
Identifier les données à ouvrir

Etalab

etalab^{gouv.fr}

13/10/2020

Table des matières

1	Introduction	3
1.1	Cadre général et opportunités	3
1.2	Un guide synthétique et pédagogique	3
2	L'open data	3
2.1	Qu'est-ce que l'open data ?	3
2.2	Dans quel intérêt publier vos données ?	4
3	L'ouverture des données	5
3.1	La communication de vos documents administratifs	5
3.2	La diffusion de vos documents administratifs	5
3.2.1	Êtes-vous concerné ?	5
3.2.2	Quelles sont vos obligations ?	5
3.2.3	Comment publier ?	6
3.2.4	Que faire si mes documents administratifs contiennent des secrets légaux ? . .	6
3.2.5	Que faire si mes documents administratifs contiennent des données à caractère personnel ?	7
4	La réutilisation des données	8
4.1	Qu'est-ce qu'une réutilisation ?	8
4.1.1	Quelles sont ses caractéristiques ?	8
4.1.2	Faut-il utiliser une licence ?	8
4.2	Les restrictions à la réutilisation des données	9
4.2.1	Le cas de la propriété intellectuelle	9
4.2.2	Le cas de la redevance	9
4.3	Les obligations du réutilisateur	10
4.4	Quels sont les bénéfices liés aux réutilisations ?	10
5	Chronologie de l'open data	10

1 Introduction

1.1 Cadre général et opportunités

En tant qu'organisme public ou privé en charge d'une mission de service public, quelles données êtes-vous tenus de rendre publiques ? Quelles sont les règles à respecter ? À quoi cela sert-il ? Ce guide vous présente le cadre général actuel en la matière afin de vous aider à l'appliquer facilement.

1.2 Un guide synthétique et pédagogique

L'objectif de ce guide est de présenter de manière simple et synthétique le cadre légal de l'open data. Afin de vous guider dans votre démarche d'ouverture, vous trouverez les questions essentielles à vous poser avant de publier des fichiers en open data.

À noter

- Ce guide n'a pas vocation à traiter des obligations prévues par des législations spéciales ;
- Les guides d'Etalab ont une visée avant tout pratique et opérationnelle. Si vous souhaitez obtenir des informations juridiques plus détaillées, nous vous invitons à consulter le [guide de publication en ligne et réutilisation des données publiques](#) proposé par la CNIL et la CADA.

2 L'open data

2.1 Qu'est-ce que l'open data ?

Il existe de nombreuses définitions de l'open data. L'objectif de ce guide n'est pas d'apporter une définition essentielle et exclusive du concept, mais de proposer une interprétation de l'open data public, qu'Etalab a pour mission de mettre en œuvre.

Qu'est-ce que l'open data public ?

Dans le cadre de ses missions de service public, l'administration produit et reçoit des documents administratifs. Ces documents administratifs peuvent contenir des informations publiques, qui peuvent elles-mêmes être représentées sous forme de données publiques. L'open data public consiste à assurer la large mise à disposition à tous de ces données, en accès libre et gratuit, sous un format numérique facilement réutilisable.

Lexique :

- **Administration** : L'administration englobe l'État, les collectivités territoriales ainsi que les autres personnes de droit public ou les personnes de droit privé chargées d'une mission de service public ([Article L300-2 du CRPA](#));
- **Document administratif** : Tout document que l'administration a pu produire ou recevoir (de la part d'une autre administration ou d'un prestataire par exemple), dans le cadre de sa mission de service public ([Article L300-2 du CRPA](#)). Ces documents peuvent correspondre à des notes de services, une base de données, une législation, un code source de logiciel, des cartes, un algorithme, etc. Un document sur lequel un tiers détient des droits de propriété n'est pas considéré comme un document administratif ;
- **Information publique** : Information contenue dans un document administratif communicable à tous ou faisant l'objet d'une diffusion publique, sur lequel des tiers ne détiennent pas de droits de propriété intellectuelle ([Article L321-2 du CRPA](#));
- **Donnée publique** : Représentation d'une information publique sous une forme conventionnelle destinée à faciliter son traitement. Cela peut être par exemple des données géographiques (adresses, références cadastrales), financières (budgets, commande publique, subventions, etc.), environnementales (émissions, vente de produits, etc.), etc.

2.2 Dans quel intérêt publier vos données ?

Au-delà du respect du cadre légal, publier vos données de manière proactive présente de multiples intérêts, notamment :

- **Valoriser votre action** : publier en open data les données que vous produisez donne de la visibilité à votre travail et vos missions ;
- **Alléger votre charge de travail** : une fois le jeu de données publié, vous n'avez plus besoin de répondre à chaque demande d'accès isolée émanant d'un citoyen ou d'une administration ;
- **Renforcer votre efficacité et améliorer les services aux usagers** : les données ouvertes par des administrations peuvent être réutilisées par d'autres services ou aboutir à des collaborations entre équipes, ce qui peut améliorer la mise en œuvre des missions de service public ;
- **Améliorer la qualité de vos données afin d'améliorer votre action et favoriser la création de nouveaux services** : les données que vous publiez seront réutilisées par des acteurs publics ou privés qui pourront les croiser avec d'autres données ou détecter des anomalies voire les corriger. Ces données pourront être utilisées par des tiers afin de créer de nouveaux services numériques.

3 L'ouverture des données

Le cadre juridique de l'open data public repose principalement sur les textes applicables en matière d'accès, de diffusion et de réutilisation des documents administratifs. Le cadre juridique relatif à l'ouverture de l'information publique a cependant considérablement évolué au fil des décennies, jusqu'à la [loi pour une République numérique](#) promulguée en 2016, qui fait de l'ouverture des données publiques par défaut la règle.

Le [livre III du Code des relations entre le public et l'administration \(CRPA\)](#) définit le cadre général de l'ouverture des données publiques. Il intègre tous les textes applicables en matière de communication, de diffusion et de réutilisation des documents administratifs.

3.1 La communication de vos documents administratifs

Le régime de droit d'accès aux documents administratifs a peu évolué depuis [la loi dite "CADA" de 1978](#) : toute administration ou délégation de service public doit communiquer à un administré le document dont il a fait la demande.

Si l'administré demande en outre la diffusion en ligne de ce document administratif, toute administration, quelle que soit sa taille, doit répondre à cette obligation. Si le document contient des données couvertes par un secret légal ou des données à caractère personnel, ces données devront au préalable faire l'objet d'une occultation ou d'une anonymisation.

3.2 La diffusion de vos documents administratifs

3.2.1 Êtes-vous concerné ?

Vous êtes concerné par la diffusion des documents administratifs si vous êtes :

- une administration centrale de plus de 50 agents ;
- une personne morale de droit privé chargée d'une mission de service public qui emploie plus de 50 agents à temps plein ;
- une collectivité territoriale de plus de 3 500 habitants et de plus de 50 agents.

3.2.2 Quelles sont vos obligations ?

Vous êtes tenus de diffuser en open data ([Article L. 312-1-1 du CRPA](#)) :

- Les documents administratifs que vous avez communiqué à des demandeurs ;

- L'inventaire des documents administratifs que vous produisez dans le cadre de vos missions de service public ;
- Les bases de données produites et reçues dans le cadre des missions de service public. Ces bases de données doivent être mises à jour régulièrement ;
- Les données dont la publication représente un intérêt économique, social, sanitaire ou environnemental.

Les documents administratifs diffusés doivent être achevés. C'est-à-dire qu'ils ont atteint leur version finale, à date : les brouillons, documents de travail, notes préalables ne sont pas considérés comme des documents achevés. Si le document administratif contient une décision, cette dernière ne doit pas être en cours de délibération mais bien prise.

Lexique : Base de données

On entend par base de données un recueil d'œuvres, de données ou d'autres éléments indépendants, disposés de manière systématique ou méthodique, et individuellement accessibles par des moyens électroniques ou par tout autre moyen ([Article L112-3 du code de la propriété intellectuelle](#)). À titre d'exemples, sont des bases de données : le registre des entreprises, l'annuaire des adresses, les données de demande de valeurs foncières, etc.

3.2.3 Comment publier ?

Les documents administratifs, informations publiques et données doivent être publiés dans un format :

- **Ouvert** : tout protocole de communication, d'interconnexion ou d'échange et tout format de données interopérable et dont les spécifications techniques sont publiques, sans restriction d'accès ou de mise en œuvre ;
- **Aisément réutilisable** : le producteur prend en considérant les connaissances et besoins du réutilisateur lors de la publication ;
- **Exploitable par un système de traitement automatisé** : la publication est optimisée pour une utilisation par un système de traitement automatisé et non pour une exploitation immédiate par des humains.

3.2.4 Que faire si mes documents administratifs contiennent des secrets légaux ?

Quels sont les documents couverts par un secret légal ?

- Les documents qui ne sont aucunement communicables. Ce sont par exemple les documents dont la diffusion porterait atteinte au secret des délibérations du Gouvernement,

- au secret de la défense nationale ou de la sûreté de l'État, etc ([Article L. 311-5 du CRPA](#)).
- Les documents dont la diffusion porterait atteinte à la protection de la vie privée, au secret médical et au secret des affaires. Les documents qui portent une appréciation ou un jugement de valeur sur une personne physique ou qui font apparaître le comportement d'une personne ([Article L. 311-6 du CRPA](#)).

Si vos documents administratifs contiennent des secrets légaux, vous êtes tenus d'occulter ces secrets par un traitement d'usage courant, sans que cette opération implique des efforts disproportionnés ou que le document soit dénaturé ou vidé de son sens. Le cas échéant, vous n'êtes pas tenu de diffuser le document administratif.

Comment occulter les données par un traitement automatisé d'usage courant ? L'occultation correspond au masquage ou au retrait des données identifiées comme confidentielles et non communicables.

3.2.5 Que faire si mes documents administratifs contiennent des données à caractère personnel ?

Lexique : Donnée à caractère personnel

Toute information relative à une personne physique identifiée ou qui peut être identifiée, directement ou indirectement, par référence à un numéro d'identification (par exemple le numéro de sécurité sociale) ou à un ou plusieurs éléments qui lui sont propres.

Le cadre juridique général proscrit la diffusion en ligne, sans anonymisation, de documents administratifs contenant des données à caractère personnel. Cependant, trois situations permettent la publication de ces documents sans avoir recours à l'anonymisation :

- Si une disposition législative spécifique autorise la publication des données sans anonymisation ;
- Si les personnes concernées ont donné leur accord à la diffusion des données sans anonymisation ;
- Si les documents administratifs figurent dans la liste prévue par le [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. Ce sont notamment les documents relatifs aux conditions d'organisation de l'administration, de la vie économique, associative, culturelle et sportive, des professions réglementées, etc.

Si votre document administratif contenant des données à caractère personnel ne correspond à aucune de ces situations, vous êtes tenus de l'anonymiser. Cette opération ne doit toutefois pas impliquer

d'efforts disproportionnés. L'anonymisation ne doit également pas dénaturer ou vider de son sens le document. Le cas échéant, vous n'êtes pas tenu de diffuser le document administratif.

Si vous souhaitez obtenir d'avantage d'informations juridiques sur l'articulation entre open data et protection des données à caractère personnel, nous vous invitons à consulter le [guide de publication en ligne et de réutilisation des données publiques](#) produit par la CNIL.

Lexique : Anonymisation des données

Processus consistant à traiter des données à caractère personnel afin d'empêcher totalement et de manière irréversible l'identification d'une personne physique. L'anonymisation suppose donc qu'il n'y ait plus aucun lien possible entre l'information concernée et la personne à laquelle elle se rattache.

4 La réutilisation des données

4.1 Qu'est-ce qu'une réutilisation ?

La réutilisation des informations publiques désigne l'utilisation des données publiques par des tiers à d'autres fins que celle de la mission de service public pour laquelle les documents ont été produits ou reçus.

4.1.1 Quelles sont ses caractéristiques ?

La réutilisation des données doit être libre, c'est-à-dire :

- Elle est gratuite. Les seuls organismes habilités à pratiquer des redevances sont l'Institut national de l'information géographique et forestière (IGN), Météo-France et le service hydrographique et océanographique de la Marine (SHOM) ([Décret n° 2016-1617](#)) ;
- Elle peut viser une autre finalité que le but initial de production du jeu de données ;
- Elle peut être réalisée par tout acteur, qu'il soit public ou privé. Une administration ne peut demander à ce que le réutilisateur ait une qualité particulière pour accéder aux données.

4.1.2 Faut-il utiliser une licence ?

- Lorsque les données sont mises à disposition gratuitement, l'usage d'une licence est conseillé, mais pas obligatoire ;
- Si les données publiées sont mises à disposition contre le paiement d'une redevance, les administrations productrices sont dans l'obligation d'apposer une licence de réutilisation.

Comme indiqué précédemment, la réutilisation des données doit être libre. La licence doit répondre aux différents critères de libre réutilisation. À ce titre, la libre réutilisation ne peut être restreinte que pour des motifs d'intérêt général. Cette restriction doit être proportionnée et ne doit pas avoir pour effet ou objectif de limiter la concurrence.

Licences de réutilisation autorisées

Dans le but d'avoir un nombre restreint de licences, la loi pour une République numérique a prévu la création d'une liste, [fixée par décret](#), de licences qui peuvent être utilisées par les administrations pour la réutilisation à titre gratuit de leurs informations publiques.

Les administrations peuvent choisir parmi cette liste de licences lorsqu'elles publient des éléments en ligne. Les administrations souhaitant recourir à une licence ne figurant pas dans la liste des licences autorisées par décret doivent au préalable demander son homologation auprès de la direction interministérielle du numérique (DINUM).

Vous pouvez consulter [la liste des licences autorisées par décret](#).

4.2 Les restrictions à la réutilisation des données

4.2.1 Le cas de la propriété intellectuelle

D'après le cadre général, l'administration ne peut se prévaloir d'un droit de propriété intellectuelle sur une base de données afin de restreindre sa réutilisation.

Cependant, une administration peut se prévaloir d'un droit *sui generis* sur une base de données, uniquement si elle valide les conditions suivantes :

- l'administration est en situation de concurrence ;
- la création de la base de données relève d'une création de l'esprit et a entraîné des investissements substantiels.

4.2.2 Le cas de la redevance

Comme indiqué précédemment, certaines administrations sont habilitées à pratiquer des redevances pour l'accès et la réutilisation des données. Certaines limitations à la réutilisation peuvent être envisagées pour les administrations qui pratiquent des redevances, afin de préserver ce modèle, notamment en interdisant la rediffusion des données achetées.

4.3 Les obligations du réutilisateur

Lorsqu'un individu réutilise un jeu de données publié, il est tenu de :

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

De plus, le réutilisateur est tenu de se conformer aux obligations légales qui découlent du [Règlement général sur la protection des données](#).

4.4 Quels sont les bénéfices liés aux réutilisations ?

L'ouverture et la réutilisation des données bénéficie aux administrations :

- Vous êtes une administration productrice et ouvrez des données : la circulation des données vous permettra d'en améliorer la qualité, à l'aune des usages qui en sont faits ;
- Vous réutilisez des données ouvertes : vous accédez aux données produites par d'autres administrations, et grâce à leur exploitation, vous pouvez élaborer des politiques publiques plus efficaces.

L'ouverture des données bénéficie également aux acteurs privés et à la société civile :

À partir de la circulation, du croisement et de l'exploitation de données publiques, des solutions existantes peuvent être améliorées et de nouveaux services peuvent être créés. La circulation des données démultiplie le potentiel d'usage des données publiques et rend possible leur réutilisation pour des finalités qui n'étaient pas envisagées lors de leur production.

5 Chronologie de l'open data

La formalisation de l'ouverture des données publiques a été progressive, au niveau national et européen :

En France :

- 1978 - [Loi n°78-753 du 17 juillet 1978, dite "loi CADA"](#) : création du droit d'accès aux documents administratifs
- 2005 - [Ordonnance n°2005-650 du 6 juin 2005 sur la réutilisation de l'information publique](#) (transposition de la directive européenne de 2003) : création du droit de réutiliser l'information publique
- 2011 - [Création d'une mission « Etalab »](#) chargée de la création d'un portail unique interministériel des données publiques

- 2015 - [Loi relative à la gratuité et aux modalités de la réutilisation des informations du secteur public](#) : création du droit de réutiliser librement les données publiques
- 2016 - [Loi pour une République numérique](#) : consécration du principe de l'open data par défaut.

Ce cadre général s'est étoffé de législations sectorielles ou territoriales, notamment en matière de transport, santé et énergie:

- 2015: [Loi pour la croissance, l'activité et l'égalité des chances économiques](#) : ouverture en open data des données de transport
- 2015 - [Loi sur la Nouvelle organisation territoriale de la République](#) : publication en open data des données des collectivités publiques de plus de 3500 habitants
- 2016 - [Loi pour la Modernisation de notre système de santé](#) : publication en open data des données de santé

En Europe:

- 2003 - [Directive européenne 2003/98/CE, dite PSI](#) : ensemble de règles concernant la réutilisation des données et documents détenus par les organismes des Etats membres de l'Union européenne
- 2007 - [Directive européenne INSPIRE](#) : obligation de publier en open data les données environnementales et géographiques
- 2013 - [Directive 2013/37/UE modifiant la directive 2003/98/CE](#) : encadrement du droit de redevance accordé aux administrations
- 2018 - [Directive 2019/1024/UE concernant les données ouvertes et la réutilisation des informations du secteur public](#) : inclusion des données des entreprises investies d'une mission de service public dans le champ de l'open data.