DATA4C+

Vers l'interopérabilité des bases de données pour être plus FAIR

Analyse des verrous juridiques à l'interopérabilité entre bases de données du Cirad, d'INRAE et de l'IRD sur le carbone du sol et les modes de gestion des sols et proposition de solutions

2022

Cloé Sigal-Guille, Pauline Corbière, Nathalie Emmanuel, Christine Le Bas, Julien Demenois

N° ANR-19-DATA-0005-01

Rapport T2.4 et T2.5 - Analyse des verrous juridiques à l'interopérabilité entre bases de données sur le carbone du sol et les modes de gestion des sols afin de proposer des solutions pour les lever

Le livrable est constitué des 3 documents suivants :

- NOTE MÉTHODOLOGIQUE - L'ouverture des données de la recherche : les identifier pour mieux les diffuser
- ANALYSE DE RISQUE - Ouverture des données de la recherche : identifier les risques et les prévenir
- RAPPORT D'INTEROPÉRABILITÉ DU SYSTÈME DATA4C+ : Quelles conséquences de l'interopérabilité sur les responsabilités du Cirad, de l'IRD et de INRAE lors de la diffusion des données.

Dans le cadre de **l'initiative « 4 pour 1000 »** et de la construction du **projet DATA4C+**

NOTE MÉTHODOLOGIQUE

L'ouverture des données de la recherche
Les identifier pour mieux les diffuser

Rédigée par Cloé Sigal-Guille
Sous la direction de Pauline Corbière et Nathalie Emmanuel.

SOMMAIRE

PRÉAMBULE

Le projet.

Le Projet DATA4C+ est un projet coordonné par le **Cirad**, en partenariat avec l'**INRAe** et l'**IRD**. Il répond à un **objectif technico-juridique de partage de données sur le carbone des sols de certains territoires.** Les résultats de DATA4C+ permettront l'estimation des potentiels de séquestration de carbone des sols de ces territoires dans de futurs projets. Ainsi, ce projet s'inscrit dans le cadre des **initiatives de sciences ouvertes et du carbone du sol (« 4 pour 1000 »)**. Afin de garantir au mieux la protection et l'utilisation des données collectées, la mise en place de bases de données permettent aux scientifiques de conserver les données générées dans le cadre de leurs expertises et de pouvoir les analyser. Il est nécessaire d'analyser ces bases de données sous l'angle juridique afin de déterminer si les données doivent ou peuvent être diffusées librement au grand public. À cette fin une première phase d'analyse juridique a été engagée afin d'**appréhender le cadre législatif français autour des données de la recherche**, et notamment, les obligations légales récentes issues de la loi n°2016-1321 pour une République numérique, dite « Loi Axelle Lemaire », du 7 octobre 2016.

L'intérêt d'une note méthodologique.

Dans un objectif de transparence et de contribution à la recherche, le législateur a cherché à inciter les organismes de recherches à partager leurs informations. Pour cela, **la loi pour une République Numérique du 7 octobre 2016** (« LRN » ci-après) a établi **un principe de diffusion par défaut**. Un périmètre des données « diffusables » est mis en place mais de nombreuses questions restent en suspens. En effet, un flou juridique est maintenu par l'absence de texte relatif au cadre juridique des données. Cette note méthodologique a pour ambition de clarifier quelque peu ces notions et de fournir les outils nécessaires à l'identification des données et à leur diffusion.

Quelques précautions à prendre.

Du fait du flou juridique qui subsiste, certains éléments textuels de la loi font **l'objet d'une interprétation** dans cette note méthodologique. Ainsi, ce n'est qu'à travers le croisement d'un faisceau d'indices que la diffusion ou non des données peut être déterminée. Il est donc impératif d'être **très attentif à l'ensemble des exceptions** qu'il existe et de prendre des précautions en cas de diffusion. Afin de pouvoir effectuer une analyse correcte, **au cas par cas**, de chacune des données, la note méthodologique doit être comprise dans son intégralité afin de ne pas omettre certaines exceptions qui pourraient générer des conflits par la suite.

Il s'agira alors de répondre à diverses problématiques. Est-ce que les données que j'ai collectées m'appartiennent ? Quel est leur statut juridique ? Dois-je les diffuser ? Puis-je les partager ?

1 COMPRENDRE LE CADRE JURIDIQUE DES DONNÉES, JEUX DE DONNÉES ET BASES DE DONNÉES.

À l'ère du numérique, les recherches scientifiques se sont **transformées**. Désormais, la recherche s'est développée vers une collecte de données numériques offrant ainsi la possibilité d'accéder à des outils interactifs, de partage et de précision. Les chercheurs alimentent ces bases de données qui vont leur permettre de conserver le fruit de leurs recherches et d'en générer des résultats. Il convient dans un premier temps de clarifier ces notions (1.1) avant d'en définir le contour juridique (1.2).

1.1 LES NOTIONS.

Qu'est-ce qu'une donnée ? Il n'existe pas de définition légale de la donnée. L'OCDE (L'Organisation de Coopération et de Développement Économiques) propose sa propre définition d'une donnée de recherche[1]. Il s'agit « *des enregistrement 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).* ».

Ainsi, *d'un point de vue scientifique*, **la typologie des données est très variée**. Elles dépendent de la discipline dans laquelle elles sont nées. Elles peuvent donc correspondre aussi bien à une donnée chiffrée (cas d'un modèle informatique) qu'à un corpus documentaire. *D'un point de vue technique*, la donnée est la représentation d'une information qui va permettre sa

[1] OCDE « Principes et lignes directrices pour l'accès aux données de la recherche financée sur fonds publics » - 2007.

communication, son traitement, ou encore son interprétation lorsqu'elle va être étudiée dans des carnets de laboratoire ou lors d'analyses préliminaires. C'est en ce sens qu'il sera déterminant de définir la nature et la typologie des données afin d'en connaitre leur potentiel d'utilisation *(cf Partie 2)*

Qu'est-ce qu'un jeu de données ? Le jeu de données correspond à un ensemble de données de différentes natures (texte, chiffre…) **non organisées entre elles**. Ce jeu de données pourra constituer, par la suite, une base de données lorsque les données seront organisées. Par exemple, un jeu de données peut correspondre à l'ensemble de résultats collectés suite à une analyse effectuée en laboratoire. Une base de données contient nécessairement un ou des jeux de données.

Qu'est-ce qu'une base de données ? Le Code de la propriété intellectuelle définit la base de données comme un « recueil d'œuvres, de données ou d'autres éléments indépendants, disposés de manière systématique ou méthodique, et individuellement accessible par des moyens électroniques ou par tout moyen »[2]. En d'autres termes, il s'agit d'une **collection d'informations, de données, regroupées afin de faciliter leur accessibilité et leur mise à jour**. La base de données est confectionnée logiquement, c'est-à-dire que les données recueillies ont un lien entre elles, ce qui va donner un sens à cette collection. La base de données permet également d'effectuer des recherches au sein des informations qu'elle contient.

1.2 LES CONTOURS JURIDIQUES

Comprendre le droit d'auteur en quelques notions.

Les droits accordés aux auteurs se décomposent en **deux séries de prérogatives aux régimes juridiques distincts.** Les droits patrimoniaux (CPI, art. L. 122-1 s.) qui permettent à l'auteur d'**autoriser** les différents modes d'utilisation de son oeuvre et de **percevoir** en contrepartie une rémunération. Les droits moraux (CPI, art. L. 121-1 s.) dont la finalité est de **protéger** la personnalité de l'auteur exprimée au travers son oeuvre

[2] Article L.112-3 du Code de la propriété intellectuelle.

→ ***Le droit moral*** est une spécialité « Made in France » qui permet à tout auteur de bénéficier de droits spécifiques qui sont **perpétuels, imprescriptible, inaliénable et intransmissible**. Quelles sont les composantes de ce droit ?

- Le droit à la divulgation = l'auteur indique s'il souhaite que son œuvre soit rendue publique ou non.

- Le droit de retrait ou de repentir = l'auteur pourra retirer ou cesser l'exploitation d'une œuvre.

- Le droit à la paternité = le nom et la qualité de l'auteur doivent être indiqués sur tous les supports. Ce dernier peut choisir d'exercer ce droit sous son propre nom ou sous un pseudonyme.

- Le droit au respect de l'œuvre = l'auteur peut autoriser ou non une modification, transformation, coupure, ou encore un changement de destination par exemple.

→ ***Le droit patrimonial*** peut être cédé contre une rémunération ou à titre gratuit selon les cas. De plus, ce droit est limité dans le temps car l'œuvre tombe dans le domaine public en principe **70 ans après la mort de l'auteur** (exception faite pour certaines œuvres et en cas de cession des droits). Quelles sont les composantes de ce droit ?

- Le droit de représentation (article L.122-2 du Code de la propriété intellectuelle) = il s'agit de l'exécution publique d'une œuvre, c'est-à-dire quand elle est diffusée à la radio par exemple ou présentée dans une galerie.

- Le droit de reproduction (article L.122-3 du Code de la propriété intellectuelle) = il s'agit de la communication indirecte au public, c'est le cas lorsque l'œuvre va être imprimée, dessinée, photographiée par exemple.

LE CONTOUR JURIDIQUE DE LA DONNÉE. Juridiquement, la donnée ne bénéficie d'aucune protection. Elle est considérée comme **une information de « libre parcours ».** Ainsi, l'établissement du créateur de la donnée peut restreindre ou non sa diffusion, sous réserve des exceptions qui peuvent limiter sa diffusion. Il s'agit des photographies illustrant l'hypothèse unique où une donnée peut être protégée par le droit d'auteur. La loi pour une République Numérique a cependant créé ce que l'on pourrait appeler **un statut juridique de la donnée de la recherche**. En effet, l'article 38 parle de « fichiers [qui] constituent des

données de la recherche ». Pour autant, la notion même de donnée de la recherche n'est pas précisément cadrée dans la loi. Il est donc nécessaire d'élaborer **une analyse au cas par cas pour identifier ce qui relève des données de la recherche ou non**. C'est l'objectif de cette note méthodologique.

De quoi s'agit-il en pratique ? Lorsque les chercheurs diffusent leur résultat de recherche en ligne, sont-ils maîtres de la réutilisation des données de leurs recherches ? Peuvent-ils autoriser ou non cette réutilisation ? Si l'on entend le terme de « données de la recherche » au sens strict (c'est-à-dire tel que décrit par la loi comme ne pouvant être protégeable) alors le chercheur n'est pas considéré comme un auteur. En effet, les données ne sont pas considérées comme des œuvres protégeables par le droit d'auteur. Ainsi, le chercheur ne pourra pas empêcher leur réutilisation, et il semblerait qu'il ne puisse pas exiger que soit appliquer son droit de paternité. En d'autres termes, il semblerait qu'il ne puisse pas exiger d'être cité lors de la réutilisation d'une seule donnée de la recherche.

LE CONTOUR JURIDIQUE DE LA BASE DE DONNÉES. La base de données est protégée par le Code de la propriété intellectuelle **de deux façons**. En effet, elle peut l'être par le biais du droit d'auteur, mais également d'un droit *sui generis* du producteur de la base.

S'agissant d'une protection de la base de données par le droit d'auteur. La structure de la base de données, c'est-à-dire la manière dont les données sont agencées, est protégée par le droit d'auteur dès lors que le critère d'originalité est satisfait.

Comprendre le critère d'originalité en droit d'auteur.

Aucune définition n'est arrêtée par la loi, et c'est le juge, au fil des années qui est venu préciser cette notion d'originalité. Cependant, le cadre de la notion n'est pas fixe et le juge continue d'en préciser les limites, jurisprudence après jurisprudence. Majoritairement, trois définitions de la notion d'originalité se distinguent. Ainsi, le juge parle « d'empreinte de la personnalité de l'auteur », de la « marque de l'apport intellectuel de l'auteur », et de « l'expression des choix libres et créatifs de l'auteur ». **Plus précisément, à quoi correspondent ces précisions ?**

1. ***L'empreinte / la marque / le reflet de la personnalité de l'auteur*** est un critère subjectif. En effet, on va s'intéresser ici à l'absence de banalités. Autrement dit, il faut que l'auteur se détache d'une conception « logique » de son œuvre. L'exemple le plus donné pour illustrer cette notion est celle de deux peintres qui peignent différemment le même paysage.
2. ***La marque de l'apport intellectuel / de l'effort créateur de l'auteur*** est quant à lui un critère un peu plus objectif. Ce critère est utilisé pour les œuvres fonctionnelles et pour les écrits plus techniques ou pratique.
3. ***L'expression des choix libres et créatifs de l'auteur*** est un critère qui provient de la jurisprudence européenne et qui a pour ambition d'être plus objectif que les précédents. Par exemple, un auteur qui n'aura pas eu le choix de modéliser son œuvre comme il l'entend, soit à cause de la technique exigée, soit à cause du modèle représenté, ne pourra faire valoir son œuvre comme étant originale.

En somme, le critère d'originalité du droit auteur est difficile à définir et dépend de l'appréciation souveraine du juge. C'est ce dernier qui aura le dernier mot en cas de conflit et qui pourra déterminer si une œuvre est originale ou pas. Il se basera sur les arguments de l'auteur, car c'est lui qui doit prouver que son œuvre est originale en cas de conflit. Par exemple, la tendance jurisprudentielle actuelle n'est pas très favorable au photographe et les juges du fond son très exigeants sur l'appréciation de la démarche artistique des photographes. *A contrario*, les créateurs de sites Internet peuvent se prévaloir de la protection des textes en démontrant son apport intellectuel, mais également de l'aspect visuel en expliquant les choix libres et créatifs effectués par rapport à des sites de modèles type.

S'agissant d'une protection par le droit *sui generis* du producteur de la base de données. Le droit *sui generis* est spécifique à l'Union européenne[3] et ne s'applique donc pas dans d'autres pays. Ce droit fait l'objet de dispositions spécifiques et ne fonctionne pas avec le droit d'auteur. Là où le droit d'auteur est automatique, on s'aperçoit au contraire que le droit des producteurs de base de données **doit faire l'objet d'une appréciation souveraine du juge**. C'est la loi du 1er juillet 1998[4] qui a instauré cette protection spécifique au droit français, « *sui*

[3] Directive 96/9/CE du 11 mars 1996 sur la protection des bases de données.
[4] Loi n°98-536 du 1er juillet 1998 concernant la protection des bases de données.

generis », au profit des producteurs de base de données. En effet, ce droit va conférer une **protection spécifique** à la personne même du producteur en considération de sa base de données et non à la base de données elle-même. Pourquoi ? Afin d'éviter toute appropriation et utilisation frauduleuse de la base de données résultants souvent d'investissements majeurs.

→ ***Qui est le producteur de la base de données ?*** D'après le Code de la propriété intellectuelle[5], il s'agit de « *la personne qui prend l'initiative et le risque des investissements correspondants* ». En d'autres termes, il s'agira de la personne, physique ou morale, qui a **financé les activités** ayant qui a conduit à la réalisation et à la mise à jour d'une base de données, et non le chercheur lui-même. Le Code de la propriété intellectuelle continue en indiquant que ce producteur « bénéficie d'une protection du contenu de la base lorsque la constitution, la vérification ou la présentation de celui-ci atteste d'une investissement financier, matériel ou humain substantiel ». En d'autres termes, cet investissement peut être de différentes natures qu'il conviendra de **prouver par tout moyen**. Le juge du fond apprécie le **caractère « substantiel » de cet investissement** au regard des coûts récurrents de gestion, d'acquisition des données, de contrôle et de maintenance de la base de données. Par exemple, le juge retient que le producteur peut bénéficier de cette protection *sui generis* dès lors qu'il a attesté, au moyen notamment de factures et de justifications diverses, avoir mis en œuvre des moyens matériels, financiers et humains considérables pour constituer la base de données et la tenir à jour en temps réel[6].

→ ***Quelle est l'étendue de la protection sui generis du producteur de la base de données ?*** Ce dernier possède 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 forme que ce soit ; **2° La réutilisation**, par la mise à la disposition du public **de la totalité ou d'une partie***

[5] Article L.341-1 du Code de la propriété intellectuelle.
[6] TGI Paris 5 septembre 2001 SA Cadremploi c/ SA Keljob et Colt Télécommunications France

qualitativement ou quantitativement substantielle *du contenu de la base, quelle qu'en soit la forme. »*[7]

De quoi est-il question ici ? Il faut imaginer le cas du détenteur d'une base qui la met à disposition sur internet pour fournir un certain service. Un tiers pourrait alors aspirer le contenu de sa base ou d'une certaine partie de celle-ci afin d'en récupérer les données et offrir un service potentiellement concurrent. Ce service ne pourrait porter que sur une certaine partie de la base de données initiale (exemple d'une base de données identifiant la composition des sols sur plusieurs pays et où, un tiers ne reprendrait que les données d'une seule zone géographique limitée pour fournir un service concurrent ou complémentaire). C'est en cela que les termes quantitativement ou qualitativement substantiels doivent être interprétés. Ils recouvrent la notion de valeur que l'on peut extraire de l'ensemble ou d'une partie significative de la base.

Le producteur peut également interdire *« … l'extraction ou la réutilisation répétée et systématique de parties qualitativement ou quantitativement non substantielles du contenu de la base lorsque ces opérations excèdent manifestement les conditions d'utilisation normales de la base de données »*[8]

De quoi est-il question ici ? Nous pouvons prendre l'exemple d'un éditeur de site de petites annonces immobilières. Il rassemble les annonces et les place dans un site en ligne, ce qui représente un coût. Mais un acteur mal honnête pourrait ouvrir un autre site de petites annonces sans payer le prix de création de la base (il va se servir du premier site pour déposer les questions et annonces et les « publier » sur son propre site, comme si elles lui revenaient). Cet acteur peut mettre des publicités sur son site. Le revenu de ces publicités lui reviendra à lui et non pas à l'acteur qui met en ligne et gère les annonces. C'est ce genre de comportement que vise à interdire **l'article L.342-2 du code de la propriété intellectuelle**. Pour autant, comme on le voit, il n'est pas question de droit d'auteur mais d'un droit économique. Les comportements déviants peuvent faire l'objet de concurrence déloyale ou de parasitisme.

[7] Article L.342-1 du Code de la propriété intellectuelle.
[8] Article L.342-2 du Code de la propriété intellectuelle.

Compte tenu de sa valeur économique, le contenu d'une base de données est susceptible d'être l'objet de convoitise. Il peut également être valorisé par la concession de licence ou de cession des droits de propriété intellectuelle. D'autres enjeux plus juridiques sont soulevés comme par exemple la question de la conformité avec la Loi pour une république numérique (LRN). En effet, cette dernière a apporté une limite au droit des administrations productrices de bases de données au sens du Code de la propriété intellectuelle[9]. Sans faire application expresse de cette disposition, le Conseil d'État a pu retenir une solution s'inscrivant dans ce mouvement d'ouverture des bases de données publiques[10]. En d'autres termes, **ces entités publiques ne peuvent plus faire obstacle à la réutilisation du contenu des bases de données qu'elles publient, sauf si ces dernières ont été produites ou reçues dans le cadre d'une activité concurrentielle**[11]. Une autre nuance doit être apportée à ce propos lorsque des données bénéficiant d'une protection spécifiques (données à caractère personnel par exemple) sont introduites dans une base de données publiée par l'entité publique. Il convient alors de toujours rester prudent lors de la publication d'une base de données.

➔ *De façon plus concise, voici un tableau permettant de récapituler les éléments importants à retenir s'agissant de la protection par le droit d'auteur et celle par le droit sui generis du producteur de la base de données.*

[9] Article 11 de la LRN
[10] CE, 8 février 2017, n°389806
[11] Article L.321-3 du CRPA

	Droit d'auteur	Droit sui generis du producteur de la base de données
Protection	Protection de la **structure** de la base de données sous réserve du **critère d'originalité** *Originalité = « Empreinte de la personnalité de l'auteur » = « Par le choix ou la disposition des matières »*[12].	Protection **du producteur** de la base de données
Objet de la protection	Protection de la **forme**, de la **structure** (aucune protection sur le contenu)	Protection du **contenu** de la base de données.
Titulaire	Auteur = personne physique qui crée l'œuvre	Producteur = il s'agit en général de l'employeur car c'est lui qui prend « l'initiative et le risque des investissements correspondants »[13].
Droits conférés	- Droit moraux - Droit patrimoniaux	- Interdire l'extraction d'une partie substantielle du contenu de la base. - Interdire la réutilisation de la base par mise à disposition du public. - Contrôler les conditions d'utilisation de la base de données.
Durée de la protection	Les droits expirent **70 ans après la mort de l'auteur**	Les droits expirent **15 ans** après le 1er janvier de l'année civile qui suit l**'achèvement** de la base de données.
Sanction	Délit de contrefaçon = peine allant jusqu'à 3 ans d'emprisonnement et 300 000€ d'amende[14].	Délit de contrefaçon = Peine allant jusqu'à 3 ans d'emprisonnement et 300 000€ d'amende[15].

[12] Article L.112-3 du Code de la propriété intellectuelle.
[13] Article L.341-1 du Code de la propriété intellectuelle.
[14] Article L.335-2 et suivants du Code de la propriété intellectuelle.
[15] Article L.343-1 du Code de la propriété intellectuelle.

Exceptions au droit d'auteur et au droit du producteur de la base de données.

Il existe des exceptions qui permettent de contourner ces protections. La loi pour une République numérique (LRN) est venue allonger la liste de ces exceptions.

→ **Exceptions au droit d'auteur**. Les exceptions au droit d'auteur sont limitativement listées à l'article L.122-5 du Code de la propriété intellectuelle. La LRN vient en rajouter deux :

- Les copies ou reproductions numériques réalisées à partir d'une source licite, en vue de l'exploration de textes et de données incluses ou associées aux écrits scientifiques pour les besoins de la recherche publique, à l'exclusion de toute finalité commerciale (10°)
- Les reproduction et représentations d'œuvres architecturales et des sculptures, placées en permanence sur la voie publique, réalisées par des personnes physiques, à l'exclusion de tout usage à caractère commercial (11°)

→ **Exception au droit du producteur de la base de données.** Désormais, l'article L.342-3 5° du Code de la propriété intellectuelle indique que le **titulaire du droit ne peut interdire les copies ou reproductions numériques de la base réalisées par une personne qui y a licitement accès, en vue de fouilles de textes et de données incluses ou associées aux écrits scientifiques dans un cadre de recherche, à l'exclusion de toute finalité commerciale**. La conservation et la communication des copies techniques issues des traitements, au terme des activités de recherche pour lesquelles elles ont été produites sont assurées par des organismes désignés par décret.

Le droit et la recherche scientifique.

Dans le cas de recherches scientifiques, qui est réellement titulaire des droits sur les données de recherche ? Au regard de ce qui a été dit, les données ne sont pas protégeables par le droit d'auteur. Cependant, les écrits scientifiques, qui sont eux protégeables, peuvent faire référence à des données résultant d'expériences ou de recherche. Ainsi, s'agissant des écrits scientifiques, ce sont les chercheurs qui sont titulaires du droit d'auteur sur les articles scientifiques sous réserve de stipulations contraires dans leur contrat et, du critère d'originalité. En effet, les œuvres créés par les chercheurs bénéficient du droit d'auteur ce qui confère à leur créateur la qualité d'auteur. Ces derniers sont titulaires d'un droit moral et d'un

droit patrimonial (ces notions sont expliquées plus haut). Si l'auteur ne cède pas son droit patrimonial à l'établissement de recherche, alors il en est le seul titulaire. En droit d'auteur, les droits naissent directement sur la tête du salarié[16].

S'agissant des bases de données, c'est différent. En vertu de Code de la propriété intellectuelle et des propos développés ci-dessus, les droits sur les bases de données peuvent être attribués, en principe, à une personne physique ou morale. En pratique c'est la personne morale qui est titulaire des droits ; en l'occurrence, les institutions de recherche par exemple. En effet, l'établissement de recherche va financer la recherche qui va conduire à la création de la base de données. De plus, c'est l'établissement qui finance le salaire des chercheurs qui collectent les données. En somme, l'établissement met en œuvre un double investissement, financier et humain, afin d'élaborer la base de données. En vertu de la loi, l'établissement est donc considéré comme le producteur de cette base de données et jouit d'un droit *sui generis*. Dans certains cas il peut y avoir plusieurs producteurs, comme dans le cadre de base de données réalisées dans le cadre de projet de recherche, on parle alors de co-indivision. Dans ce cas-là, il est conseillé aux partenaires de mettre en place un accord aux fins de déterminer les contours de l'indivision.

Comprendre le régime de l'indivision en quelques notions.

C'est une situation juridique dans laquelle plusieurs personnes détiennent des droits de même nature sur un bien ou un ensemble de biens, sans que leurs parts respectives soient réellement individualisées. De manière plus schématique, chacun a une part du bien mais tout le monde le possède dans sa globalité. L'indivision a trois caractéristiques :

- **C'est un droit individuel** : c'est-à-dire que chaque indivisaire est propriétaire d'un droit propre.

- L'indivision n'est pas forcément organisée : le processus de décision sur le bien indivis n'a pas à être organisé par avance. Cependant, il est conseillé de mettre en place une convention dès le départ afin d'éviter des situations de blocage lors de la vente d'un bien, ou tout simplement, lors de la diffusion d'une base de données.

[16] Article L.111-1 alinéa 3

- **C'est un droit précaire** : cela signifie que tout indivisaire a le droit de demander le partage du bien à tout moment, qu'elle que soit son importance au sein de l'indivision. Autrement dit, on ne peut pas forcer un indivisaire à céder sa part, tout comme on ne peut pas l'empêcher de partir.

Conclusion.

L'absence de définitions juridiques des données de recherche et des éléments qui les entourent accentue l'importance de distinguer les notions de données et de base de données. En effet, ces deux notions possèdent des cadres juridiques différents, ne répondent pas aux même règles et ne bénéficient pas la même protection. A la lumière de ce qui vient d'être développé, il est nécessaire d'analyser le cadre contractuel des données afin d'en déterminer les usages possibles.

2 LA TYPOLOGIE DES DONNÉES EN FONCTION DE LEUR ORIGINE.

Afin de déterminer si une donnée est diffusable ou non, il faut raisonner en deux temps. En effet, l'enjeu sur la diffusion d'une donnée est important à cerner car une fois cette dernière communiquée au public, elle peut être librement réutilisée. Premièrement, il est nécessaire de déterminer l'origine de la donnée. Puis, dans un second temps, il convient d'effectuer une analyse au cas par cas afin de déterminer si la donnée est par principe diffusable.

Comme nous le verrons tout au long de cette note méthodologique, la Loi pour une République numérique (LRN) impose une obligation de diffusion des données pour les administrations françaises mais aussi pour les acteurs privés et publics qui remplissent une mission de service public. **Ainsi, avant tout propos, comment définir la notion de service public ?** La notion de service public est difficile à cerner du fait de 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. Plusieurs finalités peuvent être poursuivies et ce service public remplit quatre fonctions principales. On peut alors distinguer les services publics à finalité d'ordre et de régulation (par exemple la défense nationale, la justice, la protection civile ou encore les ordres professionnels), ceux ayant pour objectif la protection sociale et sanitaire (comme par exemple la sécurité sociale ou bien le service hospitalier), ceux à vocation éducative et culturelle (il s'agit ici de l'enseignement, de la recherche ou encore du service public audiovisuel), et enfin, ceux à caractère économique.

Par 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). Nous pouvons donner un autre exemple en s'attachant au cas de l'INRAE (Institut national de recherche pour l'agriculture, l'alimentation et l'environnement). Cet institut a pour mission de réaliser, d'organiser et de coordonner tous travaux scientifique et technologique dans les domaines de l'agriculture, de l'alimentation, de la forêt, et bien d'autres encore. Ces projets sont mis en place à son initiative ou à la demande de l'État. Dans

ce dernier cas, l'institut accompli une mission de service public et doit partager ses résultats avec le ministère de l'agriculture.

La LRN impose-t-elle un délai pour diffuser les données ? Oui, la loi va imposer divers délais en fonction des documents à publier[17]. Elle renvoie au Code des relations entre le public et l'Administration (CRPA) [18] pour encadrer ces délais de diffusion.

- Six mois après la promulgation de la loi pour les documents administratifs faisant l'objet de procédures prévues par le CRPA.
- Un an après la promulgation de la loi pour les répertoires publiés par les administrations chaque année. Une nouvelle version à jour doit être publiée annuellement.
- Enfin, une date est fixée par décret, et au plus tard deux après la promulgation de la loi pour l'ensemble des autres documents (par exemple, les bases de données).

Dans cette partie, intéressons-nous d'abord à l'origine de la donnée ou, autrement dit, au contexte dans lequel elle a été générée. Plus précisément, il faut se poser plusieurs questions : la donnée a-t-elle fait l'objet d'un contrat (contrat de prestation de service, contrat de collaboration de recherche, ou autre) ? Qui en est le créateur et donc le diffuseur potentiel ? Le développement qui suit est accompagné d'un logigramme, inséré en Annexe, et a pour objet d'offrir une clarification sur l'origine de la donnée.

→ **Annexe 1 : Logigramme « Typologie des données ».**

[17] Article 8 de la LRN
[18] Articles L.312-1-1 et L.312-1-3 du CRPA

2.1 DONNÉES GÉNÉRÉES DANS LE CADRE D'UNE ACTIVITÉ QUI N'EST PAS ENCADRÉE PAR UN CONTRAT.

Premièrement, intéressons-nous au cas d'une donnée générée dans le cadre d'une activité de recherche, mais qui n'a pas fait l'objet de contrat ou, qui a fait l'objet d'un contrat, mais cette donnée est si ancienne qu'il est impossible de le retrouver.

De quoi s'agit-il en pratique ? Vous n'avez pas signé de contrat avec un tiers ou un établissement public ou privé avec lequel vous collaborez dans le cadre d'une activité de recherche ou la fourniture d'une prestation de service. Dans ce cas-là, il est difficile de déterminer à qui appartient la donnée collectée. Qui devra décider de sa diffusion ou non ?
Autre cas possible. Il s'agit par exemple d'une base de données qui comprend des données anciennes dont il vous est impossible de déterminer l'origine. Cependant, vous souhaitez diffuser votre base de données. **La question est donc de savoir si oui ou non ces données peuvent être diffusables ?**

2.1.1 Données dont l'origine est inconnue.

Premièrement, l'origine de la donnée est totalement inconnue. C'est une donnée qui figure dans les archives depuis de nombreuses années et il est difficile, voire impossible de remonter jusqu'à son origine. Dans ce cas-là il y a deux possibilités :

1/ La donnée est sous forme numérique. Dans ce cas-là, il semblerait que la diffusion de ces données, sous réserve qu'elles aient été **générées après le 7 octobre 2016**, entrent dans le champ d'application de la LRN et peuvent être sujette à **l'obligation de diffusion par défaut** si elles correspondent aux cas mentionnées par la loi *(cf Partie 4)*. Dans le cas contraire, si leur diffusion n'est pas obligatoire, ces données restent **communicables** selon la volonté de l'établissement ou de la personne qui possède les droits sur ces données. En effet, la LRN exige que les données ou documents administratifs partagés se trouvent dans **un format numérique** afin d'en faciliter leur **réutilisation**[19]. Leur communication peut également être **interdire** dans certains cas cités par la loi *(cf Partie 4)*, il convient donc d'être vigilant et d'effectuer **une analyse au cas par cas** des données avant toute diffusion.

[19] Article 3 de la LRN et article L.312-1-1 du CRPA

→ **Pourquoi seules les données générées après le 7 octobre 2016 peuvent être sujette à une obligation de diffusion ?** La LRN **n'est pas rétroactive**. Ainsi les données générées avant le 7 octobre 2016 ne peuvent pas être obligatoirement diffusables, même si elles se trouvent sous forme numérique. La diffusion est donc **facultative mais est encouragée** par le législateur. Ainsi, si les données sont très anciennes et qu'elles ne semblent pas être des données confidentielles ou stratégiques, l'établissement peut opter pour la diffusion de ces données en prenant le risque de ne pouvoir en retracer l'origine et l'autorisation de diffusion.

→ Quid de leur réutilisation. Quoi qu'il en soit, que l'établissement diffuse des données parce qu'il en est obligé ou parce qu'il le souhaite (dans le cas où la communication est facultative et pas interdite), il semblerait que l'établissement doive communiquer ces données *« dans un standard ouvert, aisément réutilisable et exploitable »*[20] *(cf Partie 5)*

Pour résumer, en cas de données générées après le 7 octobre 2016, la communication semblerait possible voire obligatoire, alors qu'elle ne sera forcément que facultative pour les données générées avant l'entrée en vigueur de la loi

Attention. Il convient de nuancer ces propos et indiquer qu'il s'agit ici d'un cas de risque faible et limité. En effet, la loi reste imprécise et l'absence de connaissance sur l'origine de la donnée implique la nécessité d'être vigilant lors de sa diffusion.

2/ La donnée est sous une autre forme. Il peut s'agir par exemple d'une donnée présente dans un document papier transmis par un laboratoire dans le cadre d'une recherche scientifique. Dans ce cas-là, la diffusion de la donnée **n'est pas obligatoire**. Elle est cependant **permise** si cette donnée est transformée en forme numérique. En effet, c'est une des **conditions de diffusion imposée par la LRN**.

La loi n'oblige pas les établissements à transformer les données se trouvant dans une autre forme pour ensuite les diffuser. L'établissement est donc **libre de transformer le format de la donnée puis de la diffuser, ou non**. Il convient de rappeler que cette règle est conditionnée par l'absence d'élément justifiant une interdiction de diffusion.

[20] Article L.300-4 du CRPA

2.1.2 Données dont l'origine est connue mais qui ne font l'objet d'aucun contrat.

Dans cette seconde hypothèse, l'origine de la donnée est connue mais l'absence de contrat amène une absence de cadre. Deux possibilités nous sont donc offertes :

1/ La donnée est financée pour plus de la moitié par un fonds public. La donnée est diffusable sous réserve des cas présentant une interdiction de diffusion *(cf Partie 4)*. La LRN a supprimé l'article 11 de la loi Valter[21] qui permettait aux établissements publics de recherche et aux universités de déroger au régime d'une diffusion obligatoire de leurs données. Ainsi, ces derniers sont dans l'obligations de partager leurs documents administratifs relevant de mission public par exemple. S'agissant des données de recherche ne correspondant pas à une mission de service public, lorsque cette dernière est financée pour plus de la moitié par un fonds public, en l'absence de contrat, la communication n'est pas obligatoire mais est **permise**. La LRN exige cependant que **la réutilisation de ces données soit obligatoire libre**.
Attention. Des **exceptions existent** et la diffusion ne pourra s'effectuer sur l'ensemble des données (par exemple des données couvertes par le secret professionnel). C'est en ce sens qu'une **analyse au cas par cas est nécessaire** dans un second temps. Cette analyse minutieuse fait l'objet de la *Partie 4 de cette Note méthodologique*.

Que faut-il comprendre de la notion de fonds public à travers la LRN ? Il s'agit du **financement accordé par une institution publique** à une organisation, un institut pour son fonctionnement normal ou pour un projet. Le code de la recherche[22] par le « dotation de l'État, des collectivités territoriales ou des établissement publics » mais également de « par des subventions d'agences de financement nationales ou pas des fonds de l'Union européenne ». En somme, dans la LRN, un fonds public est compris comme émanant d'une institution française, telle l'État, les collectivités locales, mais également de l'Union européenne. Ainsi, en pratique, **une recherche financée pour plus de la moitié par un fonds public européen est sujette à l'application de la LRN sur ses données collectées**.

[21] Loi n°2015-1779 du 28 décembre 2015 relative à la gratuité et aux modalités de la réutilisation des informations du secteur public.
[22] Article 17 de la LRN modifiant l'article L.533-4 II du Code de la recherche.

2/ La donnée est financée pour plus de la moitié par un fonds privé. Dans le cas où la donnée proviendrait d'une collecte engagée dans le cadre d'une mission de service public, elle doit **être obligatoirement publiée**[23]. Ce ne sera pas le cas dans l'hypothèse contraire. La recherche étant financée pour plus de la moitié par un fonds privé, les données collectées qui ne répondent pas à une mission de service public dérogent à l'obligation du principe de diffusion. Ainsi, si ces données ne correspondent pas aux cas d'exceptions, l'établissement qui les a collectées pourra effectuer une diffusion facultative, à condition que les données se trouvent sous forme numérique et qu'elles soient achevées, c'est-à-dire lorsqu'elle est stable et n'est pas susceptible de faire l'objet de modification. À la différence des données financée pour plus de la moitié par un fonds public, **la réutilisation** de données financées pour plus de la moitié par un fonds privé **n'a pas à être libre**. Le législateur **encourage** la possibilité de réutilisation de ces données mais ne l'exige pas.

Que faut-il comprendre de la notion de fonds privé à travers la LRN ? Il s'agit du **financement accordé par des personnes de droit privés**. Cela peut être des entreprises mais également des fondations. En effet, les fondations sont des personnes morales de droit privé à but non lucratif. Elles sont créées afin d'accomplir une œuvre d'intérêt général[24]. Il existe différentes sortes de fondations. Quel que soit le type de fondations qui finance un projet, le fonds attribué correspond à du mécénat et est compris comme étant un fonds privé.

De quoi s'agit-il en pratique ? Les travaux de recherche qui vous ont permis de collecter les données que vous souhaitez diffuser peuvent avoir été financés par des acteurs publics et des acteurs privés. Il faut identifier quelle est la part publique et privé du financement octroyé. Ainsi votre recherche est-elle financée en plus grande partie par des fonds publics (financement de la région Midi Pyrénées, de l'Union européenne, de l'ANR), ou plutôt par des fonds privés (financement par des fondations, par des entreprises privées) ? Dans le premier cas, la diffusion des données que vous avez générées sera obligatoire, alors que dans le second

[23] Article L.300-2 du CRPA *« [...] ainsi que par les autres personnes de droit public ou les personnes de droit privé chargées d'une telle mission »*
[24] Article 18 de la loi n°87-571 du 23 juillet 1987 sur le développement du mécénat.

cas, la diffusion sera facultative sous réserve que le projet ne fasse pas l'objet d'une mission de service public.

2.2 DONNÉES GÉNÉRÉES DANS LE CADRE D'UNE ACTIVITÉ ENCADRÉE PAR UN CONTRAT.

Les données peuvent être générées **dans le cadre d'activités encadrées par des contrats**. Il convient d'effectuer une articulation entre le droit commun des contrats et le droit spécial relatif à la recherche et à l'innovation. Il s'agit alors de faire la différence entre les informations scientifiques appropriables et celles qui ne le sont pas. Dans le premier cas, les données peuvent être protégées par des droits de propriété intellectuelle. Dans le second cas, il s'agit des données brutes, des idées ou encore de simples concepts. Quels sont les divers contrats auxquels vous pouvez être confrontés et que faire des données générées ?

2.2.1 Contrat de prestation de service.

Le contrat de prestation de service correspond à un « louage d'ouvrage »[25]. Il se définit comme étant un contrat par lequel une partie, le prestataire de service, s'oblige envers une autre partie, le bénéficiaire du service, à exécuter un travail déterminé contre une rémunération. Ce contrat va encadrer les conditions dans lesquelles le prestataire va fournir un service au client. Les contrats de prestation de service représentent tous les contrats par lesquels l'établissement réalise des travaux pour le compte de tiers.
Il convient d'identifier **plusieurs hypothèses** s'agissant de la diffusion des données collectées dans l'exercice de ce type de contrat.

1/ Le contrat fait l'objet d'une clause traitant des données collectées. Il convient d'effectuer une précision sur ce point. En pratique, les données sont souvent inclues dans la définition

[25] Article 1710 du Code civil.

des résultats faisant l'objet du contrat, elles sont donc considérées comme des résultats de la prestation.

Ces dernières appartiennent en principe au bénéficiaire du service. Introduire une clause au sein du contrat de prestation peut permettre d'identifier si l'organisme qui a collecté les données peut les utiliser ou non, par le biais d'une clause notamment. Il faut donc se référer au contrat.

- La clause prévoit explicitement la possibilité de diffusion et/ ou la réutilisation des données. Dans cette hypothèse, votre établissement **peut diffuser** les données sous réserve des interdictions mentionnées par la loi.

  → **En pratique**, la précision de ces clauses sont rarissimes, il convient alors de prendre des précautions particulières en cas de diffusion.
- La clause prévoit explicitement l'interdiction de la diffusion et/ou de la réutilisation des données. Dans cette hypothèse, la diffusion des données collectées dans le cadre de la prestation de service **n'est pas envisageable**, ce qui sera le plus souvent le cas dans le cadre de prestations.

2/ Le contrat ne fait pas l'objet d'une clause traitant des données collectées. Il semblerait que les résultats générés dans le cadre d'une prestation de services appartiennent par principe **au prestataire** et par analogie les données générées. En effet, l'absence de clause sur les données, générées lors de la réalisation d'une prestation de service conduisent les parties à se retrouver dans une situation assez floue juridiquement. La jurisprudence indique qu'en cas d'absence de clause de cession à la fois explicite, claire et valable, le montant payé par le commanditaire à son prestataire ne prend en compte que le travail d'exécution ainsi que, le cas échéant, le support matériel de l'œuvre qui sera livrée. En revanche, les droits de propriété intellectuelle relatifs à l'œuvre en elle-même n'en fait pas partie. Bien que les données ne bénéficient pas de protection par un droit de propriété intellectuelle en tant que telle, on pourrait en déduire par analogie que les données n'en font pas partie non plus.

Est-ce que le prestataire de service peut diffuser les données générées en l'absence de clause ? En considération de ce qui a été dit, en l'absence de clause, la diffusion des données

semble possible pour le prestataire. Cependant, il convient de prendre des précautions lors de cette diffusion afin de ne pas risquer d'entrer en conflit avec le commanditaire.

Est-ce que le donneur d'ordre peut diffuser les données générées ? En cas d'absence de clause, la jurisprudence, en matière de logiciel, estime que le commanditaire est dans l'impossibilité d'exploiter l'œuvre et de l'adapter ; donc il ne peut pas réutiliser les données et les diffuser comme bon lui semble. En matière de données de la recherche, le juge ne s'est pas encore prononcé. En suivant la même logique, nous pourrions envisager le fait que le commanditaire ne puisse pas exploiter les données comme il l'entend. Des précautions sont également à prendre sur ce point.

Toutefois, il convient de préciser qu'il faut prendre ces informations avec des précautions. Le prestataire ne pourra pas diffuser toutes les données. Il ne pourra pas diffuser, par exemple, les données qui porteraient atteinte à un droit de la propriété intellectuelle ou encore les données personnelles ou sensibles. Il conviendra également d'être vigilant quant à la diffusion de données stratégiques ou ayant une forte valeur économique ou permettant un avantage concurrentiel (il s'agit d'un peu de bon sens).

De plus, en cas d'absence de diffusion, certaines données administratives collectées dans le cadre d'une mission de service public peuvent être communiquées aux citoyens qui en font la demande.

Quels sont donc les précautions à prendre lors de la rédaction du contrat de prestation de service s'agissant de la diffusion des données ?

- **Prévoir une clause autorisant expressément le prestataire à diffuser les données générées** dans le cadre de la prestation. Cette clause devrait être proposée par défaut, à charge pour le donneur d'ordre de la négocier si besoin.
- Prévenir une exception de confidentialité. Si vous souhaitez diffuser une donnée, alors il faudra mettre en place une clause dans le contrat vous permettant de passer outre le secret professionnel. En effet, ce dernier correspond à un cas d'exception qui interdise la diffusion d'une donnée *(cf Partie 4)*.

- Dans le cadre d'une prestation de service présentant une importance pour l'établissement ou dans le cadre d'une longue prestation : formuler un plan de gestion qui aborderait la diffusion et la réutilisation des données.

De quoi s'agit-il en pratique ? (Exemple du Cirad) Les contrats de prestation de service mis en place par le Cirad font l'objet d'une clause attribuant la propriété des résultats au donneur d'ordre. Ainsi, les données collectées dans le cadre de ce contrat sont attribuées au donneur d'ordre. Le Cirad ne peut donc pas les diffuser si le donneur d'ordre n'a pas autorisé cette diffusion ou s'il n'a pas cédé ses droits.

2.2.2 Contrat de collaboration de recherche

Le contrat de collaboration de recherche est un acte juridique qui va permettre de créer une collaboration entre un laboratoire et un partenaire. Ce contrat doit être rédigé de façon à prévenir tout litige, tout en respectant les règles juridiques, administratives et comptables.
Il convient d'identifier plusieurs hypothèses s'agissant de la diffusion des données collectées dans l'exécution de ce type de contrat.

1/ Le contrat mentionne la question de la titularité des données et de leur diffusion. Il convient d'identifier plusieurs cas.

- La titularité des données est partagée entre les collaborateurs. Les données seront en copropriétés. Il faudra se référer au contrat afin de savoir si une répartition de la propriété a été établie.
- La titularité des données est attribuée à un des collaborateurs seulement. Les données sont la propriété de ce collaborateur. Il sera obligé de les diffuser si la une mission de service public. Dans le cas contraire, la diffusion des données est facultative, voire interdite si les données sont la propriété des autres collaborateurs. Si la recherche est financée pour plus de la moitié par un fonds public, en cas de communication des données, la réutilisable doit être obligatoirement libre. Cette obligation disparait lorsque le financement est pour la moitié d'origine privé.

2/ Le contrat ne mentionne pas la question de la titularité des données et de leur diffusion. Dans ce cas-là, les données seront obligatoirement diffusables seulement si la recherche correspond à une mission de service public. La communication reste néanmoins permise et est encouragée par le législateur quand elle n'est pas obligatoire. Si elle a lieu, le législateur exige seulement que la réutilisation soit libre lorsque la recherche est financée pour plus de la moitié par un fonds public.
Attention cependant car en l'absence d'accords signés, la titularité des données n'est pas strictement définie et il existe donc toujours un risque au moment de la diffusion de ces dernières.

Quels sont donc les précautions à prendre, lors de la rédaction du contrat de collaboration de recherche s'agissant de la diffusion des données ?
- Procéder à un recensement des connaissances et des données apportées par chacune des parties.
- S'interroger sur l'appartenance des résultats : différencier les résultats communs et les résultats propres.
- Établir une obligation de confidentialité réciproque afin de protéger les données de chacune des parties.

2.2.3 Contrat de consortium

Le contrat de consortium est constitué par des accords conclus entre plusieurs entreprises dans l'objectif d'obtenir et d'exécuter ensemble un projet commun sur un marché de fourniture de biens ou de services. Ce type de contrat permet d'éviter la création d'une société entre les parties. Cette volonté doit être expressément exprimée.
Il convient d'identifier plusieurs hypothèses s'agissant de la diffusion des données collectées dans l'exécution de ce type de contrat.

1/ Le contrat a établi un cadre concernant la titularité des données collectées. Dans ce cas-là il faudra se référer au contrat qui va distinguer selon que les données en question font l'objet d'un résultat commun ou propre.

- *Les données sont un résultat commun*. Elles appartiennent à toutes les parties du contrat car ils ont tous contribué à leur élaboration. Elles sont donc co-titulaire des données. Il conviendra d'établir une répartition de cette propriété en déterminant la quote-part des données qui revient à telle ou telle partie. *A contrario*, la diffusion des données sera soumise à l'accord de toutes les parties co-titulaires.
- *Les données sont un résultat propre*. Elles appartiennent à une seule des parties. Cette dernière peut les diffuser sous réserve des cas soumis à une interdiction de communication. Cependant, l'entreprise ne pourra diffuser que les données qui lui appartiennent, elle est interdite de diffuser les données appartenant seulement à l'une des autres parties.

La diffusion est obligatoire lorsque le contrat de consortium a pour objet une mission de service public. Dans le cas contraire, la diffusion des données est facultative mais pas interdite, sous réserve toujours des cas interdits mentionnés par la loi.

2/ Le contrat n'a pas établi de cadre concernant la titularité des données collectées. Dans ce cas-là, il faut appliquer strictement la LRN. Dans le cas où la recherche qui a conduit la collecte des données répond d'une mission de service public, l'établissement devra obligatoirement diffuser les données qui lui appartiennent. *A contrario*, si la recherche ne répond pas d'une mission de service public, alors la diffusion est facultative. Elle sera néanmoins interdite dans le cas où les données en question sont la propriété d'un autre établissement.
Attention cependant car en l'absence d'accords signés, la titularité des données n'est pas strictement définie et il existe donc toujours un risque au moment de la diffusion de ces dernières.

Quels sont donc nos conseils, les précautions à prendre, lors de la rédaction du contrat de la prestation de service s'agissant de la diffusion des données ?
- S'interroger sur l'appartenance des résultats : résultats communs et résultats propres.
- Effectuer une répartition en quote-part de la titularité des données correspondant à des résultats communs.

En identifiant dans un premier temps l'origine et la titularité des données, vous pouvez identifier si votre cas répond aux règles instituées par la LRN. **Dans un second temps, il faut déterminer si les données sont de nature diffusable ou si elles font l'objet d'une interdiction**.

3 QU'EST-CE QUE L'*OPEN DATA* ?

Il est important de comprendre la notion d'*Open Data* dans cette note méthodologique car c'est un terme qui peut paraitre difficile à cerner. De plus, la LRN y fait souvent référence, pourtant elle ne définit pas cette notion. Afin de faciliter la lecture des développements qui vont suivre, il est nécessaire de comprendre ce que l'on entend par « *Open Data* ». Ainsi, commençons par étudier la notion même (3.1) avant d'en expliquer son usage dans la recherche (3.2).

3.1 LA NOTION DE L'*OPEN DATA*.

L'*Open Data* correspond à des données auxquelles l'accès est totalement public et libre de droit, au même titre que l'exploitation et la réutilisation de ces données *(cf Partie 5)*. Plus largement, l'Open Knowledge Foundatation donne en 2005 une définition de *l'Open Data* en indiquant que les critères essentiels sont la disponibilité, la réutilisation et la distribution, ainsi que la participation universelle. Ces trois critères sont l'essence de *l'Open Data* car ils permettent l'interopérabilité entre les données.

Qu'est-ce que l'interopérabilité ? Cette notion désigne la capacité de différentes entreprises ou systèmes, à travailler ensemble. En somme, l'interopérabilité est la capacité d'articuler différents ensembles de données. Néanmoins, les bases de données restent indépendantes.

L'*Open Data* est une démarche de communication sans attendre la demande d'un utilisateur. De plus, cette communication doit se faire sans restriction ni technique (c'est-à-dire dans un format ouvert), ni juridique, ni financière injustifiée.

Des textes européens et nationaux définissent l'*Open Data* en identifiant 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'exceptions[26].

[26] Directive 2013/37/UE dite « PSI » pour Public Sector Information (en Europe)

Les « données ouvertes » ont été définies le 3 mai 2014 par la Commission générale de terminologie et de néologie comme les « données qu'un organisme met à la disposition de tous sous forme de fichiers numériques afin de permettre leur réutilisation »[27]. Les données ouvertes ne sont donc pas nécessairement des données collectées par l'Administration de l'État et peuvent provenir d'organismes publics tout comme privés. Cependant, il convient de préciser que la LRN s'est intéressée plus précisément à la question de l'*Open Data* sous l'angle des données publiques. C'est pour cette raison que l'on retrouve l'ensemble de ces nouvelles règles principalement au sein du Code des relations entre le public et l'administration (« CRPA » ci-après).

L'*Open Data* permet alors une **plus grande transparence de l'action publique** et garantit un meilleur **contrôle démocratique** facteur d'un mouvement de progrès scientifiques. En effet, il est évident qu'une confrontation directe entre les droits de propriété intellectuelle et l'Open Data peut subvenir. Les premiers récompensant l'inventeur en lui offrant un monopole sur sa découverte, le second l'obligeant à en faire bénéficier tout le monde. En réalité, ce contrôle démocratique illustre un équilibre trouvé grâce à l'utilisation de la propriété intellectuelle qui va encourager l'innovation par l'attribution d'un titre, mais qui en limite les pouvoirs en faveur du plus grand nombre, et en l'espèce, de la recherche. La limitation du monopole conféré par le droit de la propriété intellectuelle et l'Open Data bénéficient donc au corps social. L'idée est que les données financées grâce à des fonds publics devraient profiter gratuitement à l'ensemble des contributeurs. D'ailleurs quels sont les rapports entre *l'Open Data* et la recherche ?

Loi « CADA » n°78-753 du 17 juillet 1978 codifiée le 19 mars 2016 dans le Code des relations entre le public et l'administration (Livre III, titre 2e du CRPA) (En France)

[27] ORF n°0103 du 3 mai 2014

3.2 L'*OPEN DATA* ET LA RECHERCHE, LES APPORTS DE LA LOI POUR UNE RÉPUBLIQUE NUMÉRIQUE.

Le mouvement en faveur de *l'Open Data* a commencé à se dessiner en France dès 2011. Les administrations centrales ont été obligées de mettre en ligne, par le biais de la plateforme data.gouv.fr, un nombre croissant de jeux de données sous Licence Ouverte, dite aussi « *Open Licence* ». L'ensemble des administrations publiques n'étaient pas obligées de diffuser leurs données, certaines bénéficiant d'une dérogation avec la loi Valter. C'est ce qu'est venue modifier la LRN en renforçant le principe de de *l'Open Data* à toutes les administrations publiques.

Quels sont les objectifs du législateur ? Ce renforcement du principe de diffusion par défaut a pour objectif de garantir la transparence de l'État, de valoriser les données publiques tout en favorisant le développement d'activités privées et l'émergence de nouveaux opérateurs économiques. L'interopérabilité des données va ainsi permettre à d'autres acteurs, publics et privés, de réutiliser les données, de les comparer avec les leurs, et d'ainsi en améliorer la qualité, l'analyse et de gagner du temps sur leurs propres recherches.

Conclusion.

Il y a deux conditions préliminaires à la diffusion des données selon les principes de *l'Open Data* :

- **Les données doivent être réalisées dans le cadre de la mission de service public de l'établissement (important pour les EPIC)**
- **Les données doivent achevées, c'est-à-dire stable et non susceptible de modification.**

4 LES OBLIGATION DE LA LOI POUR UNE RÉPUBLIQUE NUMÉRIQUE.

Lorsque la nature de la donnée est identifiée, c'est-à-dire lorsque l'on sait si elle émane d'un contrat ou non et, si elle a une origine publique ou privée (*cf Partie 2*), il convient d'analyser les données au cas par cas aux fins d'identifier si elles font l'objet d'une **diffusion obligatoire** (4.1), **facultative** (4.3) ou **interdite** (4.2). La LRN n'est pas précise sur quelques notions, et certaines questions restent soulevées. Il convient donc de toujours prendre des précautions lors de la diffusion des données et bases de données.

Que faut-il entendre par « communication » des données ? Le terme de communication peut être compris de façon large lorsque les données sont obligatoirement diffusables ou, de façon plus particulière lorsque c'est un citoyen qui en fait expressément la demande lorsque les données ne font pas l'objet d'une diffusion publique. Ainsi, lorsque la communication est large, cela signifie que toute personne aura accès aux données car elles sont diffusées publiquement. *A contrario*, si cette communication est particulière, cela signifie que seule la personne qui l'a demandé aura accès aux données.

Que faut-il entendre par « diffusion » des données ? Il s'agit de la mise en ligne sur des plateformes dédiées des données mentionnées par la LRN. Cependant, il convient de préciser que le texte de la LRN ne fait pas mention du terme précis « en ligne », devant caractérisé une diffusion sur internet. Il semblerait alors qu'il soit possible de comprendre une diffusion au sens large des données. Il pourrait donc être considérée qu'une telle diffusion ait été faite par le biais d'une base de données commerciale.

→ **Annexe 2 : Logigramme « Diffusion des données selon la LRN ».**

4.1 UNE COMMUNICATION OBLIGATOIRE PAR PRINCIPE.

4.1.1 Qui est concerné ?

Initialement, la loi Valter avait institué une transparence au sein des administrations publiques, leur imposant la diffusion de leurs données. L'article 11 permettait aux universités et aux établissement de recherche de déroger à cette règle. La LRN de 2016 est venue abroger cet article et impose désormais **la diffusion par défaut aux établissement pratiquant des recherches répondant à une mission de service public**. Même lorsqu'une recherche ne fait pas l'objet d'une mission de service public, ces établissements de recherche sont **encouragés** par le législateur à diffuser le résultat de leurs travaux, qu'ils soient financés pour plus de la moitié par un fonds public ou privé. La loi émet néanmoins une différence entre ces deux sortes de financement, car lorsque l'établissement décide de diffuser, si la recherche était financée pour plus de la moitié par un fonds public, alors **la réutilisation des documents partagés doit être libre**. Cela ne sera pas exigé lorsque la recherche est financée pour plus de la moitié par un fonds d'origine privé.

La question est intéressante en ce qui concerne **les Établissements Publics à caractère Industriel et Commercial** (« EPIC » ci-après) car certains des financements sont d'origine privée, notamment lorsqu'il s'agit de recherches effectuées en prestation de service. Cela étant dit, ces EPIC restent des établissements publics et ne font plus l'objet de dérogation. Ainsi, il convient d'analyser au cas par cas les données que ces EPIC devront, pourront ou ne devront pas diffuser.

La LRN[28] permet cependant aux collectivités territoriales de moins de 3 500 habitants et aux personnes morales chargées d'une mission de service public, dont le nombre d'agents ou de salariés est inférieur à un seuil fixé par le pouvoir règlementaire[29], de déroger à la règle de diffusion par défaut.

[28] Article 6 LRN et L.312-1-1 du CRPA.
[29] Décret n°2016-1922 du 28 décembre 2016 relatif à la publication en ligne des documents administratifs.

4.1.2 Quels sont les documents et données concernés ?

Le droit pour toute personne d'accéder aux « documents administratifs » existe depuis 1978. La LRN du 7 octobre 2016 va venir élargir ce champ d'application.

Avant tout, que sont les « documents administratifs » ? Initialement, il s'agit des documents 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 mission de service public. La LRN reprend cette définition[30] en intégrant, d'une part, les « codes sources »[31], et d'autre part, les règles définissant le traitement algorithme ainsi que les principales caractéristiques de sa mise en œuvre au bénéfice de l'usager qui fait l'objet d'une décision individuelle prise sur le fondement de ce traitement algorithme[32]

La grande innovation de la LRN est de consacrer une ouverture sectorielle et « par défaut » de certains documents administratifs. L'article L.312-1-1 du CRPA (Code des relations entre l'Administration et le public) indique que les administrations citées plus haut sont tenues de :
- Diffuser les principaux documents produits ou détenues par les administrations contenant des informations publiques et figurant dans un répertoire mis à jour chaque année ;
- Diffuser les bases de données produites ou reçues par chaque administration ;
- Diffuser les données, mises à jour de façon régulière, dont la publication présente un intérêt économique, social, sanitaire ou environnemental.

Il faut préciser que cette diffusion est limitée car seuls les documents existant sous forme électronique sont concernés.

Le législateur[33] indique également qu'il s'agit de tous les documents, quelle que soit leur date, leur lieu de conservation, leur forme et leur support, qui sont produits ou reçus, dans le cadre

[30] Article 6 de la LRN
[31] Avis par la Commission d'accès aux documents administratifs (CADA) du 8 janvier 2015 + Article L.300-2 CRPA al 1er
[32] Article L.311-3-1 CRPA
[33] Article L.300-3 du CRPA et la loi CADA.

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. En addition de cela, le juge[34] a précisé qu'il s'agit de toutes les mesures *« régissant l'organisation du service public »*, en exigeant une relation étroite avec la délivrance d'un service public.
Ainsi, les données de référence doivent également être mises à disposition[35]. Un décret[36] est venu préciser que cette mise à disposition en vue de faciliter leur réutilisation constitue une mission de service public relevant de l'État.

Des exemples sont ainsi donnés, de façon non limitative par le législateur[37]. Il s'agit de dossiers, de rapports, d'études, de comptes rendus, de procès-verbaux, de statistiques, de directives, d'instructions, de circulaires, de notes et réponses ministérielles, de correspondances, d'avis, de précisions, des décisions, ou encore de codes sources comme l'a précisé le juge[38].

En somme, deux composantes sont à retenir pour qualifier les données produites par des établissements de recherche et d'enseignement en « documents administratifs » :

- **Les données sont collectées par ces établissements dans le cadre d'une mission de service public. Ces dernières sont communicables à toute personne qui en fait la demande, sauf exceptions légales** *(cf 4.2)***.**
- **Les données sont achevées et ne sont pas encore accessibles au public. Elles correspondent à des données brutes, à des données élaborées ou encore à des métadonnées, excluant alors les documents dits préparatoires.**

Qu'est-ce que les « données brutes » ? Il s'agit de données préliminaires issues d'une expérimentation, d'un procédé ou encore d'une enquête. Elles peuvent être considérées comme des données communicables de la recherche.

[34] Tribunal des Conflits 15 décembre 2008 Voisin c/ RATP req n°3662, Lebon
[35] Article 14 de la LRN
[36] Décret n°2017-331 du 14 mars 2017 relatif au service public de mise à disposition des données de référence et codifié à l'article L.321-4 du CRPA.
[37] Article L.300-2 du CRPA
[38] Tribunal administratif de Paris décision du 10 mars 2016

Quelles sont donc les conséquences de la qualification des données en « document administratif » ? Premièrement, la règlementation française va imposer une diffusion en ligne par défaut des documents administratifs sous forme électronique. De plus, il va y avoir une liberté d'accès pour toutes personnes qui en font la demande. Même lors d'une telle demande, les données doivent être analysées au cas par cas afin de ne pas porter atteinte à certains droits. Il faut faire preuve de prudence avant chaque diffusion afin de ne pas porter atteinte à des droits, des secrets ou autres engagements.

Ces diffusions sont limitées dans deux cas : si les documents correspondent à des cas présentés dans les articles L.311-5 et L.311-6 du CRPA ou s'ils ne sont pas disponibles sous forme électronique.

Qu'advient-il des écrits scientifiques tels les revues ou publication générées par les chercheurs ? La LRN[39] applique la même règle pour les données générées par des recherches. Ainsi, le chercheur possède désormais le droit de rendre ses recherches publiques malgré le fait que ses écrits aient été cédé à un éditeur. En effet, selon les termes de la loi, le critère qui rend possible le dépôt est le caractère public du financement de l'activité de recherche et non de la publication elle-même. L'interprétation littérale de la loi nous permet de penser que lorsque cette condition est remplie, alors le dépôt est possible même en cas de rémunération par l'éditeur.

Solution pour être sûr de ne pas se retrouver en conflit avec un éditeur. Il est possible de publier le contenu de l'article sur des archives ouvertes, telle que l'archive ouverte pluridisciplinaire *HAL*.

Attardons-nous quelque peu sur certaines données mentionnées dans le logigramme *(Annexe 2)* **pour plus de compréhension.**

[39] Article 17 de la LRN modifiant l'article L.533-4 I du Code la recherche.

1/ Les données achevées dont la publication présente un intérêt général. Cet intérêt général peut être identifié de trois manière selon la LRN.

- **Un intérêt économique**. Nous pouvons ici donner l'exemple des documents administratifs faisant état des subventions attribuées par l'État.
- **Un intérêt sanitaire**. Nous pouvons ici donner l'exemple des documents qui vont permettre d'avancer dans la recherche médicale ou qui font état d'une recherche sur l'atteinte de certains produits sur les populations.
- **Un intérêt environnemental**. Dans cette catégorie, nous pouvons traiter l'hypothèse des émissions de substances dans l'environnement. En d'autres termes, il s'agit des éléments qui pourraient avoir des incidences sur l'air, l'eau, le sol, etc… La Convention d'Aarhus vient fixer le contour de cette notion en indiquant 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 est repris par les législations européennes et françaises qui vont imposer aux Administrations une obligation de diffusion des documents lorsque ces derniers sont achevés. Dans le cas où ils ne le seraient pas, l'établissement sera tout de même obligé de fournir ces documents non achevés au citoyen qui en fait la demande.

  Il existe des d'exceptions. En effet, la communication peur être refusée :

  - En cas d'impact sur la conduite de la politique extérieure de la France, la sécurité publique ou la défense nationale ;
  - En cas d'impact sur des procédures juridictionnelles ;
  - En cas d'atteinte à des droits de propriété intellectuelle.

Quelques exemples :

* Enquête environnementale suite à une déclaration de mortalité massive d'abeilles[40] : la communication est obligatoire

* Enquête sur une épidémie de légionellose[41] : la communication a été refusée car les données recueillies n'étaient pas relatives à l'environnement.

[40] Avis CADA n°20130750 du 28 mars 2013.
[41] Avis CADA n°20090310 du 26 février 2009.

2/ La base de données. Le texte précise qu'il s'agit des bases de données *« mises à jour de façon régulière, qu'elles produisent ou qu'elles reçoivent, et qui ne font pas l'objet d'une diffusion publique par ailleurs »*. En interprétation de ce texte, il est possible de considérer qu'il s'agit des bases de données constituées dans le cadre d'une recherche financée pour plus de la moitié par un fonds public, ou répondant à une mission de service public. Il semblerait que cette de base de données puisse être diffusée même si elle n'est pas complète. Son fonctionnement quant à lui doit être achevé. Les informations que les chercheurs pourraient apporter à la base de données après sa diffusion doivent apparaitre par le biais d'une mise à jour régulière.

Attention cependant car le texte n'est pas précis et il convient de prendre des précautions lors de la diffusion.

3/ Les données géographiques. La directive INSPIRE *(cf Bibliographie où se trouve un cadre explicatif de cette Directive)* a imposé une obligation de publication et donc de diffusion de ces données. Elles doivent être rendues accessibles au public.

Comment ? Par le biais d'une publication de l'ensemble des données mais aussi des métadonnées correspondantes. Ces dernières sont également partagées entre les établissements**.**

À quoi correspondent-elles ? Il s'agit des données nécessaires au répertoire sur le territoire (exemple ; hydrographie); des données générales complémentaires (exemple : altimétrie, géologie); et enfin des données thématiques telles que vocation des sols, santé et sécurité des personnes, bâtiments, etc …

La diffusion de telles données est obligatoire par principe et incombe à l'établissement qui les a collectées. Dans le cas où de telles données sont réutilisées par un autre établissement pour une autre recherche, ce dernier n'est pas obligé de les publier[42]. Le Code de l'environnement indique en effet que *« lorsque plusieurs copies identiques d'une même série de données géographiques sont détenues par plusieurs autorités publiques ou en leur nom, le [chapitre faisant référence à la transposition de la directive] s'applique uniquement à la version de référence dont sont tirées les différentes copies ».*

[42] Article L.127-1 du Code de l'environnement transposant l'article 4-2 de la directive INSPIRE.

La diffusion de données géographiques peut être limitée dans le cas où il existe, par exemple, des droits de propriété intellectuelle ou lorsqu'il s'agit de données à caractère personnel (application du RGPD). Dans ce cas-là, l'autorisation préalable du titulaire des droits est nécessaire avant toute diffusion.

4/ Demande de communication. La LRN met en place un droit de communication au bénéfice de tout citoyen qui souhaite accéder à un document administratif qui ne fait pas l'objet d'une publication publique[43]. L'article L.311-2 du CRPA précise ensuite les conditions dans lesquelles une telle communication peut être autorisée. Ainsi seuls les documents achevés, sauf dans certains cas, peuvent faire l'objet d'une telle communication individuelle. Nous pouvons extraire du texte trois situations qu'il convient d'expliquer.

- **La demande de communication porte sur des données appartenant à votre établissement**. Dans ce cas, votre établissement doit répondre favorablement à la communication des données au citoyen qui en fait la demande. Cependant, une analyse au cas par cas des données doit être effectuée, et l'établissement ne peut pas communiquer des données qui porteraient atteinte à des droits ou des secrets faisant l'objet des exceptions de diffusion par défaut de la LRN *(cf 4.2)*.
- **La demande de communication porte sur des données n'appartenant pas à votre établissement.** Dans ce cas, votre établissement n'a pas à partager les données qui ne lui appartiennent pas. Cependant, il est tenu d'informer le titulaire de ces demandes. Ce dernier devra les communiquer à l'intéressé[44].
- **La demande de communication porte sur des données à caractère personnel**. Il faut ici distinguer ce qui émane de la demande de communication. Si la demande émane d'une personne tierce à ces données à caractère personnel, alors la communication devra être refusée. Cependant, si la demande émane de la personne concernée par ces données, alors l'établissement ne peut refuser la communication, notamment s'il s'agit d'une décision individuelle concernant cette personne[45]. La communication des données à caractère personnel doit donc se faire à l'attention unique de la personne concernée.

[43] Article L.311-1 du CRPA
[44] Article L.311-2 alinéa 6 du CRPA.
[45] Article L.311-3.

5/ Données relatives à un risque impactant les populations. Ces données peuvent être par exemple des documents d'analyse des dangers sanitaires pour les populations. Ce sont généralement des études qui vont permettre d'identifier des causes variées pouvant être dangereux pour une population. Ces données doivent être analysées au cas par cas de manière à ne pas porter atteinte à un secret présentant une exception à la diffusion par défaut instituée par la LRN.

6/ Le répertoire. La diffusion du répertoire doit se réaliser dans son intégralité.

4.2 UNE COMMUNICATION STRICTEMENT INTERDITE DANS CERTAINS CAS.

La LRN institue une diffusion par principe des données collectées par les établissements et personnes que nous avons identifiés plus haut. Cependant, comme il est rappelé à plusieurs reprises dans cette note, il est important de prendre des précautions lors de la diffusion des données car certains sont protégées par des secrets ou des droits personnels, et peuvent amener des conflits. L'analyse au cas par cas des données est donc nécessaire. Afin de mener à bien cette analyse, il convient de tenter d'éclaircir les cas d'exceptions identifiées par la LRN.

4.2.1 Données faisant l'objet d'un secret protégé par la loi.

L'interdiction de la diffusion de telles données est mise en place par les articles L.311-5 et L.311-6 du Code des relations entre le public et l'Administration (CRPA). C'est une exception au principe de communication instaurée par la LRN.

Le CRPA énumère les secrets protégés par la loi[46]. L'appréciation du refus de communication se fait au niveau du contenu des documents. Le Conseil d'État indique qu'il convient de chercher si, eu égard au contenu des documents demandés, leur divulgation risque effectivement de porter atteinte à un secret protégé par ces dispositions[47]. Il faut distinguer

[46] Article L.311-5 2° du CRPA.
[47] CE, 22 février 2013, Fédération chrétienne des témoins de Jéhovah France.

les secrets absolus et opposables à toute personne ; et les secrets relatifs et opposables aux tiers mais pas aux personnes intéressées.

4.2.1.1 Les secrets absolus et opposables à toute personne

1/ Secret de défense nationale. Le Code pénal[48] est venu classifier certaines données comme étant secrètes. Cette classification peut aller au-delà de données purement militaires. La diffusion de tout ou partie de ces informations sans autorisation de l'émetteur est considérée comme un délit et peut faire l'objet d'une sanction pénale. Certains documents peuvent, sans pour autant être classifiés comme secret défense, porter la mention « diffusion restreinte ». Cette mention va limiter la communication à un nombre limité de personnes qui sont destinées à en connaitre le contenu du fait de leurs fonctions.

2/ Secret relatif à la conduite extérieure de la France. Dans cette hypothèse la Commission d'accès aux documents administratifs (CADA) indique qu'il faut s'attacher plus au contenu du document même qu'à sa forme et son contexte de communication.
Quels sont les types de documents qui correspondent à cette catégorie ? Il s'agit en particulier, de documents ayant servi de base à des négociations internationales ou retraçant de telles négociations, de documents analysant une situation internationale ou le comportement d'un État, ou encore de documents traduisant la politique extérieure de la France

3/ Secret relatif à la sûreté de l'État. La sûreté de l'État comprend plusieurs possibilités qui sont énumérées par la loi. Cette notion est donc applicable dans le cas de sécurité publique, des biens de l'établissement, des personnes ou encore, des systèmes d'information de l'établissement.
→ 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. ;*

[48] Article 413-9 du Code pénal.

→ 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. ;*

→ 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'information19 de l'établissement *Exemples : diffusion des bilans annuels des failles de sécurité des systèmes d'information, architecture des systèmes d'information.*

4/ Secret lié au potentiel scientifique et technique de la nation. Il s'agit de l'ensemble des biens matériels et immatériels propres à l'activité scientifique fondamentale et appliquée au développement technologique de la nation française. Ces éléments constituent des intérêts fondamentaux de la nation et sont définis par la Code pénal à l'article 410-1. La diffusion publique de tels éléments est donc proscrite par la loi pénale mais aussi par la LRN

→ *Pour en savoir plus sur ces éléments, un guide a été rédigé par la SGDSN*
https://www.ssi.gouv.fr/uploads/2018/05/guide_protection_scientifique_technique_nation_anssi-pa-049_v1.pdf

4.2.1.2 Les secrets relatifs et opposables aux tiers mais pas aux personnes intéressées

1/ Secret industriel et commercial et secret des affaires. Cette catégorie doit faire l'objet d'une analyse au cas par cas. La diffusion sera interdite par exemple en cas de secret médical, de l'instruction, de secret bancaire et fiscal, ou encore de correspondances (liste non limitative). Cependant, il existe des exceptions et le secret peut être levé dans le cas d'une autorisation de la personne concernée par l'information, et sous réserve que soient préservés la protection des personnes, la santé publique ainsi que l'ordre public et le bon déroulement des procédures de justice.

Qu'est-ce que le secret des affaires ? Il s'agit d'une notion qui n'est pas clairement définie en droit français mais dont un certain nombre d'éléments doivent être pris en compte.

Au niveau européen, les parlementaires se sont réunis autour de trois conditions permettant d'indiquer qu'une information constitue un secret d'affaire[49].

- 1 | Les informations sont secrètes dans leur globalité ou dans la configuration. Elles ne sont généralement pas connues de personnes appartenant aux milieux qui s'occupent normalement de ce genre d'informations, 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.

Qu'est-ce que le secret industriel et commercial ? C'est une forme de secret d'affaire en réalité. En France, la CADA a tenté de définir les secret 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 »* :

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

Qu'advient-il de la diffusion des données faisant l'objet d'un tel secret ? En dehors des personnes intéressées[50], 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. La loi réprime par des sanctions pénales le fait de divulguer des informations couvertes par « les secrets » cités ci-dessus.

La CADA a précisé que *« l'ensemble des pièces administratives et comptables relatives aux missions de service public assurées par des établissements publics à caractère industriel et*

[49] Directive 206/943 du 8 juin 2016

[50] Article 6 de la LRN

commercial constituent des documents administratifs »[51]. En somme, les documents sont communicables à toute personne qui en fait la demande une fois les mentions protégées par le secret occulté. Une telle communication n'est pas obligatoirement diffusable à tout public.

Il convient de préciser que les administrations publiques font bien partie des entités pouvant élaborer des données protégées par un secret d'affaire, et doivent donc analyser les données qu'elles souhaitent diffuser en ce sens également[52]

2/ Secret médical et relatif à la vie privée des individus. Ce secret fait l'objet d'une protection particulière par le Règlement Général de la Protection des Données (RGPD). **Vous pouvez retrouver l'analyse de cette catégorie ci-dessous (4.2.3).**

4.2.2 Données protégées par un droit de propriété intellectuelle.

1/ Les données protégées par le droit d'auteur. Le code de la propriété intellectuelle[53] protège certaines données et documents administratifs par le biais du droit d'auteur. Cette protection peut limiter la diffusion de ces documents[54].

Quels sont les documents qui peuvent être susceptibles d'être protégés par le droit d'auteur ? Les études et travaux issus de la recherche scientifique sont appréhendés comme des œuvres écrites et sont donc susceptibles d'être protégés par le droit d'auteur. Les scientifiques qui publient dans des revues, par exemple, sont amenés à opérer une cession des droits d'auteurs au profit des éditeurs.

En principe, la protection par le droit d'auteur interdit l'établissement à publier ou à diffuser les données protégées, sauf avec le consentement préalable de l'auteur. Cependant, ce principe souffre de quelques exceptions. Afin de savoir si le document protégé par un droit d'auteur peut être diffusé ou non, il convient d'envisager plusieurs hypothèses :

---

[51] Avis CADA n°20175033 du 11 janvier 2018

[52] Article L.311-6 1° du CRPA *« le cas où il s'agit d'une administration mentionnée au premier alinéa de l'article L.300-2 CRPA est soumise à concurrence ».*

[53] Article L.111-2 du Code de la propriété intellectuelle.

[54] Article L.311-4 du CRPA *les documents administratifs sont « communicables […] sous réserve des droits de propriété littéraire et artistique ».*

- **L'œuvre est un logiciel créé au sein d'un établissement public**. Dans cette hypothèse, c'est à l'établissement de décider s'il souhaite publier ou non ce logiciel. En somme, la diffusion est permise mais n'est pas obligatoire.
- **L'œuvre appartient en tout ou partie à un tiers.** Dans le cas où l'œuvre a été créée grâce à des fonds publics, la communication des données est obligatoire en cas de demande de communication d'un citoyen. Elle n'est que facultative s'agissant d'une diffusion au grand public. *A contrario*, si l'œuvre émane d'une recherche financée par des fonds privés, alors la diffusion sera seulement facultative et sous condition de l'accord préalable de l'établissement.
- **L'établissement public dispose d'un droit d'exploitation des œuvres.** Si l'établissement dispose de ces œuvres aux fins de répondre à une mission de service public, alors la diffusion des données n'est que facultative mais n'est pas interdite.
- **Les œuvres appartiennent aux scientifiques.** Dans cette hypothèse, l'établissement ne possède aucun droit sur l'œuvre protégée. La diffusion est donc interdite.

2/ Les données incorporées dans une base de données protégées par le droit *sui generis* du producteur de la base. Les bases de données peuvent être protégées par le droit d'auteur mais également par un droit *sui generis* du producteur de la base. Cette protection se retrouve dans le Code de la propriété intellectuelle[55]. Concernant la diffusion de ces bases de données, il faudra différencier selon que la diffusion concerne :

- **Une extraction de données non substantielle**. Dans ce cas, la diffusion est possible et l'accès à la base de données doit donc être autorisé, sous réserve cependant de l'existence d'une protection spécifique des données. Le producteur de la base de données ne peut s'opposer à sa réutilisation.
- **Une extraction de données substantielle**. Ce cas-là est plus complexe. En effet, la base de données peut être diffusée au sein de l'établissement même. S'agissant des établissements publics, la diffusion est possible sous demande. *A contrario*, elle est interdite en ce qui concerne les établissements privés. Ces derniers pourront cependant permettre un accès à la base en cas de demande d'autorisation et sous réserve de données

[55] Article L.342-3 du Code de la propriété intellectuelle.

protégées par un autre droit. En somme, que l'établissement soit public ou privé, il n'est pas obligé de diffuser, en ligne, ses bases de données au public. Cependant, en cas de demande particulière d'un citoyen, l'établissement public ne pourra s'opposer au partage de la base de données, alors que l'établissement privé pourra lui s'y opposer.

3/ Les données protégées par un droit de propriété industrielle. La LRN a laissé un grand flou concernant le droit de la propriété industrielle. Il est difficile de savoir ce qui peut ou pas être diffusé. Afin d'être le plus méticuleux possible, il convient de se référer aux divers contrats qui régissent ces données. Si des contrats sont passés en prestation, collaboration ou encore en consortium avec d'autres acteurs, il faudra vérifier si une cession des droits est faite, et à qui appartient les données. Cela permettra de savoir si notre établissement peut les diffuser ou non. Il semble important de préciser que toutes atteintes à de tels droits sont susceptibles de sanction, il convient donc d'être prudent lors de la diffusion de ces données.

4.2.3 Données à caractère personnel.

Depuis la mise en place du Règlement Général pour la Protection des Données (RGPD), les données à caractère personnel suivent un régime particulier et bénéficient d'une protection renforcée.

→ Un guide explicatif sur les données à caractère vous est proposé en annexe. Il vous permettra de comprendre ce qu'est une donnée à caractère personnel, et vous indiquera les précautions à prendre lors du traitement de ces données *(cf Annexe 3)***.**

Dans la cadre d'une diffusion ou communication de documents administratifs, il convient d'établir une analyse particulière sur ce type de données afin de ne pas porter atteinte aux droits des personnes concernées.

1/ Données relatives à une personne physique. Il faudra distinguer selon que la personne concernée a donné son consentement ou non au traitement des données. Le consentement doit remplir des formalités strictes *(cf Annexe 3)*. Si les formalités ne sont pas remplies et que

le consentement a pourtant été donné, à l'oral par exemple, le traitement de la donnée ne sera pas considéré comme consenti.

- Le consentement a été donné par la personne concernée et est valide. Le traitement est légal et il est possible de diffuser cette donnée si cette diffusion répond à la finalité pour laquelle le consentement a été opéré.
- Le consentement n'a pas été fourni par la personne concernée. Dans cette hypothèse, la diffusion est en principe interdite car la personne n'a pas consenti au traitement même de sa donnée. Cependant, il existe des cas où il est possible de traiter une donnée puis de la diffuser sans que la personne concernée n'est consentie à ce traitement. C'est le cas lorsqu'il existe une règlementation, ou un contrat qui répond d'une mission de service public, et qui justifie donc la collecte des données. C'est également le cas lorsqu'il existe un intérêt légitime à la collecte. Dans cette dernière hypothèse, l'établissement doit être en mesure de prouver cet intérêt légitime.

Il faut cependant rappeler que ces traitements de données, consentis expressément ou non par les personnes concernées, doivent respecter les exigences de finalité, licéité, durée et d'information misent en place par le RGPD *(cf Annexe 3)*.

2/ La donnée est une photographie où des personnes sont reconnaissables. C'est la loi Informatique et Liberté et le RGPD qui viennent encadrer cette situation. Tout comme pour le premier point, il faut distinguer selon que la personne concernée a donné son consentement préalablement à la prise de la photographie ou non. Si la personne concernée a expressément consenti par écrit alors l'établissement est en droit de la diffuser. Cette diffusion n'est pas obligatoire. *A contrario*, si aucun consentement n'a été exprimé, ou si ce consentement ne répond pas aux formalités exigées par le RGPD, alors la diffusion de cette photographie est interdite.

Quelles sont les solutions offertes à mon établissement qui a besoin de diffuser des données à caractère personnel ? Il vous suffit d'anonymiser les données protégées. Cette anonymisation doit se faire préalablement à la communication du document en question. La

CNIL[56] a en effet indiqué que les données à caractère personnel peuvent faire l'objet d'une diffusion si ces dernière sont anonymisées préalablement à la communication.
De plus, un décret d'application de la LRN[57] a communiqué une liste limitative de données à caractère personnel qui peuvent être communiquées sans besoin d'anonymisation.

4.3 UNE COMMUNICATION FACULTATIVE LORSQU'ELLE N'EST PAS EXIGÉE PAR LA LOI.

Le principe est celui de la diffusion. Cependant, comme nous l'avons constaté tout au long de cette note, de nombreuses exceptions existent et limitent l'obligation de diffusion imposée par la LRN. Certaines diffusions ne seront, en effet, pas obligatoires, et c'est alors à l'établissement de décider s'il souhaite diffuser les données en question. La LRN encourage la diffusion des documents qui ne font pas l'objet d'une communication obligatoire. Ainsi, nous pouvons identifier, en complément de ceux qui ont déjà été invoqués, des hypothèses où la diffusion des données sera facultative.

1/ Une diffusion qui dépend de la politique de l'établissement. Certains établissements vont souhaiter plus facilement partager le fruit de leurs recherches. C'est le cas notamment des établissements qui ont pour habitude de collaborer avec d'autres établissements par exemple. D'autres projets seront plus farouchement protégés en fonction des apports qu'ils ont à offrir à la science.

2/ Les données antérieures à la loi pour une République numérique. En droit, seule la loi pénale peut être rétroactive. Ainsi, nous pouvons déduire que la LRN ne s'applique que pour les documents administratifs élaborés une fois la loi entrée en vigueur, c'est-à-dire le 7 octobre 2016. Une diffusion de base données, par exemple, peut conduire à publier des

[56] Délibération n°2018-101 du 15 mars 2018 portant avis sur projet de décret pris pour l'application de l'article 6 de la loi N°2016-1321 du 7 octobre 2016 pour une République Numérique
[57] Décret n°2018-1117 du 10 décembre 2018 relatif aux catégories de documents administratifs pouvant être rendus publics sans faire l'objet d'un processus d'anonymisation.

données antérieures à cette loi. Cela n'est pas interdit, cette diffusion est facultative et encouragée.

3/ La reproduction ou copie de données à certaines conditions. Au cours de recherches, des établissements sont susceptibles d'utiliser des données ou des documents confectionnés par d'autres entités. Ces documents vont être reproduits ou copiés et peuvent faire l'objet d'une diffusion en complément des nouvelles recherches effectuées. Cette diffusion n'est pas interdite, elle est néanmoins facultative. Il convient également d'indiquer que cette diffusion ne pourra s'effectuer qu'à certaines conditions, c'est-à-dire en respectant, par exemple, les droits de propriété intellectuelle qui peuvent exister.

4/ Les données statistiques. Les établissements peuvent mettre en place des statistiques grâce aux données qu'ils auraient collectée par le biais de leurs recherches ou en collaboration avec d'autres établissements. Ces statistiques ne sont pas obligatoirement diffusables mais ne font pas l'objet d'interdiction. L'établissement pourra donc les diffuser selon sa convenance en prenant garde cependant de respecter le « secret statistique » ou d'obtenir une dérogation de la part du Comité du secret statistique.

Qu'est-ce qu'un secret statique ? Ce secret est défini comme la loi[58] et il est indiqué qu'il est interdit, pendant une durée de soixante-quinze ans, toute communication de données ayant trait à la vie personnelle et familiale, et plus généralement, aux faits et comportements d'ordre privé recueillis au moyen d'une enquête statistique. Des dérogations peuvent être accordées sur avis du Comité statistique.

[58] Loi n°57-711 du 7 juin 1951 modifiée.

5 LA LIBRE RÉUTILISATION INSTITUÉE PAR LA LRN

La LRN[59] ne fait pas que rendre la diffusion de données obligatoire par défaut, elle en autorise également la réutilisation libre et gratuite[60], à toutes fins, mêmes commerciales. Pour ce dernier point, la LRN indique en effet *« à d'autres fins que celles de la mission de service public pour les besoins de laquelle les documents ont été produits ou reçus »*[61]. Les administrés ne devront toutefois pas altérer ou dénaturer le sens des données publiques[62].

Quand est-il de la réutilisation des données spécifiques à la recherche ? La LRN exige une **« réutilisation libre »** des données financées **pour plus de la moitié par un fonds public**. Pourtant, le CRPA indique que les données de la recherche peuvent être réutilisées si les établissements le choisissent. En effet, l'article L.322-1 précise *« sauf accord de l'administration »*. Mais à défaut du choix d'une licence comme la CC0, le plancher de conditions fixées par l'article va continuer à s'appliquer.
Pour la doctrine, l'article 30 de la LRN, relatif à la réutilisation des données, doit s'interpréter ainsi *« Sauf si les établissements de recherches en ont décidé autrement, notamment en utilisant une licence comme la CC0, alors la réutilisation de données de recherche publiées est libre, sous réserve de mentionner la source et de ne pas les altérer ou les dénaturer ».*

La LRN laisse place à beaucoup d'interprétations. *Ainsi, que devons-nous retenir ?* La réutilisation doit être obligatoirement libre lorsque les données diffusées (obligatoirement ou facultativement) sont financées pour plus de la moitié par un fonds public.

Quand est-il de la réutilisation des bases de données ? Les bases de données protégées par des droits de propriété intellectuelle (pour rappel : droit d'auteur et droit *sui generis* du producteur) ne peuvent être réutilisables sous le fondement de l'article 30. Elles peuvent l'être sous le fondement de l'exception *Text et data Mining* prévue à l'article 38, relatif à la réutilisation de copie ou reproduction de documents.

[59] Article 30 de la LRN
[60] Article 11 de la LRN
[61] Article L.321-1 du CRPA
[62] Article L.322-1 du CRPA

Cependant, il convient de préciser que l'application de l'article 30 est très incertaine car le droit de producteur des bases de données est souvent aléatoire dans sa mise en œuvre. La jurisprudence est en effet fluctuante sur la question de la protection des bases de données. En conséquence, il n'est pas simple *a priori* de déterminer si telle ou telle base bénéficie ou non de la protection.

Quelles sont les limites à la libre réutilisation ? La loi, va cependant limiter le périmètre d'application en excluant certains types de données : libre sauf si les données « sont protégées par un droit spécifique ou une réglementation particulière ». Ces droits correspondent à ceux que nous avons évoqué plus haut. Il s'agit par exemple des données à caractère personnel.

Conclusion.

La loi pour une République numérique du 7 octobre 2016 est venue préciser les conditions de la réutilisation de la donnée. Cette dernière est réutilisable lorsque :

- **Elle est issue d'une activité de recherche financée au moins pour moitié par des fonds publics, ou effectuée dans le cadre d'une mission de service public ;**
- **Elle n'est pas protégée par un droit spécifique ;**
- **Elle est rendue publique par le chercheur ou l'établissement.**

CONCLUSION SPÉCIFIQUE AU SYSTÈME DATA4C+

Le système DATA4C+ a la particularité de regrouper plusieurs bases de données provenant d'organismes différents et soumis à des règles qui leurs sont spécifiques.

Quelle est la difficulté ? Au sein du système et même d'une seule base de données, il est possible de retrouver des données qui font l'objet d'une diffusion obligatoire, facultative ou encore interdite. De plus, certaines données sont antérieures à 2016 et d'autres non. L'INRAe, du fait de son activité, peut collecter des données qui vont être attachées à une mission de service public alors qu'a contrario, le Cirad peut lui collecter des données générées à la suite d'une prestation de service.

Comment réussir à répondre aux exigences de la LRN avec un système présentant tant de complexité ? La solution serait de mettre en place une technique qui permettrait de diffuser certaines données, au grand public, ou juste à certains scientifiques ; alors que d'autres données resteraient elles strictement confidentielles et ne seraient donc pas partagées. Il pourrait s'agir des données dont la diffusion est interdite, mais également de données dont vous estimerez risquée leur communication, du fait d'enjeux scientifiques, environnementaux, économiques ou encore politiques par exemple. Ainsi, il convient de mettre en place une stratégie d'analyser des risques autour de la diffusion des données de DATA4C+ afin de d'identifier spécifiquement les données que vous ne souhaitez pas partager. Si vous vous positionnez dans une volonté et une idéologie de partage maximum des données pour favoriser la Sciences, il faut alors prendre des précautions autour des données interdites, en favorisant l'anonymisation de certaines (s'agissant des données à caractère personnel par exemple). Il faudra également choisir une licence qui permettra la réutilisation libre de ces données, surtout si elles sont financées pour plus de la moitié par un fonds public la volonté serait de partager aux maximum les résultats des recherches et faire avancer la Sciences, il va également falloir trouver une licence qui va permettre la réutilisation libre des données générées par un financement majoritairement constitué de fonds publics.

GUIDE DES BONNES PRATIQUES À METTRE EN PLACE.

Étape 1 | Sur l'origine des données :

- Suis-je collecteur des données ?
- Sinon, ai-je l'autorisation du fournisseur ?
    - Si oui : examiner le contrat (que puis-je faire ? diffuser ? réutiliser ?)
    - Si non : prendre contact avec le fournisseur.

Étape 2 | Réexaminer chaque type de données (personnelle, etc.) et m'assurer que j'ai bien le droit de diffuser.

Étape 3 | Si je crée une base de données :

- Anonymiser les données personnelles.
- Séparer les données par type pour pouvoir en diffuser certaines et pas d'autres (gestion des droits d'accès) + prévoir un outil de suivi des données avec les informations essentielles (diffusables, non diffusables, issus de tel ou tel projet).

Étape 4 | La base est réalisée en partenariat : m'entendre avec les autres parties sur les conditions de diffusion des données et de la base. Il est conseillé de faire un data management plan (plan de gestion des données) avec eux ou de définir les règles de réutilisation des données dans l'accord de consortium.

Étape 5 | Je peux diffuser, est ce que je dois mettre en place une licence de réutilisation ?
La loi pour une République Numérique rend les le recours aux licences libres pour la diffusion des données de la recherche (telles Open Licence ou encore ODbl) inutiles car elle instaure une libre réutilisation. On parle d'ailleurs de « licence à l'état natif » car les dispositions de la loi se déclenchent automatiquement par l'acte de publication. Désormais, seule la licence CC0 est compatible avec le régime légal des données de recherche.

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

- Respecter l'intégrité des données (absence d'altération, absence de dénaturation du sens) ;
- Faire mention de la source des données ;
- Veiller à ce que l'indication de la date de dernière mise à jour soit bien présente.

Ainsi, les réutilisations commerciales ne peuvent pas être empêchées.

Comment diffuser pour ne pas nuire à la réutilisation des données ?

Il faut mettre en place un standard ouvert. Cette notion est définie par la loi pour la confiance dans l'économique numérique qui indique que l'« on entend par standard ouvert tout protocole de communication, d'interconnexion ou d'échange et tout format de données interopérable et dont les spécifications techniques sont publiques et sans restriction d'accès ni de mise en œuvre. » [63]

L'État se voit confier une nouvelle mission : celui du service public de la donnée. Il est chargé de mettre à disposition, en vue de faciliter leur réutilisation, les principaux jeux de données présentant le plus fort impact économique ou social. Pour cela, ce nouveau service, piloté par la mission Etalab, a ouvert en avril 2017. Il offre une infrastructure nationale autour de quelques grandes bases de "données de référence".

Conclusion.

Quelles sont les précautions à prendre lors de la diffusion des données ?

- **Effectuer une analyse au cas par cas des données afin de vérifier qu'elles ne correspondent pas aux éléments cités dans la partie où la diffusion est interdite. Si des données correspondent à ces cas, il faudra envisager des solutions afin de les rendre diffusables, par exemple en anonymisant des données personnelles.**
- **Vérifier que les données sont achevées et sous format électronique.**

[63] Article 4 de la loi n° 2004-575 du 21 juin 2004 pour la confiance dans l'économie numérique (LCEN).

- **Informer l'ensemble des scientifiques sur la responsabilité de l'établissement et sur sa prise de décision concernant la diffusion de certaines données ou non. En effet, lorsqu'il y a une publication scientifique et que l'éditeur impose le dépôt de certaines données dans un entrepôt spécifique, même 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, et sous quelles conditions elles sont déposables.**

ANNEXES

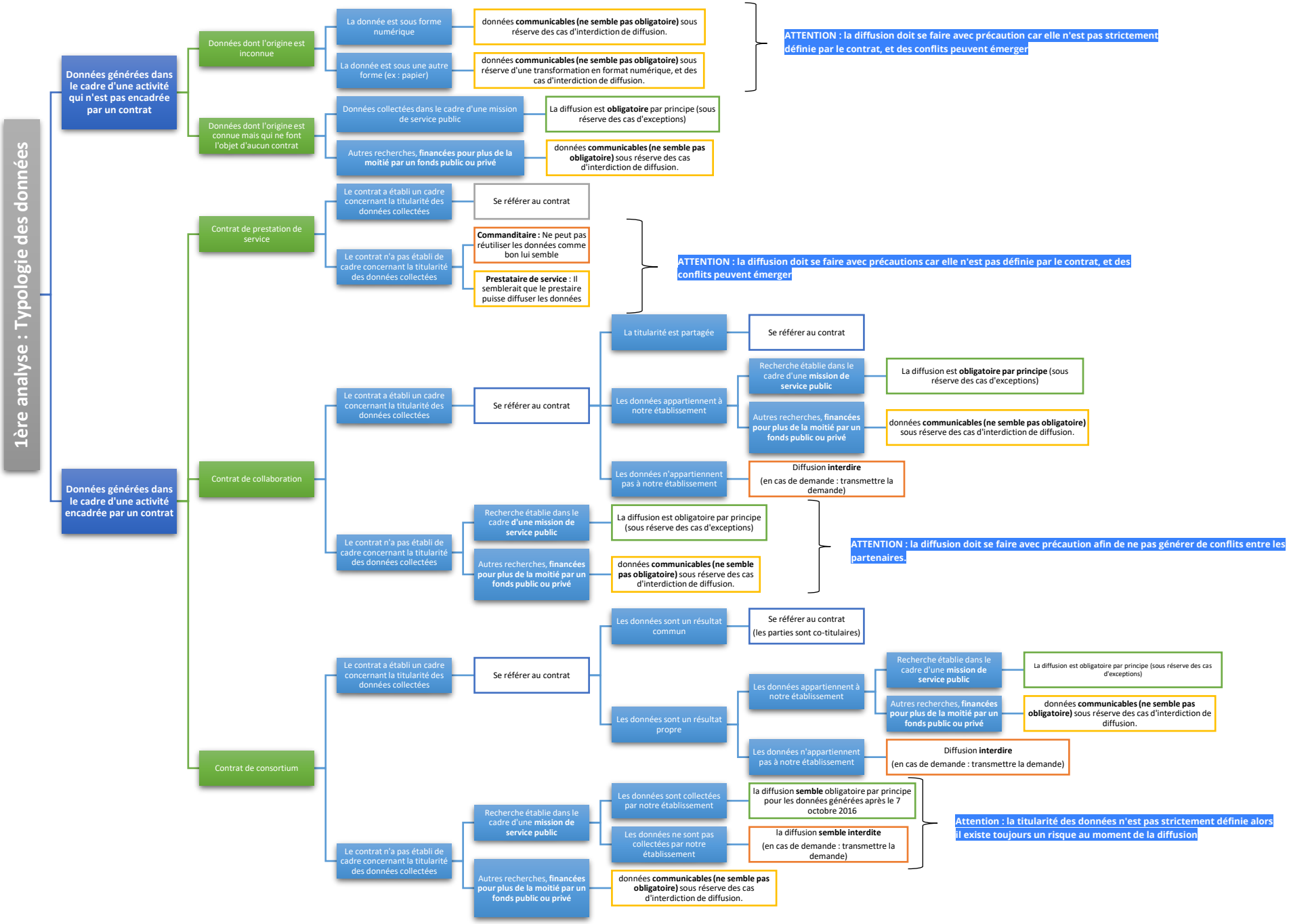
Annexe 1 : Logigramme « Typologie des données »

Annexe 2 : Logigramme « Diffusion des données selon la LRN ».

Annexe 3 : Quelques précisions sur la bonne pratique à avoir dans la gestion des données à caractère personnel (dispositions du RGPD)

Annexe 4 : Bibliographie

ANNEXE 1 : LOGIGRAMME « TYPOLOGIE DES DONNÉES ».



ANNEXE 2 : LOGIGRAMME « DIFFUSION DES DONNÉES SELON LA LRN » (4 pages)

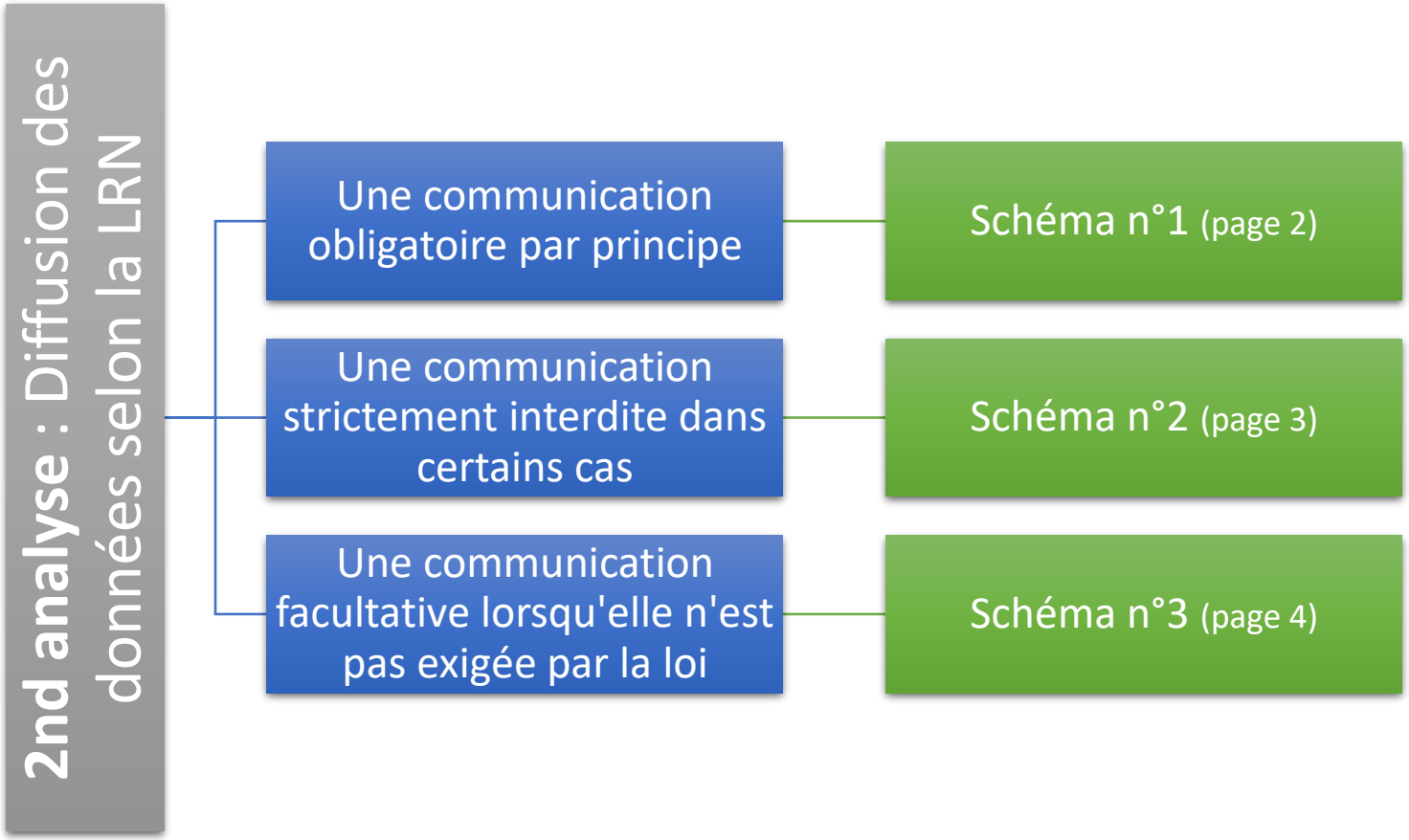


SCHÉMA N°1 : LA DIFFUSION OBLIGATOIRE PAR PRINCIPE

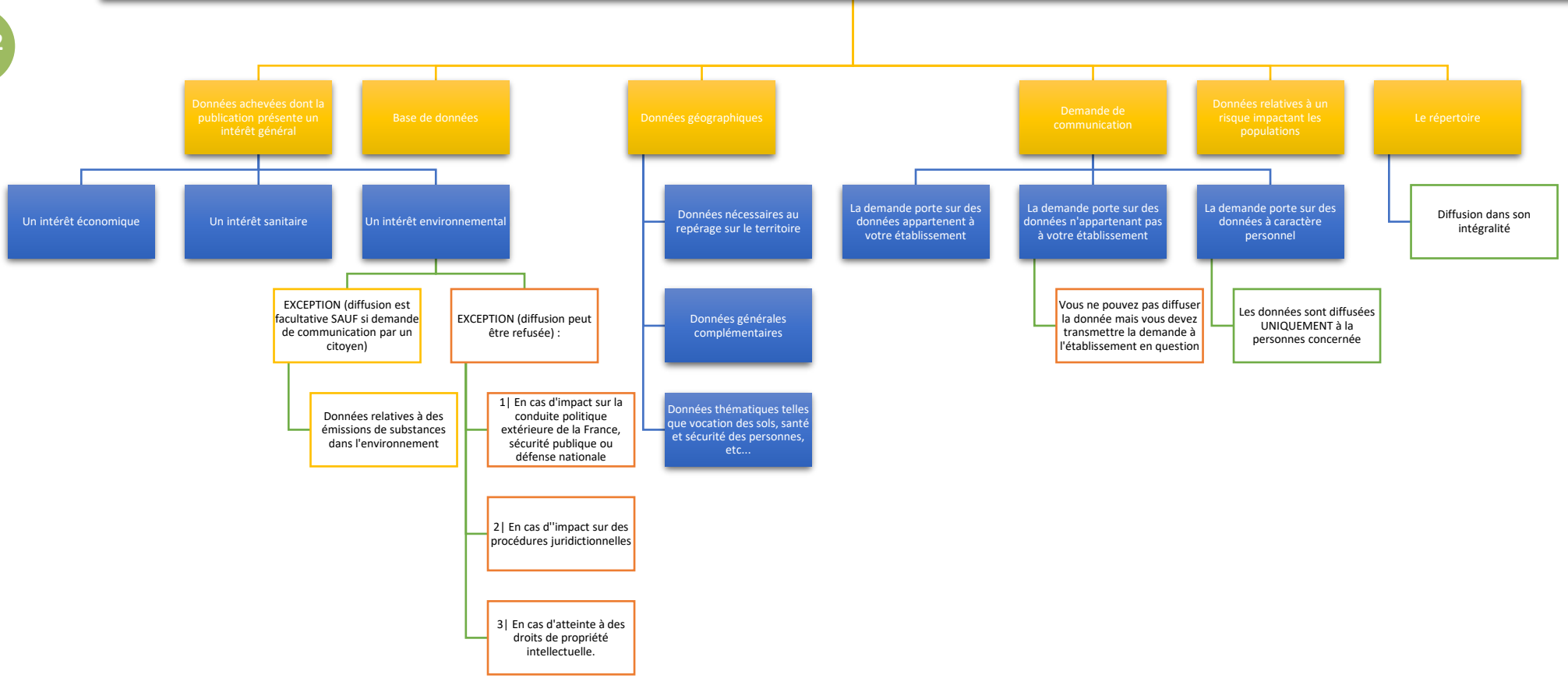


SCHÉMA N°2 : UNE COMMUNICATION STRICTEMENT INTERDITE DANS CERTAINS CAS

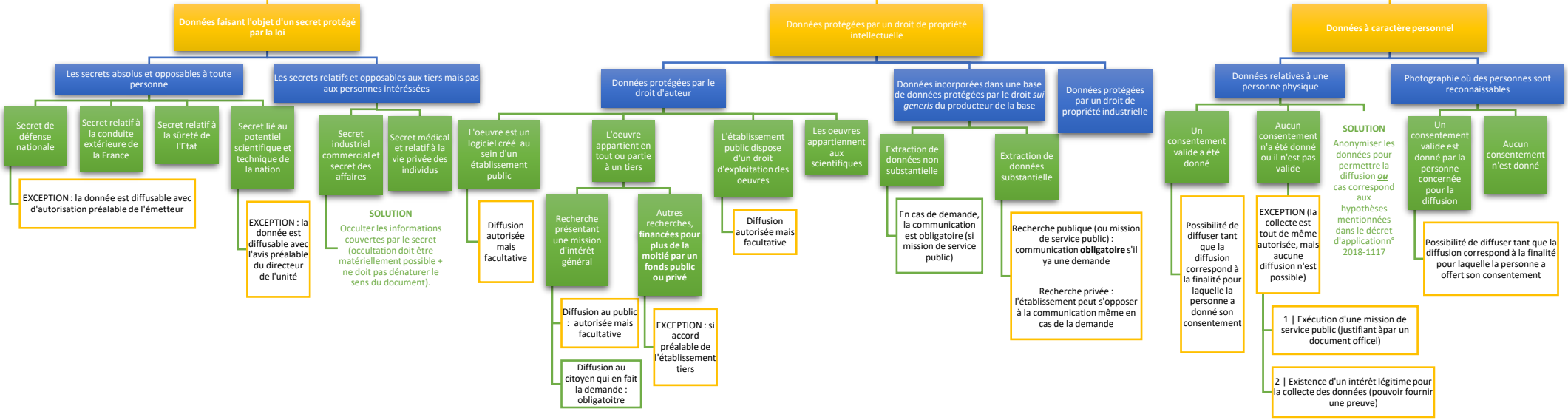


SCHÉMA N°3 : UNE COMMUNICATION FACULTATIVE LORSQU'ELLE N'EST PAS EXIGÉE PAR LA LOI

- En fonction de la politique de l'établissement
  - Diffusion ou non de toutes les communications facultatives (cadres blancs et jaunes sur les schémas 1 et 2)
- Données antérieures à la LRN
  - = Toutes les données et docuements administratifs élaborés avant le 7 octobre 2016
  - CONDITION : données sous forme numérique et achevées
- Reproduction ou copie de données dans certaines conditions
  - CONDITION : respect des droits personnels ou des secrets encadrant ces données
- Données statistiques
  - CONDITION : respecter le "secret statistique" ou obtenir une dérogation du Comité du secret statistique

ANNEXE 3 : GESTION DES DONNÉES À CARACTÈRE PERSONNEL.

Le Règlement Général sur la Protection des Données (RGPD) a été mis en place par le Parlement européen en 2016 afin de responsabiliser les organismes publics et privés qui traitent leurs données. Quels sont les bons réflexes à adopter afin de traiter convenablement ces données ?

Qu'est-ce qu'une donnée à caractère personnel ? Le RGPD indique qu'il s'agit de *« toute information se rapportant à une personne physique identifiée ou identifiable (ci-après dénommée «personne concernée»); est réputée être une «personne physique identifiable» une personne physique qui peut être identifiée, directement ou indirectement, notamment par référence à un identifiant, tel qu'un nom, un numéro d'identification, des données de localisation, un identifiant en ligne, ou à un ou plusieurs éléments spécifiques propres à son identité physique, physiologique, génétique, psychique, économique, culturelle ou sociale »*[1]

Cette définition est très large. En effet, elle ne rend pas compte du statut de certaines données personnelles qui doivent faire l'objet d'une protection spécifique (la santé, les condamnation, bancaire...) par rapport aux données courantes (nom, prénom…).

Cette définition générale doit être complétée par une définition des données sensibles. Le RGPD indique qu'il s'agit *« [du] traitement des données à caractère personnel qui révèlent les origine raciale ou ethnique, les opinions politiques, les convictions religieuses ou philosophiques ou l'appartenance syndicale, ainsi que le traitement des données génétiques, des données biométriques aux fins d'identifier une personne physique de manière unique, des données concernant la santé ou des données concernant la vie sexuelle ou l'orientation sexuelle d'une personne physique sont interdits »*[2]

Il peut s'agir également *« [du] traitement des données à caractère personnel relatives aux condamnations pénales et aux infraction ou aux mesures de sûreté »*[3]

[1] Article 4-1 du RGPD.
[2] Article 9 du RGPD
[3] Article 10 du RGPD

Quelles sont les précautions à prendre lors du traitement de vos données à caractère personnel ?

Un responsable de traitement doit choisir une (et une seule) base de licéité pour fonder le traitement de données parmi les six choix offerts par l'art 6 du RGPD et que s'il avait choisi le consentement pour tout ou partie du traitement de données, il devait mettre fin au traitement de données si la personne concernée retirait son consentement ou si le consentement se révélait invalide.

Quelques précisions sur la mise en place du RGPD.

La mise en place du consentement n'a pas vraiment changée avec le RGPD, c'est plus la manière de traiter le consentement et la détermination de l'âge pour consentir qui ont évolué. De plus, en matière de traitements de données à des fins scientifiques, il est conseillé de se référer au cadre légal et règlementaire spécifiquement adopté par chaque État membre en matière de traitement de données à des fins scientifiques. D'ailleurs, l'article 89 du RGPD indique la possibilité d'alléger les conditions du traitement des données personnelles en matière scientifique.

Comment diffuser des données à caractère personnel sans risquer de porter atteinte aux droits des personnes concernées ?

Il est possible d'avoir besoin de diffuser une base de données, par exemple, qui pourrait comprendre des données à caractère personnel. Afin de permettre la diffusion de cette base de données, il est possible d'anonymiser les données. Cela signifie que les personnes ne seront plus identifiées ou identifiables. Le RGPD ne s'applique pas aux données anonymisées.
Cependant, la CNIL a indiqué qu'une anonymisation parfaite n'existe pas. Il convient donc de diffuser ces données avec précaution même si elles sont anonymiser.

La pseudonymisation n'est pas une solution à retenir car elle ne permet pas d'occulter le caractère personnel aux données traitées car le responsable de traitement va conserver un numéro qui va être relié à d'autres données permettant d'identifier le traitement.

Quels sont les points clés à retenir avec la mise en place du RGPD ?

- **Principe de finalité des traitements de données à caractère personnel**. Cette finalité doit être déterminée, explicite et légitime[4]. La finalité est nécessairement liée à l'activité de l'entreprise. Chaque finalité identifiée correspond à une seule et unique base légale.
- **Principe de minimisation des données**. Les données doivent être adéquates, pertinentes et limitées à ce qui est strictement nécessaire à la finalité poursuivie. En somme, seules peuvent être collectées les données strictement nécessaires à la personne poursuivie.
- **Principe de licéité**. Les données doivent être traitées de manière licite et loyale. Pour cela, un consentement doit être préalablement demandé à chacun des personnes faisant l'objet du traitement, préalablement à ce dernier.
- **Principe de durée de conservation limité des données**. Le traitement des données doit être déterminé dans un espace-temps, c'est-à-dire qu'il faut établir une durée de conservation dès la mise en place du traitement.
- **Garantir la sécurité et la confidentialité des données**. Il faut empêcher que les données ne soient déformées, endommagées ou encore accessibles par des tiers non autorisés.
- **Des droits accordés aux personnes faisant l'objet du traitement**. Le RGPD permet un droit d'information, un droit d'accès, de rectification, d'effacement et un droit d'opposition. Il rajoute de nouveaux droits tels que la limitation des traitements et la portabilité.

De quelle manière doit être obtenu le consentement ?

« […] par une déclaration ou un acte clair que des données à caractère personne face l'objet d'un traitement » répond le RGPD[5]

Trois critères sont à retenir. Le consentement doit être :

- Le consentement doit être une manifestement libre.
- Cette manifestation doit être spécifique et éclairée. « Spécifique » signifie que la personne ne consent qu'à une seule et unique finalité, ce qui exclut un consentement global. « Éclairée » signifie que l'information est donnée à la personne concernée ; c'est elle qui va voir si le consentement est valable.

[4] Article 5 du RGPD
[5] Article 4 11° du RGPD.

- Cette manifestation doit être un acte positif clair. Par exemple, il est permis d'obtenir un consentement par le biais de cases à cocher, mais ces dernières ne doivent pas être pré-cochées car sinon ce consentement ne sera pas considéré comme valable.

Quid des données à caractère personnel et de la recherche.

L'article 89 du RGPD prévoit le cas des données à caractère personnel dans le cas de la recherche scientifique. Un tel traitement doit faire l'objet de la mise en place de mesures techniques et organisationnelles, notamment pour assurer le principe de minimisation. Par exemple, il est recommandé d'appliquer la pseudonymisation sur les données à caractère personnel chaque fois que cela est possible. C'est le cas notamment où l'identification des personnes, leur identité, n'est pas nécessaire pour le traitement.

ANNEXE 4 : BIBLIOGRAPHIE

TEXTES LÉGISLATIFS ET RÈGLEMENTAIRES

- Directive 96/9/CE du 11 mars 1996 sur la protection des bases de données.

- Directive 2007/2/CE du Parlement Européen et du Conseil du 14 mars 2007 établissant une infrastructure d'information géographique dans la Communauté européenne (INSPIRE).

Quelques informations complémentaires sur la Directive INSPIRE.

OBJET : Faciliter l'accès aux données géographiques servant pour la recherche en matière environnement.
→ Faciliter la prise de décision concernant les politiques et les conditions susceptibles d'avoir une incidence directe ou indirecte sur l'environnement.
→ Mise en commun de ces données pour toutes les autorités publiques + sans obstacle (gratuité et interopérabilité).
→ **Art 1er** : « fixer des règles générales destinées à établir l'infrastructure d'information géographique dans la Communauté européenne »

APPLICATION :
- Aux données détenues par les autorités publiques ou en leur nom.
- A l'utilisation dans le cadre d'une mission de service public.
- Aux données détenues par des PP ou des PM si elles en font la demande + à certaines conditions.

EXIGEANCE D'INTEROPERABILITE :
- Il va falloir prévoir des règles de mise en œuvre afin que les données sont partagées sous le même format et donc qu'elles soient lisibles et accessibles pour tout le monde.
- Services en réseau : « spécifications et critères de performances minimale arrêtés d'un commun accord afin de garantir l'interopérabilité des infrastructures mises en place par les États membres ».
- Portail communautaire exploité par la Commission, ainsi que par des points 'accès qu'ils décident d'ouvrir.

CONDITIONS DES DONNEES (article 4) :
1. Données liées à une zone où un État membre détient et/ou exerce sa compétence.
2. Format électronique.
3. Détenues par soit :
- Une autorité publique, après avoir été produites ou reçues par une autorité publique, ou bien gérées ou mises à jour par cette autorité et rentrant dans le champ de ses missions publiques.
- Un tiers à la disposition duquel le réseau a été mis → art 12
- Concerne un ou plusieurs thèmes figurant aux annexes I, II et III.

- Directive 2013/37/UE dite « PSI » (Public Sector Information).

- Directive 206/943 du 8 juin 2016.

- Règlement (UE) 2016/679 du Parlement Européen et du Conseil du 27 avril 2016 relatif à la protection des personnes physiques à l'égard du traitement des données à caractère personnel et à la libre circulation de ces données et abrogeant la directive 95/46/CE (règlement général sur la protection des données) « RGPD » [version annotée par Paul-Olivier Gilbert, Président de l'AFCDP].

- Loi n°57-711 du 7 juin 1951 modifiée.

- Loi n°78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés (actualisation du 17 juin 2019).

- Loi « CADA » n°78-753 du 17 juillet 1978 (codification le 19 mars 2016 dans le CRPA).

- Loi n°98-536 du 1er juillet 1998 concernant la protection des bases de données.

- Loi n°2004-575 pour la confiance dans l'économie numérique (LCEN) du 21 juin 2004

- Loi n°2015-1779 du 28 décembre 2015 dite « Valter » relative à la gratuité et aux modalités de la réutilisation des informations du secteur public.

- Loi n°2016-1321 du 7 octobre 2016 pour une République numérique

Quelques précisions sur la loi pour une République Numérique du 7 octobre 2016

La loi comporte trois volets.

1/ Le premier volet comprend des mesures sur l'ouverture des données publiques, la création d'un service public de la donnée. Il introduit la notion de données d'intérêt général, pour optimiser l'utilisation des données aux fins d'intérêt général. Une partie est également dédiée au développement de l'économie du savoir, avec la possibilité pour les chercheurs de publier librement leurs articles scientifiques dans un délai de six à douze mois. Le Sénat a voté en faveur de la facilitation de l'ouverture et de la réutilisation des données des administrations ainsi que des décisions des juridictions administratives et judiciaires. La diffusion de ces données est circonscrite aux données dont la publication présente un intérêt économique, social, sanitaire ou environnemental

2/ Le deuxième volet traite de la protection des citoyens dans la société numérique. Il affirme le principe de neutralité des réseaux et de portabilité des données. Il établit un principe de loyauté des plateformes de services numériques. Le consommateur dispose en toutes circonstances d'un droit de récupération de ses données. Le texte introduit également de nouveaux droits pour les individus en matière de données personnelles (droit à l'oubli numérique pour les mineurs, testament numérique pour donner des directives aux plateformes numériques, confidentialité des correspondances privées). Un amendement adopté par l'Assemblée nationale prévoit une peine de deux ans d'emprisonnement et une amende pouvant aller jusqu'à 60 000 euros pour le fait de transmettre ou de diffuser sans le consentement exprès de la personne l'image ou la voix de celle-ci, prise dans un lieu public ou privé, dès lors qu'elle présente un caractère sexuel (phénomène dit "revanche pornographique" ou "revenge porn").

3/ Le troisième volet est consacré à l'accès au numérique pour tous avec notamment la couverture mobile, l'accessibilité aux services numériques publics, l'accès des personnes handicapées aux services téléphoniques et aux sites internet. Il prévoit aussi le maintien de la connexion internet pour les personnes les plus démunies. Le Sénat a adopté en première lecture un amendement qui oblige les

opérateurs de télécommunications à s'engager, via des conventions avec les collectivités, pour l'installation du très haut débit.

- Décret n°2016-1922 du 28 décembre 2016 relatif à la publication en ligne des documents administratifs.

- Décret n°2017-331 du 14 mars 2017 relatif au service public de mise à disposition des données de référence (codifié à l'article L.321-4 du CRPA).

- Décret n°2018-1117 du 10 décembre 2018 relatif aux catégories de documents administratifs pouvant être rendus publics sans faire l'objet d'un processus d'anonymisation.

- Code de l'environnement

- Code de la propriété intellectuelle

- Code des relations entre le public et l'Administration

- Code de la recherche

- Code pénal

AVIS

- CNIL | Délibération n°2018-101 du 15 mars 2018 portant avis sur projet de décrit pris pour l'application de l'article 6 de la loi n°2016-1321 du 7 octobre 2016 pour une République numérique (demande d'avis n°17001751)

- CADA avis n°20090310 du 26 février 2009

- CADA avis n°20130750 du 28 mars 2013

- CADA avis n°20144578 du 08 janvier 2015

- CADA avis n°20175033 du 11 janvier 2018

JURISPRUDENCE

- TGI Paris 5 septembre 2001 SA Cadremploi c/ SA Keljob et Colt Télécommunications France

- Tribunal des conflits décision du 15 décembre 2008, Voisin c/ RATP req n°3662, Lebon

- Tribunal administratif de Paris décision n°1508951 du 10 mars 2016

- Tribunal de commerce de Lyon Ordonnance de référé du 22 octobre 2001, Sarl Avant-Première Design Graphique c/ Sarl Adgensite

- CE, 22 février 2013, Fédération chrétienne des témoins de Jéhovah France

- CE, 8 février 2017, n°389806

SCHÉMA / TABLEAU

- « Qui a les droits ? Qui a les droits ? Qui a les droits de faire quoi ? » Frédérique Bodignon, Romain Boistel, Delphine Du Pasquier – École des Ponts Paris Tech (2018) [logigramme sur la diffusion des données].

- « Évolution et intégration du droit des données « sol » » A. Bispo, M. Rabut et C. Gascuel pour le SP (Gis Sol) (27 novembre 2020)

- INC « Tableau synthétique des principes dispositions de la loi n°2016-1321 du 7 octobre 2016 pour une République numérique au 06/06/2017 » https://www.inc-conso.fr/sites/default/files/pdf/Tableau-economie-numerique-INC.pdf

DOCTRINE / ARTICLES

- « Ouverture des données de la recherche. Guide d'analyse du cadre juridique en France » Nicolas Bécard, Céline Castets-Renard, Gauthier Chassang, Martin Dantant, Laurence Freyt-Caffin, Nathalie Gandon, Caroline Martin, Andrea Martelleti, Alexandra Mendoza-Caminade, Nathalie Morcrette, et al. (5 juin 2017)

- « La #LoiNumérique en 15 points clés » Ministère de l'économie, des finances et de la relance (10 novembre 2016)

- « Avis sur les enjeux éthiques et déontologiques du partage et de la gestion des données issues de la recherche » Comité consultatif d'éthique pour la recherche agronomique

- « Régime juridique applicable aux traitements poursuivant une finalité de recherche scientifique (hors santé) » CNIL

- « Quel cadre juridique pour la Science Ouverte ? Un aperçu des évolutions récentes » scinfolex.com (calimaq) (23 novembre 2018)

- « Quel statut pour les données de la recherche après la loi numérique ? » scinfolex.com (calimaq) (3 novembre 2016)

- Webinaire avec Lionel Maurel, juriste, bibliothécaire et Directeur Adjoint Scientifique du CNRS (InSHS) http://www.cerpeg.fr/cerpeg/index.php/menu-accueil-ressources/outils-numeriques/443-donnees-et-propriete-intellectuelle

- [DALLOZ] Note d'information du 9 octobre 2013 relative à la production et l'utilisation de l'information géographique dans les services déconcentrés, mise en œuvre des dispositions résultant de la transposition de la directive européenne INSPIRE, programme Géo-IDE

- [DALLOZ] Fascicule 109-40 : « Données publiques – production interne et collecte sur le secteur privé » Thierry Piette-Coudol (4 mai 2020)

- [DALLOZ] Recueil Dalloz 2017 p.583 « Open Data et propriété intellectuelle | État des lieux au lendemain de l'adoption de la loi pour une République numérique » Tristan Azzi.

- [DALLOZ] Répertoire IP/IT « Communication des documents administratifs Cont. Adm. – Champ d'application du régime général du droit à communication » Alexandre LALLET ; Pearl NGUYEN DUY (juin 2019).

- « Le nouveau régime juridique de la réutilisation des informations publique » 16 mai 2017 – Siafdroit.hypothses.org

- [DALLOZ] Larcier Le règlement général sur la protection des données – Chapitre 1 : Les changements dans la mise en œuvre de la protection des données qui intéressent le secteur de la santé – Section 4 : Le rôle du consentement en matière de traitements de données à des fins scientifiques (2018).

- [DALLOZ] Praxis Cyberdroit – Chapitre 351 : Œuvres libres Christiane Féral-Schul 2020-2021 – Section 6 : Libre accès / Open Access

- [DALLOZ] Hors collection Contrats sur la recherche et l'innovation – Chapitre 003 : Cadre juridique des contrats sur la recherche et l'innovation (Etienne Vergès) 2018-2019 – Section 1 : Relations entre les contrats sur la recherche et l'innovation et le droit commun des contrats.

- [Decideo.fr] Annabelle Richard et Guillaume Morat, Pinsent Masons France LLP le 25 avril 2017 « Open Data : ce qu'il faut retenir de la loi Lemaire » https://www.decideo.fr/Open-Data-ce-qu-il-faut-retenir-de-la-Loi-Lemaire_a9297.html

- OCDE « Principes et lignes directrices pour l'accès aux données de la recherche financée sur fonds publics » 2007

Dans le cadre de **l'initiative « 4 pour 1000 »** et de la construction du **projet DATA4C+**

ANALYSE DE RISQUE

Ouverture des données de la recherche : identifier les risques et les prévenir.

Rédigée par Cloé Sigal-Guille

Sous la direction de Pauline Corbière (Cirad), Tiphaine Chevallier (IRD) et Christine Le Bas (INRAe).

Sommaire

Préambule

Le projet.

Le Projet DATA4C+ est un projet coordonné par le Cirad, en partenariat avec l'INRAe et l'IRD. Il répond à un objectif technico-juridique de partage de données sur le carbone des sols de certains territoires. Les résultats de DATA4C+ permettront **l'estimation des potentiels de séquestration de carbone des sols de ces territoires dans de futurs projets**. Ainsi, ce projet s'inscrit dans le cadre des initiatives de sciences ouvertes et du carbone du sol (« 4 pour 1000 »). Afin de garantir au mieux la protection et l'utilisation des données collectées, la mise en place de bases de données permettent aux scientifiques de conserver les données générées dans le cadre de leurs expertises et de pouvoir les analyser. Il est nécessaire d'analyser ces bases de données sous l'angle juridique afin de déterminer si les données doivent ou peuvent être diffusées librement au grand public. À cette fin **une deuxième phase d'analyse juridique** a été engagée afin d'appréhender les risques gravitant autour de la diffusion des données de la recherche ; diffusion obligatoire par défaut pour certaines données depuis la loi n°2016-1321 pour une République numérique, dite « Loi Axelle Lemaire », du 7 octobre 2016.

L'intérêt d'une analyse de risque.

La loi pour une République numérique a mis en place une **ouverture des données par défaut**. Certaines données vont donc être obligatoirement diffusées au grand public sur des plateformes dédiées. Les données de recherche scientifique ne font pas exception. De plus, lorsque les données ne sont pas obligatoirement diffusables, le législateur souhaite tout de même mettre l'accent sur les bénéfices pour la Science d'une telle diffusion et encourage donc le partage. Cependant, des risques gravitent autour d'une telle ouverture. En effet, certaines atteintes à des droits ou la récupération frauduleuse des donnés peuvent conduire des conséquences financières ou psychologiques. Il est donc nécessaire de **prendre des précautions** préalablement à la diffusion des données présentent dans le système DATA4C+. Cette analyse de risque est l'aboutissement de réflexions allant dans ce sens.

1 MISE EN CONTEXTE

Aux fins d'identifier les risques de la diffusion des données et d'en prévenir les atteintes, une analyse de risque est engagée. Cette analyse de risque a pour ambition de répondre aux questionnements subsistants autour de la diffusion des données de la recherche, mais également autour de leur absence de diffusion. En effet, la loi pour une République numérique du 7 octobre 2016 a mis en place un système de diffusion de la donnée publique par principe, répondant à la volonté du législateur d'inscrire la recherche française dans un modèle d'Open Data. Cependant, de nombreux enjeux entourent la recherche en matière de sol, et la diffusion de telles données oblige les établissements à s'adapter et à réfléchir à des stratégies pour permettre la valorisation de leurs recherches d'un côté, et la protection d'individus ou de secrets d'un autre côté. **En somme, quels sont les risques à diffuser et à ne pas diffuser des données dans certains cas ?**

Le système DATA4C+ manipule plusieurs types de données ce qui oblige une analyse au cas par cas de chacune d'elles pour en identifier leur obligation ou non diffusion, voire l'interdiction *(cf Note méthodologique)*.

1.1 LA MÉTHODOLOGIE SUIVIE.

Le dictionnaire Larousse définit le risque comme le *« fait de s'engager dans une action qui pourrait apporter un avantage, mais qui comporte l'éventualité d'un danger »*. En ce qui concerne le projet DATA4C+, le risque se situe dans la diffusion des données des recherches engagées dans le projet. L'avantage correspond, par exemple, au fait de faire bénéficier la communauté scientifique des résultats produits, et de permettre l'élargissement des recherches. Le danger correspond, par exemple, au fait que l'établissement fasse l'objet de sanctions en répréhension de la diffusion de données qui n'auraient pas dues être communiquées. Aux fins de comprendre les risques auxquels s'exposent le projet DATA4C+ et d'identifier ceux qui pourront être pris, l'analyse de risque va suivre une méthodologie qui ambitionne de se rapprocher de ce qui est susceptible de se produire. Il convient de rappeler que les probabilités de survenance du risque restent cependant **théoriques** et peuvent

différencier en pratique. Ainsi, il convient de considérer le risque (R) comme **la variable de sortie qui nous intéresse**. Elle est le produit d'une probabilité d'occurrence d'un évènement (aléa =A) par les conséquences de cet évènement (enjeu = E). Cela nous conduit à la **multiplication** suivante : R = A x E. De plus, nous pouvons distinguer différents types (i) de risques (Ri). Nous pouvons dont considérer que **ces risques s'additionnent** R = Σ(i) Ri.

Vocabulaire :

« Aléa » (A) désigne la probabilité d'occurrence d'un évènement. Il s'agit par exemple de la probabilité qu'un individu porte plainte car la diffusion des données aurait porté atteinte à ses droits (diffusion de données à caractère personnel par exemple).

→ **Prudence :** la probabilité d'occurrence est difficile à évaluer. Elle est identifiée subjectivement, et il est possible que la pratique soit ainsi différente. Ainsi, il faut prendre des précautions avec cette variable.

« Enjeu » (E) désigne les conséquences de l'évènement produit et conduisant dès lors à la survenance du risque. Il s'agit par exemple d'une sanction infligée par la Commission nationale de l'information et des libertés (ci-après la « CNIL ») en réparation des préjudices subi par une personne du fait de la diffusion de ses données à caractère personnel.

« Risque » (R) désigne la multiplication de l'Aléa et de l'Enjeu et indique le type de sanction à laquelle s'exposent les établissements en diffusion ou non les données liées au projet DATA4C+. Ce Risque est un indicatif permettant d'identifier si la diffusion des données est envisageable ou non.

« Ensemble de risques » Σ(i) désigne l'addition des types risques (i) pouvant se produire pour un seul évènement. Cette variable est à multiplier avec la variable « type de risque » (définies ci-dessous) car il est possible que plusieurs risques, mêmes faibles s'additionnent et puissent conduire à l'augmentation du Risque initial.

« Type de risque » (Ri) désigne l'addition des enjeux de l'Ensemble des risques Σ(i). Multipliez avec cette variable, elle permet d'identifier s'il y a une probabilité que le Risque initial augmente en fonction du nombre de risques présents pour un seul et même évènement. Ainsi, la somme de ces variables désigne un Risque (R) augmenté ou non.

En fonction des parties et des données traitées, les risques vont évoluer. Mais afin d'identifier les résultats potentiels de ces différentes multiplications, voici deux tableaux qui vont vous permettre d'appréhender un peu le vocabulaire qui va être utilisé. Le *Tableau 1* permet d'identifier la teneur du risque et d'indiquer s'il est acceptable de diffuser les données ou non [R = A x E]. Le *Tableau 2* permet d'identifier l'ampleur du risque en fonction de l'addition des divers types de risques [R = Σ(i) Ri].

Aléa (A)	**Enjeu (E)**	**Risque (R)**
Bas	Bas	Faible
	Moyen	Faible
	Élevé	Intermédiaire
Moyen	Bas	Faible
	Moyen	Intermédiaire
	Élevé	Maximum
Élevé	Bas	Intermédiaire
	Moyen	Maximum
	Élevé	Maximum

[Tableau 1 : Évaluation de la teneur d'un seul risque]

Vocabulaire Tableau 1 :

« Bas » désigne une faible probabilité de survenance de l'aléa. Pour illustrer cette notion, nous pourrions indiquer que la probabilité que le risque se concrétise est inférieure à 25%.

« Moyen » désigne une probabilité plus importante de la survenance de l'aléa mais il est difficile de clairement l'identifier (différents facteurs sont à prendre en compte). Pour illustrer cette notion, nous pourrions indiquer que la probabilité que le risque se concrétise se situe entre 25% et 75%.

« Élevé » désigne une forte probabilité de survenance de l'aléa. Pour illustrer cette notion, nous pourrions indiquer que la probabilité que le risque se concrétise est supérieure à 75%.

Comment pouvons-nous appréhender la variable de l'aléa ? Il y a peu de jurisprudence en la matière, néanmoins certains points présentent plus de précisions comme c'est le cas pour les données à

caractère personnel. Par exemple, la CNIL fait de plus en plus attention aux atteintes portées aux droits individuels ; et la probabilité qu'elle intervienne en cas de diffusion est plus élevée.

Comment pouvons-nous appréhender la variable de l'enjeu ? L'enjeu est évalué en fonction de l'intensité des sanctions ou des conséquences en cas de survenance du risque. Par exemple, lorsque la sanction risquée est un simple avis d'une autorité alors le risque sera « Bas ».

→ **Précision :** la probabilité qu'un risque survienne diminue avec le temps. En effet, l'analyse de risque doit être lue en ayant en tête qu'un facteur temps peut jouer sur les résultats obtenus. Autrement dit, plus le temps d'écoule entre la collecte de la données et sa publication, plus le risque s'amoindrie, et il semble plus facile de la diffuser.

« Faible » désigne un risque limité voire nul, il est donc possible de diffuser les données.

« Intermédiaire » désigne un risque de moindre importance mais à prendre en compte tout de même. Ainsi, la diffusion des données peut se faire mais avec précaution. Autrement dit, il conviendrait de trouver des moyens afin de limiter ce risque (exemple : diffusion limitée à certaines personnes).

« Maximum » désigne un risque important. Il est déconseillé d'effectuer la diffusion des données.

Ensemble de risques Σ(i) (= quantité de risques pour un seul évènement).	**Type de risque (Ri)** (= addition des enjeux de ses risques)	**Augmentation du risque.** (= augmentation du risque ou non)
Bas	Faible	Faible
	Intermédiaire	Faible
	Maximum	Intermédiaire
Moyen	Faible	Faible
	Intermédiaire	Intermédiaire
	Maximum	Maximum
Élevé	Faible	Intermédiaire
	Intermédiaire	Maximum
	Maximum	Maximum

[Tableau 2 : Évaluation de la teneur du risque final en fonction du nombre de risques potentiels pour un seul et même évènement]

Vocabulaire Tableau 2 :

« Bas » désigne un nombre faible de risque pouvant s'appliquer au cas évalué.

« Moyen » désigne un nombre négligeable mais à prendre en compte. Cette variable est plus difficile à identifier.

« Élevé » désigne un nombre important de risques à prendre en compte.

[Pour la variable « Type de risque ». La variable « Risque est défini plus haut »]

« Faible » désigne une globalité faible de l'ensemble des risques liés à la survenance du risque [exemple : seulement des avis seront données].

« Intermédiaire » désigne une globalité de risques à prendre en compte, et susceptibles de faire augmenter le risque final en fonction de l'association avec la variable **Σ(i).**

« Maximum » désigne une globalité de risques importants conduisant à l'augmentation du risque final.

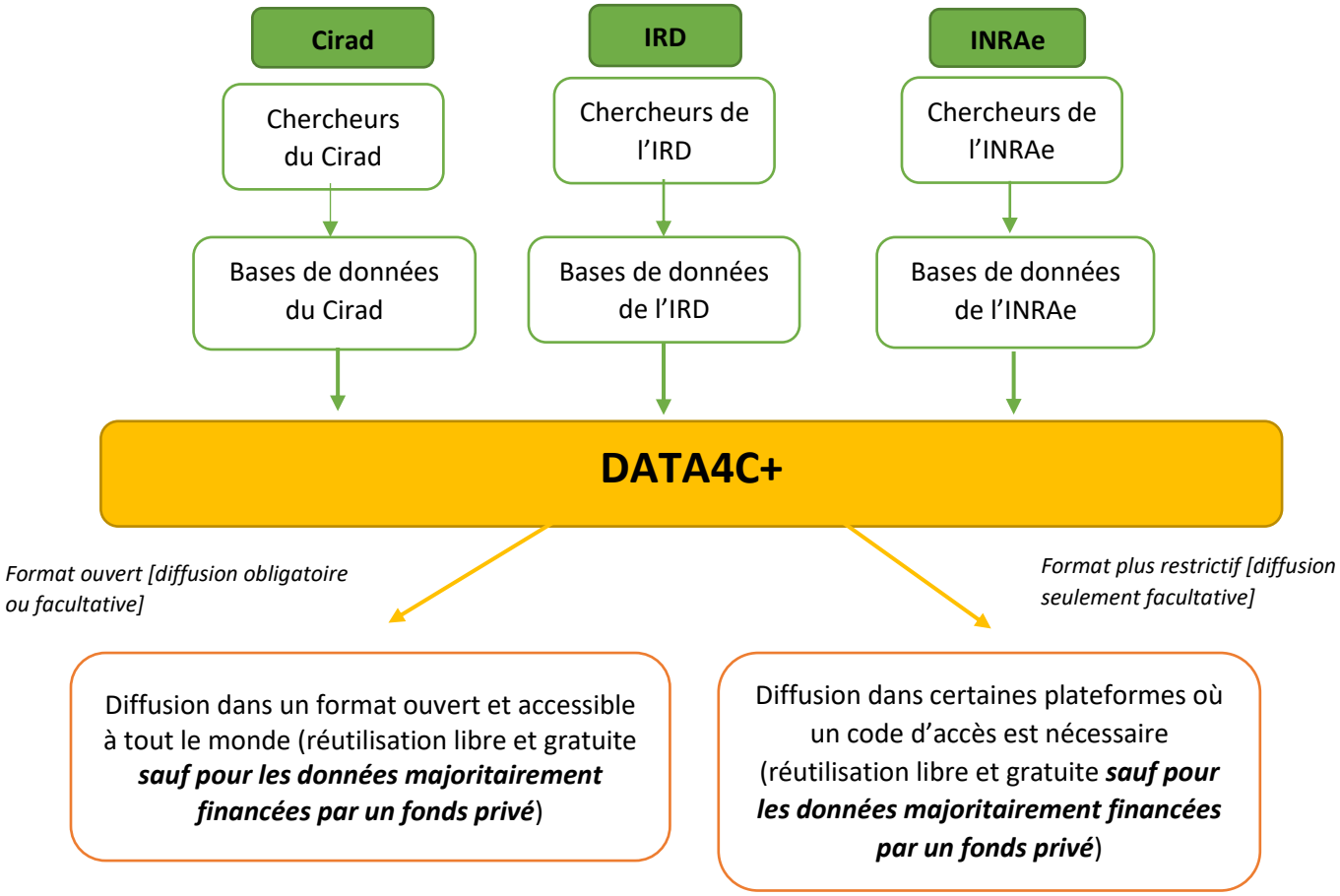
1.2 Les Spécificités du Système DATA4C+.

Une fois la méthodologie mise en place, il convient désormais de définir les contours du système spécifique à DATA4C+ de manière à pouvoir identifier les flux de données (1.2.1), et les acteurs pouvant interagir dans ces flux (1.2.2).

1.2.1 Le flux des données.

Le système DATA4C+ englobe **plusieurs bases de données**. Chacune d'elle est conçue de manière **autonome** par le Cirad, l'IRD et l'INRAe. Ces organismes vont y intégrer les données collectées par leurs chercheurs lors de missions internes ou extérieures au projet. Les données transférées au système DATA4C+ correspondent à celles collectées sur les territoires d'Outre-Mer de la France. Ainsi, lorsque des données collectées dans des pays étrangers sont présentes dans la base de données de l'une des entités (notamment dans Val Sol), l'organisme ne transfère **que les données relatives aux territoires correspondants**.

L'objectif du système DATA4C+ est de réunir tous les résultats de recherche portant sur la composition des sols en Outre-mer de manière à créer **une zone de partage bénéfique aux recherches actuelles ou futures**.



[Figure 1 : Flux de données dans le système DATA4C+]

1.2.2 Les acteurs.

Chacun des flux peut faire interagir différents acteurs en eux.

Les acteurs émetteurs et récepteurs du flux (partie verte et jaune du graphique ci-dessus) :

- ***Administrateur :*** Les personnes pouvant modifier le système (par exemple, rajouter des données, les supprimer, ou encore changer certain fonctionnement dans le système).
- ***Utilisateurs participants au projet (chercheurs, juristes, personnel des organismes) :*** Les personnes pouvant accéder au système mais ne pouvant pas le modifier. Ils peuvent observer les données et demander leur modification si besoin dans l'objectif de pérenniser le projet.

- ***Développeur :*** Les personnes qui ont créé le système et qui le maintienne en service. Ils ont la main sur les algorithmes et peuvent donc modifier le système en cas de bogue.
- ***Sous-traitant :*** personnel de soutien pouvant modifier le système aux fins de pouvoir l'améliorer ou corriger des anomalies.

Les acteurs seulement récepteurs du flux (partie orange du graphique ci-dessus) :

- ***Autres chercheurs / professionnels effectuant des recherches en lien avec les données :*** Les chercheurs extérieurs aux personnes ayant déjà accès à la base et souhaitant utiliser les données qu'il y a dessus.
- ***Autres citoyens :*** Les personnes ayant accessibilité aux données par le biais de la règlementation de la loi pour une République numérique (7 octobre 2016).
- ***Entreprise ou organisme ayant une activité commerciale*** : Toute personne morale qui souhaite accéder à des données librement réutilisables à des fins commerciales.
- ***Pirate informatique :*** Toute personne physique ou morale qui souhaiterait s'emparer des données alors qu'on ne lui a pas autorisé l'accès au système.

Les personnes ayant accès au système DATA4C+ :

→ Actuellement : quatre personnes.

→ Sur la durée : tous les membres du projet, ainsi que les membres de l'ANR.

→ Lors d'une diffusion : certaines données pourraient être accessibles sans besoin d'identification.

1.2.3 Les mesures de contrôle.

- **Techniques :** L'accès au système nécessite une identification. Ainsi, chaque Utilisateur du système possède un identifiant et un mot de passe fourni par un Administrateur.
- **Organisationnelles :** Le mot de passe et l'identifiant ne peut être obtenu qu'à la suite d'une demande effectuée à l'Administrateur responsable de la gestion et de la fourniture des accès.
- **Légales :** Aucune mesure légale pour le moment.

  → Proposition : Possibilité de mettre en place un contrat de confidentialité déterminant le rôle des individus et des possibilités qu'ils ont ou non de divulguer des informations.

1.2.4 Les risques liés à une diffusion involontaire des données.

La présente analyse de risque se centre sur l'identification des probabilité qu'un risque survienne en cas de diffusion **volontaire** des données présentes dans le système DATA4C+. Il est cependant important de mentionner que d'autres risques existent et qu'ils peuvent être indépendants des acteurs du système. Par exemple, la diffusion involontaire de données peut être effectuée et peut avoir diverses causes. L'exploitation d'une vulnérabilité du système, ou l'exécution d'un évènement redouté peuvent être une explication à cela.

Cette partie vous permet d'identifier certaines causes de risques indépendants d'une volonté de diffusion. Les exemples donnés ne sont pas exhaustifs, et ne sont donnés qu'à titre informatif. Le reste de l'analyse de risque sera consacré aux risques liées à un acte volontaire des acteurs du système.

Qu'est-ce qu'une vulnérabilité ? Une vulnérabilité est définie comme étant **une faiblesse dans les mécanismes de protection des données** (elle peut être technique, organisationnelle ou légale) **du système** ou encore **d'une absence de mécanisme qui peut ultimement résulter en un bris de confidentialité**. Elle inclue par exemple :

- La vulnérabilité introduite de façon intentionnelle lors de la conception du système ;
- Les erreurs dans la conception ou le design du système ;
- Les erreurs d'implémentation dans la réalisation du système ;

Qu'est-ce qu'un évènement redouté ? Il s'agit d'un évènement qui **se produit en conséquence de l'exploitation d'une ou de plusieurs vulnérabilités** et qui peut **créer un risque dans la protection des données du système.**

À des fins d'illustrations, le tableau ci-dessous permet d'identifier *quelques exemples* de vulnérabilités et leurs conséquences en cas d'exploitation. Nous avons inséré une colonne « scénarios pertinents » pour que vous puissiez imager l'exploitation de la vulnérabilité. Ainsi, la conséquence de la vulnérabilité sera la potentialité d'exécution de l'évènement redouté, conduisant à la production du risque.

Vulnérabilité	Scénarios pertinents	Évènements redoutés	Risques
Transmission non sécurisée de données	Un partenaire ne respecte pas les politiques de sécurité et entre des informations personnelles dans des champs qui ne sont pas prévus à cet effet	Diffusion de données à caractère personnel relatives à l'identité d'un individu ou à la localisation de sa parcelle	Risques de sanctions car des données à caractère personnel dont diffusées.
Échec de la détection d'évènements inhabituels / Usurpation d'identifiant et de mot de passe afin d'obtenir des accès non autorisés	Un criminel obtient les données permettant de localiser des denrées rares dans les sols	Utilisation des données à des fins non autorisées : - Revente de données - Réutilisation frauduleuse des données (non-respect des droits personnels)	Risque de sanction car des données interdites à la diffusion font l'objet de communications Risque d'atteinte à la réputation du projet (perte de confiance)
Erreurs fonctionnelles / Panne du serveur	Un incendie s'est déclaré au lieu d'hébergement du serveur	Les serveurs ne fonctionnent plus et des données sont perdues - Les serveurs ne fonctionnement plus correctement et des erreurs limitent la sécurité du système	- Risque que les utilisateurs ne puissent plus accéder au systèmes. - Risque que le système soit accessible au public - Risque de voir une diffusion des données s'effectuer.

Ce tableau reste illustratif et ne représente qu'une théorie des évènements susceptibles de se produire à l'absence de toute volonté des acteurs du système. La pratique peut être différente car ces vulnérabilités peuvent ne pas être prévisibles. **C'est également pour cette raison que la majeure partie de l'analyse de risques se focalise sur les événements plus « prévisibles » car ils font l'objet d'un acte positif, c'est-à-dire la décision d'une diffusion des données ou non.**

La méthodologie de l'analyse de risque ainsi que le fonctionnement du système sont désormais définis. Il convient de s'intéresser plus particulièrement aux risques liés à la diffusion des données, mais également à leur réutilisation. L'analyse de risque est ainsi divisée en plusieurs parties :

2 | Les risques relatifs à la diffusion de données dont la communication est obligatoire.

3 | Les risques liés à la diffusion de données dont la communication n'est pas obligatoire.

4 | Le risque lié à la diffusion de données dont la communication est interdite.

5 | Les risques liés à l'atteinte portée à la réutilisation des données.

2 LES RISQUES RELATIFS À LA DIFFUSION DE DONNÉES DONT LA COMMUNICATION EST OBLIGATOIRE

La diffusion de données obligatoirement communicables présente des risques très limités voire nuls. En effet, tout naturellement, le législateur ne prévoit pas de sanction lors de la diffusion de données qu'il oblige à la communication.

Cependant, il convient d'être **vigilant**, car il existe toujours un risque de divulguer des données qui n'auraient pas dû l'être. Cela peut être due à une méconnaissance du cadre juridique par exemple. Pour éviter ces risques, il convient de prendre des précautions en effectuer une **analyse au cas par cas de toutes les données présentes dans la base de données**. Cette analyse permettre d'identifier les données obligatoirement communicables, celles qui ne le sont pas, et surtout celles qui sont interdites à la communication. Pour cela, et afin de connaitre toutes les règles et leurs exceptions, nous vous invitons à lire ***la note méthodologique rédigée dans le cadre du projet DATA4C+***.

En l'absence de risques directs liés à la diffusion de données seulement obligatoires à la communication, il convient de de s'intéresser de manière plus précise à l'absence de diffusion de données obligatoirement communicables. Autrement dit, quels sont les risques qu'encourent l'organisation si des données obligatoirement communicables ne sont pas diffusées ?

Concernant l'absence de diffusion de données obligatoirement communicables. L'absence de diffusion peut se présenter sous plusieurs formes :

- L'établissement ne diffuse pas des données qui sont obligatoirement communicables du fait de leur nature. Cette absence de diffusion est effectuée en connaissance de la réglementation ou en méconnaissance de cette dernière. Elle peut donc être volontaire ou involontaire.
- L'établissement refuse de partager un document avec un citoyen qui en fait la demande, alors qu'il doit répondre d'une obligation de partager une telle information, sous réserve des cas d'interdiction de diffusion.

Il semblerait que **la Commission d'accès aux documents administratifs (ci-après « CADA »)** est compétente exclusivement en cas de manquement. Elle possède des pouvoirs gradués en fonction des atteintes et différencie les atteintes portées à la diffusion des données et à leur réutilisation. Concernant le point qui nous intéresse ici, la CADA semble ne pouvoir donner qu'un avis lorsque l'absence de diffusion de données obligatoirement communicable est notable.

Pour en savoir plus sur les pouvoirs de la CADA.

Plus précisément, elle dispose de quatre moyens pour remplir sa mission de vérification de communication des documents administratifs *(pour connaitre la définition d'un document administratif selon la loi, se référer à la Note méthodologique)* :

- Elle émet des avis « lorsqu'elle est saisie par une personne à qui est opposé un refus de communication ou un refus de publication d'un document administratif […], un refus de consultation des documents d'archives publiques […], ou une décision défavorable en matière de réutilisation d'informations publiques »[1].
- Elle conseille « les autorités mentionnées à l'article L.300-2 du Code sur toutes questions relatives à l'application des titres Ier, II, et IV du livre III du Code et titre Ier du Livre II du Code du patrimoine »[2].
- Elle propose toutes modifications des dispositions législatives ou règlementaires relatives au droit d'accès aux documents administratifs ou au droit de réutilisations des informations publiques[3].
- Elle peut infliger des sanctions à l'auteur d'un manquement aux règles de **réutilisation** (non de diffusion) des informations publiques[4]. Les sanctions sont prévues à l'article L. 326-1 du CRPA ;

➔ **En résumé,** s'agissant d'une absence de diffusion d'un document administratif, la CADA rend seulement un avis.

[1] Article L.342-1 du Code des relations entre l'Administration et le public (ci-après « CRPA »).
[2] Article R.342-1-1 du CRPA.
[3] Article R.342-5 du CRPA.
[4] Article L.342-3 du CRPA.

Il faut également indiquer qu'en cas d'avis de la CADA demandant à l'établissement de régulariser la diffusion des données obligatoirement communicables, ce dernier devra **mettre en place des moyens permettant cette diffusion**. En effet, la diffusion de données oblige l'établissement à mettre en place **une analyse au cas par cas** de chacune des données présentes dans sa base de données. ***Pourquoi ?*** Car certaines données qui ne sembleraient pas obligatoirement communicables peuvent en réalité l'être (exemple de données à caractère personnelle relatant d'un intérêt environnemental tel que mentionné dans le Code de l'environnement). De même, certaines données obligatoirement communicables à première vue pourraient sembler ne pas l'être (certaines exceptions prévues par la loi pour une République numérique du 7 octobre 2016). L'établissement se verrait donc obligé d'attacher **du personnel** à cette tâche, voire de mettre en place des **équipements et moyens techniques** (attribution d'un bureau, d'un ordinateur, d'un accès à la base de données). Enfin, cette analyse peut être **longue**. L'établissement devra faire un travail sur un pas de temps contraint alors qu'il aurait pu anticiper ce travail. En somme, en plus du **risque de l'émission d'un risque de la CADA**, il existe également **un risque organisationnel**.

ÉVALUATION DU RISQUE.

Aléa (A) : Moyen = il semblerait que la CADA puisse vérifier ou soit saisie par un citoyen qui se confronterait à un refus de communication.	**Enjeu (E) : Bas** = ce ne sont que des avis.	**Risque (R) Faible** = l'absence de communication est un risque qui peut être pris.
Ensemble des risques Σ(i) : Bas = les seuls risquent sont la saisie de la CADA par un citoyen, ou son autosaisie.	**Types des risques (Ri) : Faible** = les risques restent tous faibles car seul un avis semble pouvoir être effectué.	**Augmentation du risque : Faible** = il n'y a pas d'augmentation du risque initial.

<u>CONCLUSION</u>

Le risque est Faible en cas d'absence de diffusion des données obligatoirement communicables.

3 LES RISQUES LIÉS À LA DIFFUSION DE DONNÉES DONT LA COMMUNICATION N'EST PAS OBLIGATOIRE

Ces données ne sont pas, du fait de leur nature, ou de leur origine, obligatoire communicables et font donc l'objet d'une **dispense de diffusion au regard de la loi**. En d'autres termes, l'établissement a **trois choix** s'agissant de ces données :

- Il peut **les diffuser dans les mêmes conditions qu'une donnée obligatoirement communicable**, et donc permettre à toute personne d'y avoir accès.
- Il peut **les diffuser de façon plus restrictive** en limitant leur accès et donc leur réutilisation. C'est le cas par exemple si l'établissement partage les données sur une plateforme dédiée à certains chercheurs, et où une identification par mot de passe est nécessaire pour accéder à cette plateforme.
- Il peut **ne pas les diffuser** et les garder secrète au sein de sa base de données.

Analysons désormais les risques potentiels présents dans chacune de ces trois situations.

→ Précision : cette analyse est effectuée sur la base qu'une vérification au cas par cas des données a été effectuée par l'établissement et a permis d'identifier les données interdites à la diffusion, lesquelles ne rentrent pas dans cette partie de l'analyse de risque *(cf 4 « Le risque lié à la diffusion de données dont la communication est interdite. »).*

3.1 CONCERNANT LA DIFFUSION DANS LES MÊMES CONDITIONS QU'UNE DONNÉE OBLIGATOIREMENT COMMUNICABLES.

Ce type de diffusion présente des risques très limités voire nuls. En effet, au regard de la loi, rien n'interdit cette diffusion et le législateur l'encoure même, il semblerait donc, qu'aucune sanction ne puisse être infligée dans l'hypothèse d'une telle diffusion.

Néanmoins, il faut rappeler qu'une telle diffusion impose la mise en place de la **garantie d'une réutilisation libre de ces données**, pour tout intéressé. Cette réutilisation peut cependant faire l'objet de Licences plus restrictives limitant l'utilisation des données à des fins non commerciales par exemple. Pour connaitre les règles et les risques subsistant autour de la

mise en place de la réutilisation, je vous renvoie donc à la lecture de la ***Partie 5 « Les risques liés à l'atteinte portée à la réutilisation des données ».***

ÉVALUATION DU RISQUE.

Aléa (A) : Bas	**Enjeu (E) : Bas**	**Risque (R) : Faible**
= le législateur encourage la diffusion de données qui ne sont pas obligatoires à la communication, donc il ne prévoit pas de sanction.	= aucune sanction légale	= la diffusion de ces données est un risque qui peut être pris.

CONCLUSION

Le risque est faible en cas de diffusion des données qui ne sont pas obligatoirement communicables.

3.2 CONCERNANT LA DIFFUSION RESTRICTIVE OU L'ABSENCE DE DIFFUSION DE DONNÉES N'ÉTANT PAS OBLIGATOIREMENT COMMUNICABLES.

Ce type de diffusion présente des risques très limités voire nuls. En effet, le législateur n'impose pas de diffusion donc il ne peut pas apposer de sanction si cette dernière n'est pas effective. En d'autres termes une diffusion restrictive des données ou son absence ne peut engager de sanctions pécuniaires ou de rappel à l'ordre d'une des autorités indépendantes de l'État.

S'agissant d'autres risques potentiels. L'absence de diffusion ne présente pas de risque particulier car les données sont conservées et restent secrètes, personne, mis à part les personnes participantes au projet y ont accès. Dans le cas d'une diffusion partielle, le risque serait que les chercheurs réutilisent les données partagées et les partagent à leur tour dans un autre format, permettant alors leur accès à un cercle de personnes plus large. Ce risque peut être limité par la signature d'un accord de confidentialité lié à la plateforme, imposant l'impossibilité ou la limitation de réutilisation. En effet, comme ces données ne sont pas

obligatoirement diffusables, si leur diffusion ne s'effectue pas dans un cadre large et ouvert, il semblerait que la réutilisation puisse être limitée car son accessibilité l'est également.

ÉVALUATION DU RISQUE.

Aléa (A) : Bas	**Enjeu (E) : Bas**	**Risque (R) : Faible**
= le législateur n'oblige pas la diffusion de ces données.	= aucune sanction légale	= l'absence de diffusion de ces données est un risque qui peut être pris.

CONCLUSION

Le risque est faible en cas de diffusion partielle ou d'absence de diffusion des données qui ne sont pas obligatoirement communicables.

4 LE RISQUE LIÉ À LA DIFFUSION DE DONNÉES DONT LA COMMUNICATION EST INTERDITE.

En comparaison avec les risques présentés jusqu'à présent, cette catégorie semble la plus génératrice de risques et donc de difficultés potentielles pour le projet DATA4C+. Nous effectuerons une analyse pour trois types de données : les données à caractère personnel, les données protégées par un droit de propriété intellectuelle et les données protégées par un secret définit par la loi. En effet, ce sont les trois principales causes d'interdiction de diffusion des données énoncées par la **loi pour une République numérique du 7 octobre 2016**[5] *(cf Note méthodologique)*.

4.1 LES DONNÉES À CARACTÈRE PERSONNEL

Selon la **Commission nationale de l'informatique et des libertés (ci-après « CNIL »)**, une donnée à caractère personnel **est *« toute information relative à une personne physique susceptible d'être identifiée, directement ou indirectement »***[6]. Peu importe que ces données soient confidentielles ou publiques, leur divulgation peut porter atteinte aux droits personnels des individus. Dans le projet DATA4C+, des données à caractère personnel peuvent être collectées et s'apparentent à :

- Prénom ;
- Nom ;
- Adresse postale (localisation de la parcelle où les données sont prélevées) ;
- Numéro de téléphone ;
- Adresse mail ;
- Photo ;
- Spécialités (agriculteur, chercheur).

[5] Loi n°2016-1321 du 7 octobre 2016 pour une République numérique dite aussi « Lemaire ».
[6] CNIL, *Une donnée à caractère personnel, c'est quoi ?*

Ces données sont protégées par le **Règlement général sur la protection des données (ci-après « RGPD »)**[7] et font donc l'objet d'une protection renforcée. Leur diffusion mais également leur collecte sont strictement encadrées par le texte.

Quel est la mission du RGPD ? Protéger la vie privée des individus en limitant la diffusion de leurs données aux communications d'informations qu'ils auront autorisés préalablement à la collecte.

Quid de l'application du texte aux données collectées antérieurement au texte.

→ La collecte des données est antérieure au RGPD + la diffusion des données est antérieure au RGPD = le RGPD ne s'applique pas.

→ La collecte des données est antérieure au RGPD + la diffusion des données est postérieure au RGPD = le RGPD s'applique.

→ La collecte des données est postérieure au RGPD + la diffusion des données est postérieure au RGPD = le RGPD s'applique.

➔ *Pour en savoir plus sur le RGPD, les droits qu'il protège et la manière de le mettre en place lors de la collecte et de la diffusion des données à caractère personnel, je vous renvoie à la* ***Note méthodologie rédigée dans le cadre du projet DATA4C+*** *ou au site de la CNIL dont je vous inscris en lien :* https://www.cnil.fr/fr/rgpd-par-ou-commencer

La Commission nationale de l'informatique et des libertés (ci-après « CNIL ») est l'autorité administrative indépendante française compétente pour vérifier de la bonne application du RGPD. Lorsque des données à caractère personnel sont diffusées sans une anonymisation préalable, le titulaire des droits personnels, ou tout autre personne qui constate une atteinte à un tel droit, peut saisir la CNIL afin qu'elle procède à une ***vérification*** et / ou à des ***avertissements*** et / ou à des ***sanctions graduées***. Ces actions de la CNIL sont graduées et prennent la formes d'étapes à suivre par cette dernière :

[7] Règlement UE 2016/679 du Parlement européen et du Conseil du 27 avril 2016.

- **Étape 1 :** Avertissement ou mise en demeure de l'entreprise fautive avec un rappel du devoir de mise en conformité des traitements de données au RGPD.
- **Étape 2 :** Injonction de cesser la violation
- **Étape 3 :** Limitation ou suspension temporaire des traitements de données (dans certains cas).
- **Étape 4 :** Sanctions administratives ou pénales *(cf Tableau ci-dessous)* en cas de non-respect aux règles du RGPD après injonction vaine de l'autorité de contrôle.

<table>
<tr><th colspan="2">SANCTION ADMINISTRATIVE[8]
(2 types de sanctions en fonction de la gravité)</th><th>SANCTION PÉNALE[9]</th></tr>
<tr><td>2% du chiffre d'affaire mondial pour les entreprises <u>ou</u> 10 millions d'euros d'amende ; applicables en cas de manquement aux :
- Obligations incombant au responsable du traitement et au sous-traitant.
- Obligations incombant à l'organisme de certification.
- Obligations incombant à l'organisme chargé du suivi des codes de conduite.</td><td>4% du chiffre d'affaire mondial pour les entreprises <u>ou</u> 20 millions d'euros d'amende ; en cas de manquements à :
- L'obligation de consentement préalable de la personne concernée (avant collecte et stockage des données).
- Aux autres droits des personnes concernées.
- Aux transferts de données à caractère personnel à un destinataire situé dans un pays tiers ou à une organisation internationale.
- A toutes les obligations découlant du droit des États membres.
- Au non-respect d'une injonction, d'une limitation temporaire ou définitive du traitement ou de la suspension des flux de donnés ordonnée par l'autorité de contrôle.</td><td><u>Objectif</u> : réprimander les violations qui ne font pas l'objet d'amendes administrative[10] Exemple : sanction en cas de détournement de la finalité des données personnelles lors du traitement de la donnée

<u>Sanction</u> : Jusqu'à 5 ans d'emprisonnement et 300 000 euros d'amende[11].</td></tr>
</table>

[8] Article 83 du RGPD
[9] Article 84 du RGPD
[10] Articles 226-16 à 226-4 du Code pénal
[11] Article 226-16 du Code pénal

Que risque concrètement l'établissement en cas de diffusion de données à caractère personnel ?

En diffusant des données à caractère personnel alors qu'il n'est pas été autorisé, l'établissement risque une sanction pouvant s'élever jusqu'à 4% du chiffre d'affaire ou 20 millions d'euros selon la somme la plus importante, et selon la gravité du manquement aux obligations *(cf tableau ci-dessus)*.

Quels sont les précautions à prendre pour éviter une sanction ?

- Demande d'accord préalable pour la collecte et le traitement des données personnelles ;
- Collecte des données uniquement pour un certain traitement défini en amont ;
- Information en cas de fuites ou de piratages des données aux personnes concernées ainsi qu'à l'autorité de contrôle ;
- Garantir le droit à la consultation et à la portabilité des données pour les utilisateurs ;
- Documentations clair sur les mesures et outils en place pour le traitement

Concrètement, que doit faire l'organisme pour minimiser les risques en cas de diffusion de données à caractère personnel ?

- *Préalablement à la diffusion des données :* anonymiser les données → cette anonymisation doit être faite avec précaution car il faut éviter toute possibilité de remonter jusqu'à l'individu titulaire des droits (réversibilité).[12]
- *Préalablement à la collecte des données :* demander à la personne concernée de consentir à la diffusion de ses données pour une finalités et dans un cadre précis → ce consentement, doit être formulé par écrit.

[12] Article 116 du Code de la protection des données.

ÉVALUATION DU RISQUE.

Aléa (A) : Élevé = la CNIL est de plus en plus regardante s'agissant de l'atteinte des données à caractère personnel.	**Enjeu (E) : Élevé** = cela peut aller de l'avis à la sanction financière + atteinte à la réputation mauvaise image de l'établissement véhiculée par la diffusion de données à caractère personnel sans autorisation)	**Risque (R) : Maximum** = nous déconseillons la diffusion de ces données.
Ensemble des risques Σ(i) : Moyen - risque d'atteinte à la personne concernée par les données et donc qu'elle saisisse l'Autorité compétente. - risque d'utilisation malveillante des données (atteinte à la réputation du projet).	**Type des risques (Ri) : Élevé** - aspect financier potentiel - réputation du projet : certains agriculteurs pourraient refuser la collecte de leurs données et bloquer les recherches sur leurs sols.	**Augmentation du risque initial : Maximum** = le risque initial est augmenté du fait de la pluralité des risques ; nous déconseillons la diffusion de ces données.

<u>CONCLUSION</u>

Le risque est Maximum ce qui signifie qu'il faut éviter la diffusion de ces données.

➔ **REMARQUE : afin de pouvoir diffuser de telles données, l'anonymisation est possible. Cependant, il convient d'être prudent car la CNIL rappelle que l'anonymisation parfaite n'existe pas et que des risques peuvent se poursuivre. Ainsi, en cas de diffusion de données anonymisées, nous pourrions estimer que le risque s'amoindri, passant à Intermédiaire, mais qu'il n'est pas faible pour autant car il est possible pour des individus de retracer les données (notamment en les comparant avec d'autres données), et d'en identifier une personne physique.**

4.2 LES DONNÉES PROTÉGÉES PAR UN DROIT DE PROPRIÉTÉ INTELLECTUELLE

Le droit de la propriété intellectuelle permet de **protéger certaines données et base de données par le biais du droit d'auteur mais également du droit *sui generis* du producteur de la base**. Ainsi, l'analyse de risque permettra de déceler les fragilités pouvant porter atteinte à un droit d'auteur mais également à un droit *sui generis* du producteur de la base de données.

Lorsque des données protégées par un droit de propriété intellectuelle sont diffusées sans l'accord préalable du titulaire de ces droits, ce dernier peut en notifier l'organisme voire saisir le juge en réparation des dommages qu'il a subi. **Quels peuvent être ces dommages ?** Ils sont multiples et il faut distinguer l'atteinte effectuée à un *droit d'auteur*, à un *droit sui generis du producteur de la base de données* ou, à un *droit de propriété industrielle*. Ces dommages peuvent néanmoins se résumer à l'atteinte de l'intégrité de l'œuvre, à la perte de gain, voire à l'atteinte du titulaire même des droits et à sa réputation.

➔ **PRÉCISION SPÉCIFIQUE À CETTE ANALYSE DE RISQUE**. Concernant une donnée protégée par un droit de propriété industrielle et commerciale, il semblerait qu'il soit possible de l'exclure dans cette analyse de risque. ***Pourquoi ?*** Dans le cas particulier du système DATA4C+, une atteinte à un droit de la propriété industrielle et commerciale (droit des marques, droit des brevets ou encore des dessins et modèles) semble limitée tant les données y faisant référence sont rares dans les bases de données attachées au système.

Notion de communication au public et droit de propriété intellectuelle :

- **CJUE[13] 31 mai 2016** : La communication au public doit s'apprécier selon les mêmes critères en matières de droit d'auteur et de droits voisins.
- **CJUE 7 mars 2013** : La communication au public est effective lorsqu'il y a la réunion de *deux actes* : l'acte matériel accompli, querellé, entre dans la qualification juridique de « communication » **+** cet acte est destiné au public.
- **CJUE 14 juin 2017** : Accomplit un acte de communication au public la personne qui joue un rôle incontournable, c'est-à-dire qui *« intervient, en pleine connaissance des*

[13] **CJUE : Cour de justice de l'Union européenne.**

conséquences de son comportement, pour donner à ses clients accès à une œuvre protégée, et ce notamment lorsqu'en l'absence de cette intervention, ses clients ne pourraient, en principe, jouir de l'œuvre diffusée ».

- **<u>PUBLIC</u> → CJUE 7 mars 2013** : Un nombre indéterminé de destinataires potentiels et qui implique par ailleurs un nombre assez important (effet cumulatif qui résulte de la mise à disposition des œuvre auprès des destinataires potentiels qui accèdent aux œuvres communiquées par le biais d'une connexion un à un.

	Données relatives à un DROIT D'AUTEUR	**Données relatives à un DROIT SUI GENERIS DU PRODUCTEUR DE LA BASE DE DONNÉES**	**Données relatives à un DROIT DE PROPRIÉTÉ INDUSTRIELLE ET COMMERCIALE**
Nature de l'atteinte	- Violation du droit moral de l'auteur. - Violation de ses droits patrimoniaux.	- Atteinte aux droits du producteur de la base de données. - Contrefaçon.	- Contrefaçon.
Sanction	- **Sanction civile** : dommages et intérêts **+** cessation de l'exploitation de l'œuvre. - **Sanction pénale**[14] : 3 ans d'emprisonnement et 300 000 euros d'amende **+** confiscation des recettes procurées par l'infraction.	- Le juge peut imposer des mesures d'urgence. - Sanction pénale[15] : 3 ans d'emprisonnement et 300 000 euros d'amende. **CONDITION** : la base de données doit être protégée.	- **Sanction civile** : dommages et intérêts - **Sanction pénale** : 3 ans d'emprisonnement et 300 000 euros d'amende.

[14] Article L.335-2 du Code de la propriété intellectuelle.
[15] Article L.343-4 du Code de la propriété intellectuelle.

Que risque concrètement l'établissement en cas de diffusion de données à caractère personnel ?

En diffusant des données protégées par un droit de propriété intellectuelle alors qu'il n'y était pas autorisé, l'établissement semble, plus probablement, risquer une amende et l'obligation de faire cesser l'atteinte *(cf tableau ci-dessus)*.

Concrètement, que doit faire l'organisme pour minimiser les risques en cas de diffusion de données à caractère personnel ?

- *Préalablement à la collecte des données :* déterminer s'il existe des droits de propriété intellectuelle attachés aux données collectées. Si cela est le cas, il faut déterminer quels sont les titulaires de ses droits et leur demander leur accord pour la collecte des données.
- *Préalablement à la diffusion des données :* demander l'autorisation au titulaire des droits de diffuser les données. Une cession de droit peut être effectuée (elle peut prendre la forme d'une Licence qui va déterminer les contours géographiques et temporels de la communication). Si cela n'est pas le cas, il faudra éviter de partager ces données.

ÉVALUATION DU RISQUE

Aléa (A) : Moyen = le juge peut être saisi par le titulaire des droits d'auteurs ou son représentant.	**Enjeu (E) : Élevé** = une sanction financière peut être demandée en réparation des atteinte subis, ainsi que l'interdiction ; ainsi que l'obligation de faire cesser l'atteinte (suppression des documents partagés).	**Risque (R) : Maximum** = nous déconseillons la diffusion de ces données sans l'autorisation préalable de l'auteur.
Ensemble des risques Σ(i) : Bas - risque d'attente à la personne concernée par les données et donc qu'elle saisisse le juge.	**Type des risques (Ri) : Élevé** - aspect financier potentiel - réputation du projet car diffuse et reprend les recherches d'autres.	**Augmentation du risque initial : Intermédiaire** = le risque initial ne semble pas augmenter.

CONCLUSION

La diffusion de données protégées par un droit d'auteur ou un droit sui generis du producteur de la base de données semble Maximum. Ainsi, la diffusion de ces données est déconseillée.

→ **Précision :** plus le temps passe, plus il semblerait que la probabilité d'une saisine du juge diminue.

→ **REMARQUE : afin de pouvoir diffuser de telles données il convient de demander l'autorisation expresse et écrite du titulaire des droits pour la diffusion. Une cession de droit peut également être faite et doit contenir des mentions obligatoires telles que :**

- **Le type du droit cédé (reproduction, représentation, adaptation ou diffusion).**
- **L'étendue, la destination, le lieu et la durée d'exploitation du droit cédé.**
- **Les modes de calcul et de paiement du droit cédé.**

4.3 LES DONNÉES CORRESPONDANT À UN SECRET PROTÉGÉES PAR LA LOI

Certaines données font l'objet de secrets protégées par la loi. Ces secrets peuvent **être absolus et opposables à toutes personnes** ou, **relatifs et opposables seulement aux tiers et non aux personnes concernées**. L'analyse de risque permet d'envisager les conséquences d'une diffusion de ces données protégées par un secret, et le niveau d'acceptabilité d'une communication.

Lorsque des données sont diffusées, cette communication peut porter atteinte à des secrets protégés par la loi. Ces secrets peuvent prendre plusieurs formes. Cette analyse de risque traitera le cas du ***« secret d'affaires »***, du ***« secret d'État »*** et du ***« secret professionnel »***.

4.3.1 LE SECRET D'AFFAIRES

Le secret d'affaires permet de protéger des **techniques ou savoir-faire** détenus par une entreprise, un organisme ou des individus dans le cadre de leur profession ou pour améliorer certaines pratiques. Le savoir-faire est défini par l'accord sur les ADPIC comme **la connaissance non immédiatement accessible au public**. Plus précisément, l'article 39.2 a de l'accord indique que les renseignements doivent *« n'être ni connus de personnes appartenant aux milieux qui s'occupent normalement du genre de renseignements en question ni leur être aisément accessibles, dans leur globalité ou dans la configuration et l'assemblage exact de leurs éléments ».*

L'article 39.2 c de l'accord subordonne la protection juridique à ce que son possesseur licite ait pris des **mesures raisonnables**, compte tenu des circonstances, pour conserver le secret. De plus, ce possesseur doit **manifester sa volonté de se réserver ces connaissances**, en en interdisant l'accès intellectuel aux tiers.

EN PRATIQUE : des mesures raisonnables sembleraient prises lorsqu'une convention ou un contrat de confidentialité est signé par un organisme de recherche qui récupère des données ou informations renfermant un savoir-faire détenu par une entreprise tierce. Cette convention va encadrer l'utilisation qui pourra être faite de ces informations ; mais également identifier qui sont les titulaires des connaissances et donc des droits attachés au savoir-faire.

L'accord sur les ADPIC n'exige pas que le secret d'affaires présente une certaine distance par rapport aux connaissances antérieures, ni même qu'il se différencie par rapport à ces connaissances. De plus, le règlement UE n°316/2014 consacre la notion de secret relatif. Par exemple, le fait qu'une connaissance soit disponible sur internet ne la rend pas nécessairement « aisément accessible », si un code est requis pour y accéder.

EN PRATIQUE : il n'y aura pas violation du secret d'affaires tant que les données restent dans le système DATA4C+, ou dans une autre base de données autorisées à recevoir ces informations, et qu'aucune divulgation n'est faite au public.

Comment identifier si l'information que je détiens fait l'objet d'un secret d'affaires ?

L'article L.151-1 du Code de commerce va mettre en place **trois critères cumulatifs**[16] :

- N'est pas, en elle-même ou dans la configuration et l'assemblage exacts de ses éléments, **généralement connue ou aisément accessible** par des personnes familières de ce type d'informations en raison de leur secteur d'activité ;
- Elle fait l'objet de la part de son détenteur légitime de **mesures de protection raisonnables**, compte tenu des circonstances, pour en conserver le caractère secret ;
- Elle revêt **une valeur commerciale**, effective ou potentielle, du fait de son caractère secret.

REMARQUE : cette dernière condition, relative à la valeur commerciale, semble trop restrictive. Nous pouvons imaginer que les entreprises tiennent également secrètes d'autres types d'informations sans valeur commerciale mais avec **une forte valeur économique** *(exemples : stratégie de l'entreprise, projets de grandes importance, projet industriel, projet de croissance externe…)*.

Qui est le détenteur d'un secret d'affaire (et qui peut donc s'opposer à la diffusion ou demander une réparation en cas de diffusion) ?

Le Code de commerce[17] indique que la personne détentrice du secret est celle qui en a **le contrôle de façon licite**. *Quels sont les modes d'obtention licite ?*

- Par le biais d'une découverte ou d'une création indépendante ;
- Par le biais d'un procédé d'ingénierie inverse, c'est-à-dire l'observation, l'étude, le démontage ou le test d'un produit ou d'un objet qui a été mis à la disposition du public ou qui est de façon licite en possession de la personne qui obtient l'information, sauf stipulation contractuelle intéressant ou limitant l'obtention du secret.

Il existe des <u>exceptions</u> à la protection du secret des affaires :

- Lorsque l'obtention, l'utilisation ou la divulgation du secret est requise ou autorisée par le droit de l'UE, les traités ou accord internationaux en vigueur ou le droit national ;

[16] Le texte reprend la définition consacrée à l'article 2 de la directive du 8 juin 2016 2016/943/UE.
[17] Articles L.151-2 et L.151-3 du Code du commerce.

- Lorsque l'obtention du secret est intervenue dans le cadre de l'exercice du droit à l'information et à la consultation des salariés ou de leurs représentants ;
- Lorsque la divulgation du secret par des salariés à leurs représentants est intervenue dans le cadre de l'exercice légitime par ces derniers de leurs fonctions, pour autant que cette divulgation ait été nécessaire à cet exercice ;
- Car une protection spécifique est accordée aux journalistes et aux lanceurs d'alerte.

EN PRATIQUE : cela signifie que le risque sera amoindri lors de la diffusion ou de la communication des données**.**

4.3.2 LE SECRET D'ETAT

Le secret d'État est une notion encadrée tardivement par la législation. L'ordonnance n°60-529 du 4 juin 1960 renonce à l'idée de poser une définition légale du secret de la défense nationale. Cette absence de définition a obligé les autorités de l'État d'apprécier le caractère secret d'une information selon son contenu, son accès ou sa fonction.

UN PEU D'HISTOIRE JURIDIQUE POUR MIEUX COMPRENDRE CETTE NOTION : *pourquoi l'ordonnance renonce à une telle définition ?* Anciennement, les secrets de la défense nationale étaient classés par le décret-loi en quatre catégories de renseignement : « d'ordre militaire, diplomatique, économique ou industriel qui, par leur nature, ne doivent pas être connus que des personnels qualifiés pour les détenir et doivent, dans l'intérêt de la défense nationale, être tenus secrets à toute autre personne ». Cette définition fut par la suite abandonnée car ne couvrait pas l'ensemble des informations dont la protection était nécessaire.

CONCERNANT LA FRANCE : avec le décret n°81-514 du 12 mai 1981[18] les autorités françaises font le choix de classer les informations intéressant la défense nationale et la sûreté de l'État qui doivent être tenues secrètes **en trois niveau de protection** : **très secret défense**, **secret défense** et **confidentiel défense**.

[18] Relatif à l'organisation de la protection des secrets et informations concernant la défense nationale et la sûreté de l'État, JO 15 mai

Quels sont les types d'objets pouvant faire l'objet d'un secret d'État ?

L'article 413-9 du Code pénal indique que présente *« un caractère de secret de la défense nationale au sens de la présente section les procédés, objets, documents, informations, réseaux informatique, données informatisées ou fichiers intéressant la défense nationale qui ont fait l'objet de mesures de classification destinées restreinte leur diffusion ou leur accès ».*

Quelle est la sanction risquée en cas d'atteinte à un secret d'État ?

L'article 413-10 du Code pénal indique qu'est puni de *« **sept ans d'emprisonnement et de 100 000 euros d'amende** le fait, par toute personne dépositaire, soit par état ou profession, soit en raison d'une fonction ou d'une mission temporaire ou permanente, d'un procédé, objet, document, information, réseau informatique, donnée informatisée ou fichier qui a un caractère de secret de la défense nationale, soit de le détruire, détourner, soustraire ou de le reproduire, soit d'en donner l'accès à une personne non qualifiée ou de la porter à la connaissance du public ou d'une personne non qualifiée »*. De plus, *« lorsque la personne dépositaire a agi par imprudence ou négligence, l'infraction est punie de **trois ans d'emprisonnement et de 45 000 euros d'amende**. »*.

4.3.3 LE SECRET PROFESSIONNEL

Le secret professionnel est un autre secret protégé par la loi. Il vise *« 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 »*[19]. La diffusion de données conduisant à la violation d'un secret professionnel est punie **d'un an d'emprisonnement et de 150 000 euros d'amende**.

➔ **Précision spécifique à cette analyse de risque**. Concernant une donnée protégée par un droit de propriété industrielle et commerciale, il semblerait qu'il soit possible de l'exclure dans cette analyse de risque. ***Pourquoi ?*** Dans le cas particulier du système DATA4C+, une atteinte à un secret professionnel semble limitée tant les données y faisant référence sont rares dans les bases de données attachées au système.

[19] Article 226-13 du Code pénal.

BILAN DES SANCTIONS POTENTIELLES

SECRET D'AFFAIRES	SECRET D'ETAT	SECRET PROFESSIONNEL
- Dommages et intérêts à la hauteur du préjudice subi (versés par le contrevenant qui savait ou aurait dû savoir qu'il obtenait, divulguait ou utilisait un secret d'affaires de manière illicite)	- 7 ans d'emprisonnement et 100 000 euros d'amende	- 1 an d'emprisonnement et 150 000 euros d'amende.

Que risque concrètement l'établissement en cas de diffusion de données à caractère personnel ?

En diffusant des données protégées par un secret protégé par la loi alors qu'il n'y était pas autorisé, l'établissement semble, plus probablement, risquer une amende et l'obligation de faire cesser l'atteinte *(cf tableau ci-dessus)*.

Concrètement, que doit faire l'organisme pour minimiser les risques en cas de diffusion de données à caractère personnel ?

- *Préalablement à la collecte des données :* déterminer s'il existe des données faisant l'objet d'un secret protégé par la loi. Si cela est le cas, il faut déterminer l'importance des informations collectées et déterminer leur niveau de protection.
- *Préalablement à la diffusion des données :* demander l'autorisation au titulaire du secret de diffuser ces données. Il est possible de cacher ces données afin de limiter le risque lors de la diffusion. Si ces précautions ne peuvent pas être faites du fait de la technicité nécessaire, il convient de ne pas diffuser des informations qui seraient sensibles pour une entreprise, ou encore un État.

ÉVALUATION DU RISQUE CONCERNANT UN ATTEINTE PORTEE À UN SECRET D'AFFAIRES.

Aléa (A) : Élevé = Les dommages et intérêts peuvent être importants en fonction du partenaire avec qui été tenu le secret.	**Enjeu (E) : Élevé** = Sanction financière	**Risque (R) : Maximum** = Nous déconseillons la diffusion de ces données.
Ensemble des risques Σ(i) : Élevé - Risque d'atteinte à l'intérêt du partenaire ou à vos propres intérêts. - Risque de nuire à la relation avec le partenaire.	**Type des risques (Ri) : Élevé** - Aspect financier potentiel - Réputation du projet : certains partenaires ou personnels pourraient refuser de poursuivre le projet, voire de prendre part dans de futurs projets.	**Augmentation du risque initial : Maximum** = Le risque initial est augmenté du fait de la pluralité des risques ; nous déconseillons la diffusion de ces données.

ÉVALUATION DU RISQUE CONCERNANT UN ATTEINTE PORTEE À UN SECRET D'ETAT.

Aléa (A) : Élevé = Nous pouvons supposer que l'État surveille et limite les atteintes qui sont portés à ses secrets potentielles, donc il semble en mesure d'agir.	**Enjeu (E) : Élevé** = Une sanction financière voire une peine de prison sont encourue [les personnes morales pourraient se voir attribuer une sanction financière car elles ne peuvent pas être emprisonnées].	**Risque (R) : Maximum** = Nous déconseillons la diffusion de ces données.
Ensemble des risques Σ(i) : Élevé - Risque d'atteinte à l'intérêt de l'État. - Risque d'utilisation malveillante des données.	**Type des risques (Ri) : Élevé** - Aspect financier potentiel - Emprisonnement - Réputation du projet et collaboration avec l'État misent en péril.	**Augmentation du risque initial : Maximum** = le risque initial est augmenté du fait de la pluralité des risques ; nous déconseillons la diffusion de ces données.

CONCLUSION

S'agissant à la fois du secret d'État et du secret d'affaires, le risque semble Maximum et il convient de ne pas diffuser les données.

➔ **REMARQUE : afin de pouvoir diffuser de telles données il convient de demander l'autorisation expresse et écrite de la personne responsable pour communiquer les données. De plus, il sera également possible de supprimer les mentions sensibles de données. Ce dernier point doit cependant être fait avec beaucoup de précaution afin que personne ne puisse voir la mention retirée.**

BILAN DES SANCTIONS RISQUÉES EN CAS D'ATTEINTE À DES DROITS LORS DE LA DIFFUSION DES DONNÉES INTERDITES À LA COMMUNICATION.

DONNÉES À CARACTÈRE PERSONNEL	DROIT DE LA PROPRIÉTÉ INTELLECTUELLE	AUTRE PROTECTION DE LA LOI
- Sanction administrative - Sanction pénale **SANCTION LA PLUS RISQUÉE** ➔ Mise en garde la CNIL ➔ Amende	- Droit auteur - Droit *sui generis* du producteur de la base de données **SANCTION LA PLUS RISQUÉE** ➔ Dommages et intérêts ➔ Amende ➔ Interdiction de diffusion	- Secret d'affaires - Secret d'État **SANCTION LA PLUS RISQUÉE** ➔ Dommages et intérêts ➔ Amende ➔ Interdiction de diffusion

5 LES RISQUES LIÉS À L'ATTEINTE PORTÉE À LA RÉUTILISATION DES DONNÉES

Chacune des parties que nous avons pu voir jusqu'à présent traite de la diffusion des données. Cependant, l'évolution principale de la loi pour une République numérique a été d'imposer la réutilisation libre et gratuite pour la diffusion des données obligatoirement communicables. La loi pose un cadre autour de cette réutilisation.

→ Pour en savoir plus, je vous renvoie à la lecture de la ***note méthodologique rédigée dans le cadre du projet DATA4C+.***

Ainsi, il convient désormais de traiter le cas des risques liés à la réutilisation des données, et plus particulièrement, aux risques liés aux failles présents dans la réutilisation mise en place par l'établissement (**5.1**), mais également les risques en cas de mauvaise réutilisation de données diffusées par autrui cette fois (**5.2**).

5.1 LES RISQUES LIÉS À LA MAUVAISE MISE EN PLACE DE LA RÉUTILISATION DES DONNÉES DIFFUSÉES.

La loi pour une République numérique du 7 octobre 2016 impose une réutilisation libre et gratuite des données. Cette réutilisation est possible à des fins commerciales comme non commerciales et ne peut être empêchées.

Afin de limiter la réutilisation de ses données, l'établissement peut invoquer un motif d'intérêt général pour limiter. Ce motif est cependant encadré par le contrôle du juge. ***En quoi consiste ce contrôle ?*** Le juge pourra être amené, dans le cadre d'un contrôle de proportionnalité, à apprécier **l'atteinte portée au principe d'égalité entre réutilisateurs** ou à **la liberté du commerce et de l'industrie en matière de réutilisation commerciale**.

Qu'est-ce que le principe d'égalité entre réutilisateurs ? Il s'agit tout simplement de l'interdiction de favoriser un réutilisateur face à un autre. **Spécifiquement au projet DATA4C+,**

les établissements ne pourront pas ouvrir et permettre la réutilisation de données pour un professionnel et l'interdire pour un autre.

→ **S'agissant d'une sanction liée à une atteinte à l'égalité entre réutilisateurs**. Il semble que le juge puisse demander à ce que soient verser des dommages et intérêts. Mais la jurisprudence ne s'est pas encore prononcée sur le sujet, donc il est difficile de savoir ce qui est réellement risqué.

Qu'est-ce que la liberté du commerce et de l'industrie ? La Déclaration des droits de l'Homme et du citoyen dispose que « la liberté consiste à pouvoir faire tout qui ne nuit pas à autrui ». Plus précisément, la liberté du commerce et de l'industrie regroupe trois domaines :

- La liberté d'entreprendre, c'est-à-dire qu'un entrepreneur peut d'établir et exercer où il souhaite.
- La liberté d'exploitation, c'est-à-dire qu'un entrepreneur peut exercer son activité comme il l'entend.
- La liberté de concurrence, c'est-à-dire qu'aucun monopole ne peut exister et que la concurrence doit permettre aux entrepreneurs d'attirer des clients d'autres entreprises.

En d'autres termes, **spécifiquement au projet DATA4C+,** la liberté du commerce et de l'industrie consiste à ne pas refuser la réutilisation à un entrepreneur qui en aurait besoin pour s'établir (par exemple, il pourrait se questionner sur la présence de matières premières pour son activité avant de localiser sa société). Néanmoins, il semble qu'une telle atteinte à la liberté du commerce et de l'industrie ne puisse être produite dans le cadre du projet DATA4C+.

→ **S'agissant d'une sanction liée à une atteinte à la liberté du commerce et de l'industrie**. Le juge peut demander une réparation du dommage subi et notamment du manque à gagner en cas de pratiques anticoncurrentielles.

ÉVALUATION DU RISQUE.

Aléa (A) : Bas	**Enjeu (E) : Moyen**	**Risque (R) : Faible**
= Il semblerait que l'atteinte semble limitée, et peu de jurisprudence traitent le sujet.	= les sanctions sont financières et se limitent à des dommages et intérêts	= le risque peut être pris.

CONCLUSION

Le risque est Faible, il peut dont être pris. Néanmoins, la volonté générale des partenaires du projet tant à ne pas limiter la réutilisation là où elle ne l'est déjà pas. Ainsi, il semblerait que ce type de risque ne soit pas pris quoi qu'il en soit.

5.2 LES RISQUES LIÉS À LA RÉUTILISATION FRAUDULEUSE DE DONNÉES DIFFUSÉES PAR AUTRUI.

Dans cette partie, l'établissement ne se positionne pas en tant que diffuseur mais **en tant que réutilisateur des données.** Même si cette partie n'intéresse pas particulièrement ou du moins directement le projet DATA4C+, il semblait intéressant de la traiter car l'objectif même de l'ouverture des données est de rendre la recherche plus fructueuse par le partage des connaissances. Ainsi, il est donc normal que les chercheurs et organismes qui donnent puissent également recevoir et apprendre autres.

Lorsque qu'une réutilisation des données apparait frauduleuse, c'est la CADA qui est compétente pour se prononcer en la matière. Elle peut infliger des sanctions à l'auteur d'un manquement aux règles de réutilisation des informations publiques[20]. Les sanctions sont prévues à **l'article L. 326-1 du CRPA.**

Type de réutilisation frauduleuse	Sanction
Réutilisation en méconnaissance des règles indiquées par la loi à des fins non commerciales.	Montant maximum de 1 500€ pouvant aller jusqu'à 3 000€ en cas de récidive (égal à une contravention de 5e classe)
Réutilisation en méconnaissance des règles indiquées par la loi à des fins commerciales.	Ne peut pas excéder 1 millions d'euros pour un premier manquement. En cas de manquement réitéré dans les cinq année, le montant de l'amende sera de 2 millions d'euros ou de 5% du chiffre d'affaire hors taxe pour les entreprises.

[20] Article L.342-3 du CRPA.

De plus, la loi exige que lorsqu'un individu réutilise des données publiées, il doit :

- Mentionner la source des données ;
- Mentionner la date de dernière mise à jour de la réutilisation ;
- Ne pas altérer le sens des données.

ÉVALUATION DU RISQUE

Aléa (A) : Élevé	**Enjeu (E) : Moyen**	**Risque (R) : Intermédiaire**
= les individus sont susceptibles de faire valoir leurs droits car ils diffusent gratuitement, et souhaitent un minimum de retour ou tout simplement que l'on ne s'approprie pas leur travail.	= diffère en fonction de l'atteinte portée, mais il s'agit d'une amende.	= le risque peut être pris si des précautions encadrent la réutilisation.

<u>CONCLUSION</u>

Le risque est Intermédiaire, des précautions sont donc à prendre.

6 CONCLUSION DE L'ANALYSE DE RISQUE.

A la lumière de l'analyse, on conclut que lorsque l'analyse au cas par cas a permis de bien identifier les catégories dans lesquelles les données se trouvent :

- **S'agissant du dévoilement des données obligatoirement communicables :** le risque semble Faible, mais il faudra pour autant de conformer à l'avis de la CADA si elle en émet un.
- **S'agissant du dévoilement ou de l'absence de dévoilement des données non obligatoirement communicables :** là encore le risque semble Faible même s'il convient toujours de prendre des précautions.
- **S'agissant du dévoilement des données par principe interdites à la communication :** le risque est généralement Maximum. La diffusion de ces données est donc déconseillée en général. Il convient de prendre les précautions indiquées dans chacune des parties.

La probabilité qu'un risque survienne dépend de plusieurs facteurs. Outre ceux qui ont été énoncé depuis le début de cette analyse de risque, il convient d'indiquer que certains facteurs peuvent **augmente**r mais également **diminuer le risque**.

→ **LE FACTEUR TEMPS :** il peut permettre de diminuer le risque. En effet, lorsque la personne titulaire des droits en cause est décédée, il parait moins probable qu'une saisine du juge ou d'une Autorité compétente soir effectuée. Attention cependant, car le risque n'est pas nul pour autant, donc il convient de toujours prendre les précautions associées à la donnée en question

→ **LE FACTEUR DU NOMBRE :** plus de données sont partagées plus le risque à prendre est grand. C'est pour cela qu'il est essentiel d'effectuer une analyse préalable la diffusion des données. Elle permettra à l'établissement de distinguer les données obligatoirement diffusables de celles qui ne le sont pas, et surtout d'identifier les données présentant une interdiction de diffusion. Aux fins de comprendre les règles qui entourent cela et de mettre en place une méthode d'analyse, je vous renvoies à ***la note méthodologique rédigée dans le cadre du projet DATA4C+.***

→ **LES DIVERSES PRÉCAUTIONS ÉNONCÉES DANS CHACUNE DES PARTIES** : permettent de limiter les risques.

BIBLIOGRAPHIE

- CNIL, *Une donnée à caractère personnel, c'est quoi ?*
  https://www.cnil.fr/fr/cnil-direct/question/une-donnee-caractere-personnel-cest-quoi

- https://www.legalplace.fr/guides/rgpd-sanction/

- https://www.sacd.fr/quelles-sanctions-pour-les-infractions-aux-droits-dauteur

- https://www.actoba.com/violation-du-secret-daffaire-la-responsabilite-du-salarie/

- Nicolas Binctin, Savoir-Faire, Janvier 2018 (actualisation février 2020), Répertoire de droit commercial, Dalloz.
  https://www-dalloz-fr.proxy.unice.fr/documentation/Document?id=ENCY%2FCOMR%2FRUB000160%2F2018-01%2FPARA%2F22&FromId=COMR_RUB000160_DOC_1

- Serge Rayne, Intérêts fondamentaux de la nation : atteintes aux – Autres atteintes à la défense nationale, Janvier 2009, Répertoire de droit pénal et de procédure pénale.
  https://www-dalloz-fr.proxy.unice.fr/documentation/Document?ctxt=0_YSR0MD1ub3Rpb24gZGUgc2VjcmV0IGQnRXRhdMKneCRzZj1zaW1wbGUtc2VhcmNo&ctxtl=0_cyRwYWdlTnVtPTHCp3MkdHJpZGF0ZT1GYWxzZcKncyRzb3J0PSNkZWZhdWx0X0Rlc2PCp3Mkc2xOYlBhZz0yMMKncyRpc2Fibz1UcnVlwqdzJHBhZ2luZz1UcnVlwqdzJG9uZ2xldD3Cp3MkZnJlZXNjb3BlPUZhbHNlwqdzJHdvSVM9RmFsc2XCp3Mkd29TUENIPUZhbHNlwqdzJGZsb3dNb2RlPUZhbHNlwqdzJGJxPcKncyRzZWFyY2hMYWJlbD3Cp3Mkc2VhcmNoQ2xhc3M9&id=ENCY%2FPEN%2FRUB000186%2F2009-01%2FPLAN043

- https://siafdroit.hypotheses.org/659

➔ Pour en savoir plus sur la réutilisation des données :

https://guides.etalab.gouv.fr/juridique/reutilisation/#qu-est-ce-qu-une-reutilisation

Dans le cadre de **l'initiative « 4 pour 1000 »** et de la construction du **projet DATA4C+**

RAPPORT D'INTEROPÉRABILITÉ DU SYSTÈME DATA4C+

Quelles conséquences de l'interopérabilité sur les responsabilités du Cirad, de l'IRD et de INRAe lors de la diffusion des données.

Rédigée par Cloé Sigal-Guille

Sous la direction de Pauline Corbière (Cirad) et Julien Demenois (Cirad).

Sommaire

Préambule

Le projet.

Le Projet DATA4C+ est un projet coordonné par le Cirad, en partenariat avec l'INRAe et l'IRD. Il répond à un objectif technico-juridique de partage de données sur le carbone des sols de certains territoires. Les résultats de DATA4C+ permettront **l'estimation des potentiels de séquestration de carbone des sols de ces territoires dans de futurs projets**. Ainsi, ce projet s'inscrit dans le cadre des initiatives de sciences ouvertes et du carbone du sol (« 4 pour 1000 »). Afin de garantir au mieux la protection et l'utilisation des données collectées, la mise en place de bases de données permettent aux scientifiques de conserver les données générées dans le cadre de leurs expertises et de pouvoir les analyser. Il est nécessaire d'analyser ces bases de données sous l'angle juridique afin de déterminer si les données doivent ou peuvent être diffusées librement au grand public. À cette fin **une troisième phase de l'analyse juridique du système** va consister à analyser l'interopérabilité des bases de données sous un pan juridique, aux fins de déterminer les responsabilités incombant à chacun des partenaires en cas de production du risque.

L'intérêt d'une analyse de risque.

La loi pour une République numérique a mis en place une **ouverture des données par défaut**. Certaines données vont donc être obligatoirement diffusées au grand public sur des plateformes dédiées. Les données de recherche scientifique ne font pas exception. De plus, lorsque les données ne sont pas obligatoirement diffusables, le législateur souhaite tout de même mettre l'accent sur les bénéfices pour la science d'une telle diffusion et encourage donc le partage. Comme il a pu être démontré, grâce à *l'analyse de risques (phase 2)*, que des difficultés gravitent autour d'une telle ouverture. Il convient désormais de s'attarder sur **le cas particulier du système DATA4C+** et d'identifier les risques encourus par chacun des partenaires. En effet, le système regroupe diverses bases de données, émanant de trois organismes indépendants et reliés par le biais **d'un contrat de consortium**. En cas de survenance du risque, est-ce une responsabilité collective ou individuelle qui pourra être retenue ?

1 MISE EN CONTEXTE

Aux fins d'analyser l'interopérabilité et les obligations incombant à chacun des partenaires du projet DATA4C+, un rapport d'interopérabilité du système est mis en place. Ce rapport a pour ambition d'identifier les responsabilités incombant à chacun des partenaires du projet. En effet, la diffusion ou l'absence de diffusion de certaines données peuvent amener des conflits, voire des sanctions. De plus, certaines données sont des connaissances et résultats propres à chacun des partenaires ; et d'autres sont des résultats communs. Cela oblige à **identifier la responsabilité qui incombe pour chacun des partenaires en cas de conflit avec des tiers.**

1.1 DEFINITIONS.

Avant de débuter ce rapport, il convient de définir les notions qui vont être évoquées dans ce documents. Aux fins de se conformer aux engagements pris par les partenaires du projet DATA4C+, nous reprenons les définitions présentent dans l'accord de consortium signé entre ces mêmes parties.

« Résultat » désigne toutes les informations et connaissances techniques et/ ou scientifiques issues de l'exécution du projet, notamment le savoir-faire, les secrets de fabrique, les secrets commerciaux, les données, les bases de données, les logiciels, les dossiers, les plans, les schémas, les dessins, les formules, et/ou tout autre type d'informations, sous quelque forme qu'elles soient, brevetables ou non et/ou brevetés ou non, et tous les droits de propriété intellectuelle en découlant, générés par une ou plusieurs parties, ou leurs sous-traitants.

« Résultat commun » désigne tous résultats développés au titre du projet conjointement par des personnels d'au moins deux parties et dont les caractéristiques sont telles qu'il n'est pas possible de séparer la contribution intellectuelle de chacune desdites parties pour la demande ou l'obtention d'un droit de propriété intellectuelle.

« Résultat propre » désigne tous les résultats obtenus par une partie seule, sans le concours d'une autre partie, c'est-à-dire sans la participation en termes d'activité inventive ou intellectuelle lors de l'exécution de sa part du projet.

1.2 METHODOLOGIE SUIVIE.

Le rapport d'interopérabilité du système DATA4C+ fait suite à la rédaction d'une analyse de risques. Pour mieux saisir les enjeux gravitant autour du système et, appréhender les risques que courent pour les partenaires en cas de diffusion ou non des données, il convient, préalablement à ce document, de prendre connaissance de cette analyse de risques.

Ainsi, les partenaires du projet pourront identifier les probabilités que leur responsabilité soit engagée en cas de diffusion ou non de données. En effet, ce rapport d'interopérabilité des bases de données du système DATA4C+ est destiné exclusivement aux partenaires du projet, le Cirad, l'IRD et l'INRAe.

Nous n'indiquerons pas la probabilité qu'un risque survienne *(cf Analyse de risques)*, nous nous contenterons d'indiquer à qui incombe la responsabilité des dommages en fonction du type de données. ***Qu'entend-on par « type de données » ?*** Il s'agit tout simplement des données faisant l'objet d'un résultat propre ou d'un résultat commun *(cf Définitions.)*.

1.3 DOCUMENTS UTILISES.

Aux fins d'identifier les responsabilités qui incombent à chacun des partenaires, les recherches ont gravité autour des règles de droit mais également de l'accord de consortium signé entre le Cirad, l'IRD et l'INRAe. Ce dernier détermine certains points de titularité des données qui nous permettent d'identifier les responsables en cas de litige, et leur capacité à se retourner contre leurs partenaires.
Les articles de l'accord du consortium ne sont pas tous utilisés, et vous pouvez retrouver ceux qui ont servi à ce rapport en *Annexes*

2 LES RISQUES RELATIFS A LA DIFFUSION DE DONNEES DONT LA COMMUNICATION EST OBLIGATOIRE.

Rappel *(cf Analyse de risques)*.

La loi pour une République numérique (ci-après « LRN ») ainsi que d'autres règlementations liées aux sols, obligent la diffusion de certaines données. Pour en savoir plus, je vous invite à lire la ***Note méthodologique rédigée dans le cadre du projet DATA4C+.*** Par ailleurs, la loi CADA du 17 juillet 1978[1] avait préalablement mis un système en place, permettant à la Commission d'accès aux documents administratifs (ci-après « CADA ») de rendre des avis en cas d'absence de diffusion de ces données-là. Concernant des sanctions liées à la LRN, le texte reste assez flou, et mis à part l'avis émanant d'autorité indépendant telle que la CADA, il semblerait que le risque de sanction en cas de diffusion de ces données soit minime voire nul.

Dans le cadre de cette partie, nous distinguerons la responsabilité des partenaires lorsque l'on se trouve dans une situation d'absence de diffusion des données obligatoirement communicables, ainsi que dans la situation d'un refus non justifié de l'organisme lors d'une demande de communication.

2.1 L'ABSENCE DE DIFFUSION DES DONNEES OBLIGATOIREMENT COMMUNICABLES.

Il conviendra de distinguer entre les données provenant de résultats propres, et celles provenant de résultats communs.

Concernant les résultats propres d'un organisme et les données le constituant. L'article 7 de l'accord de consortium stipule que seul l'organisme détient la propriété des résultats propres et, par analogie, des données générées. En d'autres termes, c'est donc à lui qu'incombe la diffusion obligatoire de ces dernières. En effet, ces données ont été collectées et intégrées dans leur base de données individuelle et non pas directement dans le système DATA4C+. En cas d'absence de diffusion de ces données, seul cet organisme en supportera les risques.

[1] Loi n°78-753 du 17 juillet 1978 dite loi CADA.

Concernant les résultats communs au projet et les données le constituant. L'article 7 de l'accord de consortium indique que les organisations sont copropriétaires des résultats, et donc des données, faisant l'objet d'un résultat commun. En d'autres termes, il leur incombe la diffusion des données obligatoirement communicables, et il semblerait que les parties supportent ensemble les risques liés à sa non diffusion.

2.2 LE REFUS NON JUSTIFIE D'UNE DEMANDE DE COMMUNICATION.

Il conviendra de distinguer entre les données provenant de résultats propres, et celles provenant de résultats communs.

Concernant les résultats propres d'un organisme et les données le constituant. En rappelant ce qui a été dit plus haut, il incombe à l'organisme de diffuser les données qui doivent être obligatoirement communicables, notamment lorsqu'il s'agit d'une demande de communication légitime. Autrement dit, les risques pèsent sur l'organisme destinataire de la demande de communication du document administratif et donc des données.

Concernant les résultats communs au projet et les données le constituant. En rappelant ce qui a été dit plus haut, il incombe aux partenaires du projet de diffuser les données qui doivent être obligatoirement communicables, notamment lorsqu'il s'agit d'une demande de communication légitime. Autrement dit, l'obligation de diffusion est à la charge de l'établissement destinataire de la demande de communication. Toutefois, ils peuvent en décider autrement dans le Plan de gestion des données (ci-après « PGD »).

Il est à noter que, nonobstant les dispositions de l'accord de consortium, les parties peuvent décider, pendant et après le projet, par contrat séparé, des règles spécifiques concernant la politique de traitement des données. Ainsi elles peuvent encadrer leur diffusion, leur réutilisation, ou toute autre action nécessaires à mettre en place en la matière.

***En conclusion*, il faudra déterminer si la donnée fait l'objet d'un résultat propre ou commun aux fins d'en dégager sa titularité et d'identifier le responsable en cas de manquement à la diffusion de données obligatoirement communicables.**

3 LES RISQUES RELATIFS A LA DIFFUSION DE DONNEES DONT LA COMMUNICATION N'EST PAS OBLIGATOIRE.

Rappel *(cf Analyse de risques)*.

La loi pour une République numérique ainsi que d'autres règlementations liées aux sols, rendent facultative la diffusion de certaines données. Pour en savoir plus, je vous invite à lire la ***Note méthodologique rédigée dans le cadre du projet DATA4C+.*** Le législateur encourage, dans le cadre du mouvement d'*open science*, la diffusion de ce type de données mais ne l'oblige pas. En d'autres termes, il n'a pas établi de « sanctions » potentielles en cas de diffusion ou non de ces données. Les risques sont faibles voire inexistants lorsqu'une analyse au cas par cas des données a été préalablement effectuée aux fins d'en déterminer leur nature.

Remarque. Le mouvement de science ouverte, dit « open science », est un mouvement dont l'objectif est de rendre universellement accessibles les résultats de la recherche scientifique (publications et données de recherche, notamment). Le législateur n'est pas le seul acteur majeur de ce mouvement. Les scientifiques, juristes ou d'autres corps de métiers ouvrent pour permettre l'ouverture des données en faveur de la recherche notamment.

Comme nous pouvons le constater, peu de risques encadrent ce type de données. En effet, lorsqu'une analyse au cas par des données a permis d'identifier leur nature, la diffusion de données dont la communication est facultative fait peser peu de risques pour les organismes. Au contraire, le législateur l'encourage. Ainsi, dans le cadre de ce document, nous traiterons le cas où **un organisme partenaire du projet diffuse des données de ce type alors que le reste des partenaires s'y sont opposés.**

Concernant les résultats propres d'un organisme et les données le constituant.

→ **S'agissant d'une diffusion effectuée par l'organisme propriétaire des résultats propres :**

L'article 7 du contrat de consortium indique que l'organisme peut utiliser ses connaissances propres et ses résultats propres comme il le souhaite. En effet, il en est le seul titulaire. Ainsi,

la diffusion de données facultatives est possible s'il s'agit de résultats propres de l'organisme qui diffuse. Les autres parties ne pourront donc pas se prononcer et s'opposer à la publication de résultats propres par l'organisme qui en a la propriété.

→ **S'agissant d'une diffusion effectuée par un organisme qui n'est pas le propriétaire des résultats propres :** L'article 9.2.1 du contrat de consortium traire cette problématique, et réduit considérablement la possibilité pour l'organisme propriétaire des résultats propres de s'opposer à leur communication. *L'ensemble de l'article est expliqué plus bas dans l'encadré qui suit.*

Concernant les résultats communs au projet et les données le constituant. L'article 7 indique que les données faisant l'objet de résultats communs sont la « propriété » de tous les organismes. En d'autres termes, il faudra impérativement **l'autorisation de chacun des partenaires préalablement à la diffusion** *(cf article 9.2.1 de l'accord de consortium expliqué dans l'encadré).* En cas de diffusion, l'accord de consortium pourrait alors être rediscuté par l'initiative d'un des partenaires, car il estimerait qu'une des parties n'a pas respecté sa part de l'accord.

Développement et explication de l'article 9.2.1, permettant à un organisme de publier des résultats, et donc des données, dont il ne serait pas exclusivement propriétaire.

Que doit faire l'organisme qui souhaite partager ? L'article 9.2.1 de l'accord de consortium indique que l'organisme qui souhaite diffuser des données faisant l'objet d'un résultat propre ou commun, ne lui appartenant donc pas exclusivement, doit demander l'accord des organismes concernés préalablement à la diffusion.

Jusqu'à quand doit-il répondre de cette obligation ? Cette obligation est toujours en vigueur jusqu'à l'expiration d'un délai de 2 ans suivant la fin ou l'expiration de l'accord de consortium. Au-delà de ce délai, les publication et communication devront tout de même s'effectuées dans le respect des obligations de confidentialité (*article 9.1 de l'accord de consortium*).

Quelles sont les possibilités de réponse de ou des organismes concernés par la demande ? Le pouvoir de décision de l'organisme propriétaire du résultat n'est pas totalement libre car il ne peut pas refuser la communication des données lorsqu'un délai de 18 mois après la première soumission du projet du demandeur est dépassé. *Quelles sont les réponses possibles ?*

- Accepter sans aucune réserve la présente demande.
- Demander à ce qu'une information confidentielle le concernant soit retirée préalablement à la publication.
- Demander à ce que des modifications de la publication soient effectuées lors que les informations partagées sont susceptibles de porter atteinte à l'exploitation industrielle et commerciale de son activité.
- Demander que la communication des résultats soit différée lorsqu'il justifie d'une cause réelle et sérieuse.

Quel est le délai de réponse ? L'organisme dispose d'un délai de 60 jours pour répondre à la demande, passé ce délai, son accord sera réputé acquis.

***En conclusion*, mêmes si les risques sont faibles voire nuls, il incombe aux partenaires de respecter leurs engagements.**

4 LES RISQUES RELATIFS A LA DIFFUSION DE DONNEES DONT LA COMMUNICATION EST INTERDITE.

Rappel *(cf Analyse de risques)*.

La loi pour une République numérique ainsi que d'autres règlementations liées aux sols, rendent interdite la diffusion de certaines données. Pour en savoir plus, je vous invite à lire la ***Note méthodologique rédigée dans le cadre du projet DATA4C+.*** Le législateur a souhaité protéger certains droits. Les données concernées sont notamment les données à caractère personnel, les données attachées à des droits de propriété intellectuelle ou encore à un secret protégé par la loi. La diffusion de ces données est en principe interdite et fait l'objet de sanction si elle se produit. Toutefois, quelques exceptions sont notables et c'est pour cette raison qu'il convient d'effectuer au préalable une analyse au cas par cas de toutes les données.

Dans cette partie, nous distinguerons la responsabilité des partenaires lorsque l'on se trouve dans une situation de diffusion de données à caractère personnel, de données protégées par un droit de propriété intellectuelle, ou encore correspondant à un secret protégé par la loi ou par un contrat.

4.1 DIFFUSION DE DONNEES A CARACTERE PERSONNEL.

Le Règlement général sur la protection des données (ci-après « RGPD ») définit les termes permettant d'identifier les responsables en cas de diffusion de données à caractère personnel[2].

Est **responsable du traitement** des données à caractère personnel *« la personne physique ou morale, l'autorité publique, le service ou un autre organisme qui, seul ou conjointement avec d'autres, détermine les finalités et les moyens du traitement ; lorsque les finalités et les moyens de ce traitement sont déterminés par le droit de l'Union ou le droit d'un*

[2] Article 4 du RGPD [règlement UE 2016/679 du Parlement européen et du Conseil du 27 avril 2016 relatif à la protection des personnes physiques à l'égard du traitement des données à caractère personne et à la libre circulation de ces données].

État membre, le responsable du traitement peut être désigné ou les critères spécifiques applicables à sa désignation peuvent être prévus par le droit de l'Union ou par le droit d'un État membre ». En d'autres termes, il s'agit de la personne, du service ou de l'organisme qui seul ou conjointement, détermine les finalités du traitement et les moyens utilisés.

Exemple : dans le cas particulier du système DATA4C+, le **Cirad** sera le responsable du traitement des données qu'il collecte lors des recherches qu'il mène. Ces données sont stockées dans ses propres bases de données avant d'être intégrées au système. L**'IRD** et **INRAe** sont également responsables de traitement pour la collecte de leurs propres données lors de leurs propres recherches.

Est **sous-traitant** *« la personne physique ou morale, l'autorité publique, le service ou autre organisme qui traite des données à caractère personnel pour le compte du responsable du traitement »*. En d'autres termes, il s'agit de la personne externe qui gère pour le compte du responsable du traitement et sous ses instructions.

Exemple : dans le cas particulier du système DATA4C+, le sous-traitant est **la personne indépendante aux trois organismes** à qui est confié une partie de l'élaboration et de la maintenance du site. Il s'agit de la partie qui n'est pas traitée par le salarié du Cirad en charge de la conception du site.

Est **destinataire** *« la personne physique ou morale, l'autorité publique, le service ou tout autre organisme qui reçoit communication de données à caractère personnel, qu'il agisse ou non d'un tiers. Toutefois, les autorités publiques qui sont susceptibles de recevoir communication de données à caractère personnel dans le cadre d'une mission d'enquête particulière conformément au droit de l'Union ou au droit d'un État membre ne sont pas considérées comme des destinataires ; le traitement de ces données par les autorités publiques en question est conforme aux règles applicables en matière de protection des données en fonction des finalités du traitement »*.

Exemple : dans le cas particulier du système DATA4C+, le **Cirad** pourra être considéré comme un destinataire lorsque **l'IRD lui partage des données pour lesquelles il est le responsable du traitement.**

→ ***Précautions à prendre*** **:** la personne concernée doit être informée de la possibilité de transférer ses données à des destinataires précis. Cette information doit se faire préalablement à la collecte des données, et doit faire l'objet d'un accord exprès de la part de la personne concernée. Cet accord doit être écrit et peut s'effectuer en cochant la case liée à la demande d'autorisation du traitement des données.

Est **tiers** de la collecte *« une personne physique ou morale, une autorité publique, un service ou un organisme autre que la personne concernée, le responsable du traitement, le sous-traitant et les personnes qui, placées sous l'autorité directe du responsable du traitement ou du sous-traitant, sont autorisées à traiter les données à caractère personnel »*. Il s'agit des **tiers autorisés** comme par exemple l'administration fiscale ou les régulateurs tels que la CNIL. Ils disposent d'un droit de communication à la personne concernée.

Exemple : dans le cas particulier du système DATA4C+, un tiers **autorisé peut exiger au Cirad, à l'IRD ou à INRAe, la communication de données à caractère personnel.**

→ ***Précautions à prendre*** **:** Il peut s'avérer difficile d'identifier si un tiers est réellement autorisé ou non à recevoir certaines données. Pour cela, la CNIL a publié un *« Guide pratique des tiers autorisés »*[3]. Elle préconise trois étapes pour permettre au responsable de traitement de vérifier la qualité de tiers autorisé :

1| ***Vérification*** de l'existence d'un fondement légale autorisant la demande de communication.

2| ***Vérification*** de la qualité de l'organisme à l'origine de la demande et du périmètre des informations ciblées.

3| ***Sécurisation*** de la communication des données ou des modalités d'accès par le tiers autorisé.

[3] Lien vers le Guide : https://www.cnil.fr/sites/default/files/atoms/files/guide_tiers_autorises.pdf

Est **une personne concernée** la personne dont on traite les données.

Exemple : dans le cas particulier du système DATA4C+, la personne concernée pourrait être **un agriculteur** qui a confié au Cirad, à l'IRD, ou à INRAe, certaines informations permettant son identification (nom, prénom, coordonnée GPS, etc…)

Concernant les résultats propres d'un organisme et les données le constituant. L'article 10 de l'accord de consortium stipule que *« chacune des parties reste responsable, dans les conditions de droit commun, des dommages que son personnel pourrait causer aux tiers à l'occasion de l'exécution de l'accord »*. En d'autres termes, en cas de divulgation de données à caractère personnel le ou les personnes concernées vont pouvoir se retourner contre l'organisme qui les a diffusées. Autrement dit, contre l'organisme qui collecte des données, assimilables à des résultats propres, en sa qualité de responsable de traitement. Ainsi, en tant que responsable de la protection des données à caractère personnel, il est tenu de se conformer à la règlementation en vigueur et, en cas de diffusion, sa responsabilité pourra être engagée.

Concernant les résultats communs au projet et les données le constituant. Les organismes définissent, dans le cadre du projet, la finalité des recherches et donc de la collecte des données personnel. En principe, l'accord de consortium ne donne pas lieu à la création d'une personne morale. Ainsi, les parties à l'accord restent responsables des données qu'ils ont collectées. Il semblerait qu'une responsabilité collective ne puisse pas être engagée.

***En conclusion*, lors de la diffusion de données à caractère personnel, les risques reposent sur le responsable du traitement des données.**

Suggestions : bonnes pratiques à mettre en place lors de la rédaction du traitement de données à caractère personnel.

1 | ***Mettre en place un système d'autorisation préalable*** à la collecte des données. Cette autorisation est à faire signer et remplir par la personne concernée par la collecte des données.

→ Important : Cette autorisation doit rappeler les finalités de la collecte. Les données ne pourront être utilisées que dans le cadre de cette finalité définie. Il convient également de préciser les destinataires, s'il y en a, des données. En l'absence de cette précision, l'IRD, par exemple, ne pourra pas partager ses données avec le Cirad et INRAe et inversement.

2 | *Effectuer une analyse au cas par cas* des données lors de la diffusion ou du partage avec les partenaires. Cela permettra d'identifier les données à caractère personnelle qui ne font pas l'objet d'autorisation. Ces données ne pourront être diffusées telles quelles.

***Rappel* des sanctions en cas de diffusion de données à caractères personnelles.**

SANCTION ADMINISTRATIVE[4] **(2 types de sanctions en fonction de la gravité)**		**SANCTION PÉNALE[5]**
2% du chiffre d'affaire mondial pour les entreprises ou 10 millions d'euros d'amende ; applicables en cas de manquement aux : - Obligations incombant au responsable du traitement et au sous-traitant. - Obligations incombant à l'organisme de certification. - Obligations incombant à l'organisme chargé du suivi des codes de conduite.	**4% du chiffre d'affaire mondial pour les entreprises ou 20 millions d'euros d'amende** ; en cas de manquements à : - L'obligation de consentement préalable de la personne concernée. - Aux autres droits des personnes concernées. - Aux transferts de données à un destinataire situé dans un pays tiers ou à une organisation internationale. - A toutes les obligations découlant du droit des États membres. - Au non-respect d'une injonction ordonnée par l'autorité de contrôle.	**Objectif :** réprimander les violations qui ne font pas l'objet d'amendes administrative[6] *Exemple : sanction en cas de détournement de la finalité des données personnelles lors du traitement de la donnée* **Sanction :** Jusqu'à 5 ans d'emprisonnement et 300 000 euros d'amende[7].

[4] Article 83 du RGPD
[5] Article 84 du RGPD
[6] Articles 226-16 à 226-4 du Code pénal
[7] Article 226-16 du Code pénal

4.2 DIFFUSION DE DONNEES PROTEGEES PAR UN DROIT DE PROPRIETE INTELLECTUELLE.

<u>Tableau</u> : Les divers droits de la propriété intellectuelle et leur lien avec le projet DATA4C+

DROITS DE PROPRIÉTÉ INTELLECTUELLE	**TITULAIRE DU DROIT**	**PROJET DATA4C+**
Brevet	La personne physique ou morale qui effectue la demande de brevet.	Il semblerait qu'une atteinte portée à un brevet soit faible car le partage de seulement quelques données ne semble pas de nature à atteindre les droits de son titulaire.
Marques	La personne physique ou morale qui dépose la marque.	Il semblerait que le droit des marques n'est pas besoin d'être traité car la circulation de données est limitée.
Dessins et modèles	La personne physique ou morale qui dépose le dessin ou modèle.	Il semblerait que le droit des dessins et modèles n'est pas besoin d'être traité car la circulation de données est limitée.
Logiciel	En principe : la personne physique ou morale qui a développé le logiciel [sous réserve d'originalité] → **dans le cas du système DATA4C+ : ce sera l'employeur du salarié qui a conçu le logiciel car ce salarié l'a conçu dans la cadre de ses fonctions ou sur instruction de son employeur.**	Les logiciels conçus par les salariés du Cirad ne peuvent pas être utilisés ou modifié par l'IRD ou INRAe. → *<u>Sauf</u> autorisation expresse ou cession de droits.*

Droit du producteur de la base données	En principe : la personne physique ou morale qui démontre un investissement financier ou humain pour permettre la conception de la base de données [sous réserve d'originalité]. → **dans le cas du système DATA4C+ : ce sera l'organisme (Cirad, IRD ou INRAe), car il est le majoritaire investisseur dans la conception de la base de données.**	Une extraction ou réutilisation de la totalité ou d'une partie **substantielle** (de façon qualitative ou quantitative) du **contenu** de la base ne peut être effectué par l'organisme qui ne possède pas les droits de propriété intellectuelle. → *Sauf autorisation expresse ou cession de droits.*
Droit d'auteur *(exemples : publication, photos, article scientifique, etc… répondant aux conditions d'originalité).*	En principe : la personne physique (ou morale dans des cas précis) qui a conçue l'œuvre [sous réserve d'originalité]. → **dans le cas du système DATA4C+ : les chercheurs sont titulaires des droits sur l'ensemble des œuvres (et données) qu'ils ont effectués, même en tant que salarié d'un des organismes *sauf* en cas de cession des droits.**	L'organisme ne peut utiliser, modifier ou partager des documents, données, ou toute œuvre protégée par un droit d'auteur.

Concernant les résultats propres d'un organisme et les données le constituant.

→ **S'agissant des droits attachés à l'organisme même :** En vertu du contrat de consortium, seul l'organisme titulaire des données susceptibles de faire l'objet d'un droit de propriété intellectuelle peut les utiliser, les modifier et les publier. Par exemple, une photographie appartenant au Cirad ne pourra pas être divulguée ou modifiée par l'IRD ou INRAe (sous réserve toutefois que cette photographe remplisse les conditions d'originalité nécessaires pour faire valoir un droit d'auteur).

Néanmoins, il est possible pour le titulaire des droits (un salarié ou un organisme) de concéder une licence à la partie qui souhaite utiliser l'œuvre protégée. D'ailleurs l'article 8 de l'accord de consortium semble prévoir une « obligation » en la matière, incitant les organismes a partagé leurs données. Cette « obligation » court jusqu'à un délai de 24 mois après la fin du

projet. Une fois ce délai dépassé, la liberté de contracter entre les organismes leur permettra de mettre en place d'autres licenses voire, d'effectuer des cessions de droit.

→ **S'agissant des droits attachés à des tiers :** en cas d'atteinte à des droits de propriété intellectuelle lors de la diffusion de données, la personne responsable est l'auteur de la publication litigieuse. En d'autres termes, si cette publication est effectuée à l'initiative d'un seul organisme, il semblerait que ce soit lui qui engage sa responsabilité.
Par exemple, il peut s'agir d'une extraction substantielle frauduleuse d'une base de données, en quantité ou en qualité, par un organisme, qui la réintègrerai dans le système DATA4C+.

Concernant les résultats communs au projet et les données le constituant.

→ **S'agissant des droits attachés aux partenaires :** L'accord de consortium prévoit que les résultats communs qui font l'objet de droit de propriété intellectuelle conduisent une co-titularité des droits. En d'autres termes, il semblerait que l'accord de tous les organismes soit exigé lorsqu'ils souhaitent concéder une licence à tiers par exemple. Pour cela il convient également de se référer aux régimes légaux de corporation ou aux accords passées entre les parties.

→ **S'agissant des droits attachés à des tiers :** en cas d'atteinte à des droits de propriété intellectuelle lors de la diffusion de données, la personne responsable est le titulaire de la publication litigieuse. En d'autres termes, si cette publication est effectuée dans le cadre du projet DATA4C+ à l'initiative du consortium, c'est l'ensemble des organisations qui seront responsables devant la loi. De plus, il est possible de prévoir que chaque partie est libre de diffuser ou d'exploiter les données relatives à des résultats communs. Dans ce cas-là, ce ne sera pas le consortium, mais bien l'organisme qui diffuse qui sera responsable.

***En conclusion*, lorsqu'une diffusion porte atteinte à des droits de propriété intellectuelle d'un tiers, c'est le titulaire de la publication litigieuse qui voit sa responsabilité engagée. Cette responsabilité peut être partagée lorsque la publication est effectuée à l'initiative de toutes les parties du consortium.**

<u>Concernant les responsabilités entre l'hébergeur et l'éditeur du site internet DATA4C+.</u> Il est essentiel de différencier le statut de l'hébergeur et celui de l'éditeur car la loi prévoit une différence de responsabilité entre ces prestataires.

→ S'agissant de l'hébergeur [DATA4C+ = OVH pour le système DATA4C+]

Définition. Il s'agit de la personne ou la société qui *assure « même à titre gratuit, pour mise à disposition du public par des services de communications au public en ligne, le stockage de signaux, d'écrits, d'images, de sons ou de messages de toute nature fournis par des destinataires de ces services »*[8]. Autrement dit, il s'agit de la société permettant la mise en ligne du site internet DATA4C+ où les données seront partagées au public.

Responsabilité. La loi indique que les hébergeurs ne sont pas soumis *« à une obligation générale de surveillance des informations qu'ils stockent, ni à une obligation générale de recherche des faits ou des circonstances révélant des activités illicites »*[9]. En effet, l'hébergeur est un simple intermédiaire technique qui n'a pas pour mission d'analyser les contenus des publications qu'il héberge. Sa responsabilité sera engagée dans deux cas seulement : si l'information a été dénoncée par un tiers comme étant manifestement illicite et, si le retrait de l'information a été préalablement ordonné par un juge[10].

→ S'agissant de l'éditeur [DATA4C+ = le consortium Cirad-INRAE-IRD]

Définition. Il s'agit de la personne ou de la société qui « édite un service de communication en ligne »[11] à titre professionnel ou non, c'est-à-dire qui détermine les contenues mis à la disposition du public sur le service qu'elle a créé. Autrement dit, dans le cadre du projet DATA4C+, il s'agit des organismes qui vont créer le site et déterminer sa finalité, en choisissant son contenu et en y publiant les données, par exemple.

Responsabilité. L'éditeur est susceptible d'engager sa responsabilité à raison des contenus hébergés. Par exemple, il peut être poursuivi pour contrefaçon si les contenus présentent une atteinte à un droit antérieur de propriété intellectuelle. Le juge considère qu'un prestataire qui publie sur un site internet est considéré comme bénéficiant du statut de l'éditeur et non celui de l'hébergeur. Sa responsabilité peut être engagée au même titre qu'un éditeur[12].

[8] Loi n°2004-575 du 21 juin 2004 pour la confiance dans l'économie numérique.
[9] Article 6-I-7° de la loi pour la confiance dans l'économie numérique.
[10] Décision n°2004-496 du Conseil constitutionnel du 10 juin 2004.
[11] Article 6 de la loi pour la confiance dans l'économie numérique.
[12] TGI (Tribunal de grande instance) de Paris, 12 mars 2010.

4.3 DIFFUSION DE DONNEES SECRETES PROTEGEES PAR UN TEXTE JURIDIQUE.

Rappel *(cf Note méthodologique et Analyse de risque)*

Certaines données font l'objet de secrets protégées par des textes légaux (lois, règlements, directives). Ces secrets peuvent **être absolus et opposables à toutes personnes** ou, **relatifs et opposables seulement aux tiers et non aux personnes concernées**. L'analyse de risques permet d'envisager les conséquences d'une diffusion de ces données protégées par un secret, et le niveau d'acceptabilité d'une communication.

Lorsque des données sont diffusées, cette communication peut porter atteinte à des secrets protégés. Ces secrets peuvent prendre plusieurs formes. Cette analyse de risques traitera le cas du ***« secret d'affaires »***, du ***« secret d'État »*** et du ***« secret professionnel »***.

***Tableau* : Les divers secrets protégés par la loi et leur lien avec le projet DATA4C+**

SECRET PROTÉGÉ	LIEN AVEC DATA4C+	RÉGIME
Secret d'affaires	Il peut s'agir, par exemple, de la protection d'un savoir-faire, de techniques, de connaissances non immédiatement accessibles au public *(définition à l'article L.151-1 du Code de commerce).* → Des techniques agricoles particulières.	**Régime légal : loi sur les secret d'affaires du 30 juillet 2018.** → Utilisation ou divulgation du secret d'affaires sans accord express de son détenteur = illicite. → Le responsable est la personne qui a diffusé l'information protégée.
Secret d'État	Un secret d'État est déterminé en fonction de son contenu, son accès ou sa fonction. Il s'agit généralement d'un secret de défense, mais d'autres	**Régime légal : loi du 29 juillet 2009.** → Utilisation ou divulgation d'un secret d'État sans accord express d'une autorité compétente = illicite.

	enjeux peuvent être protégeables. → Données relatives à la présence de denrées intéressants la défense nationale.	→ Le responsable semble être la personne qui a diffusé l'information protégée.
Secret professionnel	Le secret professionnel lie certaines professions précises (avocat, médecins, etc…). Il n'en est donc pas question dans le projet.	Code pénal : 1 an d'emprisonnement et 150 000 euros d'amende.

<u>Concernant les résultats propres et communs du projet et les données les constituant</u>. L'article 10 de l'accord de consortium indique que l'organisme est seul responsable devant les tiers pour les dommages qu'il a causés. En d'autres termes, cela signifie que l'organisme semble responsable des données faisant l'objet d'un résultat qu'il diffuse. Ainsi, le risques que cette communication porte atteinte à un secret protégé par la loi semble lui incomber à lui seul.

***En conclusion*, lorsqu'un secret protégé par la loi fait l'objet d'une divulgation, l'organisme responsable semble être celui qui a procédé à la diffusion.**

4.4 DIFFUSION DE DONNEES CONFIDENTIELLES PROTEGEES PAR UN CONTRAT OU ACCORD.

Rappel *(cf Note méthodologique)*

Les organismes peuvent être amenés à conclure des **accords de confidentialité** avec des partenaires ; ou bien, des **contrats de prestation de service** dans le cadre recherche pour des acteurs privés par exemple. Dans les deux cas, les contrats et accords peuvent prévoir que les données générées ne pourront pas faire l'objet de diffusion.

Dans le cadre d'un accord de confidentialité, cela s'explique par le contenu de la donnée. Sa divulgation pourrait porter préjudice à l'activité du partenaire car elle évoquerait un élément de sa situation par exemple.

Dans le cadre d'une prestation de service, le contrat peut prévoir que les données générées appartiennent au donneur d'ordre. Ainsi, la diffusion des données par l'organisme prestataire de service n'est pas possible sans l'accord du prestataire de service.

Concernant les résultats propres d'un organisme et les données les constituant. Lors de la conclusion d'un contrat ou d'un accord, les parties prévoient des obligations à respecter durant toute la durée de la relation, voire au-delà dans certains cas. Lorsqu'une obligation de confidentialité est exigée, la diffusion de ces informations peut engager la responsabilité de la partie violant cette obligation. Dans le cadre **d'un accord de confidentialité**, l'organisme partie à cet accord est responsable de la diffusion des données, même si cette diffusion est effectuée par l'un de ses partenaires. Par exemple, si le Cirad avait conclu avec une entreprise un accord de confidentialité sur des données qu'il a collectées ; s'il advient que l'IRD diffuse ses données protégées, alors il semblerait que ce ne soit pas l'IRD qui soit responsable de l'atteinte portée à l'accord, mais bien le Cirad.

S'agissant désormais **d'un contrat de prestation de service**, si le contrat indique que les données appartiennent au donneur d'ordre, alors l'organisme prestataire de service ne pourra les utiliser et les partager sans son accord. Il est possible que le contrat prévoit que l'organisme puisse utiliser ces données à des fins de recherche, sans pour autant autoriser leur diffusion. Par exemple, lorsque le Cirad effectue des prestations de service et que les données sont caractérisées comme étant la propriété de l'entreprise avec laquelle le contrat est signé ; si le contrat le prévoit, le Cirad pourra les utiliser dans le cadre de ses propres recherches, et non

les partager avec l'IRD ou l'INRAe, même si des projets sont communs. En cas de diffusion de ces données, même si elle a été effectuée par un organisme autre que le Cirad, il semblerait que ce soit ce dernier qui reste responsable en cas d'atteintes portées à son cocontractant.

Concernant les résultats communs du projet et les données les constituant. Nous pouvons envisager le cas où, dans le cadre du projet, l'IRD, le Cirad et INRAe signent tous les trois un contrat de prestation de service, ou un accord de confidentialité, avec un tiers. Dans ce cas-là, ils s'engagent individuellement à respecter les obligations contenues dans le contrat. *Pourquoi individuellement ?* Tout simplement parce que l'accord de consortium qui les lie dans le cadre du projet DATAC4+ ne permet pas la création d'une personnalité morale. Autrement dit, chaque organisme reste responsable individuellement en cas d'atteinte portée au tiers. Ainsi, il semblerait qu'en cas de diffusion par le consortium de données confidentielles ou dont la propriété est attribuée au donneur d'ordre, chaque organisme voit sa responsabilité engagée. Néanmoins, lorsque la diffusion de données confidentielles est effectuée par un seul organisme, alors il est à prévoir sa seule responsabilité soit engagée. En effet, l'article 10 du contrat de consortium prévoit que chaque organisme est responsable des dommages qu'il cause au tiers. Il semblerait alors que la responsabilité soit différente selon que la diffusion soit effectuée par l'ensemble des organismes ou bien par un seul.

Prenons l'exemple d'un contrat signé, dans le cadre du projet, par le Cirad, l'IRD et INRAe avec une entreprise. Les trois organismes s'engagent à ne pas diffuser les données. Dans le cas où les données seraient tout de même diffusées, le nombre d'organismes inquiétés sera différent selon la nature de la diffusion. Si la diffusion a lieu sur le site internet du projet DATA4C+, alors l'IRD, le Cirad et INRAe sont susceptibles de voir leur responsabilité engagée. A contrario, si la diffusion est effectuée lors d'une publication de INRAe, alors seul INRAe semble pouvoir être inquiété.

En conclusion, en cas de résultat propre, seul l'organisme lié par le contrat pourra voir sa responsabilité engagée, quel que soit l'initiateur de la diffusion. *A contrario*, dans le cadre de résultats communs, il faudra identifier l'initiateur de cette diffusion, et sa nature, afin de déterminer si l'ensemble des organismes pourront être inquiétés ou non.

4.5 DIFFUSION DE DONNEES TRANSMISES LEGALEMENT PAR UN ORGANISME QUI NE MENTIONNE PAS QUE LEUR COMMUNICATION EST LIMITEE VOIRE INTERDITE.

Lorsque les bases de données sont constituées par les organismes, elles regroupent toutes sortes de données : communicables ou non. Au moment de la mise en commun des données dans le système DATA4C+, ou lors d'un simple partage de la base de données, l'organisme devrait prévenir son ou ses partenaires de la présence de données qu'il ne faudrait pas diffuser. Par exemple, il peut s'agir de données à caractère personnel dont la communication entre les trois organismes a été autorisée (autorisation préalable déterminant la finalité de la collecte des données), mais pas sa diffusion au public.

Qui est responsable si, en cas d'absence de mention de la nature de ces données protégées, un organisme les diffuse ? En cas de litige dû à un résultat commun, l'ensemble des organismes voient leur responsabilité engagée. Il convient d'indiquer que l'article 10, même s'il stipule que les organismes sont individuellement responsables envers les tiers, précise également que **les parties de l'accord ne peuvent se retourner entre elles**. En effet, l'article met en place une garantie du fait des connaissances propres, résultats et autres informations (10.2). Plus précisément, il indique que *« les parties reconnaissent que les connaissances propres, les résultats et les autres informations communiquées par l'une des parties à une autre partie dans le cadre de l'exécution de l'accord sont communiquées en l'état, sans aucune garantie de quelque nature qu'elle soit. ces connaissances propres, ces résultats et ces autres informations sont utilisés par les parties dans le cadre de l'accord à leurs seuls frais, risques et périls respectifs, et en conséquence, aucune des parties n'aura de recours contre une autre partie, ni ses sous-traitants éventuels, ni son personnel, à quelque titre que ce soit et pour quelque motif que ce soit, en raison de l'usage de ces connaissances propres, ces résultats et ces autres informations, y compris en cas de recours de tiers invoquant l'atteinte à ses droits de propriété intellectuelle. »*

***En conclusion*, en considération de ce qui a été dit tout au long de ce rapport, l'organisme qui diffuse une donnée protégée semble être responsable des dommages qu'il cause au tiers. Il ne pourra pas se retourner contre son partenaire même s'il a failli dans sa mission d'information sur la nature des données qu'il partage au projet.**

5 LES RISQUES LIES A LA DIFFUSION INVOLONTAIRE DE DONNEES.

Rappel *(cf Analyse de risques)***.**

Une diffusion involontaire est caractérisée comme indépendante de la volonté du Cirad, de l'IRD et de INRAe. Autrement dit, des données vont être divulguées alors que les organismes ne le souhaitent pas ou qu'ils n'avaient pas préparés cette communication au public (par exemple, ils n'avaient pas encore anonymiser les données à caractère personnel). Ainsi, nous pouvons constater que des risques peuvent émaner de cette diffusion involontaire. Cette dernière peut être due à l'exploitation d'une vulnérabilité par un individu ou un organisation extérieurs au projet. Cette diffusion peut conduire à la communication de données à enjeux majeurs, du fait de leur interdiction légale à la communication, ou du fait de leur sensibilité.

La diffusion est involontaire mais peut tout de même engager la responsabilité de l'organisme qui a collecté ces données, car il a failli dans son **devoir de protection des données**.

Par exemple, le piratage au sein d'une seule organisation (par le biais du compte d'un chercheur) peut permettre à l'individu malveillant d'accéder à l'ensemble des données présentes dans le système DATA4C+, et donc d'accéder à toutes les données collectées, sensibles ou non.

L'article 10 de l'accord de consortium indique que *« les parties renoncent mutuellement à se demander réparation des préjudices indirects qui pourraient survenir dans le cadre de l'accord »*. En d'autres termes, la divulgation de certaines données semble pouvoir permettre aux tiers d'engager des poursuites contre plusieurs entités.

Premièrement, les tiers devraient pouvoir se retourner contre chaque membre du consortium. En effet, l'accord de consortium ne conduit pas à la création d'une personnalité morale regroupant l'ensemble des organismes. En d'autres termes, chaque organisme restera responsable individuellement, même si une donnée a été réalisée avec la collaboration du collectif. Il n'en demeure pas moins, que les responsabilités de chacun des organismes peuvent être engagées simultanément.

Deuxièmement, les tiers pourraient également se retourner contre l'organisme qui héberge le site regroupant les données. En effet, la diffusion involontaire des données peut résulter d'une faille du système. L'hébergeur a pour vocation de fournir une prestation durable d'hébergement et donc de stockage des informations fournies par son client. Cette définition est initiée par la loi[13] et confirmée par la jurisprudence[14] et la doctrine[15]. L'hébergeur est donc assujetti à une obligation générale de prudence et de diligence de manière à ne pas permettre la création de vulnérabilité dans le site.

Dans ce cas-là, on pourrait valablement en interpréter que l'éditeur est également susceptible d'engager sa responsabilité. En effet, en tant que codeur et créateur du système, il semblerait qu'il puisse, sans le savoir, laisser place à des vulnérabilités qui pourraient fragiliser la sécurité du système.

En tout état de cause, en cas de diffusion de données, il semblerait que le tiers contacte en priorité les organismes chargés de veiller à la non divulgation de leurs données et, engage leur responsabilité.

Est-ce que ces organismes pourraient ensuite se retourner contre l'éditeur et l'hébergeur ?

Dans le cas particulier du système DATA4C+, il semblerait que les données déposées sur le site internet ne soient que des données autorisées à la diffusion. En effet, un travail en amont devrait pouvoir identifier toutes les données interdites à la diffusion dans chacune des bases de données des organismes ; et les retirer du partage commun. En somme, la responsabilité de l'hébergeur semble écartée en pratique.

Concernant l'éditeur, le Cirad fait appel à un sous-traitant pour coder le système. En cas de vulnérabilité émanant de ces travaux, il conviendra de se référer au contrat de sous-traitance conclu avec ce dernier.

***En conclusion*, en cas de diffusion involontaire, il semblerait que l'ensemble des responsabilités des organismes puissent être engagées.**

[13] Loi pour la confiance dans l'économie numérique, LCEN, loi n°2004-575 en date du 21 juin 2004.
[14] Cour d'Appel d'Aix-en-Provence arrêt du 22 janvier 2008.
[15] Fauchoux, Deprez et Bruguière, « Droit de l'internet', Lexis-Nexis.

6 LES RISQUES RELATIFS A L'ABSENCE DE MISE EN CONFORMITE DES DONNEES ANCIENNES A UN FORMAT INTEROPERABLE.

Rappel *(cf Note méthodologique)*

Un des apports de la loi pour une République numérique est d'exiger, pour les données obligatoirement diffusables, leur garantie d'une réutilisation libre et gratuite. Pour cela, un décret[16] fixe les licences que les établissements peuvent utiliser lors de la communication des documents répondant à cette exigence. De plus, les organismes qui partagent des données, de façon facultative, sont également encourager à ouvrir l'accès à ces informations. Le législateur va imposer un format facilitant cette réutilisation. En d'autres termes, cela va rendre les bases de données interopérables afin qu'elles puissent être disponibles et accessibles à tout le monde. On parle de « standard ouvert »

La loi du 21 juin 2004 pour la confiance dans l'économie numérique[17] définie **la notion de « standard ouvert »** en indiquant qu'il s'agit de *« tout protocole de communication, d'interconnexion ou d'échange et tout format de données interopérable et dont les spécifications techniques sont publiques et sans restriction d'accès ni de mise en œuvre »*. Plus précisément, il s'agit des formats lisibles par des logiciels libres, par opposition à ceux qui ne le sont que par des logiciels propriétaires. Cette exigence pourrait constituer un point d'appui légal permettant à la CADA et à la CNIL de soumettre un avis à l'organisme, et à lui demander de se conformer aux exigences de dorme imposée par la loi.

Le paragraphe 1 de l'article 5 de la directive du 20 juin 2019, qui devait être transposée au plus tard le 17 juillet 2021, crée l'obligation d'accompagner la diffusion des documents de leurs métadonnées, répondant à des normes formelles ouvertes. La loi pour une République numérique, en s'appuyant sur les principes FAIR (Findable, Accessible, Interoperable, Reusable) exige donc un partage des données permettant leur accessibilité libre, mais également leur réutilisation.

[16] Décret n°2017-38 du 27 avril 2017 relatif aux licences garantissant la libre réutilisation des données.
[17] Article 4 de la loi n°2004-575 du 21 juin 2004 pour la confiance dans l'économie numérique.

Il convient de préciser que ces exigences sont obligatoires pour les données collectées à partir du 7 octobre 2016, date d'entrée en vigueur de la loi. Pour les données antérieures (et donc ce qui nous intéresse dans cette partie), la loi ne semble pas obligé la mise en place d'un format standard ouvert. En effet, le législateur semble seulement fortement l'encourager. Ainsi, la loi reste assez vague s'agissant des données collectées antérieurement au 7 octobre 2016.

Concernant plus précisément le projet DATA4C+. Un grand nombre de données ont été collectées antérieurement à cette date. La volonté majoritaire des partenaires au projet est de participer au mouvement d'ouverture de la science, dans l'objectif de faire profiter aux scientifiques de leurs études et résultats. Ainsi, lors de la diffusion de ces données, il est préconisé de les communiquer dans un format qui sera ouvert, et qui pourrait permettre leur réutilisation. La loi ne semblant pas s'appliquer pour ce type de données anciennes, il est possible pour les partenaires de préciser le contour de leur réutilisation en précisant « à des fins non commerciale » par exemple.

***En conclusion*, la loi pour une République numérique reste assez floue concernant les données anciennes. Il semblerait que le législateur n'impose pas de format et ne semble donc pas mettre en place de sanction. Néanmoins, il conviendrait de diffuser ces données dans un même format que le reste des diffusions, afin de permettre une homogénéité d'accessibilité et de réutilisation.**

7 CONCLUSION

Premièrement, le projet DATA4C+ et l'accord de consortium entre les parties ne créent pas de personnalité morale regroupant le Cirad, l'IRD, et INRAe. Autrement dit, même si des responsabilités collectives peuvent être engagées, chaque organisme est responsable des atteintes portées aux tiers, et chaque organisme peut se voir attribué une sanction individuellement.

→ ***Précaution.*** Lorsque des diffusions communes sont effectuées sur le site hébergeant le système DATA4C+, il est important que les trois organismes soient mentionnés. En effet, le tiers pourra ainsi demander à ce que la responsabilité des trois organismes soit engagée, et non pas seulement de l'organisme mentionné sur le site. Il convient de mentionner que si tel était le cas, cet organisme ne semble pas pouvoir se retourner ensuite contre ses partenaires.

Deuxièmement, une analyse au cas par cas des données doit être effectuées. Cette analyse est essentielle afin de déterminer qu'elles sont les données à diffuser et celles qui ne le sont pas. Pour cela, il convient également de vérifier les autorisations effectuées lors de la collecte des données, et identifier le paramètre de diffusion qui est autorisé à l'organisme.

→ ***Précaution.*** Aux fins de se protéger au mieux de tout contentieux, il convient de prévoir des autorisations préalables d'utilisation et de diffusion des données lors de la collecte. Cela est d'autant plus nécessaire en ce qui concerne les données à caractère personnel (se conformer au RGPD).

Troisièmement, en considération de l'ensemble des recherches et études qui ont été effectuées sur le cas particulier du système DATA4C+ et des bases de données qu'il regroupe, il semblerait que la diffusion des données présente un risque faible. Il faudra cependant au moins anonymiser les données à caractère personnelle, car elles n'ont pas donné lieu à des autorisations préalablement de diffusion pour la plupart. Cette diffusion devra se faire avec des précautions et devra impliquer les trois organismes, aux fins qu'ils puissent être solidaires entre eux en cas de litige.

Dans le cas où des questions subsisteraient, notamment pour des données précises, il conviendra de prendre contact avec le service juridique compétant.

ANNEXES

ARTICLES DU CONTRAT DE CONSORTIUM.

ARTICLE 7 - PROPRIÉTÉ

7.1 CONNAISSANCES PROPRES

À l'exception des stipulations ci-après, l'ACCORD n'emporte aucune cession ou licence des droits de la PARTIE détentrice sur ses CONNAISSANCES PROPRES.

Sous réserve des stipulations de l'article 8 ci-après, rien dans le présent ACCORD n'interdit à la PARTIE détentrice d'utiliser de quelque manière que ce soit ses CONNAISSANCES PROPRES pour elle-même ou avec tout tiers de son choix.

7.2 RÉSULTATS PROPRES.

Les RESULTATS PROPRES sont la propriété de la PARTIE qui les a générés.

Les éventuels BREVETS NOUVEAUX et les autres titres de propriété intellectuelle sur lesdits RESULTATS seront déposés à ses seuls frais, à son seul nom et à sa seule initiative.

7.3 RÉSULTATS COMMUNS.

Les PARTIES ayant généré des RESULTATS COMMUNS en sont par principe copropriétaires.

Les PARTIES ayant généré des RÉSULTATS COMMUNS en sont par principe copropriétaires.

Toutefois, les PARTIES à l'origine d'un RÉSULTAT COMMUN pourront se concerter afin d'en attribuer la propriété à l'une ou plusieurs d'entre elles.

Les PARTIES COPROPRIÉTAIRES signeront, par acte séparé et avant toute exploitation, un accord définissant la répartition des quotes-parts définies à hauteur de leur contribution ainsi que les droits et obligations s'y rapportant et reprenant pour ce qui concerne les RÉSULTATS COMMUNS brevetables et/ou les droits d'auteur les principes exposés ci-dessous.

Dans le cas où des RÉSULTATS COMMUNS seraient générés en partie par le personnel d'une structure commune de recherche (de type « UMR »), les tutelles de ladite structure seront considérées comme une seule PARTIE COPROPRIÉTAIRE. Il est entendu que lesdites tutelles feront leur affaire de la répartition entre elles de la quote-part de copropriété qui leur est attribuée, conformément à la convention régissant la structure.

7.3.1 *RÉSULTATS COMMUNS brevetables.*

7.3.2 *RÉSULTATS COMMUNS relevant du droit d'auteur hors logiciels*

Un règlement de copropriété entre les indivisaires définira les droits détenus par les PARTIES COPROPRIÉTAIRES concernées notamment au regard de la spécificité des RÉSULTATS COMMUNS obtenus et des conditions d'accès et d'utilisation qu'elles souhaitent se réserver.

ARTICLE 8 – UTILISATION / EXPLOITATION

8.1 CONNAISSANCES PROPRES

8.1.1 *Aux fins d'exécution du PROJET*

Pour la durée du PROJET, les PARTIES concèdent sans contrepartie financière un droit d'utilisation de leurs CONNAISSANCES PROPRES aux autres PARTIES sur demande écrite de celles-ci lorsqu'elles leur sont nécessaires pour exécuter leur PART DU PROJET.

8.1.2 *Aux fins d'exploitation des RÉSULTATS*

Pendant la durée du PROJET et 24 mois après son terme et sous réserve des droits des tiers et des éventuelles restrictions figurant à l'Annexe 2, chaque PARTIE s'engage à concéder aux autres PARTIES et/ou à leurs AFFILIÉS, par acte séparé et sur demande écrite, une licence sur ses CONNAISSANCES PROPRES lorsqu'elles sont nécessaires à l'exploitation, par la PARTIE ou l'AFFILIÉ qui en fait la demande, de ses RÉSULTATS ou des RÉSULTATS sur lesquels elle a obtenu des droits d'exploitation.

La PARTIE détentrice s'engage à concéder lesdites licences à des conditions commerciales normales pour le secteur d'application considéré.

Ces droits seront non exclusifs, non cessibles et sans droit de sous licence sauf accord préalable et écrit de la PARTIE détentrice.

8.2 RÉSULTATS

8.2.1 *Utilisation – Exploitation de ses RÉSULTATS par une PARTIE*

Chaque PARTIE est libre d'exploiter ses RÉSULTATS sous réserve des droits des autres PARTIES prévus à l'article 8.2.3 ci-après.

8.2.2 *Utilisation – Exploitation des RÉSULTATS COMMUNS par les PARTIES COPROPRIÉTAIRES*

Les PARTIES COPROPRIÉTAIRES et leurs AFFILIÉS disposent d'un droit non exclusif d'exploitation industrielle et/ou commerciale, directe et indirecte des RÉSULTATS COMMUNS.

En cas d'exploitation effective par une PARTIE et/ou ses AFFILIÉS, celle-ci donnera lieu à une compensation financière, forfaitaire ou proportionnelle, qui sera équitable eu égard aux contributions respectives des PARTIES COPROPRIÉTAIRES. Toutefois, aucune compensation ne sera due entre industriels en cas d'exploitation directe par l'un d'entre eux.

L'accord de toutes les PARTIES COPROPRIÉTAIRES est nécessaire en cas d'exploitation exclusive.

Pour les RÉSULTATS COMMUNS consistant en des logiciels, l'accord des autres PARTIES COPROPRIÉTAIRES est nécessaire en cas de diffusion des codes sources.

8.2.3 *Utilisation – Exploitation de RÉSULTATS par les PARTIES non détentrices autres que les PARTIES COPROPRIÉTAIRES*

Sauf accord entre les PARTIES concernées, les droits prévus au présent article 7.2.3 seront non exclusifs, non cessibles et sans droit de sous licence.

8.2.3.1 Aux fins d'exécution du PROJET

Pour la durée du PROJET, les PARTIES concèdent un droit d'utilisation de leurs RÉSULTATS aux autres PARTIES sur demande écrite de celles-ci lorsqu'ils leur sont nécessaires pour exécuter leur PART DU PROJET. Cette concession se fait sans contrepartie financière.

8.2.3.2 Aux fins d'exploitation des RÉSULTATS

Chaque PARTIE s'engage à concéder aux autres PARTIES et/ou à leurs AFFILIÉS, une licence sur ses RÉSULTATS lorsqu'ils sont nécessaires à l'exploitation, par la PARTIE ou l'AFFILIÉ qui en fait la demande, de ses RÉSULTATS.

À cette fin, pendant la durée du PROJET et 24 mois après son terme, chaque PARTIE détentrice s'engage sur demande écrite à concéder par acte séparé aux autres PARTIES une licence à des conditions justes et raisonnables.

8.2.3.3 A des fins de recherche interne

Les PARTIES concèdent un droit d'utilisation de leurs RÉSULTATS aux autres PARTIES à des fins de recherche interne exclusivement.

Cette demande devra être faite par acte séparé et sur demande écrite pendant la durée du projet ou [•X] mois après son terme.

Cette concession se fait sans contrepartie financière.

La PARTIE détentrice ne peut s'y opposer, sauf intérêts légitimes.

ARTICLE 9 – CONFIDENTIALITÉ – PUBLICATIONS

9.1 CONFIDENTIALITÉ

9.1.1 Chacune des PARTIES, pour autant qu'elle soit autorisée à le faire, transmettra aux autres PARTIES ses seules INFORMATIONS CONFIDENTIELLES qu'elle juge nécessaires à la réalisation du PROJET.

Aucune stipulation de l'ACCORD ne peut être interprétée comme obligeant l'une des PARTIES à communiquer ses INFORMATIONS CONFIDENTIELLES à une autre PARTIE.

9.1.2 La PARTIE qui reçoit une INFORMATION CONFIDENTIELLE (ci-après désignée la « PARTIE RÉCIPIENDAIRE ») d'une autre PARTIE (ci-après désignée la « PARTIE ÉMETTRICE ») s'engage, pendant la durée de l'ACCORD et pendant les cinq (5) ans qui suivent la fin de l'ACCORD, quelle qu'en soit la cause, à ce que les INFORMATIONS CONFIDENTIELLES émanant de la PARTIE ÉMETTRICE :

a) soient protégées et gardées strictement confidentielles ;

b) ne soient communiquées qu'aux seuls membres de son personnel, à ses AFFILIÉS ou à ses sous-traitants ayant à en connaître pour la réalisation du PROJET et sous réserve qu'ils soient tenus d'obligations de confidentialité au moins aussi strictes que celles résultant des présentes ;

c) ne soient utilisées par lesdites personnes visées au b) ci-dessus que dans le but défini par l'ACCORD ;

d) ne soient copiées, reproduites ou dupliquées totalement ou partiellement qu'aux fins de réalisation du PROJET.

Toutes les INFORMATIONS CONFIDENTIELLES et leurs reproductions, transmises par une PARTIE à une autre PARTIE, resteront la propriété de la PARTIE ÉMETTRICE sous réserve des droits des tiers et devront être restituées à cette dernière ou détruites sur sa demande, à l'exception d'une copie qui pourra être conservée à des seules fins d'archivage.

En tout état de cause, la PARTIE RÉCIPIENDAIRE reste responsable envers la PARTIE ÉMETTRICE du respect par ses AFFILIÉS et sous-traitants des obligations prévues au présent article 9.1.2.

9.1.3 La PARTIE RÉCIPIENDAIRE n'aura aucune obligation et ne sera soumise à aucune restriction eu égard à toutes les INFORMATIONS CONFIDENTIELLES dont elle peut apporter la preuve :

a) qu'elles sont entrées dans le domaine public préalablement à leur divulgation ou après celle-ci mais dans ce cas en l'absence de toute faute de la PARTIE RÉCIPIENDAIRE ;

b) qu'elles étaient licitement en sa possession avant de les avoir reçues de la PARTIE ÉMETTRICE ;

c) qu'elles ont été reçues d'un tiers autorisé à les communiquer ;

d) que leur utilisation ou communication a été autorisée par écrit par la PARTIE ÉMETTRICE ;

e) qu'elles ont été développées de manière indépendante et de bonne foi par des personnels de la PARTIE RÉCIPIENDAIRE n'ayant pas eu accès à ces INFORMATIONS CONFIDENTIELLES.

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.

9.1.4 Sans préjudice des articles 7 et 8, il est expressément convenu entre les PARTIES que la communication par les PARTIES entre elles d'INFORMATIONS CONFIDENTIELLES, au titre de l'ACCORD, ne peut en aucun cas être interprétée comme conférant de manière expresse ou implicite à la PARTIE RÉCIPIENDAIRE un droit quelconque, notamment de propriété intellectuelle (sous forme d'une licence ou par tout autre moyen) sur les INFORMATIONS CONFIDENTIELLES.

9.2 PUBLICATIONS – COMMUNICATIONS

9.2.1 Dans le respect des stipulations de l'article 9.1, tout projet de communication, notamment par voie de publication, présentation sous quelque support ou forme que ce soit, relatif au PROJET, aux RÉSULTATS COMMUNS ou intégrant les RÉSULTATS PROPRES des autres PARTIES, par l'une ou l'autre des PARTIES, devra recevoir, pendant la durée de l'ACCORD et

les deux (2) ans qui suivent son expiration ou sa résiliation, l'accord préalable écrit des autres PARTIES.

Ces autres PARTIES feront connaître leur décision dans un délai maximum de soixante (60) jours calendaires à compter de la date de notification de la demande, cette décision pouvant consister :

- à accepter sans réserve le projet de communication ; ou
- à demander que les INFORMATIONS CONFIDENTIELLES leur appartenant soient retirées du projet de communication ; ou
- à demander des modifications, en particulier si certaines informations contenues dans le projet de communication sont de nature à porter préjudice à l'exploitation industrielle et commerciale des CONNAISSANCES PROPRES et/ou RÉSULTATS ; ou
- à demander que la communication soit différée si des causes réelles et sérieuses leur paraissent l'exiger, en particulier si des informations contenues dans le projet de publication ou de communication doivent faire l'objet d'une protection au titre de la propriété industrielle.

Toutefois, aucune des PARTIES ne pourra refuser dans ce cas son accord à une publication ou communication au-delà d'un délai de dix-huit (18) mois suivant la première soumission du projet concerné.

En l'absence de réponse d'une PARTIE à l'issue de ce délai de soixante jours (60) calendaires, son accord sera réputé acquis.

À l'issue du délai des deux (2) ans, toute publication ou communication se fera dans le respect des obligations de confidentialité stipulées à l'article 9.1 ci-avant.

Ces communications devront mentionner le concours apporté par chacune des PARTIES à la réalisation du PROJET, ainsi que l'aide apportée par l'ADEME.

9.2.2 Sous réserve du respect des stipulations de l'article 9.1 relatives à la confidentialité, les termes de l'article 9.2.1 ne pourront faire obstacle :

- ni à l'obligation qui incombe à chacune des personnes participant au PROJET de produire un rapport d'activité à ou aux organisme(s) dont elle relève ;
- ni à la soutenance de thèse des chercheurs participant au PROJET; cette soutenance, organisée dans le respect de la réglementation universitaire en vigueur. Cette soutenance pourra être organisée à huis clos à chaque fois que cela est nécessaire ;

– ni aux dépôts par une ou plusieurs PARTIES d'une demande de brevet découlant uniquement de leurs RÉSULTATS ;

– ni à la publication ou communication par une PARTIE de ses RÉSULTATS PROPRES.

ARTICLE 10 – RESPONSABILITÉS – ASSURANCES

10.1 DISPOSITIONS GÉNÉRALES

10.1.1 RESPONSABILITE A L'EGARD DES TIERS

Chacune des PARTIES reste responsable, dans les conditions du droit commun, des dommages que son personnel pourrait causer aux tiers à l'occasion de l'exécution de l'ACCORD.

10.1.2 RESPONSABILITE ENTRE LES PARTIES

10.1.2.1 Dommages corporels

Chacune des PARTIES prend en charge la couverture de son personnel conformément à la législation applicable dans le domaine de la sécurité sociale, du régime des accidents du travail et des maladies professionnelles dont il relève et procède aux formalités qui lui incombent.

Chaque PARTIE est responsable, dans les conditions de droit commun, des dommages de toute nature causés par son personnel au personnel de toute autre PARTIE.

10.1.2.2 Dommages aux biens

Chaque PARTIE est responsable, dans les conditions de droit commun, des dommages qu'elle cause du fait ou à l'occasion de l'exécution de l'ACCORD aux biens mobiliers ou immobiliers d'une autre PARTIE.

10.1.2.3 Dommages indirects

Les PARTIES renoncent mutuellement à se demander réparation des préjudices indirects (perte de production, perte de chiffre d'affaires, manque à gagner, etc.) qui pourraient survenir dans le cadre de l'ACCORD.

10.2 GARANTIES ET RESPONSABILITES DU FAIT DES CONNAISSANCES PRORES, RÉSULTATS ET AUTRES INFORMATIONS

Les PARTIES reconnaissent que les CONNAISSANCES PROPRES, les RÉSULTATS et les autres informations communiquées par l'une des PARTIES à une autre PARTIE dans le cadre de

l'exécution de l'ACCORD sont communiquées en l'état, sans aucune garantie de quelque nature qu'elle soit.

Ces CONNAISSANCES PROPRES, ces RÉSULTATS et ces autres informations sont utilisés par les PARTIES dans le cadre de l'ACCORD à leurs seuls frais, risques et périls respectifs, et en conséquence, aucune des PARTIES n'aura de recours contre une autre PARTIE, ni ses sous-traitants éventuels, ni son personnel, à quelque titre que ce soit et pour quelque motif que ce soit, en raison de l'usage de ces CONNAISSANCES PROPRES, ces RÉSULTATS et ces autres informations, y compris en cas de recours de tiers invoquant l'atteinte à ses droits de propriété intellectuelle.

10.3 ASSURANCES

Chaque PARTIE doit, en tant que de besoin et dans la mesure où cela est compatible avec ses statuts, souscrire et maintenir en cours de validité les polices d'assurance nécessaires pour garantir les éventuels dommages aux biens ou aux personnes qui pourraient survenir dans le cadre de l'exécution de l'ACCORD.

ARTICLE 11 – DURÉE DE L'ACCORD

L'ACCORD entre en vigueur à la DATE D'EFFET

Il est conclu pour une durée de 30 mois

Toute prolongation donnera lieu à l'établissement d'un avenant signé des PARTIES.

Les stipulations des articles 7, 8, 9 et 10 demeureront en vigueur, pour la durée qui leur est propre si une telle durée est précisée, nonobstant l'expiration ou la résiliation de l'ACCORD.