

## Module 4

# Partage et valorisation des données : questions juridiques

Paulette Lieby [0000-0002-9289-9652](tel:0000-0002-9289-9652)



# BiGEst

## Questions juridiques

avec Frédéric Lamotte [0000-0003-4234-1172](tel:0000-0003-4234-1172)  
et Lionel Maurel [0000-0002-8667-5900](tel:0000-0002-8667-5900)

**Avertissement** : *Les aspects juridiques ne sont pas mon coeur de métier, aussi je n'aurai pas de réponse immédiate à toutes les questions que vous pourriez poser.*

*Ces notes ont été conçues comme un guide d'exploration du cadre juridique français et européen. Toute erreur est ma responsabilité seule.*

- Définition des « données de la recherche »
- Présentation du cadre juridique

- définition «standard» ([OCDE 2007](#)) : « 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 les résultats de la recherche »
- néanmoins, définitions moins restrictives ([ref.](#)) :
  - données d'observation, données expérimentales, données computationnelles ou de simulation, données dérivées ou compilées, et données de référence (au delà du «factuel», INIST)
  - aussi, les données ne sont pas seulement des données «nécessaires à la validation des résultats» ([ref.](#)) :
    - typiquement, en bioinformatique, les données produites sont plus nombreuses que celles strictement nécessaires pour valider un résultat
    - ces données gagnent en valeur précisément si elles peuvent être partagées

- d'ailleurs la dernière mise à jour du document OCDE ([2021](#)) complète la définition des données de la recherche avec :
  - « *autres objets numériques pertinents au regard de la recherche financés sur fonds publics : métadonnées, algorithmes, flux de travail, modèles et logiciels [...] utilisés dans le cadre de travaux de recherche et de développement* »
- définition légale, directive 2019/1024 sur les données publiques (art. 2) :  
« *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* »
- question : **quel est le statut des objets matériels ?**



- quel est le statut des données par rapport à leur ouverture ?
- quelles sont les questions que vous vous posez quant au partage de vos données,
- et comment y répondez vous

- *a priori*\*, les **données issues d'une activité de la recherche** sont soumises à une obligation de partage

**Les données sont soumises à un principe d'ouverture par défaut et de libre utilisation (Loi République numérique 2016 LPRN)**

\*«*a priori*» : l'objet d'une partie de ce module est de définir les contours de cette obligation de partage

Les données de la recherche sont des informations publiques :

- elles sont soumises à un principe d'ouverture par défaut et de libre utilisation (*Loi Lemaire - LPRN 2016*)
- elles sont soumises à un principe de gratuité (*Loi Valter 2015*)
- elles sont protégées contre les risques d'accaparement



retour sur la définition

- *Les « documents administratifs » visés par la LPRN sont tous les documents quels que soient leur date, leur lieu de conservation, leur forme et leur support qui sont 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** (cf. EPICs)*
- qualificatif important : on parle ici des documents **achevés**

## Principe : Les données de la recherche sont protégées contre les risques d'accaparement

- quels que soient les droits cédés à un éditeur, la partie *concernant les données* est nulle si elle va contre l'ouverture des données

## Principe : Les données de la recherche sont soumises à un principe de gratuité

- seule une liste fermée d'administrations peuvent encore valablement fixer des redevances de réutilisation :  
IGN, Météo France
- mais une articulation est possible avec le dépôt de brevets et d'autres formes de valorisation

## Principe d'ouverture par défaut et de libre utilisation (LPRN)

- la LPRN supprime le principe du régime dérogatoire où chaque établissement devait déterminer les modalités de la réutilisation des données
- il y a des protections possibles pour certains types de données, certains cas
  - il s'agit du **principe de l'exception**, fixé dans la loi et par la commission européenne :  
**« les données sont aussi ouvertes que possible et aussi fermées que nécessaire »**

- le **droit d'auteur**, pour les œuvres de l'esprit
  - les **publications scientifiques** : les chercheurs sont considérés comme des tiers par rapport à l'administration. Ils décident de la manière de diffuser, de partager leur œuvre.
  - l' **image**, qui sous certaines conditions d'originalité, peut aussi être considérée comme une œuvre :
    - la ligne de partage est fixée par l'**originalité**
  - les **logiciels**
  - les *bases de données*
  - etc...
- Note : une *exception aux règles du droit d'auteur* est applicable à l'exploration automatisée de textes et de données (Text and Data Mining ou TDM). ([La fouille de textes et de données à des fins de recherche](#))

## Exception : la propriété intellectuelle appartenant à des tiers 2.

- le **droit *sui generis*** (aussi nommé droit de production)
  - concerne exclusivement les **bases de données**
  - peut les protéger même à défaut d'originalité de celles-ci
  - l'objectif étant de protéger les investissements réalisés
- les bases de données sont donc susceptibles de recevoir une double protection, le droit d'auteur (originalité de la structure) et le droit *sui generis*



- **les projets partenariaux et les droits des tiers**
  - dans le cas de **recherche partenariale** (acteurs publics et privés), quand une entreprise privée partage des données
  - l'accès, la transmission, etc. sont fixés par des règles, par notamment les accords de consortium et les contrats
- néanmoins, la réutilisation des données est libre si :
  - elles sont issues d'une activité de recherche financée **au moins pour moitié** par des fonds publics

- L'ancienne *Loi Informatique et Liberté* devenue en 2018 le [RGPD](#) ordonne que les données à caractère personnel ne **doivent pas** être soumises à l'ouverture
  - Il est **obligatoire** de recueillir le **consentement**
  - Il est obligatoire de **pseudonymiser** ou d' **anonymiser** les données
    - Note : anonymisation différente de pseudonymisation -- la 1re ne tombe pas sous le RGPD, mais la 2nde oui
- Il existe néanmoins des **dérogations pour la recherche**, en organisant un cadre pour le partage sécurisé des données
- s'applique au *personnes physiques vivantes*

## *Exception : les données sensibles, les secrets protégés*

- Données sensibles qui ne sont pas des données personnelles
  - données de biodiversité, etc...
- Secret médical, secret des affaires, secret des procédés, secret militaire, ...

**sortent de l'ouverture par défaut**

« aussi ouvert que possible ;  
aussi fermé que nécessaire » :  
les principes FAIR

- les principes FAIR sont un *ensemble de principes directeurs* visant à rendre les données **trouvables, accessibles, interopérables et réutilisables**
- ces principes fournissent des orientations pour l'ouverture des données scientifiques et sont pertinents pour toutes les parties prenantes de l'écosystème numérique
- ces principes permettent la valorisation des données ; ils garantissent la transparence

**F**indable   
**A**ccessible   
**I**nteroperable   
**R**eusable 

- les licences pour encadrer le partage et la réutilisation des données
- techniquement, suite à la LPRN, une licence n'est pas nécessaire là où aucun droit ne s'exerce (par exemple, si le droit d'auteur ne s'exerce pas)
  - néanmoins, il est (fortement) recommandé d'en utiliser une dans tous les cas
- **pour clairement afficher les droits afférents**
- en effet,

<< Les titulaires peuvent décider de ne pas exercer leurs droits, mais **l'absence de formalisation de ce choix serait source d'insécurité juridique pour le projet et sa pérennité.** Une mauvaise gestion des droits pourrait conduire à des actions en contrefaçon et à la paralysie des projets de communs numériques. Il est donc important d'identifier tous ces titulaires (en limitant lorsque cela est nécessaire l'injection d'éléments protégés par des droits dont les contributeurs ne sont pas titulaires, et qu'ils ne pourront donc pas céder ou licencier) >>  
([Doctrines juridiques appliquées aux communs numériques](#))

- LPRN : une liste de licences [Licences - data.gouv.fr](https://data.gouv.fr/licences)
  - etalab : équivalent CC-BY 2.0
  - les licences CECILL pour les logiciels
  - ...
- les licences Creative Commons
  - toutes n'ont pas de jurisprudence spécifique en France mais leur validité n'est pas remise en cause
- Note : les licences CC 4.0 sont internationales et ne nécessitent pas/plus de transcriptions nationales



Le cas des données de la santé :

- le RGPD n'autorise pas le transfert de données hors EU
  - concerne les données sensibles, mais on peut se poser la question des données en général
- les entreprises américaines sont soumises au Cloud Act (Clarifying Lawful Overseas Use of Data Act). Votée en mars 2018, cette loi permet aux États-Unis d'accéder plus facilement aux données stockées sur des serveurs situés hors des États-Unis.
- mais : le cas du Health Data Hub (serveurs Microsoft aux Pays-Bas)
- 14 octobre 2020 : [décision](#) du conseil d'état « *Aucune donnée personnelle ne peut être transférée en dehors de l'Union européenne dans le cadre du contrat conclu avec Microsoft* »
- et pourtant : [Google et Thales s'allient dans le « Cloud de confiance »](#) (d'accord, on ne parle pas de données sensibles, néanmoins...)

## dans le contexte du partage des données

- en France
  - Loi Lemaire - LPRN 2016 (*ouverture des données*)
  - Décret n° 2021-1572 du 03/12/21 (*Obligations plus claires et plus fortes aux établissements d'enseignement supérieur et de recherche : ceux-ci sont appelés pour la première fois à définir une politique de conservation et de communication des résultats bruts de leurs travaux scientifiques, via notamment la mise en œuvre de PGDs*)
  - Loi Valter 2015 (*gratuité*)
  - RGPD 2018 (*protection des données sensibles*)
  - plus la loi Jardé : elle encadre expérimentations impliquant la personne humaine (ne concerne pas directement le partage des données)
- en Europe
  - [Directive on open data and the re-use of public sector information](#), **Open Data Directive**, 16 July 2019
  - [Regulation on a framework for the free flow of non-personal data in the EU](#) (EU) 2018/1807 (*free movement of non-personal data*)
  - GDPR 2018
- notons la prééminence de la LPRN

#DécretIntégrité

## Décret n° 2021-1572 du 3 décembre 2021 relatif au respect des exigences de l'intégrité scientifique

- **Article 2** "les établissements publics et fondations reconnues d'utilité publique (...) promeuvent la diffusion des publications en accès ouvert et la mise à disposition des méthodes et protocoles, des données et des codes sources associés aux résultats de la recherche afin d'en garantir la traçabilité et la reproductibilité. Ils incitent à la publication des résultats de recherche dits négatifs"
- **Article 6** "les établissements publics et fondations reconnues d'utilité publique (...) définissent une politique de conservation, de communication et de réutilisation des résultats bruts des travaux scientifiques menés en son (*sic*) sein. A cet effet, ils veillent à la mise en œuvre par leur personnel de plans de gestion de données et contribue (*sic*) aux infrastructures qui permettent la conservation, la communication et la réutilisation des données et des codes sources"
- **Article 5** précise que les résultats bruts en question "sont constitués des données produites au cours du processus de recherche, ou à défaut des données traitées et scientifiquement validées, ainsi que des codes sources utilisés dans le traitement de ces données"

du réseau [consortium Couperin](#)



## *Des principes sur lesquels s'appuie la politique publique de science ouverte*

- [Deuxième Plan national pour la science ouverte](#)
- [Rapport Bothorel sur la politique publique de la donnée, des algorithmes et des codes sources](#)



- [INRAE](#) ← excellent !!
- [Pasteur](#)
- [CIRAD](#)
- [Ponts et Chaussées](#)
- [Ouvrir la science](#)
- un outil de gestion des données soumises au RGPD : [DAISY](#) (ELIXIR-LU)

- points d'entrée excellents et compréhensifs :
  - [Doctrine juridique appliquée aux communs numériques](#)
  - [Ouverture des données de la recherche. Guide d'analyse du cadre juridique en France](#)
- autres :
  - [Aspects juridiques et éthiques – DoRANum](#)
  - [Ouverture des données de la recherche : de quoi parle-t-on ?](#)
  - [À qui appartiennent les données ?](#) excellent résumé de la présa de Lionel Maurel
  - [Les principes FAIR – DoRANum](#)
  - [En route vers la science ouverte : Gérer les données de sa recherche](#)
  - [Guide de bonnes pratiques sur la gestion des données de la Recherche](#)
  - [OCDE : Recommandation du Conseil concernant l'accès aux données de la recherche financée sur fonds publics](#)
  - [Guide d'application de la loi pour une République numérique](#)



- CNIL
  - [Recherche \(hors santé\) : les principales références juridiques](#)
  - [Quelles formalités pour les traitements de données de santé à caractère personnel ?](#)
- EU :
  - [Opinion 3/2019 concerning the Questions and Answers on the interplay between the Clinical Trials Regulation \(CTR\) and the General Data Protection Regulation \(GDPR\)](#)
- [Comment démarrer la mise en place du RGPD pas à pas](#)
- [Impact du Règlement Général sur la Protection des Données \(RGPD\) en Recherche Clinique](#)
- [Synoptique juridique de la recherche en santé \(oct 2020\)](#)

## Commentaires ? Questions ?

Remerciements :

- toute l'équipe de formation @ IFB @ BiGEst
- ELIXIR-CONVERGE