Ouverture des données de la recherche. Guide d'analyse du cadre juridique en France

Nicolas Becard, Céline Castets-Renard, Gauthier Chassang, Martin Dantant, Laurence Freyt-Caffin, Nathalie Gandon, Caroline Martin, Andrea Martelletti, Alexandra Mendoza-Caminade, Nathalie Morcrette, et al.

▶ **To cite this version:**

Nicolas Becard, Céline Castets-Renard, Gauthier Chassang, Martin Dantant, Laurence Freyt-Caffin, et al.. Ouverture des données de la recherche. Guide d'analyse du cadre juridique en France. [Rapport de recherche] Comité pour la science ouverte. 2017, 45 p. hal-02791224

HAL Id: hal-02791224
https://hal.inrae.fr/hal-02791224

Submitted on 5 Jun 2020

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.

Open licence - etalab

Ouverture des données de recherche

Guide d'analyse du cadre juridique en France

Contenu sous licence ouverte

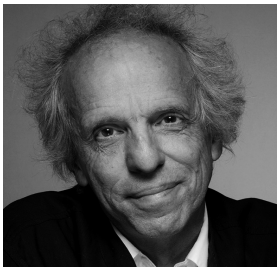
Le présent guide est issu des réflexions d'un groupe de travail inter-organismes animé par l'INRA. Il ne prétend pas à l'exhaustivité et est fourni uniquement à titre d'information. Il ne saurait en tout état de cause se substituer aux politiques d'établissements, au respect des dispositions législatives ou réglementaires et au respect de la jurisprudence applicable en la matière. Ce guide peut évoluer.

Membres du groupe de travail : BECARD Nicolas (INRA), CASTETS-RENARD Céline (UT1), CHASSANG Gauthier (Inserm, Membre de la Plateforme Genotoul Societal), DANTANT Martin, FREYT-CAFFIN Laurence (Irstea), GANDON Nathalie (co-animatrice, INRA), MARTIN Caroline (Agreenium), MARTELLETTI Andrea (stagiaire INRA, M2 droit et Informatique), MENDOZA-CAMINADE Alexandra (UT1), MORCRETTE Nathalie (co-animatrice, INRA), NEIRAC Claire (Cirad), avec la participation d'Inno[3] (Benjamin JEAN, Laure KASSEM).

Avec le soutien du Comité pour la science ouverte

V2 - Décembre 2017

Préface



Alain Beretz,
Directeur général de la recherche et de l'innovation

Public money? Public data! De la nécessité d'un guide juridique

Le Code de la recherche stipule que « la recherche publique a pour objectifs le développement et le progrès de la recherche dans tous les domaines de la connaissance », notamment à travers « l'organisation de l'accès libre aux données de la recherche ». En même temps, la France a joué un rôle pionnier dans la protection des données personnelles en publiant le 6 janvier 1978 la loi relative à l'informatique, aux fichiers et aux libertés, plus connue sous le nom de loi informatique et libertés. Celle-ci régule avec prudence la façon dont les données personnelles peuvent circuler sur le réseau. Et elle s'applique bien entendu aux données scientifiques, aussi légitime que soit la recherche qui les a produites.

Ces deux textes résument à eux seuls la complexité de la question de l'accès ouvert aux données de la recherche. Le législateur souhaite clairement, et de plus en plus, faire de l'accès ouvert le principe par défaut pour toutes les données produites par la recherche publique. *Public money? Public data!* Cette volonté est particulièrement explicite dans la loi pour une République numérique d'octobre 2016. Mais en parallèle, le secret médical, le secret des affaires, le droit d'auteur ou le règlement européen sur la protection des données personnelles (RGDP) s'imposent à tous et pondèrent, réduisent voire annulent toute possibilité d'ouverture des données. Il n'y aura pas d'ouverture qui serait irrespectueuse des individus et des règles légitimes de protection collectives.

Frédérique Vidal, la ministre de l'Enseignement supérieur, de la Recherche et de l'Innovation, a décidé de faire de la science ouverte une de ses priorités. Or, pour passer de l'approche générale à l'implémentation de la science ouverte, il faut affronter des nombreux questionnements, notamment techniques, documentaires, économiques, juridiques. Dès lors, la publication par le Comité pour la science ouverte de ce guide juridique marque une étape essentielle dans la maturation par les établissements et par les acteurs de la recherche de la question de l'ouverture des données.

À la lecture de ce guide, on comprend vite que le sujet ne se résume pas en une opposition binaire entre ouverture et fermeture. Je remercie les auteurs pour le temps et l'expertise qu'ils ont mobilisés afin produire ces pages limpides, dont nous espérons qu'elles guideront les politiques d'ouvertures les plus ambitieuses, et les plus raisonnables, les plus fertiles et les plus respectueuses. Les données de la recherche appartiennent au patrimoine de l'humanité et sont un bien commun, coûteux et précieux. Grâce à ce guide, vous saurez jusqu'où pouvez et devez les ouvrir.

Sommaire

Préambule 7

Qu'est-ce que l'Open Data ? 7

Notion d'Open Data 7

Open Data des données de la recherche 10

1-Les données que j'exploite sont-elles concernées par l'Open Data ? 11

1.1 Communication obligatoire 11

1.1.1 Données géographiques 11

1.1.2 Données relatives à des émissions de substances dans l'environnement 12

1.2 Communication interdite par principe 13

1.2.1 Données présentant des risques pour la protection du secret de la défense nationale 13

1.2.2 Données présentant des risques pour la sûreté de l'État, la sécurité publique, la sécurité de l'établissement 13

1.2.3. Secrets professionnels 14

1.3 Communication sous conditions 16

1.3.1 Données présentant des risques pour la protection du potentiel scientifique et technique de la nation 16

1.3.2 Le cas des zones à régime restrictif (ZRR) 17

1.3.3. Données protégées par le droit d'auteur et autres droits de propriété intellectuelle 17

1.3.3.1 Droit d'auteur 17

1.3.3.2 Autres droits de propriété intellectuelle 19

1.3.4 Données relatives aux personnes, à la vie privée 20

1.3.5 Données statistiques 21

1.3.6 Données liées à un contrat avec un tiers non soumis à une obligation de service public 21

1.4 Conclusion 22

2-Comment diffuser les données ? 23

2.1 Les grands principes 23

2.1.1 Aspects techniques de l'Open Data 23

2.1.2 Aspects juridiques : la licence de diffusion 24

2.2 Les différentes modalités de diffusion 25
2.2.1 Les pratiques par discipline 25
2.2.2 La politique d'établissement 25
2.2.3 Le répertoire des données publiques 26
2.2.4 La demande d'accès formulée par un tiers 26

Logigramme de communicabilité d'une donnée 28

Annexes 30

Fiche 1. Les bases de données en bref 31
Fiche 2. Les données personnelles 33
Fiche 3. Les données statistiques 37
Fiche 4. Convention d'Aarhus sur l'information en matière d'environnement 38
Fiche 5. Tableau comparatif des licences gratuites ODBL et Etalabab 40
Fiche 6. Mise en place d'une licence par la voie du contrat électronique 42
Fiche 7. Archives 43

Préambule

« La recherche publique a pour objectifs le développement et le progrès de la recherche dans tous les domaines de la connaissance ; la valorisation des résultats de la recherche ; le partage et la diffusion des connaissances scientifiques ; le développement d'une capacité d'expertise ; la formation à la recherche et par la recherche et l'organisation de l'accès libre aux données de la recherche »[1].

Ce guide sur l'ouverture des données de recherche (ou Open Data) a pour vocation d'accompagner les agents des établissements concernés (établissements d'enseignement et organismes de recherche) dans une démarche d'ouverture raisonnée des données de recherche en tentant de répondre aux questions les plus courantes auxquelles ils pourront être confrontés, que cette démarche soit volontaire et réponde aux objectifs de l'établissement ou qu'elle soit imposée par la réglementation. Il est précisé que le cadre légal est cité, lorsqu'il existe. L'attention du lecteur est toutefois attirée sur le paysage très mouvant du droit sur ce sujet et sur la nécessité de se référer à la politique de son établissement en matière d'Open Data.

Qu'est-ce que l'Open Data ?

Notion d'Open Data

De façon générale, l'expression « Open Data » peut être définie comme une démarche de communication des documents ou données publics, afin qu'ils soient diffusés de manière structurée selon une méthode garantissant leur libre accès et leur réutilisation par tous, sans restriction technique, juridique ou financière injustifiée.

L'Open Data est régie en Europe par une directive 2013/37/UE dite « PSI » pour *Public Sector Information* et, en France, par la loi dite « CADA[2] » n° 78-753 du 17 juillet 1978 (modifiée à plusieurs reprises) codifiée depuis le 19 mars 2016 dans le Code des relations entre le public et l'administration (Livre III, titre 2e du CRPA). Ces textes définissent, d'une part, **un droit d'accès individuel aux documents administratifs** et, d'autre part, **un droit de réutilisation pour tous des informations qui y sont contenues, sous réserve d'exception détaillées dans ce guide.**

Les « documents administratifs » visés par la loi française[3] 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'État, les collectivités territoriales ainsi que par les autres personnes de droit public ou les personnes de droit privé chargées d'une telle mission.

La loi donne les exemples suivants : dossiers, rapports, études, comptes rendus, procès-verbaux, statistiques, directives, instructions, circulaires, notes et réponses ministérielles, correspondances, avis, prévisions, codes sources[4] et décisions. Cette liste n'est pas limitative.

1 Article L112-1 du Code de la recherche.
2 Commission d'accès aux documents administratifs.
3 Article L300-2 du Code des relations entre le public et l'Administration.
4 Voir en ce sens le jugement du tribunal administratif de Paris 10/03/2016 sur le code source du logiciel de calcul des impôts.

Cette définition est celle reprise, presque mot à mot, par le Code du patrimoine pour désigner les archives, bien que ces dernières soient soumises à un régime spécifique.

Cf. Fiche 8 sur les archives

Pour la qualification des données produites des établissements de recherche et d'enseignement en « documents administratifs », il y a lieu de retenir deux composantes :

1. Les données produites par les établissements de recherche et d'enseignement ***dans le cadre de leur mission de service public*** sont considérées comme des documents administratifs et sont donc communicables à toute personne qui en fait la demande, sauf exceptions légales.
2. Lorsque la loi parle de « documents administratifs », cela englobe également les ***données produites par ces établissements, quelles qu'elles soient :*** données brutes[5] , données élaborées et métadonnées. Cependant, la loi précise que ne peuvent être accessibles au public que les **documents « achevés ».** Par conséquent, tous les documents préparatoires ne sont pas communicables. On peut en déduire que les cahiers de laboratoire sont exclus de la réglementation sur l'ouverture des données (cf. également infra la définition de l'OCDE des données de recherche).

Ce guide n'aborde pas tous les types de données mais uniquement celles que les équipes de recherche produisent le plus couramment. Certaines données (données de transport, données essentielles des marchés publics, etc.) qui font l'objet d'une obligation de diffusion ne sont pas abordées ici.

Quelles sont les conséquences de la qualification des données en « document administratif » ?

1. Le principe posé par la réglementation française est celui de la ***diffusion en ligne par défaut*** des documents administratifs disponibles sous format électronique[6].
2. La réglementation affirme également la ***liberté d'accès*** aux documents administratifs pour toute personne qui en fait la demande. Si les données ne font pas déjà l'objet d'une diffusion publique, tout un chacun (citoyen, association, entreprise, administration, etc.), peu importe sa nationalité, peut demander à l'établissement/l'organisme qui a produit les données la simple consultation de celles-ci en vue d'une utilisation privée ou interne (c'est-à-dire non communiquée au public, pour les besoins propres de l'utilisateur) ou encore leur mise en accès public.

Pour ces deux principes, il y a cependant des exceptions liées à certains types de données, rendant nécessaire l'analyse, au cas par cas, de la nature et du contenu des données dont la communication est sollicitée. Les différentes étapes qui sont décrites dans ce guide fournissent une méthode pour effectuer cette analyse.

5 Les données préliminaires sont des données préparatoires, préalables, nécessaires à la mise en place d'une expérimentation, d'un procédé, d'une enquête, etc. : il ne s'agit pas de données de recherche. Les données brutes sont des données issues d'une l'expérimentation, d'un procédé, d'une enquête, etc. : il peut s'agir de données de recherche communicables..

6 Article L 312-1-1 du Code des relations entre le public et l'Administration.

3. Pour certaines données, en sus du droit d'accès aux documents, il existe un droit de ***réutilisation des informations publiques***[7], réutilisation qui par principe est **gratuite**[8]. L'instauration de ce droit à la réutilisation d'informations publiques témoigne de la volonté des institutions européennes et françaises d'amplifier le mouvement d'Open Data, dont l'objet est l'accès libre aux informations du secteur public au bénéfice du citoyen en vue de :
   - garantir la transparence de l'État ;
   - valoriser les données publiques ;
   - favoriser le développement d'activités privées et l'émergence de nouveaux opérateurs économiques.

Parmi les données accessibles, certaines peuvent être réutilisées par tout acteur qui souhaite en faire un usage différent de celui répondant à la mission de service public initiale. Aucune restriction d'usage ne peut alors être apportée par l'Administration. Cependant, toutes les données accessibles ne sont pas forcément réutilisables (cf. la méthode d'analyse ci-après).

Aussi, pour la r**éutilisation de données issues d'une activité de recherche**, la loi n°2016-1321 pour une République numérique du 7 octobre 2016 prévoit qu'elle est libre si[9] :

- ces données sont issues d'une activité de recherche financée au moins pour moitié par des fonds publics ;
- et ces données ne sont pas protégées par un droit spécifique ;
- et ces données ont été rendues publiques par le chercheur ou l'établissement.

Il est à noter que cette disposition s'inscrit dans un article du Code de la recherche sur les écrits scientifiques, l'objectif étant de réguler les relations entre les chercheurs et les éditeurs.

Pour l'analyse des modalités de réutilisation, cf. 2.1.2.

La distinction entre droit d'accès et droit de réutilisation est très importante. En effet, une personne qui a droit à l'accès à des données n'a pas nécessairement le droit de les réutiliser. Cependant, dans le cadre de l'Open Data, les documents ouverts sont de facto réutilisables

Il convient d'être particulièrement prudent lorsqu'il y a une publication scientifique et que l'éditeur impose le dépôt des données dans un entrepôt spécifique. Si c'est bien le scientifique qui décide du contenu de la publication, c'est, en revanche, à l'établissement de décider quelles données seront ouvertes, où elles doivent être déposées et sous quelles conditions. Les décisions d'ouverture des données se prennent au niveau de l'établissement et non pas au niveau de l'agent.

7 Article L321-1 du Code des relations entre le public et l'Administration – les informations publiques sont celles qui figurent dans les documents administratifs publics ou communiqués..

8 Article L324-1 du Code des relations entre le public et l'Administration.

9 Article L533-4 II du Code de la recherche.

Open Data des données de recherche

Les droits et principes définis ci-dessus en matière d'Open Data s'appliquent pleinement aux données de la recherche. Les règles de réutilisation des données sont aujourd'hui les mêmes pour tout document administratif, les établissements et organismes de recherche ne bénéficiant plus de la possibilité de fixer leurs propres conditions de réutilisation des données.

En matière de données de la recherche, il existe une définition (qui n'a pas de valeur légale) donnée par l'OCDE en 2007 dans un document relatif aux principes et lignes directrices pour l'accès aux données de recherche financée sur fonds publics. Selon l'OCDE, les *« données de la 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. Ce terme ne s'applique pas aux éléments suivants : carnets de laboratoire, analyses préliminaires et projets de documents scientifiques, programmes de travaux futurs, examens par les pairs, communications personnelles avec des collègues et objets matériels (par exemple, les échantillons de laboratoire, les souches bactériennes et les animaux de laboratoire tels que les souris). »*

En conclusion, les deux conditions préliminaires pour diffuser les données selon les principes de l'Open Data sont :

- **des données réalisées dans le cadre de la mission de service public de mon établissement (ceci est particulièrement important pour les EPIC) ;**
- **des données achevées.**

MAIS d'autres conditions doivent aussi être réunies, c'est pourquoi il faut vérifier la nature des données selon les étapes ci-dessous.

1- Les données que j'exploite sont-elles concernées par l'Open Data ?

Les données produites par les établissements de recherche et d'enseignement dans le cadre de leur mission de service public sont considérées comme des documents administratifs et sont donc **communicables à toute personne qui en fait la demande, sauf exceptions légales.**

Cette communication peut être obligatoire (1.1), interdite (1.2) ou soumise à certaines conditions (1.3) : cela dépend de la nature des données. Cependant, passé un certain délai (défini par le Code du patrimoine), même les documents qui ne sont pas communicables ou qui sont communicables seulement à certaines conditions deviennent accessibles au public (Cf. Fiche 7 sur les archives).

1.1 Communication obligatoire

Certaines données géographiques et environnementales doivent obligatoirement être ouvertes au public.

1.1.1 Données géographiques

Pour les données géographiques **qui sont disponibles sous format électronique,** c'est la directive européenne INSPIRE[10] qui s'applique. Elle impose aux autorités publiques :

- de rendre accessibles au public leurs données géographiques en publiant sur internet ces données et les métadonnées correspondantes ;
- de partager leurs données environnementales géographiques entre elles.

Ces données géographiques se décomposent en trois groupes principaux :

- données nécessaires au repérage sur le territoire,
  Exemple : hydrographie ;
- données générales complémentaires,
  Exemples : altimétrie, géologie ;
- données thématiques telles que bâtiments, vocation des sols, santé et sécurité des personnes, services d'utilité publique et services publics, etc.

Attention à ne pas confondre la communication obligatoire de jeux de données et le droit d'alerte institué par la loi n°2016-1691 du 9 décembre 2016, qui dit que : « Un lanceur d'alerte est une personne physique qui révèle ou signale, de manière désintéressée et de bonne foi, un crime ou un délit, une violation grave et manifeste d'un engagement international régulièrement ratifié ou approuvé par la France, d'un acte unilatéral d'une organisation internationale pris sur le fondement d'un tel engagement, de la loi ou du règlement, ou une menace ou un préjudice graves pour l'intérêt général, dont elle a eu personnellement connaissance. Les faits, informations ou documents, quel que soit leur forme ou leur support, couverts par le secret de la défense nationale, le secret médical ou le secret des relations entre un avocat et son client sont exclus du régime de l'alerte […]. ». Une telle diffusion d'information ne constitue pas une démarche d'Open Data puisque l'information s'adresse avant tout aux autorités publiques et que sa diffusion n'entraîne pas automatiquement un droit à réutilisation.

Pour en savoir plus

Guide* La Directive Inspire pour les néophytes, *page 10

Site internet naturefrance.fr

10 Directive INSPIRE 2007/2/CE du Parlement européen et du Conseil du 14 mars 2007.

1.1.2 Données relatives à des émissions de substances dans l'environnement

Concernant les données environnementales, le principe d'accès à ces données est fixé par la Convention d'Aarhus[11], qui déclare que « *toute personne a le droit d'être informée, de s'impliquer dans les décisions et d'exercer des recours en matière d'environnement* ». Ce principe a été repris par la législation européenne[12], puis par la législation nationale[13].

Pour en savoir plus :
Fiche 5. Convention Aarhus

Parmi les données environnementales, le cas des informations relatives à des émissions de substances dans l'environnement[14] est particulier puisque ces informations peuvent être communiquées à celui qui en fait la demande même si le document administratif dans lequel elles figurent n'est pas achevé.

Ces données recouvrent les émissions de substances susceptibles d'avoir des incidences sur l'air, l'eau, le sol, etc.

Exemples : enquête environnementale suite à une déclaration de mortalité massive d'abeilles[15] => communication obligatoire
Enquête sur une épidémie de légionellose[16] => communication refusée (car les données recueillies n'étaient pas relatives à l'environnement)

En outre, les cas de refus de communication sont extrêmement limités. Le refus de communiquer les données n'est justifié que si :

- la conduite de la politique extérieure de la France, la sécurité publique ou la défense nationale sont en cause ;
- des procédures juridictionnelles peuvent être impactées ;
- la communication porte atteinte à des droits de propriété intellectuelle.

En conclusion :

Les données concernées sont des données géographiques informatisées

=> l'établissement doit les diffuser.

L'établissement possède des données relatives à des émissions de substances dans l'environnement

=> létablissement doit les communiquer à ceux qui les demandent, sauf cas exceptionnels.

11 La Convention internationale d'Aarhus a été signée le 25 juin 1998.

12 Directives européennes 2001/42/CE du 27 juin 2001 ; 2003/4/CE du 28 janvier 2003 ; 2003/35/CE du 26 mai 2003 ; 2000/60/CE du 23 octobre 2000.

13 Loi n° 2005-1319 du 26 octobre 2005, décret n°2002-1187 du 12 septembre 2002, décret n° 2006-578 du 22 mai 2006 codifiés dans le Code de l'environnement, art. L124-1 & s.

14 Article L124-5 du Code de l'environnement.

15 Avis CADA n° 20130750 du 28/03/2013.

16 Avis CADA n° 20090310 du 26/02/2009.

1.2 Communication interdite par principe

1.2.1 Données présentant des risques pour la protection du secret de la défense nationale

Selon les dispositions du Code pénal[17], les procédés, objets, documents, données ou fichiers informatisés intéressant la défense nationale peuvent faire l'objet d'une classification au titre de la protection du secret de la défense nationale (très secret-défense, secret-défense, confidentiel défense). Cette classification et son niveau doivent apparaître sur le document, quel qu'en soit le support, afin d'éviter toute méprise[18]. Cette classification concerne, au-delà du domaine militaire, tous les champs d'activités nécessitant de limiter l'accès à la connaissance d'un contenu, pour ne pas créer des préjudices graves pour la sécurité nationale. La diffusion de tout ou partie de ces informations sans autorisation de l'émetteur est considérée comme un délit.

À titre d'exemple, peuvent être classifiés au titre de la protection du secret de la défense nationale :

- les documents relatifs aux plans gouvernementaux ;
- les documents relatifs aux opérateurs d'importance vitale ;
- les études financées par le ministère de la Défense ou l'Union européenne pour lesquelles il a été demandé une classification au titre de la défense.

Exemple : la réalisation d'un projet de recherche classifié, mené en collaboration avec le ministère de la Défense ne pourra donner lieu à communication.

Certains documents peuvent, sans être classifiés, porter la mention « diffusion restreinte », limitant la communication à un nombre limité de personnes destinées à en connaître le contenu du fait de leurs fonctions.

1.2.2 Données présentant des risques pour la sécurité de l'État, la sécurité publique, la sécurité de l'établissement

L'impératif de protection de la sécurité publique ou de la sécurité des biens et des personnes d'un établissement peut dans certains cas faire obstacle à la communication d'informations et de données.

Les motifs de refus du droit d'accès, compte tenu de ces risques, sont :

- la sécurité publique
  Exemples : la liste des laboratoires ayant de l'anthrax ne doit pas être ouverte pour prévenir les attentats tels que celui des « enveloppes piégées » envoyées suite aux attentats du 11 septembre 2001, la liste des réservoirs d'eau potable ne doit pas être ouverte pour prévenir des attentats à la contamination, etc. ;
- la sécurité des biens de l'établissement
  Exemples : la liste des laboratoires dont les recherches peuvent être soumises à contestation (OGM, expérimentation animale) pour prévenir des actions malveillantes, le plan complet des sites pouvant révéler les points de vulnérabilité pour prévenir des infractions, etc. ;

17 Article 413-9 du Code pénal.
18 Article 41 de l'Instruction générale interministérielle n°1300.

- la sécurité des personnes
  Exemples : diffusion sur internet d'informations permettant de fabriquer une bombe artisanale, liste de personnels d'installations sensibles, etc. ;
- la sécurité des systèmes d'information[19] de l'établissement
  Exemples : diffusion des bilans annuels des failles de sécurité des systèmes d'information, architecture des systèmes d'information.

Avertissement :
La liste des secrets n'est pas exhaustive. Par exemple, le secret des correspondances n'est pas traité par le guide.

1.2.3 Secret professionnel

L'article L311-5 2° h du Code des relations entre le public et l'administration indique que les documents administratifs dont la diffusion porterait atteinte aux autres secrets protégés par la loi ne sont pas communicables.

Une analyse au cas par cas est un prérequis.

Parmi ces secrets, le plus fréquemment rencontré par les établissements publics de recherche et d'enseignement supérieur est le secret professionnel.

La loi a édicté l'obligation du secret professionnel pour un certain nombre de professions dont les membres sont amenés, dans l'exercice habituel et normal de leur activité, à recueillir des informations confidentielles au sujet de personnes ou d'intérêts privés.

Le secret professionnel comprend notamment :

- **le secret médical**, qui s'impose à tous les médecins. Il couvre tout ce qui est venu à la connaissance du médecin dans l'exercice de sa profession, c'est-à-dire non seulement ce que lui a confié son patient, mais aussi ce qu'il a vu, entendu ou compris[20] => en cas de demande de communication de ce type de données, on ne peut pas faire droit à cette demande sans l'accord exprès de la personne concernée. *Exemple : un projet de recherche soumis au code de la santé publique, donc impliquant la personne humaine ;*
- **le secret de l'instruction ;**
- **le secret bancaire et le secret fiscal.**

Toutefois, le secret professionnel peut être levé sur autorisation de la personne concernée par l'information et sous réserve que soient préservés :

- la protection des personnes
  Exemple : révélation de maltraitances ;
- la santé publique
  Exemple : révélation de maladies nécessitant une surveillance ;
- l'ordre public et le bon déroulement des procédures de justice
  Exemples : dénonciation de crimes ou de délits, témoignages en justice.

19 Instruction interministérielle sur les systèmes d'information sensibles n°901.

20 Article 4 du Code de déontologie médicale, articles R.4127-4 et L.1110-4 du Code de la santé publique.

Cas particulier du « secret des affaires »

Le secret des affaires est une notion qui n'est pas clairement définie (notamment par le droit français) mais un certain nombre d'éléments sont à prendre en compte.

En Europe

Les parlementaires européens viennent de s'accorder au travers d'une directive[21] qui énonce 3 conditions à réunir pour qu'une information constitue un secret des affaires :

1. les informations sont secrètes en ce sens que, dans leur globalité ou dans la configuration et l'assemblage exacts de leurs éléments, elles ne sont pas généralement connues de personnes appartenant aux milieux qui s'occupent normalement du genre d'informations en question, ou ne leur sont pas aisément accessibles ;
2. les informations ont une valeur commerciale parce qu'elles sont secrètes ;
3. les informations ont fait l'objet, de la part de la personne qui en a licitement le contrôle, de dispositions raisonnables, compte tenu des circonstances, destinées à les garder secrètes.

En France

La CADA[22] a tenté de définir les secrets industriels et commerciaux : il s'agit d'éléments sensibles ayant notamment un impact sur l'environnement concurrentiel de l'établissement et de ses partenaires. Selon cette autorité, *« la communication des documents contenant des informations dont la divulgation pourrait porter atteinte au secret industriel et commercial est réservée aux seuls intéressés. La notion de secret industriel et commercial recouvre trois catégories de données »*[23] *:*

- **le secret des procédés ;**
- **le secret des informations économiques et financières ;**
- **le secret des stratégies commerciales ou industrielles.**

Ainsi en dehors des personnes intéressées, les documents comportant des mentions ou informations couvertes par le secret industriel et commercial ne sont communicables **qu'après occultation de ces mentions.** L'occultation doit être matériellement possible et le sens du document ne doit pas être dénaturé[24].

La loi réprime par des sanctions pénales le fait de divulguer des informations[25] couvertes par « les secrets » cités ci-dessus.

Les données couvertes par le secret des affaires (y compris donc les secrets industriels et commerciaux) peuvent être des données développées par l'établissement public (procédé, secret), mais aussi des données appartenant à des partenaires industriels qui ont été transmises à un établissement public à l'occasion, par exemple, d'une collaboration.

Pour en savoir plus :

Documents CADA
- Le secret en matière commerciale et industrielle
- Les documents couverts par le secret en matière commerciale et industrielle.

21 Directive 2016/943 du 8 juin 2016.

22 Commission d'accès aux documents administratifs.

23 Cette définition a été reprise par la loi pour une République numérique du 7 octobre 2016 (article L311-6 Code des relations entre le public et l'Administration).

24 Voir avis CADA n°20134817 du 19/12/2013 ; JP du CE 04/01/1905 David.

25 « La révélation d'une information à caractère secret par une personne qui en est dépositaire soit par état ou par profession, soit en raison d'une fonction ou d'une mission temporaire, est punie d'un an d'emprisonnement et de 15 000 euros d'amende. » Article 226-13 du Code pénal.

En conclusion

Il faut impérativement vérifier que les informations dont la diffusion est envisagée ne contiennent pas d'éléments pouvant :

1. **relever du secret de la défense nationale ;**
2. **présenter des risques pour la sécurité publique ou celle de l'établissement ;**
3. **relever du secret professionnel (médical, instruction, etc.) ou être couvert par la confidentialité (affaires, etc.).**

L'interdiction de diffusion est le principe.

En cas de question

Dans les cas 1 et 2, je contacte le fonctionnaire de sécurité et/ou la direction générale de l'établissement.

Dans le cas 3, je contacte le service juridique et/ou la personne responsable de l'accès aux documents administratifs de mon établissement.

1.3 Communication sous conditions

1.3.1 Données présentant des risques pour la protection du potentiel scientifique et technique de la nation

Le potentiel technique est constitué de l'ensemble des biens matériels et immatériels propres à l'activité scientifique fondamentale et appliquée et au développement technologique. Les éléments essentiels du potentiel constituent des intérêts fondamentaux de la nation définis dans le Code pénal[26] .

Un dispositif réglementaire rénové en 2012[27] a pour objectif la protection de ce potentiel scientifique et technique de la nation (PPST). Il vise à protéger l'accès aux savoirs et aux savoir-faire et aux technologies des établissements privés ou publics localisés sur le territoire national, lorsque leur détournement ou leur captation pourraient :

- soit porter atteinte aux intérêts économiques de la nation,
- soit renforcer des arsenaux militaires étrangers ou affaiblir les capacités de défense de la France,
- soit contribuer à la prolifération des armes de destruction massive (nucléaire, chimique ou biologique) et de leurs vecteurs,
- soit favoriser les actions malveillantes (terroristes) sur le territoire national ou à l'étranger.

Si mon laboratoire est confronté à l'un de ces risques, on parle d'unité protégée. Dans ce cas, il convient de solliciter l'avis préalable du directeur de l'unité avant diffusion, qui peut interroger le fonctionnaire de sécurité ou d'autres services compétents, afin de vérifier que la diffusion des données ne porte pas préjudice aux potentiels scientifiques et techniques de mon unité, de mon établissement ou de la nation.

26 Article 410-1 du Code pénal.
27 Circulaire interministérielle du 7/11/2012.

1.3.2 Le cas des zones à régime restrictif (ZRR)

Dans l'hypothèse où le ministère de tutelle de mon établissement a identifié un besoin supplémentaire de protection en raison du niveau de risque lié à l'activité de l'unité, une zone à régime restrictif (ZRR[28]) peut être créée.

L'accès physique et **numérique** à ces zones est soumis à autorisation. La diffusion de données issues de ces zones doit au préalable avoir été expressément autorisée par le responsable de la ZRR (ou son suppléant) qui peut interroger le fonctionnaire de sécurité ou d'autres services compétents.

Il est rappelé que si les données partagées (dans le cadre d'une autorisation d'accès aux informations accordée à un tiers) se situent toujours au sein de la ZRR, alors **il est obligatoire d'effectuer au préalable une demande d'autorisation d'accès comparable à celle d'un accès physique.**

En cas de doute, il convient de se référer à la politique de l'établissement et de contacter le fonctionnaire de sécurité défense.

1.3.3. Données protégées par le droit d'auteur et autres droits de propriété intellectuelle

Lorsque les informations à diffuser sont contenues dans des « documents » sur lesquels des tiers – non chargés d'une mission de service public – détiennent des droits de propriété intellectuelle[29], leur diffusion est soumise à l'accord préalable de ces tiers.

1.3.3.1 Droit d'auteur

Lorsque les informations que je souhaite communiquer relèvent du droit d'auteur, je ne peux les communiquer que sous réserve de respecter les droits du(es) auteur(s)[30]. Par ailleurs, ces informations ne sont pas réutilisables par des tiers sans l'accord de(s) l'auteur(s).

Comment savoir s'il y a un droit d'auteur sur les informations que je souhaite diffuser ?

Les éléments ou données (textes, interviews, musiques ou sons, images, discours, représentations graphiques, etc.) :

- qui existent sous une forme concrète (leur existence est perceptible par les sens),
- ET dont la forme est originale

constituent une œuvre de l'esprit protégée par le droit d'auteur.

Seul l'auteur a le droit exclusif d'exploiter son œuvre, de façon directe ou par l'intermédiaire d'un tiers *(exemple : par l'intermédiaire d'un éditeur)*. L'auteur peut donc discrétionnairement autoriser ou interdire l'exploitation de son œuvre et, en cas d'autorisation, il peut imposer des conditions d'exploitation.

Les idées exprimées dans un écrit peuvent être reprises librement. On dit qu'elles sont« de libre parcours ».
En effet, le droit d'auteur protège la façon dont ces idées sont exprimées et non le contenu des idées.

La qualification d'œuvre de l'esprit pour des données photographiques recueillies de façon automatique (par exemple un appareillage) lors d'une expérimentation n'est pas tranchée par les tribunaux mais on peut penser qu'une photographie prise automatiquement par un appareillage n'est pas originale.

28 « Les ZRR sont, aux termes de l'article 413-7 du code pénal, constituées de locaux et de terrains clos dans lesquels l'accès et la circulation sont réglementés afin d'assurer la protection des installations, du matériel ou du secret des recherches, études ou fabrications ».

29 Article L321-2 du Code des relations entre le public et l'Administration.

30 Article L311-4 du Code des relations entre le public et l'Administration.

Avertissement
Les œuvres (textes, images, etc.) glanées sur internet sont soumises au droit d'auteur. Même si aucune indication n'existe sur le site, l'accord de l'auteur est indispensable préalablement à toute réutilisation.

Dans le contexte de l'Open Data des données de recherche, quatre cas principaux peuvent se présenter.

1) *Les œuvres appartiennent aux scientifiques*

C'est le cas des œuvres réalisées par les chercheurs et enseignants-chercheurs.

Exemple : publications scientifiques.

Elles ne sont pas diffusables sans l'accord des scientifiques qui en sont les auteurs.

2) *L'établissement public dispose du droit d'exploitation des œuvres*

Dans la mesure où cela est nécessaire à l'accomplissement d'une mission de service public, le droit d'exploitation des œuvres réalisées par les agents autres que les chercheurs et enseignants-chercheurs est cédé de plein droit à la personne publique.

C'est aussi le cas des œuvres dites « collectives » c'est-à-dire les œuvres commandées par l'établissement et réalisées selon les directives précises données par l'établissement.

Exemple : encyclopédie.

Elles sont diffusables par l'établissement sans accord préalable des auteurs dès lors que cette diffusion est strictement nécessaire à l'accomplissement d'une mission de service public[31].

Il existe notamment une exception légale qui permet de reprendre certains éléments d'une œuvre sans autorisation de l'auteur. Il s'agit du droit de courte citation (article L122-5 du Code de la propriété intellectuelle). J'ai donc le droit de citer une œuvre sous réserve :
1. que la citation soit courte et non substantielle par rapport à l'œuvre
2. ET que soient cités le (ou les) auteur(s) et la source.

3) *Les œuvres appartiennent en tout ou partie à un tiers*

a. *Si ce tiers est un établissement public :*

Si un citoyen me demande l'accès à une œuvre appartenant à un établissement public autre que mon établissement, je dois transférer sans attendre cette demande audit établissement.

Si je veux diffuser une œuvre appartenant à un autre établissement public, je dois lui demander de le faire.

b. *Si ce tiers est un établissement ou une personne privé(e) :*

Si l'œuvre appartient à un tiers privé et a été réalisée en dehors d'une mission de service public, seul le tiers peut en autoriser l'accès et la réutilisation.

Son autorisation préalable est indispensable.

4) *L'œuvre est un logiciel réalisé dans le cadre d'un établissement public*

Quel qu'en soit l'auteur (chercheur, enseignant-chercheur, autre agent salarié), si l'œuvre est un logiciel réalisé dans le cadre de la mission, c'est l'établissement public qui est propriétaire des droits sur le logiciel[32].

En conséquence, c'est l'établissement qui décide de sa diffusion.

31 Article L131-3-1 du Code de la propriété intellectuelle.
32 Article L113-9 du Code de la propriété intellectuelle.

1.3.3.2 Autres droits de propriété intellectuelle

Lorsque les informations que je souhaite diffuser sont contenues dans des « documents » sur lesquels des tiers – non chargés d'une mission de service public – détiennent des droits de propriété intellectuelle[33], leur réutilisation est soumise à l'accord préalable de ces tiers.

Ces droits de propriété intellectuelle regroupent les marques, les dessins et modèles, les brevets et le droit des producteurs de bases de données.

1) Marques, dessins et modèles

Si les données que je veux diffuser contiennent des marques ou des dessins protégés, je ne peux pas les réutiliser sans prendre certaines précautions et je dois prendre contact avec mon service juridique.

2) Brevets, éléments brevetés

Le texte du brevet est accessible au public au maximum 18 mois après le dépôt de la demande de brevet[34], donc rediffusable à compter de cette date.

Toutefois, la reproduction du produit ou du procédé breveté n'est pas autorisée si le titulaire du brevet n'a pas donné son accord car cela constituerait une contrefaçon. Par exception, la vérification expérimentale que le brevet fonctionne est permise (exemption de recherche).

3) Bases de données

Si les données sont structurées dans une base de données, celui qui souhaite accéder aux données a le droit de faire une **extraction non substantielle, soit qualitativement, soit quantitativement, s'il a licitement accès à la base.** Le producteur de la base de données ne peut pas s'opposer à ce droit[35].

En revanche, dès lors que l'extraction devient substantielle, l'accord du producteur est requis. L'extraction est qualitativement substantielle dès lors qu'elle concerne des données déterminantes de l'objectif poursuivi par la base de données. Elle est quantitativement substantielle lorsqu'elle comporte un volume important de données.

Le producteur est celui qui a fait les investissements financiers, matériels ou humains substantiels pour la création de la base et son enrichissement. Les tribunaux sont très sélectifs et ne prennent pas en compte les investissements réalisés pour créer les données. En revanche, comptent les moyens consacrés à la recherche d'éléments existants, à leur vérification et correction et à leur rassemblement dans la base.

Plusieurs cas peuvent se présenter :

- **le producteur de la base de données est mon établissement.** En principe, l'établissement doit communiquer la base de données au citoyen qui en fait la demande. Depuis la loi pour une République numérique, c'est un principe de diffusion plutôt que de communication sur demande qui doit s'appliquer et l'établissement ne peut plus faire valoir son droit de producteur de base de données

La notion de service public est difficile à cerner en l'absence de définition légale. C'est la jurisprudence administrative qui définit cette notion. Sur la base d'un faisceau d'indices, la qualification de service public se déduit de la qualité de l'organisme et de la mission d'intérêt général exercée par celui-ci.
Exemple : France Télécom accompli certaines missions dans le cadre d'un service public (entretien du réseau de communication) et d'autres missions dans un cadre privé (vente de prestations, de matériels).

Pour en savoir plus sur les bases de donnée et l'identification du producteur :
Fiche 1 :
Les bases de données en bref

33 Article L321-2 du Code des relations entre le public et l'Administration.
34 Article L612-21 du Code de la propriété intellectuelle.
35 Article L342-3 du Code de la propriété intellectuelle.

pour s'opposer à la diffusion[36]. Pour déterminer et analyser les conditions et limites de l'accès à la base et de la réutilisation des données, je prends contact avec mon service juridique ;

- **le producteur de la base de données est un autre établissement public.** Il faut transmettre la demande d'accès du citoyen à cet établissement ;
- **le producteur de la base de données est une personne physique ou morale de droit privé (hors mission de service public).** Il faut transmettre la demande d'accès du citoyen à cette personne qui pourra s'opposer à la diffusion ou la réutilisation de sa base de données.

1.3.4 Données relatives aux personnes, à la vie privée

Si les informations que je souhaite diffuser sont relatives à des personnes physiques et permettent leur identification directe ou indirecte, alors il s'agit de données à caractère personnel[37].

Lors de la constitution d'un jeu de données qui va inclure des données personnelles, je ne peux collecter ni réaliser aucun traitement de ces données sans avoir, **au préalable**, respecté les « formalités CNIL »[38] imposées par la loi 78-17 du 6 janvier 1978 modifiée.

Les formalités imposées par la loi diffèrent en fonction de la nature des données traitées et de la finalité poursuivie par le traitement.

Pour en savoir plus :
Fiche 2. Les données personnelles

Je dois être très vigilant à l'égard du respect des personnes et de leur vie privée[39]. Je dois me rapprocher du correspondant Informatique et libertés (CIL) ou, à défaut, des services juridiques de mon établissement lorsqu'une demande d'accès est formulée pour des documents ou jeux de données contenant des données à caractère personnel ou si je souhaite diffuser de telles informations.

À condition d'avoir respecté les « formalités CNIL », les documents comprenant des données à caractère personnel ne sont communicables et réutilisables que **si au moins l'une des trois conditions suivantes est remplie :**

1. le consentement des personnes concernées a été recueilli après leur bonne information sur la finalité et les modalités de la communication ou de la réutilisation des données les concernant ;
2. les données sont anonymisées[40] (c'est-à-dire non identifiantes ou ne permettant pas, compte tenu de leur niveau d'agrégation, d'identifier à nouveau les personnes concernées) ;
3. la réutilisation est autorisée par un texte législatif ou réglementaire.

36 Article L321-3 du Code des relations entre le public et l'administration.
37 Article 2 de la loi n° 78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés.
38 Commission nationale Informatique et Libertés (CNIL).
39 Article 9 du Code civil.
40 Le processus d'anonymisation est en principe à la charge de l'administration qui détient les données demandées. D'après l'article 40 du décret n°2005-1755 du 30 décembre 2005, l'administration peut refuser d'anonymiser les données si cette opération entraîne pour elle des « efforts disproportionnés ».

Lorsqu'une demande d'accès ou de réutilisation est formulée pour des données à caractère personnel communicables, la communication doit être encadrée par un contrat ou une licence adaptée aux risques de mauvaises réutilisations des données. La licence ou le contrat utilisé doit notamment garantir l'intégrité, la sécurité et la confidentialité des données transmises. Le demandeur doit en outre s'engager à ne pas tenter d'identifier à nouveau les personnes concernées par les données ayant été anonymisées.

Outre l'application des textes sur la protection des données personnelles, la diffusion des données peut également être limitée par le cadre règlementaire national.

Exemple : dans le cadre d'un projet de recherche sur les caractéristiques génétiques d'un groupe humain aux fins d'étudier le flux migratoire d'une population au Moyen Âge, je ne peux diffuser que les résultats globaux et totalement anonymisés. En effet, les dispositions du Code civil qui visent à protéger l'intégrité de la personne humaine interdisent la communication des résultats individuels[41]. La diffusion de ces données est sanctionnée pénalement[42].

1.3.5 Données statistiques

Ces données proviennent des statistiques publiques qui regroupent l'ensemble des productions issues[43] :

- des enquêtes statistiques dont la liste est arrêtée chaque année par un arrêté du ministre chargé de l'économie ;
- de l'exploitation, à des fins d'information générale, de données collectées par des administrations, des organismes publics ou des organismes privés chargés d'une mission de service public.

Ces données peuvent contenir des données personnelles et des données sur les entreprises relevant du secret des affaires (parts de marché, etc.).

Les données statistiques peuvent être diffusées à condition de respecter le « secret statistique » ou en obtenant une dérogation de la part du Comité du secret statistique.

Pour en savoir plus :
Fiche 4. Données statistiques

1.3.6 Données liées à un contrat avec un tiers non soumis à une obligation de service public

Les règles exposées dans ce guide concernant la diffusion et la réutilisation des documents administratifs sont des règles d'ordre public qui prévalent sur les conditions que les parties peuvent fixer dans un contrat.

Pour rappel, et concernant les données acquises par l'établissement lors de relations avec des tiers privés, l'Administration n'est pas tenue :

- de communiquer ou de diffuser les documents réalisés en exécution d'un contrat de prestation de services exécuté pour le compte d'une ou plusieurs personnes déterminées ; attention : l'expression « prestation de service » est entendue au sens large, il s'agit de tous les contrats par lesquels l'établissement réalise des travaux (analyse, expérimentation, recherche,

L'exception relative aux prestations de services est peu documentée. Il n'existe pas d'avis de la CADA concernant cette exception appliquée aux établissements de recherche et d'enseignement supérieur.

41 Articles 16-10 et suivants du Code civil .
42 Articles 226-25 et suivants du Code pénal.
43 Source : INSEE.

etc.) pour le compte d'un tiers. Cette exception est atténuée par le fait que, depuis la loi pour une République numérique du 7 octobre 2016[44], il est prévu que la réutilisation de données issues d'une activité de recherche **qui ont fait l'objet d'un écrit scientifique** est libre si :

- ces données sont issues d'une activité de recherche financée au moins pour moitié par des fonds publics ;
- ces données ne sont pas protégées par un droit spécifique ;
- ces données ont été rendues publiques par le chercheur ou l'établissement ;

- de communiquer ou de diffuser des données protégées par le secret des affaires (ou tout autre secret protégé par la loi) ;
- de conférer des droits de réutilisation à des données sur lesquelles existent des droits de propriété intellectuelle de tiers.

Pour les données échangées ou recueillies lors d'un partenariat public-privé, il est conseillé d'établir un contrat détaillant l'usage que l'on peut faire des données.

Notamment, il est recommandé :

- de 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 ;
- *a minima* de prévoir une exception à la confidentialité des informations échangées dans le cadre du partenariat ;
- de formuler un plan de gestion des données, qui abordera la problématique de la diffusion et de la réutilisation des données.

1.4 Conclusion

L'ensemble des développements de ce point 1 sont reportés sur la carte heuristique « Logigramme de la communicabilité d'une donnée » que vous trouverez en p. 28.

À titre d'exemple, le contrat UnicANR, utilisé dans le cadre de partenariats financés par l'ANR, contient la clause suivante (qui préserve les facultés d'Open Data des partenaires publics) : « Dans le cas où la communication d'INFORMATIONS CONFIDENTIELLES est imposée par l'application d'une disposition légale ou réglementaire ou dans le cadre d'une procédure judicaire, administrative ou arbitrale, cette communication doit être limitée au strict nécessaire. La PARTIE RÉCIPIENDAIRE s'engage à informer immédiatement et préalablement à toute communication la PARTIE ÉMETTRICE afin de permettre à cette dernière de prendre les mesures appropriées à l'effet de préserver leur caractère confidentiel. » La clause peut également indiquer que ne sont pas confidentielles les informations que la partie publique est légalement tenue de communiquer.

44 Article L 533-4 II Code de la Recherche

2. Comment diffuser les données ?

2.1. Les grands principes

2.1.1 Aspects techniques de l'Open Data

Les préconisations élaborées en 2007 par le groupe de travail Open Government Data définissent les bonnes pratiques relatives aux données mises en Open data. Ces données doivent être :

- complètes et de qualité : données brutes accompagnées de leurs métadonnées ;
- accessibles et exploitables directement : format numérique non propriétaire, open source de préférence.

Ainsi, pour être considérées « ouvertes » au sens du mouvement « Open Data », les données publiques doivent être, dans la mesure du possible et sous réserve de certains choix laissés à l'appréciation des établissements publics :

- **complètes :** toutes les données sont mises à disposition, à l'exception des données qui sont sujettes à des limitations. *Exemple : données personnelles ;*
- **primaires :** les données sont ouvertes telles que collectées à la source, avec le moins de modifications et la plus grande granularité possible ;
- **opportunes :** les données sont mises à disposition aussi rapidement que nécessaire pour préserver leur valeur ;
- **accessibles :** les données sont accessibles au plus grand éventail d'utilisateurs possible et pour des usages aussi divers que possibles ;
- **exploitables :** les données sont exploitables par ordinateur ou lisibles par des machines. Les données sont structurées pour permettre le traitement automatisé ;
- **non discriminatoires :** les données sont non discriminatoires, c'est-à-dire accessibles à tous, sans aucune obligation préalable ni inscription et en principe sans coût ;
- **non propriétaires :** les données sont accessibles dans un format sur lequel aucune entité ne dispose d'un contrôle exclusif. *Exemple : format de fichier texte non propriétaire permettant une réutilisation facile et ne nécessitant pas l'achat d'une licence ;*
- **libres de droits :** les données sont libres de droit. Elles ne sont pas soumises au droit d'auteur, au droit des marques ou au secret commercial. Des règles raisonnables de confidentialité, de sécurité et de priorité d'accès peuvent être admises ;
- **permanentes :** les données doivent être rendues accessibles en ligne avec un système d'archivage, de suivi des modifications et de mention de la dernière mise à jour.

La réglementation fait écho à certaines de ces recommandations en obligeant à la diffusion de documents achevés et en recommandant que la diffusion électronique soit effectuée dans un standard ouvert et librement réutilisable.

Par ailleurs, la réglementation[45] oblige les administrations à diffuser :

- les bases de données qu'elles produisent ou qu'elles reçoivent et qui ne font pas l'objet d'une diffusion publique par ailleurs ;
- les données dont la publication présente un intérêt économique, social, sanitaire ou environnemental.

Bien entendu, l'application des recommandations purement techniques est laissée à l'appréciation de l'établissement.

2.1.2 Aspects juridiques : la licence de diffusion

Si le contrôle de la nature des données a démontré la possibilité de leur diffusion en Open Data, il reste un dernier point à préparer : le choix d'une licence.

Une licence de diffusion des données protège le fournisseur des données. En effet, elle permet de limiter la responsabilité du fournisseur des données lors de la réutilisation (clause de limitation ou d'exclusion de responsabilité).

La loi pour une République numérique du 7 octobre 2016 indique que la licence doit être choisie sur une liste de licences fixée par décret[46] ou selon une licence homologuée par l'État (à la demande de l'établissement).

Deux grands types de licences sont proposés :

- licences permissives (pas ou peu de contraintes pour le réutilisateur) : Licence ouverte et, pour les logiciels : BSD, MIT, Apache, CeCILL-B ;
- licences copyleft (le réutilisateur est contraint de partager tout ou partie de son travail sur les informations publiques lorsqu'il rediffuse) : ODbL et, pour les logiciels : MPL, GPL, CeCILL. Cependant, pour utiliser ces licences copyleft, il faut satisfaire aux exigences de l'article L323-2 du Code des relations entre le public et l'Administration. Ces exigences sont les suivantes : les restrictions à la réutilisation doivent être proportionnées, dictées par des motifs d'intérêt général et ne peuvent avoir pour objet ou pour effet de restreindre la concurrence.

Que l'établissement choisisse ou non une licence, la loi indique que tout « réutilisateur » doit (sauf si l'établissement y renonce) [47] :

- **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.**

Ainsi, les réutilisations commerciales ne peuvent pas être empêchées.

45 Article L312-1-1 du Code des relations entre le public et l'Administration.
46 Article D323-2-1 du Code des relations entre le public et l'Administration.
47 Article L322-1 du Code des relations entre le public et l'Administration.

Licence gratuite ou payante ?

Depuis la loi dite Valter du 28 décembre 2015[48], c'est le principe de gratuité de la réutilisation qui est consacré.

L'accès payant aux données n'est désormais possible que pour les établissements qui sont tenus de couvrir par des recettes propres une part substantielle des coûts liés à l'accomplissement de leurs missions de service public[49].

Exemple : Météo France, INSEE[50]

Pour savoir si votre établissement est concerné, vous pouvez contacter le service juridique.

2.2 Les différentes modalités de diffusion

L'ouverture des données peut être volontaire ou sur demande d'un tiers, c'est pourquoi il est proposé d'aborder les modalités d'ouverture suivantes : prise en compte de pratiques par discipline *(2.2.1)* et existence d'une politique d'établissement *(2.2.2)*. Dans tous les cas, l'établissement doit mettre en place un répertoire recensant les données publiques à disposition du citoyen *(2.2.3)*. Enfin, nous parlerons des réflexes à avoir lors d'une demande d'accès formulée par un tiers *(2.2.4)*.

2.2.1 Les pratiques par discipline

Certaines disciplines scientifiques proposent des entrepôts reconnus permettant l'hébergement des jeux de données (ex : NCBI, EBI pour la bioinformatique) pour leur mise en accès libre, notamment à la communauté scientifique. Ces entrepôts disciplinaires peuvent tout à fait constituer une façon de rendre accessible les données.

Toutefois, l'obligation d'Open Data faite aux établissements publics par la réglementation prescrit d'ouvrir les données à **tout public** (citoyen, entreprise, organisme d'origine française ou étrangère). Il faut donc vérifier que l'entrepôt de données ne restreint pas l'accès aux seuls scientifiques.

Si c'est le cas, il convient d'envisager un dépôt complémentaire dans un entrepôt totalement ouvert (cf. ci-dessous).

> **Attention**
> ***Je ne dois pas choisir d'entrepôt de données par moi-même sans me conformer à la politique de mon établissement ou sans m'être rapproché de ma hiérarchie. En effet, certains entrepôts ou hébergeurs peuvent avoir des conditions d'utilisation de nature à confisquer le droit d'exploitation ultérieure des données ou imposer des restrictions de réutilisation des données qui seraient incompatibles avec le droit français.***

2.2.2 La politique d'établissement

Afin de répondre aux enjeux de l' Open et du Big Data, un établissement peut se doter d'une offre de services en direction des équipes de recherche pour les aider à rentrer dans le mouvement de l'Open Data. À l'occasion de cette démarche de création de services en faveur du partage des données, les établissements peuvent mettre en place :

- un annuaire des sources de données ;
- un entrepôt de données : également appelé « data repository », il s'agit d'un réservoir constitué majoritairement de données de recherche, brutes ou élaborées, qui sont organisées de façon à pouvoir être retrouvées.

> ***Les EPIC n'ont pas l'obligation légale d'autoriser la réutilisation des données qui ne rentrent pas dans le cadre de leur mission de service public.***

48 Loi 2015-1779.
49 Décret 2016-1036 du 28/07/2016.
50 Voir notamment le décret 2016-17 sur le paiement autorisé pour l'accès à certaines bases.

Ce type d'entrepôt devrait être en mesure d'héberger des données provenant de divers projets ou de fédérer des systèmes d'informations, projets ou plateformes[51]. Attention ! La mise à disposition en Open Data peut être un enjeu de sécurité pour l'établissement en fonction des cloisonnements sur son réseau informatique. En cas de doute, contacter votre responsable de sécurité des systèmes d'information (RSSI) ou correspondant SSI.

À défaut d'entrepôt d'établissement disponible, je peux ouvrir les données en les mettant à disposition dans un entrepôt de référence de données conformément à la politique de mon établissement.

Je dois me renseigner auprès de mon établissement pour connaître le répertoire des informations publiques (Exemple : à l'INRA, il s'agit de ProdINRA).

2.2.3 Le répertoire des données publiques

La réglementation oblige chaque établissement à disposer d'un répertoire des informations publiques (= documents publics réutilisables) mis à jour annuellement. Tout dépôt de données en Open Data doit donc être mentionné dans ce répertoire.

Les conditions de réutilisation des données publiques ainsi répertoriées doivent également être rendues publiques.

2.2.4 La demande d'accès

En cas de demande d'accès à certaines données créées ou collectées par l'établissement, l'Administration dispose d'un délai d'un mois pour y répondre[52]. L'absence de réponse dans ce délai par l'administration vaut décision implicite de rejet.

Dès lors, le demandeur peut exercer un recours devant la Commission d'accès aux documents administratifs (CADA). La CADA rend un avis (qui sera publié sur son site) qui ne lie pas l'établissement.

Si l'établissement décide de ne pas se conformer à l'avis rendu par la CADA, le demandeur peut alors saisir le juge administratif pour obtenir la communication du document sollicité.

Si je suis confronté à une demande d'ouverture de données, je dois donc contacter immédiatement le service juridique et/ou la personne responsable de l'accès aux documents administratifs (PRADA) désignée dans mon établissement[53] de mon établissement qui m'aidera à formaliser la réponse.

Sous réserve de la communicabilité des données (cf. 1), je peux me contenter de fournir au tiers demandeur les informations en l'état. Je n'ai pas l'obligation de les retraiter ou encore de modifier le format dans lequel les données sont disponibles. Cependant, la loi précise que les données sont mises à disposition dans un format ouvert et aisément réutilisable[54]. Aussi,

51 Open Science, Groupe de travail INRA données expérimentales.

52 Article R311-13 du Code des relations entre le public et l'Administration.

53 PRADA/ voir liste pour les organismes : https://www.data.gouv.fr/fr/datasets/liste-des-prada-referencees-sur-le-site-de-la-cada-juillet-2017/

54 Article 10 de la loi n°78-753 dite « CADA » modifié par la loi n°2015-1779 du 28 décembre 2015.

dans la mesure du possible, les modalités de mise à disposition et les formats doivent favoriser l'exploitation des informations par les utilisateurs.[55]

Si des données communicables sont mêlées à des données non communicables, telles que des données personnelles ou couvertes par les secrets visés au 1.2.3, une analyse fine de la nature et du contenu de ces données doit être faite.

Sous réserve qu'il soit possible de disjoindre les données communicables de celles non communicables ou d'occulter les données non communicables, la communication pourra se faire uniquement **après cette occultation ou disjonction[56].**

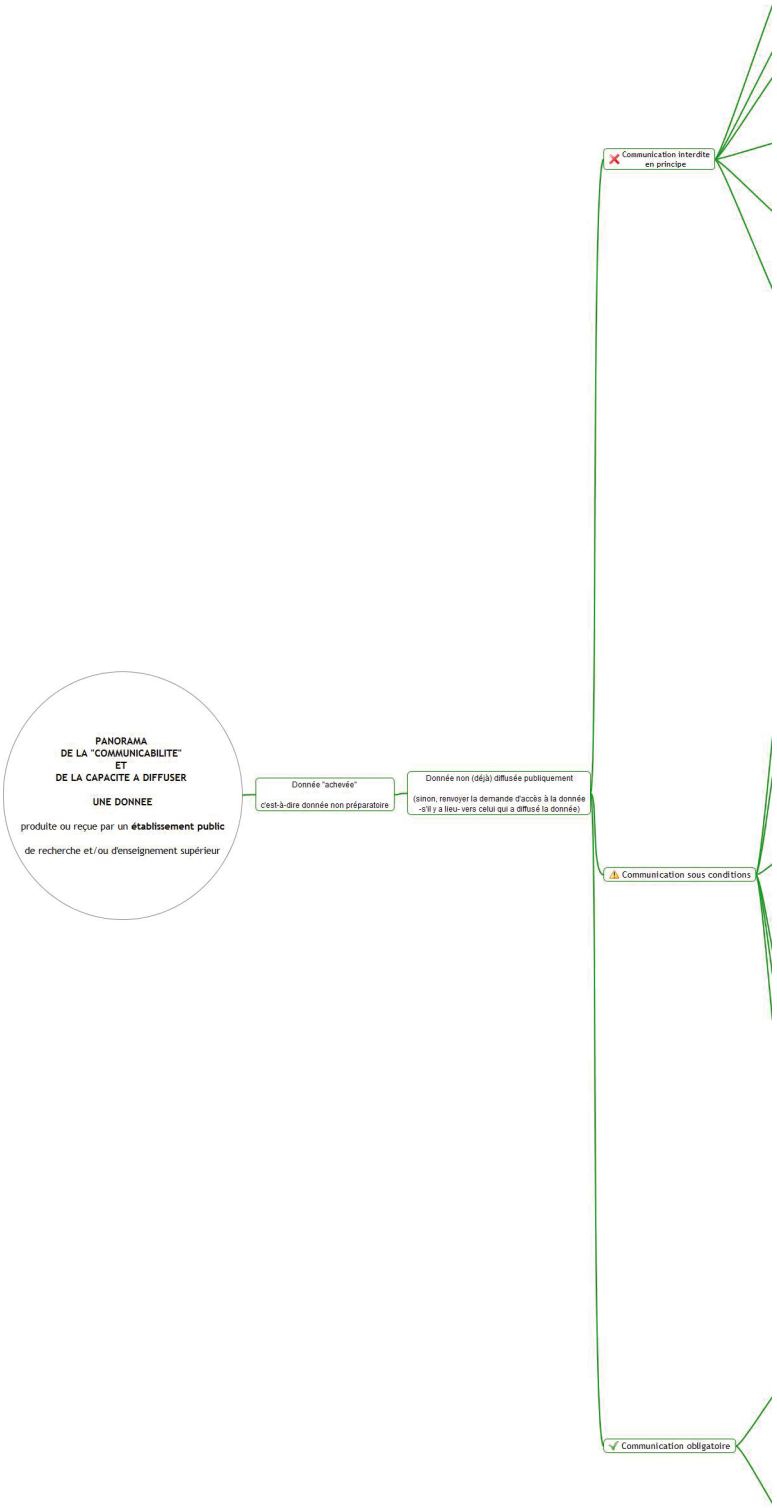
Aussi, il convient de saisir le service juridique de mon établissement qui m'aidera à faire une réponse qui se doit de respecter un certain formalisme et d'être circonstanciée.

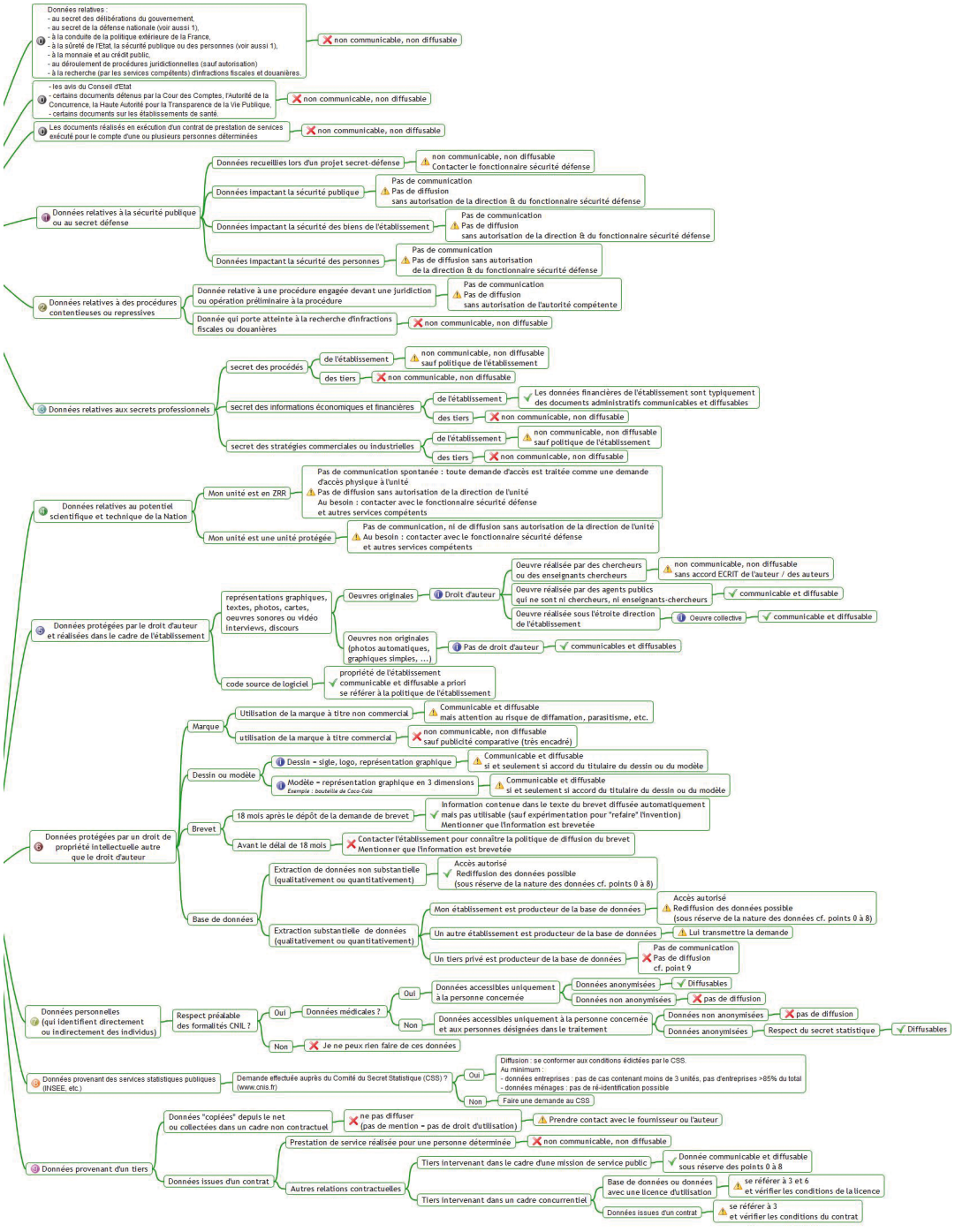
V2 - Décembre 2017

55 Pour en savoir plus sur le régime juridique applicable en cas de détention ou codétention d'une information dans l'exercice d'une mission de service public administratif (SPA) et/ou d'une mission de service public industriel et commercial (SPIC) ; consultez le « Cahier pratique : Ressources de l'immatériel » de l'APIE, page 8.

56 Article L 311-7 du Code des relations entre le public et l'Administration.

Logigramme de communicabilité d'une donnée




Données relatives :
- au secret des délibérations du gouvernement,
- au secret de la défense nationale (voir aussi 1),
- à la conduite de la politique extérieure de la France,
- à la sûreté de l'Etat, la sécurité publique ou des personnes (voir aussi 1),
- à la monnaie et au crédit public,
- au déroulement de procédures juridictionnelles (sauf autorisation)
- à la recherche (par les services compétents) d'infractions fiscales et douanières.
non communicable, non diffusable
- les avis du Conseil d'Etat
- certains documents détenus par la Cour des Comptes, l'Autorité de la Concurrence, la Haute Autorité pour la Transparence de la Vie Publique,
- certains documents sur les établissements de santé.
non communicable, non diffusable
Les documents réalisés en exécution d'un contrat de prestation de services exécuté pour le compte d'une ou plusieurs personnes déterminées
non communicable, non diffusable
Données relatives à la sécurité publique ou au secret défense
Données recueillies lors d'un projet secret-défense
non communicable, non diffusable
Contacter le fonctionnaire sécurité défense
Données impactant la sécurité publique
Pas de communication
Pas de diffusion
sans autorisation de la direction & du fonctionnaire sécurité défense
Données impactant la sécurité des biens de l'établissement
Pas de communication
Pas de diffusion
sans autorisation de la direction & du fonctionnaire sécurité défense
Données impactant la sécurité des personnes
Pas de communication
Pas de diffusion sans autorisation
de la direction & du fonctionnaire sécurité défense
Données relatives à des procédures contentieuses ou répressives
Donnée relative à une procédure engagée devant une juridiction ou opération préliminaire à la procédure
Pas de communication
Pas de diffusion
sans autorisation de l'autorité compétente
Donnée qui porte atteinte à la recherche d'infractions fiscales ou douanières
non communicable, non diffusable
Données relatives aux secrets professionnels
secret des procédés
de l'établissement
non communicable, non diffusable
sauf politique de l'établissement
des tiers
non communicable, non diffusable
secret des informations économiques et financières
de l'établissement
Les données financières de l'établissement sont typiquement des documents administratifs communicables et diffusables
des tiers
non communicable, non diffusable
secret des stratégies commerciales ou industrielles
de l'établissement
non communicable, non diffusable
sauf politique de l'établissement
des tiers
non communicable, non diffusable
Données relatives au potentiel scientifique et technique de la Nation
Mon unité est en ZRR
Pas de communication spontanée : toute demande d'accès est traitée comme une demande d'accès physique à l'unité
Pas de diffusion sans autorisation de la direction de l'unité
Au besoin : contacter avec le fonctionnaire sécurité défense
et autres services compétents
Mon unité est une unité protégée
Pas de communication, ni de diffusion sans autorisation de la direction de l'unité
Au besoin : contacter avec le fonctionnaire sécurité défense
et autres services compétents
Données protégées par le droit d'auteur et réalisées dans le cadre de l'établissement
représentations graphiques, textes, photos, cartes, oeuvres sonores ou vidéo interviews, discours
Oeuvres originales
Droit d'auteur
Oeuvre réalisée par des chercheurs ou des enseignants chercheurs
non communicable, non diffusable
sans accord ECRIT de l'auteur / des auteurs
Oeuvre réalisée par des agents publics qui ne sont ni chercheurs, ni enseignants-chercheurs
communicable et diffusable
Oeuvre réalisée sous l'étroite direction de l'établissement
Oeuvre collective
communicable et diffusable
Oeuvres non originales (photos automatiques, graphiques simples, ...)
Pas de droit d'auteur
communicables et diffusables
code source de logiciel
propriété de l'établissement
communicable et diffusable a priori
se référer à la politique de l'établissement
Données protégées par un droit de propriété intellectuelle autre que le droit d'auteur
Marque
Utilisation de la marque à titre non commercial
Communicable et diffusable
mais attention au risque de diffamation, parasitisme, etc.
utilisation de la marque à titre commercial
non communicable, non diffusable
sauf publicité comparative (très encadré)
Dessin ou modèle
Dessin = sigle, logo, représentation graphique
Communicable et diffusable
si et seulement si accord du titulaire du dessin ou du modèle
Modèle = représentation graphique en 3 dimensions
Exemple : bouteille de Coca-Cola
Communicable et diffusable
si et seulement si accord du titulaire du dessin ou du modèle
Brevet
18 mois après le dépôt de la demande de brevet
Information contenue dans le texte du brevet diffusée automatiquement
mais pas utilisable (sauf expérimentation pour "refaire" l'invention)
Mentionner que l'information est brevetée
Avant le délai de 18 mois
Contacter l'établissement pour connaître la politique de diffusion du brevet
Mentionner que l'information est brevetée
Base de données
Extraction de données non substantielle (qualitativement ou quantitativement)
Accès autorisé
Rediffusion des données possible
(sous réserve de la nature des données cf. points 0 à 8)
Extraction substantielle de données (qualitativement ou quantitativement)
Mon établissement est producteur de la base de données
Accès autorisé
Rediffusion des données possible
(sous réserve de la nature des données cf. points 0 à 8)
Un autre établissement est producteur de la base de données
Lui transmettre la demande
Un tiers privé est producteur de la base de données
Pas de communication
Pas de diffusion
cf. point 9
Données personnelles (qui identifient directement ou indirectement des individus)
Respect préalable des formalités CNIL ?
Oui
Données médicales ?
Oui
Données accessibles uniquement à la personne concernée
Données anonymisées
Diffusables
Données non anonymisées
pas de diffusion
Non
Données accessibles uniquement à la personne concernée et aux personnes désignées dans le traitement
Données non anonymisées
pas de diffusion
Données anonymisées
Respect du secret statistique
Diffusables
Non
Je ne peux rien faire de ces données
Données provenant des services statistiques publiques (INSEE, etc.)
Demande effectuée auprès du Comité du Secret Statistique (CSS) ? (www.cnis.fr)
Oui
Diffusion : se conformer aux conditions édictées par le CSS.
Au minimum :
- données entreprises : pas de cas contenant moins de 3 unités, pas d'entreprises >85% du total
- données ménages : pas de ré-identification possible
Non
Faire une demande au CSS
Données provenant d'un tiers
Données "coplées" depuis le net ou collectées dans un cadre non contractuel
ne pas diffuser
(pas de mention = pas de droit d'utilisation)
Prendre contact avec le fournisseur ou l'auteur
Données issues d'un contrat
Prestation de service réalisée pour une personne déterminée
non communicable, non diffusable
Autres relations contractuelles
Tiers intervenant dans le cadre d'une mission de service public
Donnée communicable et diffusable
sous réserve des points 0 à 8
Tiers intervenant dans un cadre concurrentiel
Base de données ou données avec une licence d'utilisation
se référer à 3 et 6
et vérifier les conditions de la licence
Données issues d'un contrat
se référer à 3
et vérifier les conditions du contrat
Données géographiques
Données géographiques
sauf données relevant des points 0 à 9
Disponibles sous format électronique
Obligation de communication et de diffusion
Disponibles sous format papier
Données communicables
Pas d'obligation de diffusion => cf. données environnementales
Données environnementales
Données relatives à des émissions de substances dans l'environnement
Données relevant du 0 ou du 1
pas communicable, pas diffusable
Données portant atteinte à des droits de propriété intellectuelle
pas communicable, pas diffusable
Autres données
communication obligatoire à tout citoyen demandeur
sauf si la communication et/ou la réutilisation porte atteinte :
- à la protection de l'environnement
- aux intérêts de la personne ayant fourni les données
- au secret statistique
Autres données environnementales
- les traités, conventions et accords internationaux, ainsi que la législation communautaire, nationale, régionale et locale concernant l'environnement ou s'y rapportant ;
- les plans et programmes et les documents définissant les politiques publiques qui ont trait à l'environnement ;
- les rapports établis par les autorités publiques ou pour leur compte relatifs à l'état d'avancement de la mise en œuvre des textes quand ces rapports sont élaborés ou conservés sous forme électronique par les autorités publiques ;
- les rapports établis par les autorités publiques sur l'état de l'environnement ;
- les données ou résumés des données recueillies par les autorités publiques dans le cadre du suivi des activités ayant ou susceptibles d'avoir des incidences sur l'environnement ;
- les autorisations qui ont un impact significatif sur l'environnement ainsi que les accords environnementaux ;
- les études d'impact environnemental et les évaluations de risques concernant les éléments de l'environnement.
Communication et diffusion obligatoires
Autres données environnementales
Communication obligatoire sauf données relevant des 0, 1, 2, 4, 7, 8, 9
Diffusion possible

Annexes

Fiche 1. Les bases de données en bref 31
Fiche 2. Les données personnelles 33
Fiche 3. Les données statistiques 37
Fiche 4. Convention d'Aarhus sur l'information en matière d'environnement 38
Fiche 5. Tableau comparatif des licences gratuites ODBL et Etalabab 40
Fiche 6. Mise en place d'une licence par la voie du contrat électronique 42
Fiche 7. Archives 43

Fiche 1. Les bases de données en bref

La protection des bases de données relève de deux droits différents : le droit d'auteur et le droit *sui generis* du producteur de base de données. Ce dernier a été instauré par la directive 96/9/CE du Parlement européen et du Conseil du 11 mars 1996, transposée par la loi n° 98-536 du 1er juillet 1998. Le Code de la propriété intellectuelle définit depuis lors la notion de base de données comme un « recueil d'œuvres, de données ou d'autres éléments indépendants, disposés de manière systématique ou méthodique, et individuellement accessibles par des moyens électroniques ou par tout autre moyen »[57].

Le droit d'auteur portant sur une base de données naît dès lors que sa structure est originale, c'est-à-dire que la disposition des éléments qu'elle inclut relève d'un choix qui « reflète l'empreinte de la personnalité de l'auteur ». Ainsi, un annuaire listant des personnes par ordre alphabétique ne caractérise pas une structure originale.

Devant les faiblesses du droit d'auteur, qui ne permet pas de reconnaitre de protection aux bases de données ayant fait l'objet d'investissements lourds mais dont la structure n'est pas originale, le législateur a fondé le droit *sui generis*.

L'objectif de cette législation est d'encourager le développement des bases de données en instaurant un droit spécifique ; le droit *sui generis* des bases de données, destiné à protéger les investissements financiers, matériels et humains consentis par les producteurs de bases de données. Ce droit protège avant tout le contenu informationnel de la base, indépendamment de la protection par le droit d'auteur. L'investissement doit néanmoins être substantiel, que ce soit quantitativement ou qualitativement.

La directive 96/9/CE cherche en effet à protéger l'investissement substantiel nécessaire à l'obtention, la vérification ou la présentation des données de la base.

Obtention : « moyens consacrés à la recherche d'éléments existants et à leur rassemblement dans ladite base. Elle ne comprend pas les moyens mis en œuvre pour la création des éléments constitutifs du contenu d'une base de données. »[1]

Vérification : « moyens consacrés, en vue d'assurer la fiabilité de l'information contenue dans ladite base, au contrôle de l'exactitude des éléments recherchés, lors de la constitution de cette base ainsi que pendant la période de fonctionnement de celle-ci. Des moyens consacrés à des opérations de vérification au cours de la phase de création d'éléments par la suite rassemblés dans une base de données ne relèvent pas de cette notion. »[1]

Présentation : « moyens visant à conférer à ladite base sa fonction de traitement de l'information, à savoir ceux consacrés à la disposition systématique ou méthodique des éléments contenus dans cette base ainsi qu'à l'organisation de leur accessibilité individuelle. »[2]

1 CJCE, 9 nov. 2004, The British Horseracing Board Ltd e/a.
2 CJCE, 9 nov. 2004, Fixtures Marketing Ltd c/ Organismos Pronostikon.

Si les données dont je dispose sont structurées dans une base de données, je dois identifier le producteur de cette base pour pouvoir procéder à l'ouverture des données. En effet, le producteur de la base de données est le seul à pouvoir autoriser l'ouverture. Pour l'identifier, je dois me poser une série de questions :

- Qui a pris l'initiative du projet ?

Que ce soit entre équipes ou au sein d'une équipe de recherche, il faudra recenser l'ensemble des employeurs et acteurs pour déterminer les participations (matérielles, intellectuelles ou financières) de chacun. En cas de contrat de consortium, de recherche ou de collaboration : prendre connaissance des termes du contrat relatif au projet, et prêter une attention particulière aux clauses de propriété intellectuelle.

- Qui a réalisé l'architecture (le modèle relationnel) de la base de données ?

Contrat avec une société privée (ESN ou SS2i) : contrôler la clause de propriété intellectuelle pour savoir si le prestataire reste propriétaire du modèle relationnel.

Typologie de l'architecture : est-elle d'une conception classique pour une base de données, ou est-elle plutôt originale par rapport à ce qui existe ? Si oui, l'auteur de la structure a-t-il cédé les droits nécessaires à son utilisation ?

57 Article L112-3 du Code de la propriété intellectuelle (CPI).

Base de données composite : si la base de données est composite et a été constituée par l'agrégation, même partielle, d'autres bases de données, il faudra respecter les licences d'utilisation des bases de données sur laquelle elle s'appuie et veiller à détenir les cessions de droits nécessaires à l'utilisation des modèles relationnels préexistants.

Note importante

Si l'article L111-1 du Code de la propriété intellectuelle reconnait la qualité d'auteur au fonctionnaire, le droit d'exploitation d'une œuvre créée par un agent de l'État est le plus souvent « cédé de plein droit à l'État » (article L131-3-1 du même code). Ce n'est en revanche pas le cas, d'après l'alinéa 4 de l'article L111-1, pour les « agents auteurs d'œuvre dont la divulgation n'est soumise, en vertu de leur statut ou des règles qui régissent leurs fonctions, à aucun contrôle préalable de l'autorité hiérarchique », ce qui vise notamment les maîtres de conférences, les professeurs d'universités, les chargés de recherches ou encore les directeurs de recherches. Ces agents ont pu être qualifiés, à l'occasion de débats parlementaires, d'« agents qui disposent dans leurs fonctions d'une grande autonomie intellectuelle, voire une indépendance de jugement, même si celle-ci s'inscrit dans une hiérarchie. »

- Et donc, qui est le producteur de la base de données ?

Une fois l'ensemble des acteurs déterminés[58], la solution la plus simple pour savoir qui aura des droits et à qui demander l'autorisation d'ouvrir la base de données serait d'attribuer, dès le début de la collaboration, soit :

- une part égale à chaque acteur : tous seront alors dans un régime proche de l'indivision, c'est-à-dire avec nécessité de concertation et unanimité pour prendre la plupart des décisions, sauf à avoir organisé l'exploitation de la base selon des règles différentes à la manière d'un règlement de copropriété ;

- une part différente pour chacun : sur la base des investissements de chacun.

Mais des négociations peuvent intervenir *a posteriori* entre les acteurs et les débats auront lieu sur le terrain de la preuve des investissements de chacun.

Dans tous les cas, il est très important de conserver des preuves de la nature et de l'importance de l'investissement réalisé en concevant un dossier bien documenté de tout élément ayant servi à l'investissement de la base de données :

- investissement financier : coût de la base de données (factures, contrats, etc.) ;

- investissement humain : le nombre de personnes chargées de la création et/ou de la mise à jour de la base de données (contrats de travail) ;

- investissement matériel : équipements nécessaires au développement de la base de données (factures).

Attention : il ne faut prendre en compte que les investissements portant sur l'obtention, la vérification ou la présentation de la base, et non pas ceux consentis pour la création des données. *Exemple : les investissements consentis pour mieux organiser la base, pour la rendre plus intelligible, sont pris en compte.*

Le producteur de la base de données a le droit d'interdire :

1. « L'extraction, par transfert permanent ou temporaire de la totalité ou d'une partie qualitativement ou quantitativement substantielle du contenu d'une base de données sur un autre support, par tout moyen et sous toute forme que ce soit ;
2. la réutilisation, par la mise à la disposition du public de la totalité ou d'une partie qualitativement ou quantitativement substantielle du contenu de la base, quelle qu'en soit la forme. »

Depuis la loi pour une République numérique du 7 octobre 2016, il est prévu que le droit des administrations sur leurs bases de données ne peut faire obstacle à la réutilisation du contenu des bases de données que ces administrations publient[59].

58 Rappel : uniquement les acteurs ayant eu un rôle d'investissement pour l'obtention, la collecte, la vérification ou la présentation de la base de données.

59 Article L321-3 du Code des relations entre le public et l'administration.

Attention :

Il existe un arrêt de cour d'appel qui a jugé que le producteur de la base de données ne peut exercer ces droits d'interdiction qu'à condition de les avoir mentionnés lors de la mise à disposition de la base . Dès lors, il faudrait mentionner de façon claire et évidente à côté de chaque base de données publiée, que le droit du producteur de bases de données est expressément réservé.

Exemple : licence interdisant l'extraction et la réutilisation sans autorisation expresse préalable.

Cet arrêt a fait l'objet d'un pourvoi en cassation qui a confirmé cette interprétation, mais qui reste critiqué. En effet, on peut douter de la portée d'une telle décision : celle-ci traitait du volet pénal de la protection du droit du producteur, qui implique d'opérer une interprétation stricte de la loi. Par ailleurs, imposer la mention expresse d'une interdiction de la part du producteur pour mettre en œuvre son droit priverait la protection légale d'une grande partie de son fondement

Cette ouverture de la base de données par son producteur devra également s'accompagner de la cession des droits du ou des auteur(s) de la structure de la base de données. En effet, l'utilisation et la diffusion de l'architecture, et donc de la base de données, sont soumises à l'autorisation préalable de l'auteur qui conservera dans tous les cas des droits moraux de paternité et d'intégrité sur son œuvre. Le concepteur de l'architecture de la base peut être à la fois coproducteur de l'ensemble de la base et auteur unique de son architecture.

Fiche 2. Les données personnelles

1 - Définitions

D'après l'article 2 de la loi n°78-17 **« Informatique et libertés »** du 6 janvier 1978, « **constitue une donnée à caractère personnel toute information relative à une personne physique identifiée ou qui peut être identifiée, directement ou indirectement,** par référence à un numéro d'identification ou à un ou plusieurs éléments qui lui sont propres. Pour déterminer si une personne est identifiable, il convient de considérer l'ensemble des moyens en vue de permettre son identification dont dispose ou auxquels peut avoir accès le responsable du traitement ou toute autre personne. » Par ailleurs, « la personne concernée par un traitement de données à caractère personnel est celle à laquelle se rapportent les données qui font l'objet du traitement. »

> **Attention**
>
> Le règlement général européen sur la protection des données à caractère personnel sera applicable le 25 mai 2018. Un projet de loi nationale est en cours d'adoption. Cette fiche sera revue en conséquence.
>
> D'ici là, les règles applicables restent celles décrites ci-dessous.

Les données sont donc considérées « à caractère personnel » dès lors qu'elles concernent des personnes physiques :

- identifiées directement : lorsque par exemple son nom apparaît dans un fichier. ;
- identifiables indirectement : lorsqu'un fichier comporte des informations telles que l'adresse I.P., le numéro d'immatriculation, le numéro de téléphone, un numéro d'identification lié à un fichier où se trouvent les données à caractère personnel (dans ce cas : données dites codées ou pseudonymisées), une photographie, des éléments biométriques, etc.

Pour déterminer si une personne est identifiable *via* les données traitées, il faut donc analyser les risques en fonction du contexte et des moyens à disposition des utilisateurs leur permettant d'identifier cette personne. Par exemple, un croisement de données peut permettre une identification indirecte de la personne concernée : « Certaines données peuvent donc constituer des données à caractère personnel si elles permettent d'identifier indirectement ou par recoupement d'informations une personne précise. Il peut en effet s'agir d'informations qui ne sont pas associées au nom d'une personne mais qui permettent aisément de l'identifier et de connaître ses habitudes ou ses goûts. »[60]

Exemple : une date de naissance associée à une commune de résidence

Les principes de protection des données à caractère personnel s'appliquent à toute opération ou tout ensemble d'opérations (appelés « traitement ») portant sur de telles données, quel que soit le procédé utilisé, et notamment à la collecte, l'enregistrement, l'organisation, la conservation, l'adaptation ou la modification, l'extraction, la consultation, l'utilisation, la communication quelle que soit la forme de mise à disposition, le rapprochement ou interconnexion, ainsi que le verrouillage, l'effacement ou la destruction et quel que soit le support, papier ou informatique.

Un traitement ne peut être mis en place qu'après avoir accompli les formalités imposées par la loi « Informatique et libertés » et ce, sous réserve du respect de la confidentialité et de la sécurité des données collectées. Les données à caractère personnel ne sont pas, en tant que telles, éligibles à la mise à disposition en Open Data. Il est nécessaire d'avoir, au préalable, obtenu le consentement éclairé (sur la finalité et les modalités de la communication des données) de la personne concernée ou bien d'avoir anonymisé[61] les données.

60 Commission nationale de l'informatique et des libertés (CNIL).

61 Cf. la fiche n°16 du guide CNIL sur la sécurité des données personnelles.

2 - Principes

La loi « Informatique et libertés » définit les principes à respecter lors de la collecte, du traitement et de la conservation des données personnelles.

- Principe de finalité : les données à caractère personnel ne peuvent être recueillies et traitées que pour un usage légitime et déterminé correspondant aux missions de l'établissement.
- Principe de proportionnalité : seules doivent être enregistrées les informations pertinentes et nécessaires à l'égard de la finalité déclarée.
- Principe de durée limitée de conservation des données : les informations ne peuvent être conservées de façon indéfinie dans les fichiers informatiques. Une durée de conservation doit être établie en fonction de la finalité de chaque fichier. Par la suite, les données doivent être supprimées ou archivées sur un support distinct (car elles ne doivent plus être utilisées). Il y a toutefois quelques exceptions, notamment pour les données conservées « en vue d'être traitées à des fins historiques, statistiques ou scientifiques »[62].
- Principe de sécurité et de confidentialité : le responsable du traitement[63] est astreint à une obligation de sécurité. Il doit prendre les mesures nécessaires pour garantir la confidentialité, l'intégrité et la sécurité des données.
- Principe de transparence : le responsable du traitement doit informer les personnes des traitements auxquels leurs données sont soumises tout en leur accordant un droit d'accès, de modification, de rectification, voire d'opposition au traitement. Le responsable du traitement doit avertir ces personnes dès la collecte des données et en cas de transmission de ces données à des tiers.

Certaines données dites « sensibles » bénéficient d'un régime spécifique. Les données sensibles sont celles qui font apparaître, directement ou indirectement :

- **les origines raciales ou ethniques ;**
- **les opinions politiques, philosophiques ou religieuses ;**
- **l'appartenance syndicale des personnes ;**
- **des informations relatives à la santé ou à la vie sexuelle.**

Par principe, la collecte et le traitement de ces données sont interdits. Cependant, dans la mesure où la finalité du traitement l'exige, ne sont pas soumis à cette interdiction :

- les traitements pour lesquels la personne concernée a donné son consentement exprès ;
- les traitements justifiés par un intérêt public après autorisation de la CNIL ou adoption d'un décret en Conseil d'État ;
- etc.[64].

Le traitement de certaines autres données à risque doit respecter un formalisme particulier : données génétiques, données relatives aux infractions pénales, aux condamnations, données comportant des appréciations sur les difficultés sociales des personnes, données biométriques, données comprenant le NIR[65], etc.

3 - Formalités préalables au traitement initial

Tout fichier ou traitement informatisé comportant des données personnelles doit être déclaré au correspondant informatique et libertés (CIL) qui, selon le type de données ou de finalité du traitement, l'inscrit au registre des traitements de l'établissement ou instruit avec le responsable de traitement la demande d'autorisation auprès de la CNIL.

Pour les traitements de données à caractère personnel courantes (cas des données ne relevant pas des données sensibles ou d'autres données à risque comme les données génétiques, les infractions pénales, les données comportant des appréciations sur les difficultés sociales des personnes, les données biométriques, les données comprenant le numéro NIR), le CIL est habilité à donner l'autorisation de la mise en œuvre du traitement et à en superviser le déroulement dans le respect de la loi. Dans l'hypothèse d'un transfert de données hors de l'Union européenne (UE) et hors pays figurant sur la liste de protection adéquate, il est très important d'avertir au préalable le CIL ou, à défaut, d'effectuer une déclaration préalable auprès de la CNIL.

62 Article 36 alinéa 1 de la loi 78-17 du 6 janvier 1978 modifiée.

63 Le responsable du traitement de données à caractère personnel est la personne, l'autorité publique, le service ou l'organisme qui détermine ses finalités (à quoi il va servir) et ses moyens (selon quelles modalités).

64 Cf. article 8 Loi 1978-17.

65 Numéro d'inscription des personnes (NIR) : également appelé numéro INSEE ou numéro de sécurité sociale.

Pour en savoir plus sur les transferts hors UE :

Guide de la CNIL sur « les transfert de données à caractère personnel hors Union Européenne »

Pour les transferts de données vers les USA avec des opérateurs privés, il existe un accord spécifique entre l'Union européenne et les États-Unis intitulé le privacy shield.

Pour certaines catégories de traitements, soit en raison des typologies de données enregistrées (ex : données sensibles), soit en raison de leurs finalités spécifiques ou des risques qu'ils comportent, la loi a prévu des formalités particulières d'autorisation préalable délivrée par la CNIL[66]. Cette autorisation peut intervenir, dans certains cas, après consultation et avis d'un comité d'experts. *Exemple : traitements de données à caractère personnel à des fins de recherche en santé ou d'évaluation des pratiques de soins.*

Pour en savoir plus, rapprochez-vous du CIL de votre établissement.

Certains traitements mis en œuvre par des organismes publics doivent donc recueillir l'autorisation préalable de la CNIL. Cette procédure concerne les traitements mis en œuvre par des organismes publics ou des organismes privés gérant un service public et qui concernent :

- la recherche en santé (exception faite des recherches couvertes par une méthodologie de référence CNIL [MR001 / MR002 / MR003] et qui en respectent scrupuleusement les dispositions par la voie d'un engagement de conformité auprès de la CNIL) ;
- l'utilisation du NIR (numéro de sécurité sociale) ou la consultation du RNIPP (lorsque les organismes ne sont pas déjà habilités) ;
- l'utilisation de données biométriques (empreintes digitales, contour de la main, iris de l'œil, etc.) ;
- le recensement de la population ;
- les téléservices de l'administration électronique ;
- la sûreté, la défense ou la sécurité publique ;
- la prévention, la recherche, la constatation ou la poursuite d'infractions pénales ou l'exécution des condamnations pénales ou des mesures de sûreté ;

Quand les principes et formalités ont été respectés, la question de la mise en Open Data des données collectées peut être instruite (cf 2.1.5 du présent guide).

4 - Les principales évolutions liées au règlement général sur la protection des données (2016/679/UE)

Les principales modifications à venir et applicables dès le 25 mai 2018 sont :

- la suppression du principe de déclaration systématique auprès de la CNIL (sauf situation explicitement soumise à autorisation) ;
- la généralisation de la désignation d'un délégué à la protection des données personnelles ;
- renforcement du principe de transparence et du droit à l'information préalable et possibilité du retrait du consentement ;
- le principe d'*accountability* : être en mesure, à tout moment, de prouver le respect du règlement, notamment par une traçabilité et documentation du respect des 6 principes relatifs au traitement ;
- un renforcement de la protection de l'intégrité (sécurité) et la confidentialité des données personnelles ;
- la *data protection by design et by default* : prendre en compte dès la conception du traitement et des outils informatiques nécessaires à sa mise en œuvre les règles de protection des données telles que fixées par la règlementation de façon à en faire une composante essentielle de l'activité ;
- l'étude d'impacts des traitements pouvant présenter un risque élevé pour la vie privée ;
- renforcement des règles relatives à l'encadrement des transferts de données personnelles en dehors de l'UE ;
- notification des failles de sécurité dans les 72 h.

66 Articles 26 et 27 de la loi du 6 janvier 1978 modifiée.

Fiche 3. Les données statistiques

Le secret statistique est une forme particulière du secret professionnel qui s'applique aux statisticiens, chargés de recueillir et d'exploiter des statistiques publiques. Son principe général est d'apporter aux personnes qui fournissent des informations utilisées pour l'établissement de statistiques publiques l'assurance que ces informations ne seront pas utilisées d'une façon susceptible de leur porter atteinte[67].

En France, le secret statistique est garanti par la loi n° 51-711 du 7 juin 1951 sur l'obligation, la coordination et le secret en matière de statistiques. Au niveau européen, la confidentialité des informations statistiques est affirmée dans plusieurs documents[68]. Le secret statistique permet d'assurer :

- aux personnes physiques que la confidentialité sur leur vie personnelle et familiale sera garantie ;
- aux entreprises que le secret commercial sera respecté : les informations transmises ne seront pas mises à la disposition de leurs concurrents.

Les données des enquêtes statistiques sont détenues par l'INSEE et les services ministériels chargés des questions statistiques. Toutefois, les données des enquêtes publiques françaises peuvent être accessibles à des fins scientifiques, après instruction d'une demande formulée auprès du comité du secret statistique. Pour veiller à l'anonymat des données agrégées qui sont publiées après retraitement, les recommandations suivantes sont à respecter :

- aucune donnée publiée ne doit concerner moins de 3 entités à la fois ;
- aucune donnée publiée ne doit concerner une seule entreprise pour plus de 85 % du total ;
- aucune donnée publiée ne doit permettre l'identification directe ou indirecte des personnes.

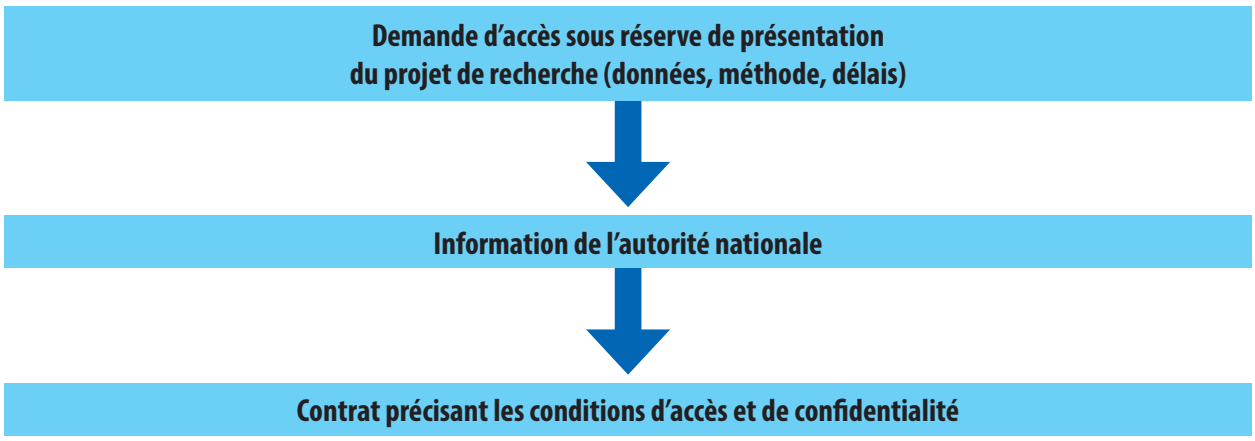
L'ensemble des recommandations et interdictions (par exemple : prohibition de diffusion de données individuelles fiscales) est disponible dans le guide du secret statistique édité par l'INSEE.

Pour ce qui concerne les données collectées au niveau communautaire par les autorités européennes, le règlement n° 322/97 organise l'établissement de statistiques européennes avec deux sources possibles :

- les données collectées par les États membres sont transmises aux autorités européennes ;
- et/ou les autorités européennes organisent des enquêtes.

Il existe un principe de diffusion des résultats statistiques obtenus.

En parallèle de la réglementation française, existent des modalités d'accès aux données collectées par l'Union Européenne à des fins scientifiques. C'est le règlement n° 831/2002 qui s'applique. Celui-ci prévoit une procédure similaire à celle qui existe en France :



67 Source : INSEE - Guide du secret statistique.

68 Article 338 du Traité UE ; Règlement n° 322/97 du Conseil du 17 février 1997 ; règlement d'application n° 831/2002.

Fiche 4. Convention d'Aarhus sur l'information en matière d'environnement

La convention d'Aarhus (Danemark) est une convention internationale signée le 25 juin 1998 et entrée en vigueur en octobre 2001. Cette convention concerne l'accès à l'information, la participation du public au processus décisionnel et l'accès à la justice en matière d'environnement.

La convention pose trois principes :

1) le développement de l'accès au public en matière d'information environnementale détenue par les autorités,
2) la participation du public aux décisions relatives à l'environnement,
3) l'accès à la justice du public pour faciliter les recours en cas de non-respect ou de contestation du respect des deux premiers principes.

Cette convention a été transposée en droit européen puis en droit français[69].

L'effet majeur de cette convention est que certaines informations doivent faire l'objet d'une diffusion publique pour que les citoyens puissent exercer leurs droits de participation et de recours en matière d'environnement.

Pour pouvoir appliquer ces droits, il s'est avéré nécessaire de définir la notion d'information en matière d'environnement dans le Code de l'environnement (article L124-2). Il s'agit de toute information disponible, quel qu'en soit le support, qui a pour objet :

- 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 ;
- les décisions, les activités et les facteurs, notamment les substances, l'énergie, le bruit, les rayonnements, les déchets, les émissions, les déversements et autres rejets ;
- l'état de la santé humaine, la sécurité et les conditions de vie des personnes, les constructions et le patrimoine culturel, dans la mesure où ils sont ou peuvent être altérés par des éléments de l'environnement, des décisions, des activités ou des facteurs mentionnés ci-dessus ;
- les rapports établis par les autorités publiques ou pour leur compte sur l'application des dispositions législatives et réglementaires relatives à l'environnement.

En matière d'accès à l'information environnementale par le public, il existe deux modalités :

- soit l'autorité publique diffuse l'information,
- soit l'autorité publique répond aux demandes des citoyens.

Sur le premier point, le droit français indique que les informations suivantes doivent faire l'objet d'une diffusion publique :

- les traités, conventions et accords internationaux, ainsi que la législation communautaire, nationale, régionale et locale concernant l'environnement ou s'y rapportant ;
- les plans et programmes et les documents définissant les politiques publiques qui ont trait à l'environnement ;
- les rapports établis par les autorités publiques ou pour leur compte relatifs à l'état d'avancement de la mise en œuvre des textes quand ces rapports sont élaborés ou conservés sous forme électronique par les autorités publiques ;
- les rapports établis par les autorités publiques sur l'état de l'environnement ;
- les données ou résumés des données recueillies par les autorités publiques dans le cadre du suivi des activités ayant ou susceptibles d'avoir des incidences sur l'environnement ;
- les autorisations qui ont un impact significatif sur l'environnement ainsi que les accords environnementaux ;
- les études d'impact environnemental et les évaluations de risques concernant les éléments de l'environnement.

69 Décret 2002-1187 du 12 septembre 2002.

Concernant les informations environnementales qui seraient contenues dans d'autres documents, la Convention d'Aarhus encourage les états à les diffuser au public. Toutefois, la transposition de cette convention dans le droit français[70] indique que ces informations sont communicables suivant le même régime que les documents administratifs (les exceptions à l'accès et à la réutilisation des documents sont quasiment les mêmes que pour les autres catégories de données publiques), sauf pour les données relatives à l'émission de substances dans l'environnement pour lesquelles les restrictions d'accès sont très réduites (autrement dit ces données sont presque systématiquement communiquées).

70 Titre II, chapitre IV du Code de l'environnement.

Fiche 5. Tableau comparatif des licences gratuites ODBL et Etalab

Attention : ces licences ne peuvent être utilisées que s'il est possible de garantir que le contenu ne contient pas de droits de propriété intellectuelle appartenant à des tiers. Cela implique de contrôler chaque élément faisant objet de la licence, ce qui en fait une garantie très difficile à donner.

<table>
<tr><th></th><th>ODBL 1.0</th><th>Etalab 2.0 (avril 2017)</th></tr>
<tr><td>Objet</td><td>Base de données
(= structure + données,
mais pas les données isolément)</td><td>Informations
= celles qui sont contenues dans des documents
adminitratifs (données, textes, etc.)</td></tr>
<tr><td>Définitions</td><td>Nombreuses définitions</td><td>Peu de définitions</td></tr>
<tr><td>Durée</td><td>Durée des droits sur la base
(= 15 ans sur le contenu)</td><td>Durée illimitée*</td></tr>
<tr><td>Périmètre</td><td>Non exclusif, monde entier</td><td>Non exclusif, monde entier</td></tr>
<tr><td>Gratuité</td><td>Oui</td><td>Oui</td></tr>
<tr><td>Utilisation commerciale permise</td><td>Oui</td><td>Oui</td></tr>
<tr><td rowspan="2">Droits d'utilisation accordés</td><td colspan="2">• Extraire
• Copier
• Diffuser
• Réutiliser
• Créer des bases dérivées</td></tr>
<tr><td>Créer des bases collaboratives
Modifier pour des raisons techniques</td><td>Modifier, adapter, transformer
Exploiter à titre commercial</td></tr>
<tr><td>Droits moraux à respecter</td><td>Paternité
Respect de l'honneur de l'auteur
Respect de l'intégrité de l'œuvre, divulgation et repentir</td><td>Paternité (nom de l'auteur sur l'œuvre)
Mention de la source (lien hypertexte possible)
Mention de la date de mise à jour</td></tr>
<tr><td>Exceptions à la licence</td><td>• Hors Europe : impossibilité d'interdire l'utilisation de la base si c'est pour un « usage loyal » (fair use)
• Europe : Extractions à des fins scientifiques ou d'illustration pour l'enseignement ou lorsque cela est requis par la loi
• Partout : impossibilité d'interdire les extractions qualitativement ou quantitativement non substantielles de la base</td><td>Pas de mentions</td></tr>
</table>

	ODBL 1.0	Etalab (au 01/04/2015)
Création dérivée	Obligation de communication au public de la base de données dérivée (effet contaminant)	Aucune obligation de communication au public de l'information dérivée
Partage de la création dérivée	*Share alike* : obligation de partager toute création dérivée sous cette licence ODBL ou licence équivalente). Sauf exceptions art. 4.5	Pas d'obligations pour le réutilisateur
Conférer des droits à un tiers sur la base initiale	Non autorisé. Le tiers doit contacter le producteur initial de la base de données.	Possibilité de transmettre l'information à un tiers, sous la licence de son choix.
Garanties apportées par le concédant	Aucune	Le concédant doit garantir que l'information : - soit ne contient de droits de propriété intellectuelle provenant de tiers, - soit que les tiers sont d'accord avec les conditions de réutilisation des informations.
Exclusion de responsabilité pour le concédant	Oui (pas forcément conforme au droit français)	Oui (conforme au droit français)
Conclusion	Choisir cette licence si l'on a une base de données et que l'on souhaite contrôler les redistributions de la base et les travaux dérivés	Choisir cette licence si l'on n'a pas besoin d'un suivi sur le devenir des données et que les données sont essentiellement distribuées en France (bien que la licence soit compatible avec une distribution à l'étranger). Cette licence ne reprend pas les droits spécifiques en matière de base de données. Cette licence est expressément compatible avec les licences CC-BY, OGL et ODC-BY.

*Sauf extinction des droits du concédant sur le contenu

Fiche 6. Mise en place d'une licence par la voie du contrat électronique

Au-delà de la licence ou du contrat électronique, tout site internet édité à titre professionnel doit d'abord contenir un certain nombre de mentions légales. Le manquement à l'une de ces obligations peut être sanctionné jusqu'à un an d'emprisonnement et 75 000 € d'amende pour les personnes physiques ou 375 000 € pour les personnes morales.

Une licence est un contrat par lequel le titulaire des droits d'auteur sur un logiciel, une œuvre audiovisuelle, une base de données, etc., définit avec son cocontractant les conditions dans lesquelles l'œuvre peut être utilisée, diffusée ou modifiée.

Dans le cas spécifique des licences, il s'agit le plus souvent de contrats d'adhésion, c'est-à-dire d'un contrat dont les termes sont imposés par une partie à l'autre. Les clauses contractuelles sont fixées à l'avance et aucune discussion n'est possible : les cocontractants sont alors libres d'adhérer ou non à la licence en acceptant les termes du contrat tels quels ou bien en ne les acceptant pas du tout.

Il doit être clairement indiqué que le contenu mis à disposition est soumis à une licence. Par ailleurs, cette licence ou un lien direct vers celle-ci doit être visible et très facilement accessible. *Exemple : apposition d'un lien vers l'adresse URL de la licence juste en dessous de la phrase indiquant que le contenu est soumis à une licence.*

L'indication de la soumission du contenu à une licence peut se faire à l'aide d'une formule telle que : « le contenu est mis à disposition selon les termes de la licence *[nom et identification de la licence]*. Toute utilisation du contenu autre que celle autorisée par cette licence est interdite. »[71]

La saisie d'une « case à cocher » et le clic sur un bouton d'acceptation sont recommandés pour pouvoir organiser la preuve de l'acceptation de la licence. En pratique, l'indication de la soumission du contenu à une licence se fait par une citation du nom de la licence choisie et de son logo :

- sur la première page du contenu lorsqu'il s'agit d'un document fixe ou non modifiable. *Exemple : page de garde d'un fichier « .pdf » ;*
- sur chaque page lorsqu'il s'agit d'un document dynamique tel qu'un site web ou une base de données.

> *Attention*
>
> Ces recommandations sont valables pour les licences gratuites. Pour les licences payantes lorsqu'elles sont exceptionnellement possibles, le formalisme est beaucoup plus lourd et je dois prendre contact avec mon service juridique.

71 Il est préférable d'utiliser une licence telle que les licences Creative Commons reprenant par ailleurs en son sein une formule telle que : « *l'exercice de tout droit sur l'œuvre mise à disposition emporte acceptation des termes de la licence.* »

Fiche 7. Archives

1 - Définitions

Définition des archives

Selon l'article L211-1 du Code du patrimoine :

« 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 sont conservées pour (article L 211-2 du Code du patrimoine) :
- la bonne gestion des dossiers et le suivi des projets,
- leur valeur de preuve,
- la sauvegarde de la mémoire.

Les documents et données sont des archives sur tout leur cycle de vie, c'est-à-dire depuis leur création jusqu'à leur archivage ou destruction.

Définition des données de la recherche selon les archivistes

Les archivistes, au sein de la section des archivistes des universités, rectorats, organismes de recherche et mouvements étudiants (AURORE) de l'Association des archivistes français, ont souhaité proposer une définition du périmètre dans lequel ils interviennent.

Les données de la recherche sont des informations, spécimens et matériaux produits, recueillis et documentés. Elles sont collectées ou exploitées à des fins de recherche et de preuves par les chercheurs et leurs équipes. À ce titre, elles constituent une partie des archives de la recherche.

Précisions

1. Les archives de la recherche englobent l'ensemble des documents et données produits ou reçus dans le cadre du processus de recherche. C'est-à-dire à la fois l'activité de recherche au sein des laboratoires et par les chercheurs et l'administration de la recherche au sein des organismes ainsi que par les fonctions venant en appui à la recherche.
2. Les données de la recherche sont en grande partie électroniques mais peuvent exister également sur d'autres supports.
3. Les données de la recherche sont soit collectées soit exploitées dans le cadre du processus de recherche.

2 - Archives publiques

Les établissements publics produisent et reçoivent des archives publiques. Selon le Code du patrimoine (articles L212-1 à L212-5), les archives publiques sont :
- imprescriptibles et inaliénables,
- soumises à des règles de gestion (durée de conservation, autorisation pour la destruction, versement des archives ayant un intérêt historique ou scientifique dans des services d'archives publics, délais de communicabilité),
- soumises au contrôle scientifique et technique de l'État sous l'égide du Service interministériel des archives de France.

3 - Délais de communicabilité des archives

La communication d'archives est soumise à des délais de communicabilité (articles L213-1 et 2 du Code du patrimoine). Les archives sont également considérées comme des documents administratifs. De ce fait,elles sont soumises à la loi CADA.

Les archives publiques sont, sous réserve des délais de communicabilité, communicables de plein droit.

Les principaux délais sont les suivants :

- **25 ans** « à compter de la date du document ou du document le plus récent inclus dans le dossier : pour les documents dont la communication porte atteinte au **secret en matière commerciale et industrielle ou au secret en matière de statistiques** sauf lorsque sont en cause des données collectées au moyen de questionnaires ayant trait aux faits et comportements d'ordre privé. »

- **25 ans** « à compter de la date du décès de l'intéressé », pour les documents dont la communication porte atteinte au **secret médical**. Si la date du décès n'est pas connue, **120 ans** « à compter de la date de naissance de la personne en cause ».

- **50 ans** « à compter de la date du document ou du document le plus récent inclus dans le dossier, pour les documents dont la communication porte atteinte **au secret de la défense nationale, à la sûreté de l'État, à la sécurité publique, à la sécurité des personnes ou à la protection de la vie privée. »**

- **75 ans** « à compter de la date du document ou du document le plus récent inclus dans le dossier » ou **25 ans** « à compter de la date du décès de l'intéressé si ce dernier délai est plus bref : Pour les documents dont la communication porte atteinte au secret en matière de statistiques lorsque sont en cause des données collectées au moyen de questionnaires ayant trait aux faits et comportements d'ordre privé. ».

- **100 ans** « à compter de la date du document ou du document le plus récent inclus dans le dossier, ou un délai de 25 ans à compter de la date du décès de l'intéressé si ce dernier délai est plus bref. Les mêmes délais s'appliquent aux documents couverts ou ayant été couverts par le secret de la défense nationale dont la communication est de nature à porter atteinte à la sécurité de personnes nommément désignées ou facilement identifiables. »

« Ne peuvent être consultées les archives publiques dont la communication est susceptible d'entraîner la diffusion d'informations permettant de concevoir, fabriquer, utiliser ou localiser des armes nucléaires, biologiques, chimiques ou toutes autres armes ayant des effets directs ou indirects de destruction d'un niveau analogue. »

Cette fiche a été complétée grâce aux relectures et avis des archivistes : Sarah Cadorel (CDSP, Sciences Po), Hélène Chambefort (Inserm), Marie-Pierre Diquelou (Inria), Magalie Moysan (Université Paris-Diderot) Lina Sbeih (INRA).