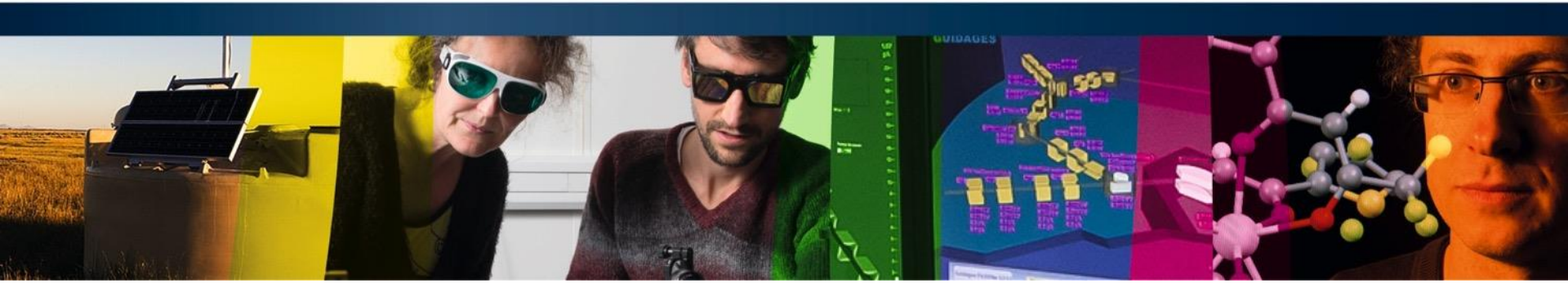
www.cnrs.fr

Questions juridiques autour de l'ouverture des données

Nathalie Le Ba
Direction des affaires juridiques
CNRS

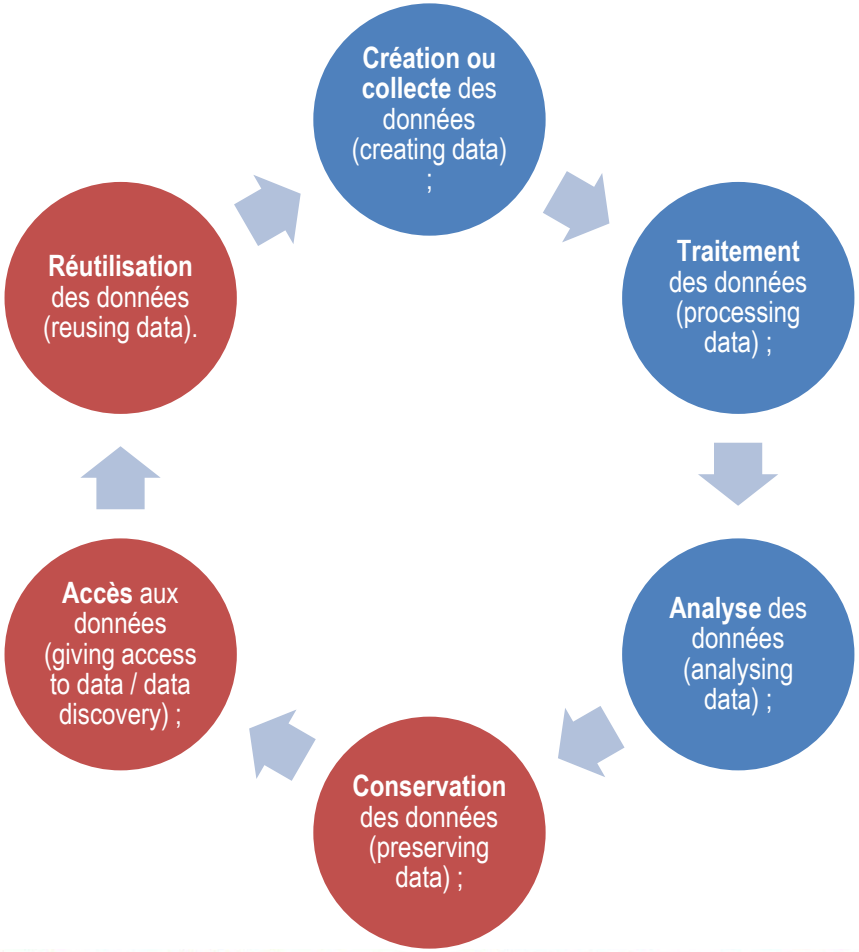


De quelles données parle-t-on?

www.cnrs.fr

- Données protégées par le droit d'auteur
- Données contenues dans une BDD de l'administration
- Données à caractère personnel
- Données brutes
- Données couvertes par le secret (professionnel, industriel)
- Données géographiques (directive INSPIRE),
- Données « dont la publication présente un intérêt économique, social, sanitaire ou environnemental
- Données archivées
- …

Cycle de la vie des données



www.cnrs.fr

La conservation des données

Conservation

- Article L211-1 Modifié par LOI n°2016-925 du 7 juillet 2016 - art. 59 : « les archives sont l'ensemble des documents, y compris les données, quels que soient leur date, leur lieu de conservation, leur forme et leur support, produits ou reçus par toute personne physique ou morale et par tout service ou organisme public ou privé dans l'exercice de leur activité »

- Les archives publiques sont : les documents qui procèdent de l'activité de l'Etat, des collectivités territoriales, des établissements publics et des autres personnes morales de droit public

- Instruction nationale CNRS – Direction des Archives de France du 15 janvier 2007 : Traitement et conservation des archives des délégations du CNRS et des archives des unités de recherche et de service

Conservation

- LRN article 6 : « les administrations mentionnées au premier alinéa de l'article L. 300-2 CRPA ne sont pas tenues de publier les archives publiques issues des opérations de sélection prévues aux articles L. 212-2 et L. 212-3 du Code du patrimoine »
  - Services d'archives pas tenus de diffuser en ligne les archives papiers et données numériques des administrations (charge trop importante)
  - Possibilité d'accès par les chercheurs aux grandes BDD des administrations qui ne sont pas librement communicables (parce que comportant des secrets protégés par la loi (secret de la vie privée en particulier)

Conservation

- Critères de viabilité d'un format pour la préservation de la donnée (selon la mission Etalab) :
  - l'utilisation large du format
  - le caractère non propriétaire du format
  - l'ouverture du format
  - la disponibilité d'outils de validation et d'analyse de ces formats
    - la compression avec ou sans perte
    - l'absence de mécanismes de protection de droits (DRM)

www.cnrs.fr

L'accès aux données

Droit d'accès des citoyens aux documents administratifs

www.cnrs.fr

- Droit personnel des citoyens d'accès aux documents administratifs
  - s'applique à tous les documents, quels qu'en soient la forme et le support, produits ou reçus par les « autorités administratives » dans le cadre de leur mission de service public.
  - Certaines restrictions au droit d'accès sont néanmoins prévues pour préserver divers secrets

- Élargissement de ce droit avec la LRN
  - codes sources
  - documents relatifs à la gestion du domaine privé de l'État

Diffusion de la donnée

- **Open data par principe** des informations publiques
  - Deux conditions :
    - les documents sont disponibles sous forme électronique
    - ils sont détenus par des administrations ou par des personnes morales soumises à la « loi CADA » dont le nombre d'agents ou de salariés est supérieur à un seuil fixé par décret

Diffusion de la donnée

www.cnrs.fr

- Les documents suivants doivent faire l'objet d'une publication en ligne :
  - tout document communiqué à une personne en ayant fait la demande
  - le répertoire des principaux documents dans lesquels figurent des informations publiques, qui doit obligatoirement être tenu par chaque administration
  - les bases de données de ces administrations,
  - les données « dont la publication présente un intérêt économique, social, sanitaire ou environnemental
  - les règles définissant les principaux traitements algorithmiques utilisés dans l'accomplissement de leurs missions

Diffusion de la donnée

- sort particulier pour la publication de documents comportant des données à caractère personnel. Trois conditions alternatives :
  - l'existence de dispositions législatives expresses,
  - l'accord des personnes concernées
  - ou la mise en œuvre d'un traitement permettant de rendre impossible l'identification de ces personnes (anonymisation).
- analyse du risque de ré-identification des personnes préalable nécessaire

Diffusion de la donnée

- Autres exceptions ou restrictions
  - le secret en matière commerciale et industrielle,
  - Secrets professionnels
  - le secret de la défense nationale,
  - des impératifs relatifs à la sécurité de l'Etat ou de l'établissement (ex. sécurité des systèmes d'information des administrations),
  - PPST
  - Droits de tiers
  - documents non encore librement communicables au regard du code du patrimoine.

L'accès aux données géographiques

- Directive EU INSPIRE, si disponible sous format électronique :
  - rendre accessibles au public les données géographiques en publiant sur internet ces données et les métadonnées correspondantes ;
  - partager leurs données environnementales géographiques entre les autorités publiques.

- Données environnementales (relatives à l'émission de substance dans l'environnement) :
  - Communication obligatoire, sur demande, sauf cas très exceptionnels

www.cnrs.fr

La réutilisation des données

Objectif de l'Open data par principe

www.cnrs.fr

- Objectifs LRN : favoriser et à faciliter la réutilisation de documents administratifs à d'autres fins que celles pour lesquelles ils ont été produits ou reçus par les administrations

- Toute communication ou publication *« se fait dans un standard ouvert, aisément réutilisable et exploitable par un système de traitement automatisé »*

La réutilisation des données

- le droit sui generis des administrations sur les bases de données publiées ne peut faire obstacle à la réutilisation du contenu de ces bases par tout tiers

- sous réserve de droits de propriété intellectuelle détenus par des tiers

Le droit sui generis BDD

www.cnrs.fr

- Droit sui generis des bases de données :

- Droit spécifique octroyé sur le contenu de la base
- En faveur du producteur
- Pouvoirs d'interdire :
  - L'extraction permanente ou temporaire de la totalité ou d'une partie qualitative ou quantitativement substantielle de la base
  - La réutilisation par mise à disposition du public de la totalité ou d'une partie qualitativement ou quantitativement substantielle de la base
  - L'extraction ou la réutilisation répétée et systématique de parties qualitativement ou quantitativement non substantielles de la base lorsque ces opérations excèdent manifestement les conditions d'utilisation normale de la base

- Le producteur doit mentionner ces droits pour pouvoir en bénéficier.

Recherches financées pour moitié sur fonds publics

www.cnrs.fr

- L'établissement ou le chercheur ne peut pas s'opposer à la réutilisation dans ce cas :
  - Article L.533-4 II code de la recherche : « Dès lors que les données issues d'une activité de recherche financée au moins pour moitié par des dotations de l'Etat, des collectivités territoriales, des établissements publics, des subventions d'agences de financement nationales ou par des fonds de l'Union européenne ne sont pas protégées par un droit spécifique ou une réglementation particulière et qu'elles ont été rendues publiques par le chercheur, l'établissement ou l'organisme de recherche, leur réutilisation est libre. »

Le principe de gratuité

www.cnrs.fr

- La réutilisation est gratuite par principe
- La tarification devient l'exception et n'est autorisée que dans de rares cas
- Licence non obligatoire, sauf dans les cas de tarification
- doit être conforme avec une liste homologuée par décret et établie par Etalab, ou bien être homologué par eux.

Les licences

- Décret n° 2017-638 du 27 avril 2017 relatif aux licences de réutilisation à titre gratuit des informations publiques et aux modalités de leur homologation
  - Liste (https://www.data.gouv.fr/fr/licences)
    - Pour les données :
      - Etalab 2.0 = totale liberté de réutilisation, y compris à des fins commerciales, des données accordée au réutilisateur
      - ODbL 1.0 = obligation de réciprocité
  - Ou homologation :besoins spécifiques qui ne pourraient être couverts par les licences figurant dans le décret
    - Auprès de la DINSIC

Les licences

- L.322-1 CRPA : Que l'établissement choisisse ou non une licence, la loi indique que tout « réutilisateur » doit (sauf si l'établissement y renonce) :
  - respecter l'intégrité des données (absence d'altération, absence de dénaturation du sens) ;
  - faire mention de la source des données ;
  - veiller à ce que l'indication de la date de dernière mise à jour soit bien présente.

- Les réutilisations commerciales ne peuvent pas être empêchées.

- Attention, licence = garantie de jouissance paisible

www.cnrs.fr

En pratique?

LRN au CNRS

www.cnrs.fr

- Données concernées :
  - données administratives détenues par exemple au sein des directions fonctionnelles ou des délégations régionales du CNRS
  - Données de recherche
  - Bases de données et logiciels

- Supports :
  - dossiers, rapports études, comptes rendus, procès-verbaux, statistiques, instructions, circulaires, notes, correspondances, avis, prévisions, codes sources ou décisions
  - Sous forme électronique

Le management des données

1. Personnes impliquées :
   - Producteurs
   - Responsable de la gestion des données au cours du projet (plan de gestion, sa rédaction, son suivi et ses modifications)
   - recenser l'ensemble des employeurs et acteurs pour déterminer les participations (matérielles, intellectuelles ou financières) de chacun

2. Définir des règles de diffusion en amont avec les partenaires
   - Ex. charte d'utilisation d'un instrument scientifique, clauses de l'accord de consortium, ou du contrat de collaboration
   - formuler un plan de gestion des données, qui abordera la problématique de la diffusion et de la réutilisation des données
   - Choisir la licence si nécessaire

Le management des données

3. Introduire l'open data dès le début du partenariat :
   - prévenir le partenaire, dans le contrat, que l'administration sera tenue de communiquer certaines données aux citoyens qui en feront la demande et que l'accès, voire la diffusion de certaines informations est une obligation légale pesant sur l'Administration ;
   - prévoir une exception à la confidentialité des informations échangées dans le cadre du partenariat ;
   - dans les contrats de marché public.

Le management des données

4. Déterminer l'existence de droits de tiers sur les données :
   - Confidentialité? Savoir-faire?
   - Droit d'auteur sur les contenus?
   - Réutilisation de données existantes?
   - Prestations externes? Notamment de stockages
     - Conditions générales d'utilisation?
     - Droits du prestataire?

5. Quel public cible envisagé?
   - grand public, communauté scientifique, secteur privé…
   - perspectives d'application ou de développement possibles à partir des données

Le management des données

www.cnrs.fr

6. Justification du principe d'exception aux conditions générales de diffusion
   - risque industriel, respect de la vie privée, priorité d'exploitation, exposition de ressources rares (biodiversité, archéologie)
   - mesures de protection (Ex. des données personnelles (anonymisation)
   - Embargo

7. Questions éthiques?

8. Y a-t-il des services à valeur ajoutée ?
   - Quelle rémunération possible?