

Qu'est-ce que l'open data ?

L'open data est le principe de mise en accès libre des données. Il implique une liberté d'accès, de réutilisation et de redistribution des données libre, immédiate, permanente et en ligne.

En France, la loi pour une République numérique du 7 octobre 2016 précise que dès lors qu'une donnée issue d'une activité de recherche, financée au moins pour moitié par des fonds publics, a été rendue publique, sa réutilisation doit être libre de droit.

Cela signifie donc que l'open data pour les données de la recherche est désormais obligatoire sauf exception liée à une protection spécifique, comme un brevet.

Quelles sont les données concernées ?

Les données concernées sont les données de recherche.

Selon l'OCDE, les données de recherche (research data) sont « des enregistrements factuels (chiffres, textes, images et sons), qui sont utilisés comme sources principales pour la recherche scientifique et sont généralement reconnus par la communauté scientifique comme nécessaires pour valider des résultats de recherche. »

Il est aussi possible de mettre en accès libre toute autres donnée analysée et/ou brute collectée ou produite durant le projet en précisant les outils et instruments utilisés et nécessaires à la validation des résultats

Afin de permettre leur réutilisation, les données mises en accès libre doivent être associées de leurs métadonnées – informations concernant le type et l'origine des données, les outils et instruments utilisés, etc.

Comment concilier open data et RGPD ?

Le principe « aussi ouvert que possible, aussi fermé que nécessaire » est au cœur de la démarche d'open data. À ce titre, pour diffuser librement les données de la recherche, il convient de concilier open data et RGPD.

En pratique, l'articulation entre manipulation de données à caractère personnel et open data devrait être actée dans un plan de gestion des données. Aujourd'hui, ce plan n'est obligatoire que dans le cadre de projets financés par l'ANR et par la Commission Européenne.

FOCUS

À l'UCA, la Bibliothèque universitaire propose un accompagnement pour la mise en place de tels plans. Les ingénieurs projets du Pôle Ingénierie de Projets de la Direction de la Recherche et de la Valorisation, qui accompagnent le montage des projets de recherche, assure le lien entre les différents acteurs du dossier dans le cadre de l'accompagnement du projet.

Le plan de gestion de données ou Data Management Plan (DMP) est un document formalisé explicitant notamment les modalités de production, de traitement, de protection, de stockage, d'archivage et de diffusion des données dans le cadre de la recherche sur projet. L'objectif est de faciliter la gestion, la réutilisation et la protection, ainsi que l'archivage à long terme des données de recherche.

Par exemple, l'Agence Nationale de la Recherche (ANR) recommande dans son modèle de plan de gestion de données de veiller « à ce que les lois sur la protection des données soient appliquées ».

Elle préconise notamment :

- d'obtenir un consentement éclairé pour la préservation et/ou le partage de données personnelles ;
- d'envisager :
 - l'anonymisation des données à caractère personnel pour la préservation et/ou le partage. Des données correctement anonymisées ne sont plus considérées comme des données à caractère personnel (cf fiche 03 [↗ CLIQUER ICI](#)) ;
 - la pseudonymisation des données à caractère personnel ;
 - le chiffrement des données, qui est considéré comme un cas particulier de pseudonymisation. La clé de cryptage doit alors être stockée séparément des données, par exemple par un tiers de confiance (cf fiche 03 [↗ CLIQUER ICI](#)) ;
- d'expliquer si une procédure d'accès spécifique a été mise en place pour les utilisateurs autorisés à accéder aux données à caractère personnel.