

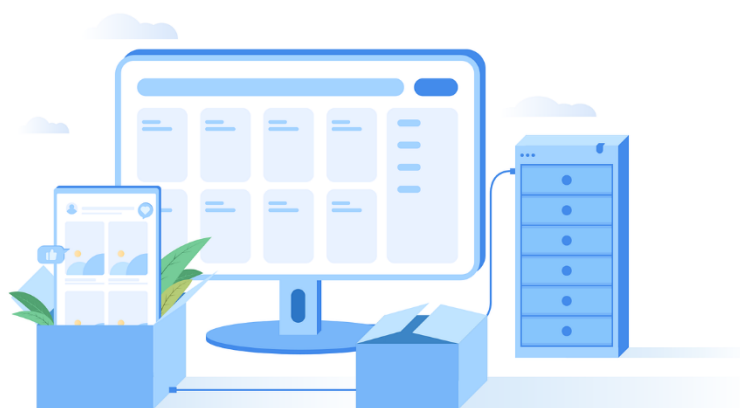
Service d'Archives Itinérant – CDG 90

FICHE ARCHIVES N°25

L'OPEN DATA

Qu'est-ce que l'open data ?

1. Définition
2. L'open data, outil de publication des données issues des services d'archives
3. Les enjeux de l'open data en matière de conservation
4. Les enjeux de l'open data en matière de communication



Lois et textes de références

Loi du 17 juillet 1978 sur la liberté d'accès aux documents administratifs (CADA).

Loi n°2015-1779 du 28 décembre 2015 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 qui entérine notamment le principe d'ouverture des données publiques en France.

Directive (UE) 2019/1024 du 20 juin 2019 sur l'open data et la réutilisation des informations du secteur public qui fixe des obligations minimales de mise à dispositions et de réutilisation des données publiques pour tous les États membres de l'Union Européenne.

Code des relations entre le public et l'administration (CRPA) qui intègre les règles fondamentales sur l'accès et la réutilisation des informations publiques (issues de la loi CADA et de la loi pour une République numérique) et codifie les obligations de publication des données publiques en France.

1. DÉFINITION

L'Open Data (aussi appelé « données ouvertes ») est une démarche de mise à disposition libre et gratuite de données produites ou collectées par des acteurs publics (et parfois privés), afin qu'elles puissent être consultées, réutilisées et partagées par tous. Ces données sont publiées dans des Formats ouverts et accompagnées de licences précisant les conditions de réutilisation.

Quatre grands principes définissent l'Open Data :

- L'**accessibilité** : les données doivent être facilement accessibles notamment en ligne ;
- La **réutilisabilité** : elles doivent pouvoir être exploitées, croisées ou transformées ;
- L'**interopérabilité** : les formats utilisés doivent être standards et lisibles par des outils variés ;
- La **transparence** : l'ouverture des données vise à renforcer la compréhension de l'action publique.

Pour de nombreux pays, un cadre juridique précis lié aux politiques de transparence administrative et de modernisation de l'action publique encadre la mise en place de l'Open Data.

QUI EST CONCERNÉ ?

Plusieurs acteurs sont soumis à l'obligation de communiquer leurs documents administratifs, et donc d'ouvrir leurs données plus largement. C'est notamment le cas pour :

- 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.

2. L'OPEN DATA, UN OUTIL DE PUBLICATION DES DONNÉES ISSUES DES SERVICES D'ARCHIVES

Les services d'archives produisent et gèrent une grande diversité de données susceptibles d'être publiées en Open Data. On peut ainsi citer les **instruments de recherche** (inventaires, répertoires, index), les **métadonnées descriptives** des Fonds et collections, les **données de contextualisation** (producteurs, Fonctions administratives, chronologies), les **données statistiques** (Fréquentation, volumes conservés, typologies documentaires). Ces données ne correspondent pas nécessairement aux documents eux-mêmes mais plutôt aux informations permettant de les identifier, de les comprendre et de les localiser.

De plus, la publication des données archivistiques en open data s'inscrit directement dans les missions traditionnelles des services d'archives : collecter et conserver la mémoire des institutions et des territoires, classer et décrire les documents, communiquer les archives au public

L'open data peut alors être vu comme un **outil complémentaire de communication**, permettant de diffuser largement des informations structurées sur les Fonds, au-delà des seuls lecteurs des salles de consultation.

LES MODALITÉS DE PUBLICATION

Dans une logique d'open data, les services d'archives publient leurs données :

- Sur des **portails dédiés** (portails open data nationaux ou territoriaux)
- Dans des **formats ouverts** (CSV, XML, JSON, RDF, etc.)
- Sous des **licences ouvertes** (exemple d'etalab), autorisant la réutilisation, parfois sous conditions

Cette publication facilite la réutilisation des données archivistiques par des chercheurs, des développeurs, des enseignants ou des acteurs culturels, qui peuvent alors créer de nouveaux outils, visualisations ou services.

3. LES ENJEUX DE L'OPEN DATA EN MATIÈRE DE CONSERVATION

PÉRENNITÉ ET QUALITÉ DES DONNÉES

L'ouverture des données pose des questions importantes en matière de conservation numérique. En effet, il est nécessaire que les données publiées soient :

- **Fiables** et correctement documentées
- **Maintenues dans le temps**, en tenant en compte de l'évolution des formats et des outils
- **Versionnées**, afin de conserver une traçabilité des modifications

Pour les services d'archives, cela implique une réflexion sur les chaînes de traitement, la normalisation des métadonnées et l'articulation entre systèmes de gestion internes et plateformes de diffusion.

DISTINCTION ENTRE DONNÉES OUVERTES ET ARCHIVES DÉFINITIVES

Un enjeu majeur consiste à bien distinguer :

- Les **données ouvertes**, conçues pour la diffusion et la réutilisation
- Les **documents d'archives**, soumis à des règles strictes de conservation, d'authenticité et d'intégrité

L'open data ne se substitue pas aux missions archivistiques classiques, mais nécessite une clarification des statuts juridiques et archivistiques des données publiées.

CONTRAINTES JURIDIQUES ET ÉTHIQUES

Les services d'archives doivent également veiller à prendre en compte la protection des données personnelles (RGPD) et les délais de communicabilité.

L'ouverture des données implique donc un travail d'analyse préalable pour éviter toute diffusion d'informations sensibles ou protégées.

4. LES ENJEUX DE L'OPEN DATA EN MATIÈRE DE COMMUNICATION ET DE VALORISATION

4.1. Les archives

ÉLARGISSEMENT DES PUBLICS

L'open data permet aux services d'archives de toucher des publics qui ne fréquentent pas forcément les institutions archivistiques, tels que les acteurs du numériques, les journalistes et médiateurs culturels, les enseignants et étudiants, le grand public intéressé par l'histoire notamment.

Les données ouvertes peuvent être réutilisées dans des projets éducatifs, culturels ou scientifiques, contribuant ainsi à une meilleure visibilité et connaissance des archives.

NOUVELLES FORMES DE MÉDIATION

Grâce à l'open data, les archives peuvent être valorisées à travers :

- Des **cartographies interactives** ;
- Des **visualisations de données** ;
- Des **applications ou sites web thématiques**.

Ces dispositifs favorisent une approche plus intuitive et pédagogique des fonds, tout en renouvelant les pratiques de médiation.

IMAGE INSTITUTIONNELLE ET TRANSPARENCE

L'open data contribue également à renforcer la transparence des institutions, leur image d'ouverture et de modernité ainsi que la confiance des usagers dans la gestion de l'information publique.

Pour les services d'archives, cela leur permet d'affirmer leur rôle central dans la gestion, la diffusion et la valorisation des données publiques.

4.2. Les collectivités

Depuis la loi CADA de 1978, le régime de droit d'accès aux documents administratifs se définit comme suit : **toute administration ou délégation de service public doit communiquer à un administré le document dont il fait la demande.**

En ce sens, si l'administré demande la diffusion en ligne de ce document administratif, toute administration, quelle que soit sa taille, doit répondre à cette obligation. Cependant, **si le document contient des données à caractère personnel ou si ces données sont couvertes par un secret légal, une occultation ou une anonymisation** de ces données devront être réalisées pour garantir le respect du RGPD¹ notamment.

QUE FAUT-IL DIFFUSER EN OPEN DATA ?

D'après l'article L.312-1-1 du CRPA (Code des Relations entre le Public et l'Administration), pour les personnes concernées par l'obligation légale, il est nécessaire de diffuser en open data :

¹ Voir la fiche pratique n°23 sur le RGPD.

- 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.

Par ailleurs, il est essentiel que les documents administratifs diffusés soient achevés au sens où ils sont présents dans leur version finale (donc pas de brouillons, documents de travail, notes préalables). Aussi, si le document administratif contient une décision, celle-ci doit être prise et non pas en cours de délibération.

COMMENT FAUT-IL PUBLIER EN OPEN DATA ?

Les Formats

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 considération 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.

L'occultation des secrets légaux

Pour les documents administratifs qui contiennent des secrets légaux, il est obligatoire de les occulter mais sans que le document se retrouve ensuite dénaturé ou vidé de son sens. Pour se faire, marquez ou retirez les données identifiées comme étant confidentielles et non communicables.

Le secret légal désigne plusieurs types de documents :

- Ceux qui ne sont aucunement communicables (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 [article L.311-5 du CRPA]) ;
- Ceux dont la diffusion porterait atteinte à la protection de la vie privée, au secret médical et au secret des affaires ;
- Ceux 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].

L'anonymisation des données

Si l'on se réfère au cadre juridique général, il est interdit de diffuser en ligne, sans anonymisation, des documents administratifs contenant des données à caractères personnels. Néanmoins, trois situations exceptionnelles existent :

- 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 se trouvent 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 (comme par exemple les documents concernant les conditions d'organisation de l'administration, de la vie économique, associative, culturelle et sportive, des professions réglementées, etc.)

LA RÉUTILISATION DES DONNÉES

La réutilisation des informations publiques signifie que les données sont utilisées par des tiers à d'autres fins que celle de la mission de service public pour laquelle les documents ont été produits ou reçus. Cette réutilisation se doit d'être **libre**, c'est-à-dire qu'elle est **gratuite**, peut viser une **autre finalité que le but initial** de production et **peut être réalisée par tout acteur** (aussi bien public que privé).

Celui qui réutilise des données est tenu de respecter les conditions de la licence sous laquelle les données publiques ont initialement été publiées. En ce sens, il existe deux principales licences : la **licence ouverte 2.0 (Licence Etalab)** ; la **licence ODbL (Open Database License)**. La première demande de faire mention de la source des données et de la date de dernière mise à jour de la réutilisation ainsi que de ne pas altérer le sens des données. La seconde suit les mêmes règles que la première et rajoute une clause dite de « partage à l'identique ».

S'il n'y a pas de mention d'une licence, il faudra appliquer les dispositions de l'article L322-1 du CRPA, qui fixe des conditions de réutilisation similaires à celles de la licence ouverte.

Le réutilisateur se doit également de respecter les obligations légales qui découlent du RGPD.