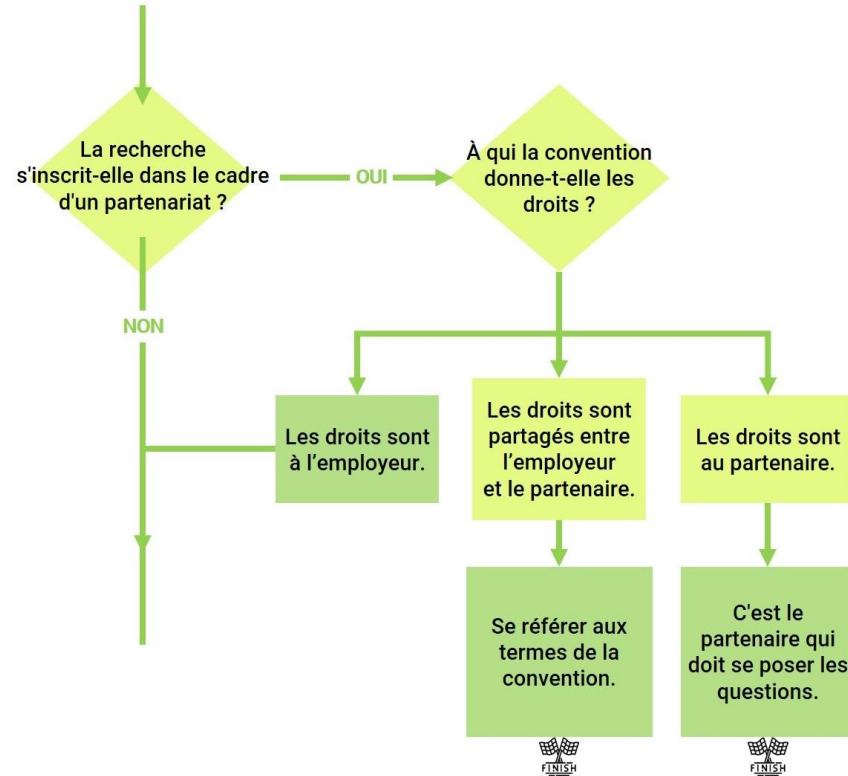
La question du matériel

On peut envisager une convention clarifiant la propriété des données lors du prêt de matériel entre établissements.



Projets partenariaux et droits des tiers

Retour au contrat et importance des accords de consortium



Propriété du code informatique

L'établissement est toujours titulaire des droits sur le code informatique

- » quel que soit le statut du chercheur
- » sauf si un contrat signé avec le chercheur stipule l'inverse

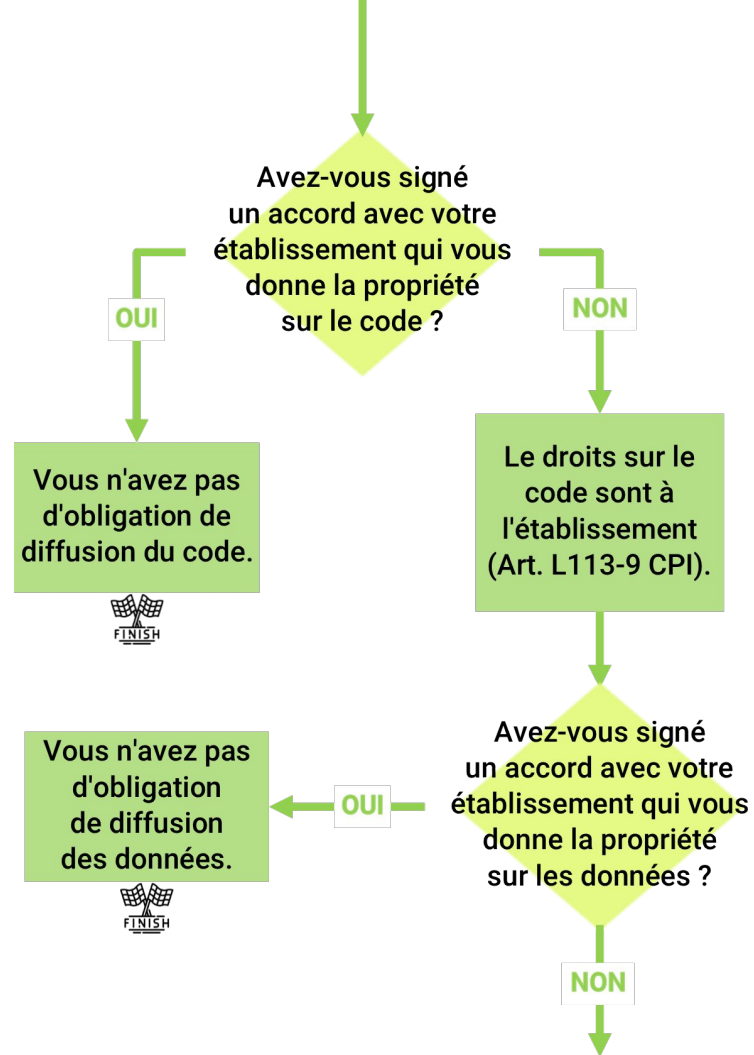
[Art. L113-9 du CPI](#)

Si le code fait l'objet d'une valorisation commerciale, les auteurs obtiennent :

- » un droit de préférence
- » une prime d'intéressement

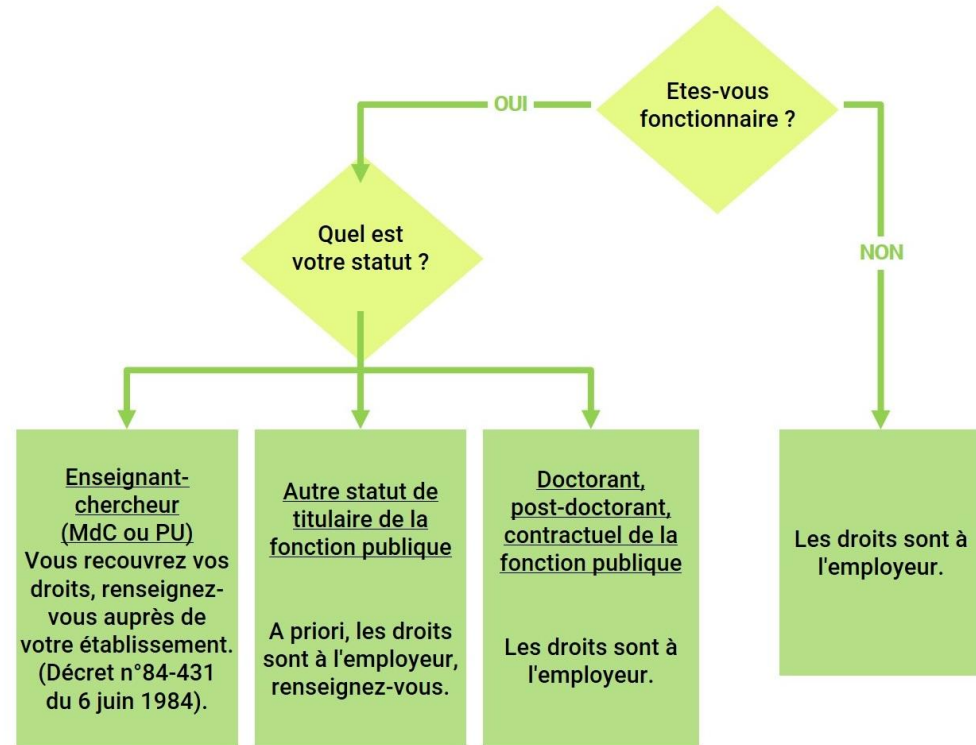
[Art. L131-3-1 du CPI](#)

[Décret n° 96-858 du 2 octobre 1996](#)



Limites du statut des chercheurs

Seuls les enseignants-chercheurs de statut conservent *a priori* leur droit d'auteur sur les données qu'ils produisent.



Limites du statut des chercheurs

Pourquoi ?

- » Les agents de l'état n'ont pas de droits sur les œuvres produites dans l'exercice de leur fonction
- » Sauf si de par leur statut ils ne sont pas soumis au contrôle préalable de l'autorité hiérarchique
- » Cette exception ne s'applique qu'aux enseignants-chercheurs de statut (corps des professeurs des universités et maîtres de conférence)

[Art. L131-3-1 du CPI](#)

[Art. L111-1 du CPI](#)

[Art. L952-2 du Code de l'Éducation](#)

[Art. 2 décret statutaire](#)

Différence entre les textes (plus restrictifs) et la pratique (plus permissive).

Limites du droit du producteur de base de données

Droit du producteur de base de données : la personne ou l'établissement qui a investi dans sa création ou sa maintenance en est titulaire.

[Art. L341-1 du CPI](#)

L'établissement peut facilement justifier le financement d'une base de données (salaires, matériel, locaux...).

→ Si les données prennent la forme d'une base de données : elles sont considérées comme produites par l'établissement.



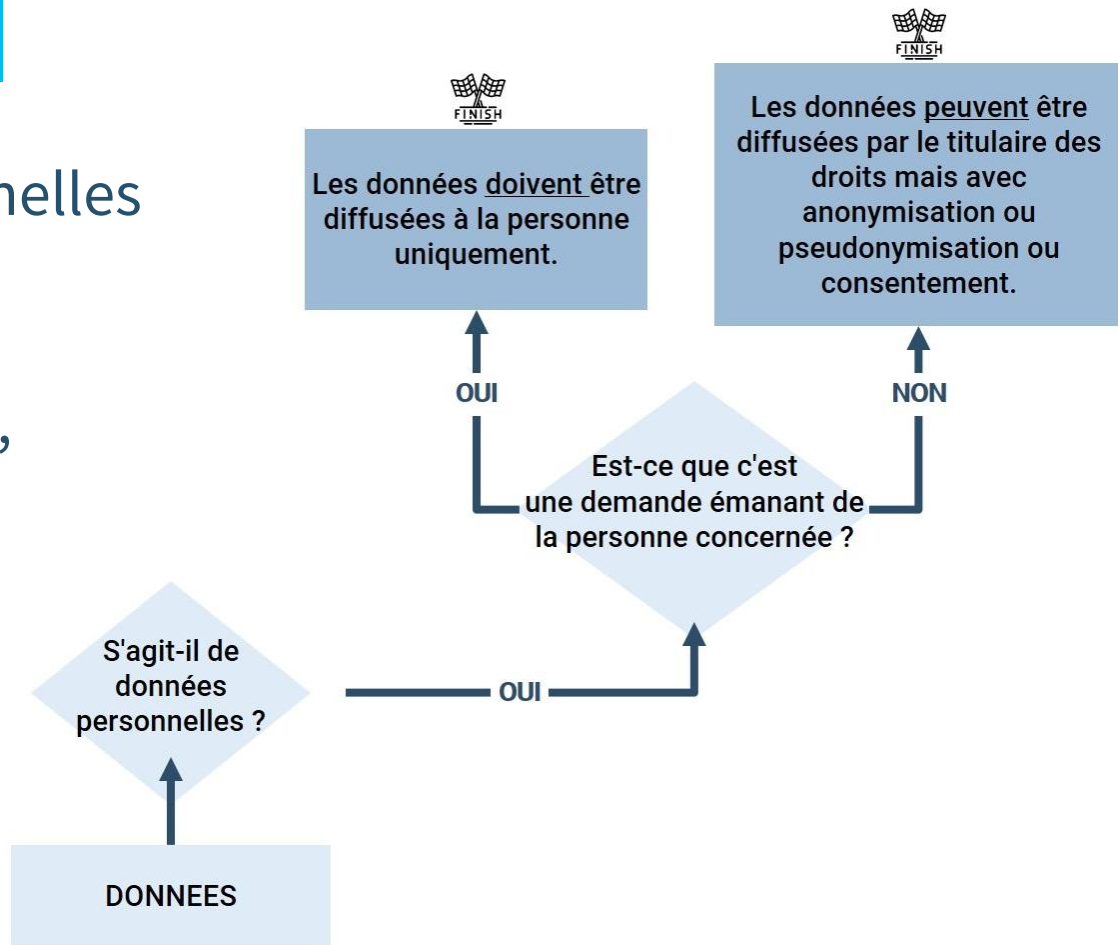
Bilan :

dans la plupart des cas, les données de la recherche et le code constituent des données publiques.

Droits et obligations de diffusion

Données personnelles

Les données personnelles doivent être communiquées à la personne concernée, à sa demande.



Données personnelles

- Les données personnelles ne peuvent pas être diffusées
- » sauf si elles sont anonymisées ou pseudonymisées
 - » ou si les personnes concernées ont donné leur consentement préalable

Règlement Général sur la Protection des Données personnelles (RGPD)
→ [Loi relative à la protection des données personnelles](#)

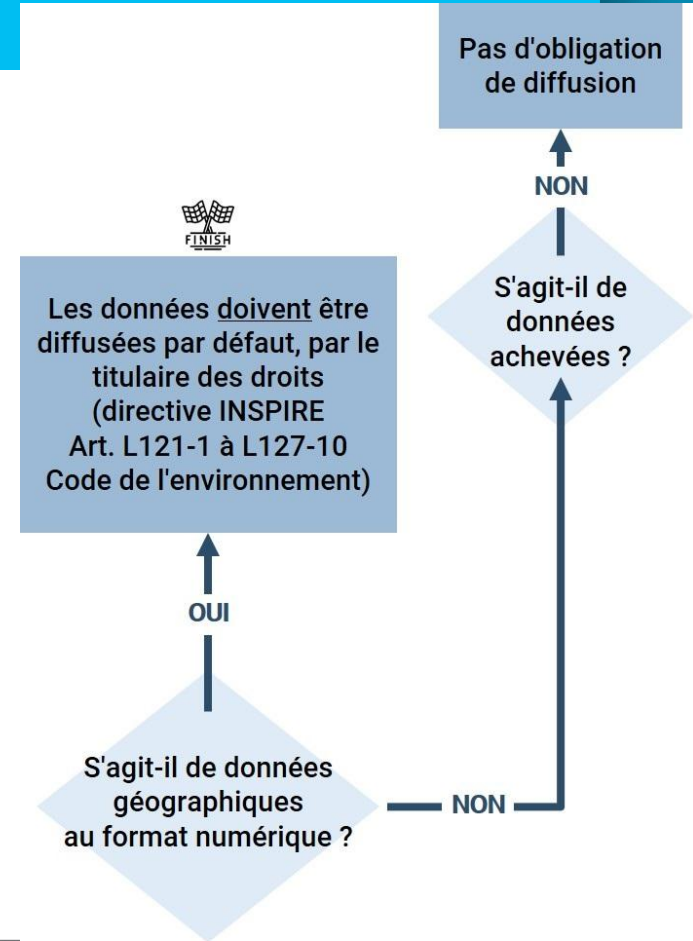


Seules les données achevées
peuvent être diffusées.

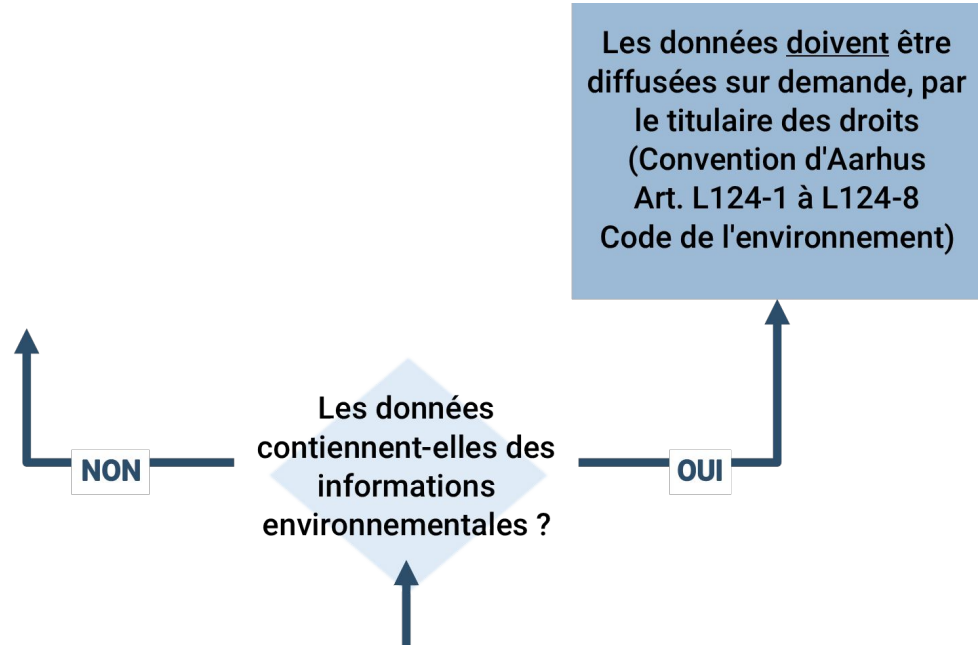
[Art. L311-2 du CRPA](#)

Les données géographiques doivent être diffusées
même inachevées.

Directive INSPIRE
[Art. L127-1 à L127-10 du Code de l'Environnement](#)



Données environnementales



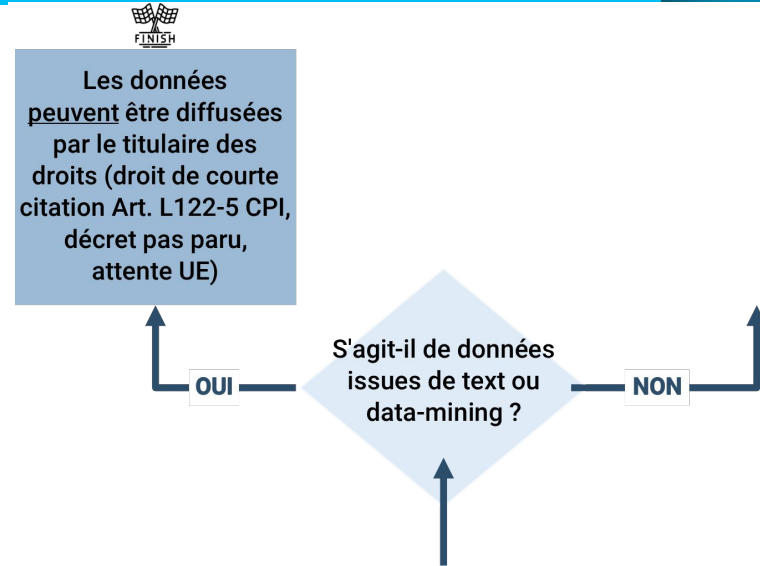
Fouille de texte et de données

Les éditeurs ne peuvent plus s'opposer à la fouille de texte ou de données associées à des publications

Mais on attend la transposition de la directive européenne Copyright

Les données issues de la fouille de texte ou de données peuvent donc être diffusées

Mais pas les textes sources (seulement des courtes citations)



[Art. L.122-5 3a° du CPI](#)

Photos

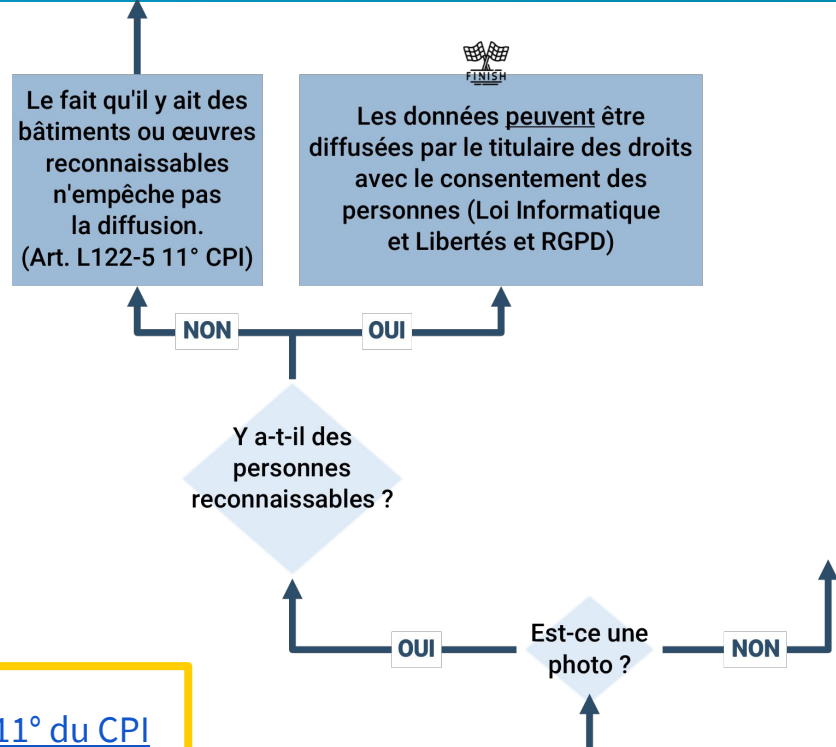
Les photos peuvent être diffusées :

- » même s'il y a des bâtiments ou des œuvres permanentes reconnaissables dessus (exception de panorama) - mais elles ne sont pas librement réutilisables

[Art. L.122-5 11° du CPI](#)

- » sauf s'il y a une personne reconnaissable dessus (il faut demander son accord)

[Art. 9 du Code Civil](#)
+ jurisprudence



Données scrapées

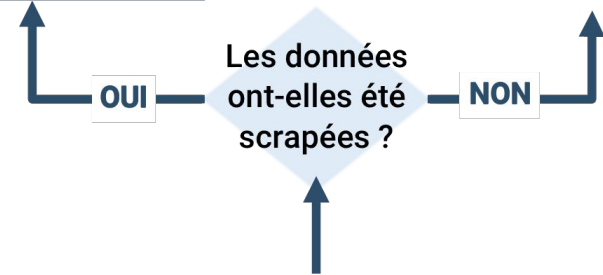
Scraping : automatisation de l'extraction des données contenues sur un site web

- » collecte des montants de loyers sur le Bon Coin
- » collecte de tweets
- » collecte de données relatives à l'auto-partage...

Souvent interdit par les conditions d'utilisation, donc pas de diffusion des données scrapées. Mais :

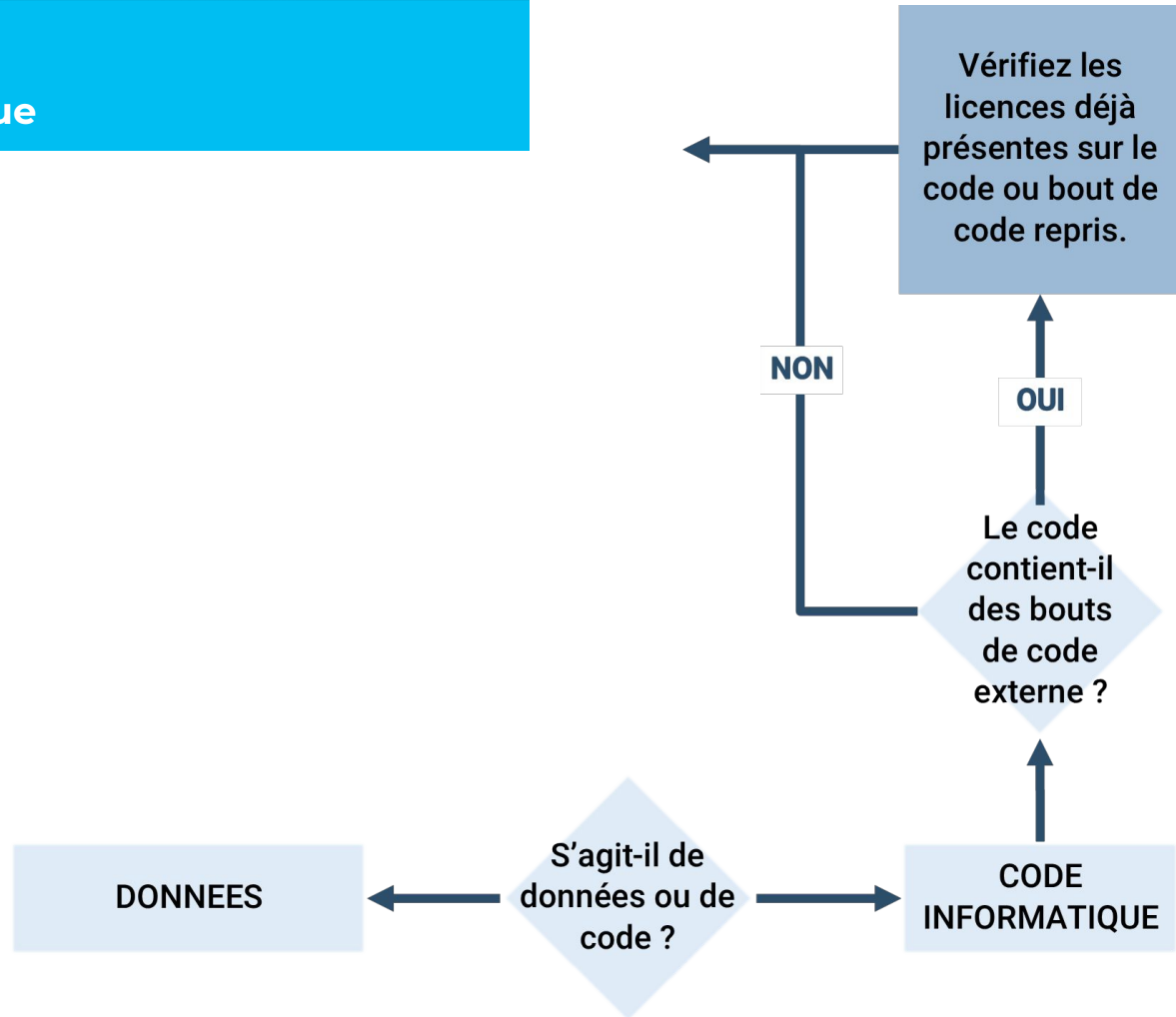
- » si le préjudice du producteur des données est nul, le risque encouru l'est aussi
- » le scraping d'une partie non substantielle des données reste possible

Les données ne peuvent être diffusées que dans le respect des CGU et dans une part non substantielle.



[Art. L.342-3 1° du CPI](#)

Pour le code informatique

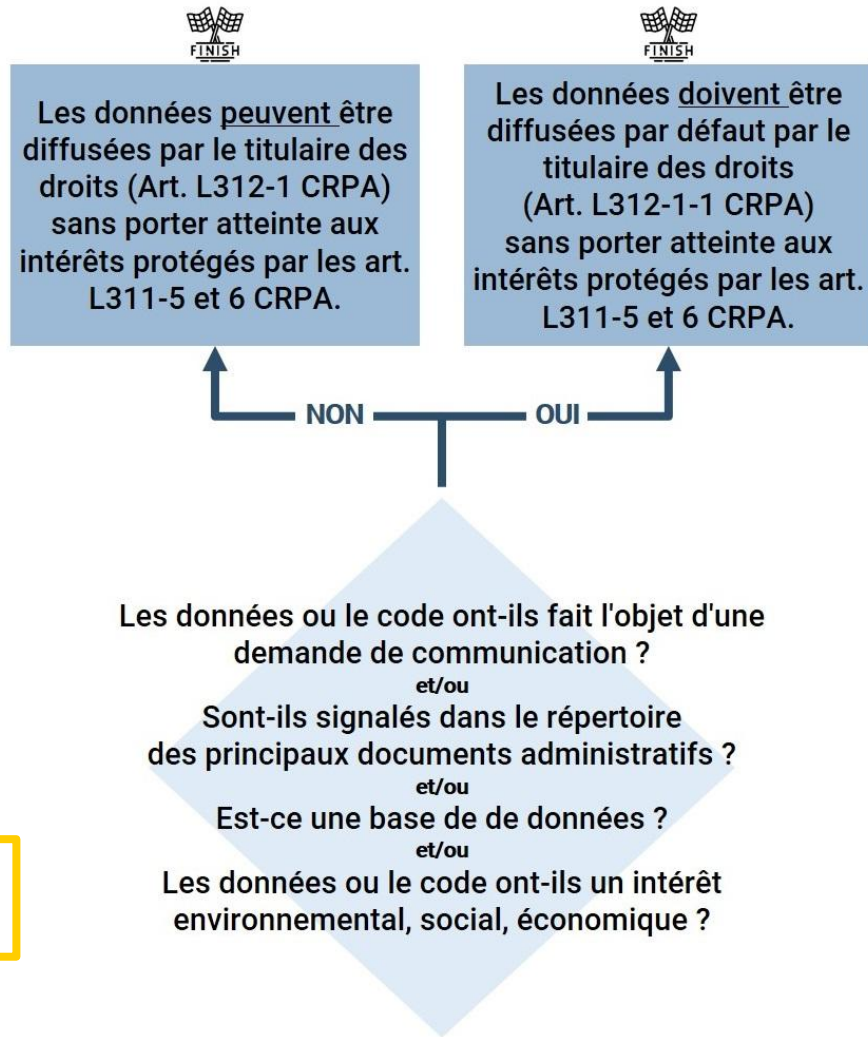


Principe d'ouverture par défaut

Peu probable que les données de recherche n'aient aucun intérêt environnementale, social ni économique.

Une fois mises en ligne, les données doivent être mises à jour régulièrement.

Principe d'ouverture par défaut : [Art. L312-1-1 CRPA](#)





Bilan :
*par défaut, la diffusion des données de la
recherche et du code est obligatoire.*

Les licences prévues par la loi

La loi prévoit deux types de licences :

[Art. D323-2-1 du CRPA](#)

Permissives

- » Permet de modifier la licence
- » Réutilisation libre y compris commerciale

Pour le code : BSD, Apache, CeCILL-B et MIT License

Pour les données :  Licence ouverte Etalab

Avec obligation de réciprocité

- » Propagation de la licence à tout le logiciel qui intègre le code
- » Peut empêcher une réutilisation commerciale

Pour le code : Mozilla Public License, GNU GPL et CeCILL

Pour les données : Open Database License

Comment choisir sa licence ?

Dans les entrepôts de données et de code : souvent une liste fermée de licences qui n'inclut pas les licences prévues par la loi.

- » Opter pour une licence CC-BY
- » Préférer une licence permissive

Licences applicables		
	Données	Code
Licences permissives	Licence ouverte Etalab	BSDDL, Apache, CeCILL-B et MIT Licence
Licences copyleft	Open Database Licence	Mozilla Public License, GNU GPL et CeCILL

« Aussi ouvert que possible ; aussi fermé que nécessaire »

F
indable



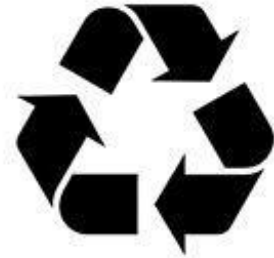
A
ccessible



I
nteroperable



R
eusable



Romain Boistel - romain.boistel@enpc.fr - @RomBoistel

Frédérique Bordignon - frederique.bordignon@enpc.fr - @freddie2310

Lionel Maurel - lionel.maurel@cnrs.fr - @Calimaq



Merci de nous avoir écoutés...