



Cadre juridique des données de la recherche

Formation CNRS 06 juillet 2016



Sommaire

1

- Données de la recherche, essai de définition

2

- Quel cadre réglementaire suivre?

3

- Principe d'open data et précautions

Chapitre 1

1

- Données de la recherche, essai de définition

2

- Quel cadre réglementaire suivre?

3

- Principe d'open data et précautions

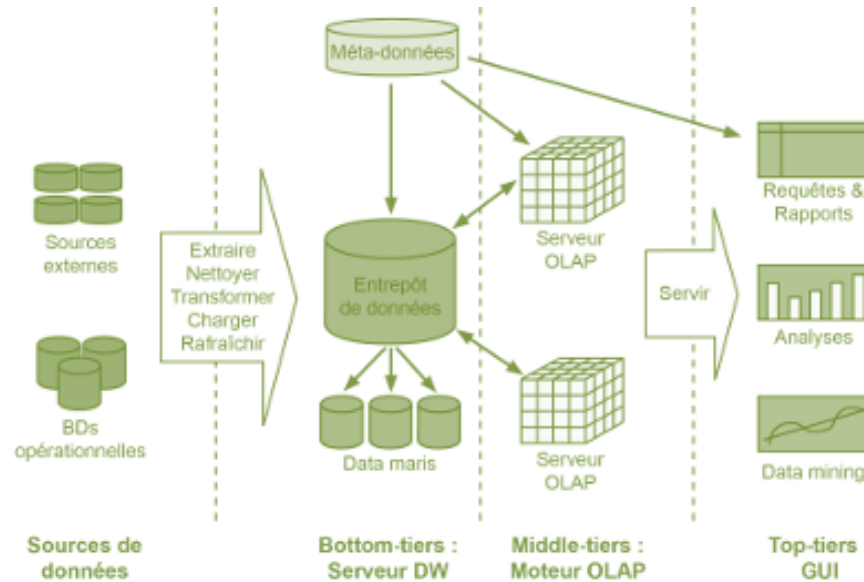
Données de la recherche, essai de définition

- Pas de définition légale => pas de régulation par un texte spécifique.
- D'un point de vue scientifique, les données vont dépendre de la discipline : allant de données chiffrées sortant d'un modèle de simulation informatique ... à un corpus documentaire.
- D'un point de vue technique, une donnée est la représentation d'une information permettant sa communication, son traitement, son interprétation.
- Définition de l'OCDE ([principes et lignes directrices pour l'accès aux données](#) de la recherche financée sur des fonds publics)
 - « *les données de recherche sont définies comme des **enregistrements factuels** (chiffres, textes, images et sons) qui sont utilisés comme sources principales pour la recherche scientifique et sont généralement reconnus par la communauté scientifique comme nécessaires pour valider des résultats de recherche. Un ensemble de données de recherche constitue une représentation systématique et partielle du sujet faisant l'objet de la recherche.* »

Données de la recherche, essai de définition

- Dans les textes législatifs, on trouve les notions de :
 - documents administratifs (produits par le secteur public),
 - informations du secteur public (contenues dans les documents administratifs),
 - données scientifiques (code de la recherche),
 - d'archives publiques.
- Dans le langage courant, on entend « données publiques ».
- Attention aux mots « administratif » et « public » qui n'ont pas forcément les mêmes sens dans la vie courante et dans les textes. On retiendra qu'il vaut mieux leur affecter la notion de « tout ce qui est produit par le secteur public ».

Données de la recherche, essai de définition



Ce qu'il faut retenir :

- ⇒ Pas de distinction entre données brutes, élaborées ou métadonnées d'un point de vue juridique.
- ⇒ **Pas de droit de propriété** dans la plupart des cas sur la donnée (données machine, etc.). Elle est considérée comme une information « de libre parcours ». A ce titre, l'établissement du producteur de la donnée peut restreindre ou non sa diffusion.
- ⇒ Mais il existe deux exceptions où une « propriété » peut s'exercer.

Données de la recherche, essai de définition

Exceptions : « données propriétaires »

- ⇒ Données sont soumises au **droit d'auteur** : textes, plans, photographies, etc. et notamment les publications scientifiques.
 - ⇒ **Condition** : **originalité** de la forme (pas de l'idée).
 - ⇒ **Conséquence** : pour utiliser ces données, **l'accord de l'auteur** est indispensable (donc attention!! au text mining) sauf exception de courte citation .
Le droit revient à l'auteur et non à l'établissement (sous réserve de conditions « d'autonomie »).
- ⇒ Données organisées en bases de données : **Droit sui generis** qui peut s'appliquer sous réserve de la preuve d'un investissement substantiel (le plus souvent financier). Le droit revient à l'investisseur (le plus souvent l'établissement).

Chapitre 2

1

- Données de la recherche, essai de définition

2

- Quel cadre réglementaire suivre?

3

- Principe d'open data et précautions à mettre en œuvre?

Contexte international de l'Open Science

Déclaration Berlin 2003 : Libre accès à l'Information Scientifique et Technique (publications et données).

OCDE 2007 : Recherche financée sur fonds public (valeur des données).

Principes Pantan 2009 : promotion de l'open science.

Groupe G8+5 : Open data for Agriculture (promouvoir les politiques et les investissements dans l'open access pour les projets financés par les pouvoirs publics...).

Research Data Alliance (faciliter le partage et l'échange des données de la recherche...).

Apparitions de nouvelles revues (Gigascience), entrepôts de données (Dryad), data papers, etc.

Recommandations de la CE en juillet 2012 (organiser le libre accès aux publications et données d'ici à 2014), H2020 et son pilote open data.



Quel cadre réglementaire suivre ?

Il existe beaucoup de textes de différentes formes (loi, circulaire, directive, etc.) qui régissent les données qui peuvent être produites par la recherche :

Texte principal : Loi n°78-753 du 17.07.1978 dite « **loi CADA** » modifiée par la loi du 28/12/15 relative à la gratuité et aux modalités de réutilisation des informations du secteur public, codifiée dans le code des relations entre le public et l'administration.

Code de la recherche : article L112-1 « e) L'organisation de l'accès libre aux données scientifiques »,

A venir : le **projet de loi Lemaire** pour une république numérique,

Autres textes : code du Patrimoine, loi Informatique et Libertés, circulaire pour la Protection du Patrimoine Scientifique et Technique, code de l'Environnement (ex art L124-2), directive Inspire, etc.

Quel cadre réglementaire suivre ?

Le cadre actuel est celui de « l'open data » découlant du texte principal (loi « CADA ») et du code de la recherche :

« Sont considérés comme **documents administratifs**, que soient leur date, leur lieu de conservation, leur forme et leur support, les documents produits ou reçus, **dans le cadre de leur mission de service public**, par l'Etat, les collectivités territoriales ainsi que par **les autres personnes de droit public** ou les personnes de droit privé chargées d'une telle mission.

Constituent de tels documents **notamment** les dossiers, rapports, études, comptes rendus, procès-verbaux, statistiques, directives, instructions, circulaires, notes et réponses ministérielles, correspondances, avis, prévisions et décisions. »

Cela rejoint la définition des archives publiques.



Quel cadre réglementaire suivre ?

Est-ce que je produis des documents administratifs ?



Oui si mon employeur est public

- ❖ Tout ce que je fait dans le cadre de ma mission peut être considéré comme un document administratif.
 - ❖ Attention ! Pour les **doctorants** : si la thèse est cofinancée ou réalisée en collaboration avec un partenaire privé => il faut se reporter au contrat.
-
- ❖ Cas particulier : je suis **chercheur ou enseignant-chercheur** : mes écrits, cartes, photographies, plans qui sont originaux et donc soumis au droit d'auteur m'appartiennent (exception – loi DADVSI 2006-961)

Mais le reste appartient bien à mon employeur.

Quel cadre réglementaire suivre ?



- Un jeu de données produit par un chercheur s'apparente à une information du secteur public. La CADA, dans certains avis, l'assimilera à un document administratif. Ce sera toujours une archive.
- La notion de **document achevé** importante pour la communicabilité. Elle rejoint, par certains côtés, les notions d'embargo sur les données.
- La réutilisation des informations publiques est soumise à la condition que ces dernières ne soient pas altérées, que leur sens ne soit pas dénaturé et que leurs sources et date de leur dernière mise à jour soient mentionnées.

Quel cadre réglementaire suivre ?



CADA modifiée par Loi Valter (28/12/2015) — Transposition directive européenne	Projet de loi Lemaire Nationale
Communication sur demande mais incitation à la mise en ligne des <u>documents achevés</u>	Communication en ligne après une première demande et les bases de donnée seraient formellement concernées, sans opposition du droit des bases de données.
Libre réutilisation par toute personne à d'autres fins que celle de la mission de service public, (abrogation exception)	
Principe de gratuité pour la réutilisation (sauf exceptions),	
Incitation à la mise sous format ouvert et librement réutilisable,	Obligation à la mise sous format ouvert,
Possibilité de choisir une licence.	Choix de la licence parmi une liste publiée par décret.

Chapitre 3

1

- Données de la recherche, essai de définition

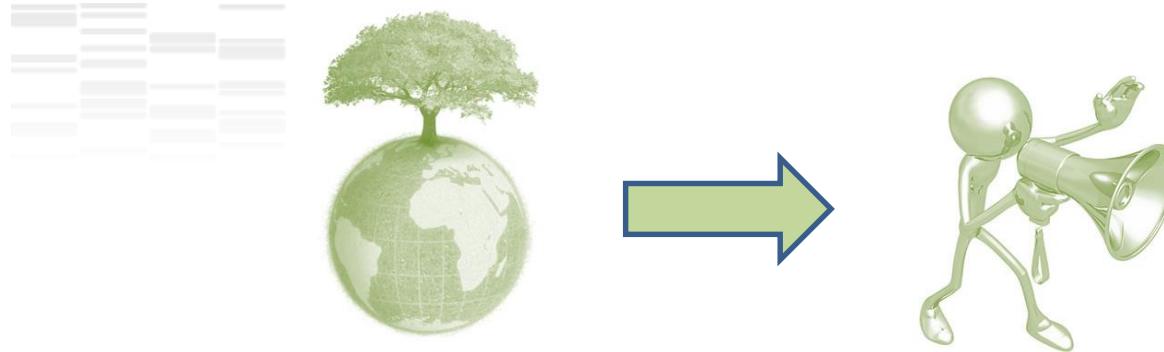
2

- Quel cadre réglementaire suivre?

3

- Principe d'open data et précautions à mettre en œuvre?

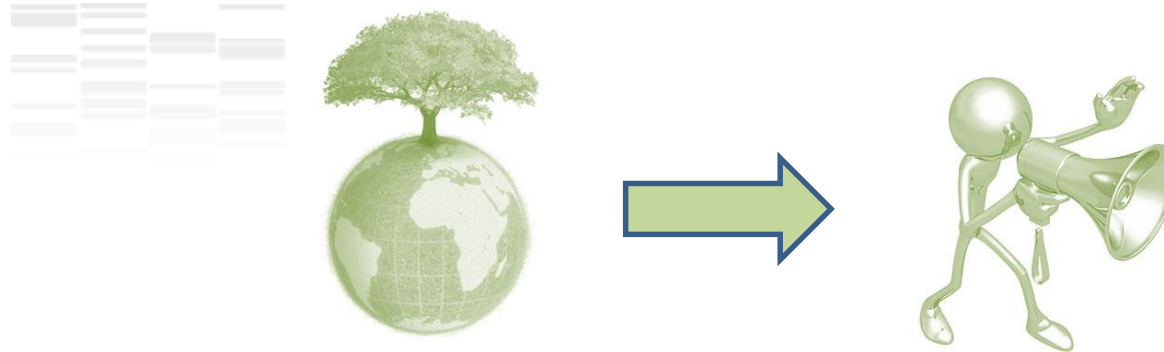
Principe d'open data et précautions



Principe d'Open Data:

- ❖ Démarche de communication des données sans attendre la demande d'un utilisateur,
- ❖ Sans restriction technique (format ouvert), juridique ou financière injustifiées.

Principe d'open data et précautions



Obligation de communication :

- ❖ Données environnementales (code de l'environnement, convention Arrhus) : par exemple l'état des éléments de l'environnement, notamment l'air, l'atmosphère, l'eau, le sol, les terres, les paysages, les sites naturels, les zones côtières ou marines et la diversité biologique, ainsi que les interactions entre ces éléments , etc. (sous réserve de la sécurité publique, secret défense ou droit de PI)
- ❖ Données géographiques (Directive Inspire)

Principe d'open data et précautions



Communication sous conditions :

- ❖ Données relatives au potentiel scientifique et technique de la nation
- ❖ Données protégées par le droit d'auteur et réalisées dans le cadre de l'établissement
- ❖ Données protégées par un droit de propriété intellectuelle autre que le droit d'auteur
- ❖ Données personnelles
- ❖ Données de santé
- ❖ Données provenant d'un tiers



Principe d'open data et précautions



Principe d'interdiction de communication :

- ❖ Les documents réalisés en exécution d'un contrat de prestation de services exécuté pour le compte d'une ou de plusieurs personnes déterminées (non publiques)
- ❖ Les données à la sécurité publique ou au secret défense : données recueillies lors d'un projet secret-défense, données impactant la sécurité publique, données impactant la sécurité des biens de l'établissement, données impactant la sécurité des personnes
- ❖ Les données relatives aux secrets professionnels : secret des procédés, secret des informations économiques et financières, secret des stratégies commerciales ou industrielles

Principe d'open data et précautions

Il faut s'astreindre à « auditer » les données avant de les organiser en base de données :



1. Sur l'origine des données :
 - Suis-je producteur des données ?
 - Sinon, ai-je l'autorisation du fournisseur ?
 - Si oui => examiner la licence d'utilisation (correspond-elle à mes besoins?)
 - Si non => prendre contact avec le fournisseur
2. Réexaminer chaque type de données (personnelle, etc.) et m'assurer que j'ai bien le droit de les diffuser (Cf. Diapo 16 à 18)

Principe d'open data et précautions



3. Si je crée une base de données :
 - Anonymiser les données personnelles
 - Séparer les données par type pour pouvoir en diffuser certaines et pas d'autres (gestion des droits d'accès)


4. La base est réalisée en partenariat : m'entendre avec les autres sur les conditions de diffusion des données et de la base => faire un data management plan (plan de gestion des données) avec eux

5. Je peux diffuser => licence de réutilisation
 - Limitation de la responsabilité
 - Attention à la future loi Lemaire qui pourrait imposer une licence
 - A titre gratuit sauf exceptions (loi valter)

Principe d'open data et précautions

Exemple de licences en attendant la loi Lemaire :

-Creative commons  : faciles à utiliser et adaptées à toutes les situations. Attention aux contraintes pour la réutilisation.

-ODBL  : de type « copyleft », les redistributions et travaux dérivés doivent se faire sous la même licence ou licence équivalente. Attention aux contraintes pour la réutilisation.

-Licence ouverte Etalab  : de type « permissive », pas de contrôle sur les redistributions et travaux dérivés.

Principe d'open data et précautions

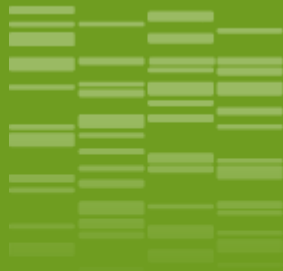
Pour vous aider :

Un guide juridique est en cours de finalisation. Il est à l'attention des chercheurs pour les aider à expertiser la nature de leurs données avant toute mise en open data.

Ce guide est issu d'un groupe de travail d'experts de divers organismes et universités.

Publication prévue pour septembre 2016,
et pour l'INRA sur le site en cours de construction de l'unité
Ingénierie Numérique : <http://www.inra.fr/ingenium>





MERCI DE VOTRE ATTENTION