

Qu'est-ce que l'open data ?



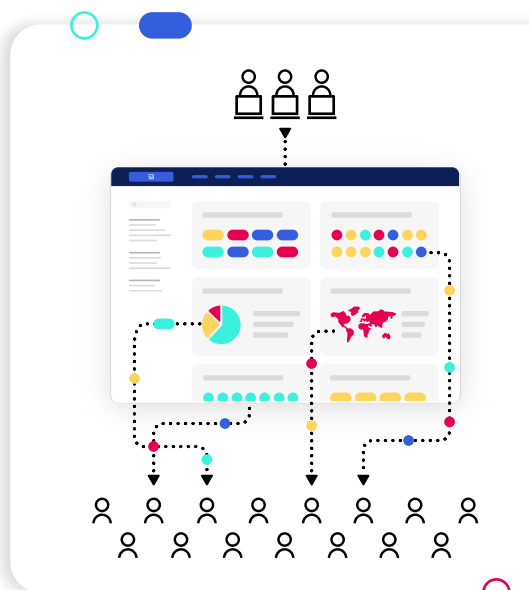
Sommaire

1	Qu'est-ce que l'open data ?	3
2	Quelles sont les réglementations en matière de données ouvertes ?	7
3	Pourquoi avons-nous besoin de données ouvertes ? Quels en sont les avantages ?	10
4	Quels secteurs peuvent ouvrir leurs données ?	12
5	Comment préparer vos données ?	17
6	Comment publier vos données en open data ?	20
7	Comment réutiliser des données publiées en open data ?	23
8	Gouvernance, outils et culture	26
9	Comment trouver des données ouvertes	29
10	Maturité des données ouvertes et exemples de bonnes pratiques	31

1. Qu'est-ce que l'open data ?

Qu'entendez-vous par “données” ?

En termes simples, les données sont des faits ou des statistiques rassemblés à des fins de référence et d'analyse. Il peut s'agir de chiffres, de textes ou d'une combinaison des deux. Les données sont normalement collectées pour permettre une analyse ou une prise de décision. Il peut s'agir de données simples, comme le comptage du nombre de ventes d'un produit particulier, le relevé d'un compteur d'électricité ou l'enregistrement des noms des employés. Il peut aussi s'agir de données plus complexes, telles que les ventes de produits par lieu ou la rentabilité de certains services.



C'est là que l'adoption de [l'open data](#) permet de libérer tout le potentiel des données. Les données ouvertes sont des données librement accessibles à tous, y compris aux entreprises, aux citoyens, aux médias et aux consommateurs. Voici quelques définitions populaires des données ouvertes :

✓ **“Les données et le contenu ouverts peuvent être librement utilisés, modifiés et partagés par n’importe qui, dans n’importe quel but.”**

[Open Definition, Open Knowledge Foundation.](#)

✓ **“Les données ouvertes sont des données auxquelles tout le monde peut accéder, utiliser ou partager.”**

[Open Data Institute](#)

✓ **“Les données ouvertes comprennent les données rendues accessibles pour réutilisation selon un spectre d’ouverture et de conditions de réutilisation.”**

[Open Data Policy Lab](#)

Qu'est-ce qu'une donnée ouverte ?

Les données sont considérées comme ouvertes si elles répondent à trois critères clés :



L'accessibilité et l'accès

Les données doivent être facilement disponibles, par exemple en ligne via un [portail de données ouvertes](#), sous une forme pratique et dans leur intégralité. L'accès aux datasets doit être gratuit, les organisations ne faisant payer que les coûts de reproduction (par exemple, lorsqu'elles impriment et publient des informations).



Réutilisation et redistribution

Les personnes qui accèdent aux données doivent être autorisées à les réutiliser et à les redistribuer, y compris en les combinant avec d'autres ensembles de données. Selon le mode d'octroi de la licence, il peut être obligatoire d'attribuer les données à leur source et de partager tout nouveau jeu de données produit.



Disponibles pour tous

Les jeux de données ouverts doivent être également accessibles à tous les groupes, qu'il s'agisse d'entreprises du secteur privé, de particuliers, de chercheurs ou d'organismes publics. Chacun doit pouvoir les utiliser, les réutiliser et les redistribuer, quel que soit le but recherché.

Il est évident que l'accès à des données précises et actualisées est essentiel pour comprendre ce qui se passe au sein d'une organisation et pour permettre aux individus de gérer leur vie quotidienne. Cependant, ces données sont souvent cloisonnées et ne sont accessibles qu'à des groupes spécifiques au sein d'une organisation. Cela signifie qu'elles n'offrent pas leur véritable valeur à tous, d'où la [nécessité de les ouvrir](#).

Open data : un bref historique

Le terme « données ouvertes » est apparu pour la première fois en 1995, en rapport avec le partage de données géophysiques et environnementales. Le concept a ensuite été codifié en 2007, lors d'une réunion d'activistes et de penseurs de l'internet à Sebastopol, en Californie. Il était fondé sur les principes du mouvement open source et se concentrait sur le partage de données ouvertes par les institutions publiques. Au fur et à mesure que le mouvement des données ouvertes prenait de l'ampleur, des pays du monde entier ont introduit une législation imposant que les données des institutions publiques (telles que les administrations gouvernementales, les agences, les municipalités et autres organismes publics locaux) soient mises à disposition en tant que données ouvertes. Voici quelques exemples de données ouvertes :

- ✓ les informations sur les performances et le budget,
- ✓ données SIG (système d'information géographique) ouvertes,
- ✓ les progrès réalisés par rapport aux objectifs publics,
- ✓ les listes de demandes, par exemple de permis.

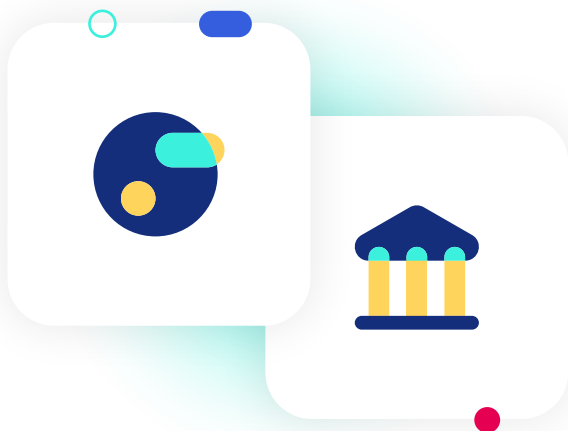
Cette première vague de données ouvertes a été suivie par d'autres règles et lignes directrices qui ont étendu les exigences en matière de données ouvertes à d'autres secteurs, tels que l'énergie, les services publics et les [opérateurs de transport](#) public. Nous constatons aujourd'hui qu'un nombre croissant d'organisations du secteur privé, qu'il s'agisse de [banques](#), d'entreprises de télécommunications ou d'acteurs industriels, ouvrent leurs données, en raison des avantages que cela représente pour leurs activités, leurs écosystèmes et leurs relations avec les clients.

Quelle est la différence entre les données ouvertes et les données du secteur public ?

Le secteur public est tenu de publier et de partager une grande partie de ses données, soit en vertu de réglementations, soit parce qu'elles sont essentielles à un gouvernement et à un contrôle démocratiques. Ces données sont partagées en tant que données ouvertes, par le biais de canaux tels que les sites web, les portails et les applications.

Cependant, les données ouvertes peuvent également être publiées et partagées par des organisations du secteur privé, afin d'accroître l'engagement avec les parties prenantes et de faire preuve de transparence. Un nombre croissant d'entreprises, dans de multiples secteurs, partagent désormais des informations sous forme de données ouvertes.

Par conséquent, les données du secteur public sont toujours des données ouvertes, mais les données ouvertes ne proviennent pas nécessairement du secteur public.



2. Quelles sont les réglementations en matière de données ouvertes ?

À l'origine, l'ouverture des données était motivée par la législation gouvernementale et visait à mettre à la disposition de tous les données créées par les organismes publics. Le raisonnement était clair. En tant que contribuables, les citoyens financent ces organismes, qu'il s'agisse de municipalités locales ou d'administrations centrales et déconcentrées, et ils ont le droit de comprendre comment ils fonctionnent grâce à une [plus grande transparence](#). Les réglementations se sont maintenant étendues à d'autres secteurs, notamment dans des domaines tels que la réalisation des objectifs de décarbonisation et de durabilité, en particulier dans les [secteurs de l'énergie](#) et des services publics locaux.

La législation en Europe



À partir des années 1970, de nombreux pays européens ont adopté des lois sur la liberté d'information, imposant aux organismes du secteur public de fournir des réponses aux questions posées par les citoyens ou les organisations, et de permettre aux personnes d'accéder aux informations détenues par les organismes publics à leur sujet. En France, la loi CADA de 1978 garantit l'accès du public aux données et précise les conditions d'exercice de ces droits, tandis que la loi britannique sur la liberté d'information est entrée en vigueur le 1er janvier 2005.

L'Union européenne a adopté pour la première fois une législation sur la réutilisation des informations du secteur public en 2003, par le biais de la directive 2003/98/CE du Parlement européen. Cette directive fixe des exigences minimales pour les États membres de l'UE en ce qui concerne la mise à disposition des informations du secteur public en vue de leur réutilisation. Elle a ensuite été révisée pour former la directive sur les informations du secteur public (ISP) en 2013, puis remplacée par la directive sur les données ouvertes en 2019. Cette dernière a étendu les réglementations pour couvrir les données pertinentes dans les secteurs des services publics et des transports, ainsi que les données de recherche résultant d'un financement public. Les États membres avaient jusqu'au 16 juillet 2021 pour transposer la nouvelle directive en droit national.



The Data Governance Act

Parmi les autres textes législatifs européens pertinents, citons [la loi sur la gouvernance des données](#), publiée en 2020, et la loi sur les données, publiée en 2022. Ces lois visent à faciliter l'accès et la réutilisation des données industrielles, à étendre l'open data et à positionner l'Europe à l'avant-garde d'une société fondée sur les données. Grâce à cet ensemble de règles, le volume de données disponibles pour la réutilisation augmentera considérablement et devrait générer 270 milliards d'euros de PIB supplémentaire d'ici 2028.

Législation aux États-Unis

L'un des objectifs de la réunion initiale de Sebastopol en 2007 était de faire pression sur le gouvernement fédéral américain pour qu'il adopte une législation sur les données ouvertes. Celle-ci est entrée en vigueur en 2009 sous la forme d'une directive sur le gouvernement ouvert (Open Government Directive). Cette directive impose aux agences fédérales d'ouvrir leurs données et de les partager via le portail Data.gov. Elle a été étendue par la loi OPEN (Open, Public, Electronic and Necessary) Government Data Act, adoptée en 2019.

The Open Government Directive

La législation au niveau des États varie, mais tous les États disposent désormais de leurs propres portails de données ouvertes, ainsi que de lois sur la liberté d'information imposant la transparence et le partage d'informations. Cela signifie qu'il existe un large éventail de données gouvernementales ouvertes disponibles à travers les États-Unis.

3. Pourquoi avons-nous besoin de données ouvertes ? Quels en sont les avantages ?

Nous vivons dans des économies et des sociétés digitalisées qui génèrent des quantités exponentielles de données au quotidien pour fonctionner et opérer. En partageant des informations vitales, les données ouvertes donnent du pouvoir aux citoyens, aux consommateurs, aux employés et aux organisations. Elles leur permettent de prendre de meilleures décisions, fondées sur des données, tout en garantissant qu'ils sont pleinement informés.

L'ouverture des données offre des avantages spécifiques sous quatre formes principales :

Une plus grande transparence

Le partage des données ouvertes aide les citoyens et les consommateurs à mieux comprendre le fonctionnement et les performances des organismes du secteur public et des entreprises privées. Cela renforce l'engagement et la confiance. La mise à disposition d'informations aide les citoyens à contrôler les performances du secteur public et fournit aux consommateurs des données sur des domaines tels que les efforts des entreprises en matière de développement durable, la qualité et la diversité de leur main-d'œuvre, par exemples.

Favoriser l'innovation et la croissance

L'exploitation des données ouvertes aide les organisations à devenir plus innovantes et plus axées sur les données. Elles peuvent créer [des services et des activités entièrement nouveaux](#), soit pour augmenter leurs revenus, soit pour favoriser le changement sociétal. En outre, les organisations peuvent améliorer la prise de décision en complétant leurs propres informations afin d'obtenir une vision plus approfondie. Cela permet d'améliorer l'efficacité, de stimuler l'innovation et d'améliorer l'ensemble des opérations

Une meilleure gestion des politiques publiques

Le partage des données dans l'ensemble du secteur public permet d'améliorer la cohésion et l'intégration des pouvoirs publics. Il supprime les cloisonnements entre les organisations et garantit que les décisions sont prises sur la base d'une image complète des informations disponibles. Cela signifie que les différentes parties de l'administration n'ont pas à collecter leurs propres données, ce qui améliore leur efficacité. Enfin, en rendant les performances transparentes et en montrant les progrès réalisés par rapport aux objectifs, les organisations seront en mesure de mesurer leur succès et de concentrer leurs efforts sur la réalisation des objectifs.

Avantages pour la culture et l'environnement

Aujourd'hui, les gens veulent de la transparence, et les données ouvertes contribuent à donner une image complète. Par exemple, en combinant des données ouvertes provenant de différentes organisations, les municipalités peuvent fournir aux visiteurs toutes les informations culturelles dont ils ont besoin pour planifier leur voyage. En rassemblant des données sur les performances environnementales provenant de différentes sources (telles que la qualité de l'air, les émissions des transports et l'efficacité énergétique), les villes peuvent mesurer leur impact sur l'environnement. Elles peuvent ensuite prendre des mesures ciblées pour améliorer la situation.

Découvrez [une variété de cas d'usage inspirants autour des données ouvertes](#) dans notre article de blog.

4. Quels secteurs peuvent ouvrir leurs données ?

Bien que le mouvement des données ouvertes ait commencé dans le secteur public et ait été initialement stimulé par les lois sur l'ouverture des données des administrations centrales et locales, les organisations de tous les secteurs peuvent tirer profit de l'ouverture de leurs données. Dans les secteurs réglementés et non réglementés, le partage permet d'améliorer l'efficacité, l'engagement avec les clients, les citoyens et les employés, l'innovation grâce au partage de nouvelles idées et à la création de nouveaux services, ainsi que la collaboration entre les différents acteurs.

Quels types de données les différents secteurs partagent-ils et quel est l'objectif de leurs stratégies d'ouverture des données ? Si l'on considère les différents secteurs, voici quelques exemples de la manière dont l'ouverture des données apporte des avantages aux organisations et à leurs parties prenantes :

Municipalités et collectivités locales

Les municipalités sont confrontées à un certain nombre de défis : elles doivent devenir plus efficaces sur le plan opérationnel, améliorer leur durabilité et, surtout, fournir les services et les expériences que leurs citoyens, de plus en plus avertis sur le plan technologique, exigent. Les données ouvertes sont essentielles pour répondre à tous ces besoins. C'est le carburant des projets de villes intelligentes, le partage des données étant au cœur de nouvelles applications telles que :



- ✓ [Les parkings intelligents](#), qui indiquent où se trouvent les places disponibles,
- ✓ La météo géolocalisée et le suivi du trafic en temps réel,
- ✓ Le suivi des performances dans des domaines tels que la prévention de la criminalité et les délais d'intervention en cas d'urgence,
- ✓ La surveillance des niveaux d'eau dans les rivières et les lacs locaux, qui permet d'alerter rapidement en cas d'inondation potentielle.
- ✓ Le partage de la gestion énergétique des bâtiments pour montrer les progrès accomplis par rapport aux objectifs environnementaux et pour maximiser l'efficacité,
- ✓ [La mesure les données relatives à la qualité de l'air](#) afin de fournir des informations clés permettant aux municipalités de planifier des améliorations, par exemple en fermant des routes à la circulation à certaines heures,
- ✓ La facilitation de la recherche des équipements locaux, la surveillance de l'affluence de certaines attractions, ainsi que la génération de visualisations et de cartes mettant en évidence les événements et les lieux culturels, permettant l'attraction et la fidélisation des citoyens et des visiteurs.

Exemples de portails open data de référence dans le secteur des municipalités et collectivités locales :

[Paris Data](#)

[Bordeaux Métropole](#)

[Hérault Data](#)

[Occitanie Open Data](#)

[Data Île-de-France](#)

[NYC \(New York\)](#)

[Bristol](#)

[Métropole Européenne de Lille](#)

[Orléans Métropole Data](#)

[Open Data Pays-Basque](#)

[Data Région Bretagne](#)

[Bruxelles Data](#)

[City of Vancouver](#)

[Eindhoven](#)

Gouvernement central/de l'État et agences fédérales/nationales

La prise de décision du gouvernement doit être basée sur des données complètes et fiables. La collecte et l'analyse de multiples datasets ouverts garantissent que :

- ✓ **Les politiques publiques sont fondées sur des données**, compréhensibles par les citoyens et peuvent être contrôlées pour vérifier si les initiatives ont produit les résultats escomptés.
- ✓ **Les agences gouvernementales renforcent l'engagement** en partageant les données avec d'autres acteurs et le public. Par exemple, les recherches menées par les agences nationales peuvent être utilisées par les entreprises pour planifier leurs activités, tandis que les données démographiques des États fournissent aux chercheurs et aux universitaires les informations dont ils ont besoin pour modéliser l'évolution de la population.
- ✓ **Tout le monde peut prendre des décisions en connaissance de cause.** Pendant la pandémie de COVID, le partage des données sur les taux d'infection par les gouvernements a permis à chacun, des prestataires de soins de santé au public, de disposer des informations nécessaires pour prendre des décisions cruciales.
- ✓ **Les villes, les municipalités et les États attirent des visiteurs**, de nouveaux résidents et des entreprises en partageant des données statistiques. Les informations peuvent même être utilisées par les enfants dans le cadre de projets scolaires.



Exemples de portails open data de référence dans le secteur des administrations centrales/de l'État et agences fédérales/nationales :

[Ministère de l'Éducation Nationale et de la Jeunesse](#)

[Ministère de l'Économie et des Finances](#)

[Ministère de la Transition Écologique](#)

[Open Data URSSAF](#)

[North Carolina Office of State Budget and Management](#)

[Caisse des Dépôts et Consignations](#)

[Ministère de la Culture](#)

[Gouvernement de Nouvelle-Calédonie](#)

[Data Ameli - CNAM](#)

[Ajman Data \(Emirat des Émirats Arabes Unis\)](#)

Les secteurs de l'énergie et des utilities

Les services publics de l'énergie et de l'eau sont confrontés à de multiples [défis autour de la décarbonisation](#), de l'efficacité et de la satisfaction des besoins changeants des clients. La numérisation des opérations et le partage de données ouvertes avec les parties prenantes sont essentiels pour :



- ✓ **Une plus grande efficacité**, par exemple en étant capable de repérer les domaines à améliorer,
- ✓ **Une plus grande ouverture**, par exemple en permettant aux citoyens de suivre facilement des questions telles que la qualité de l'eau,
- ✓ **Une collaboration plus étroite au sein de leurs écosystèmes**. En mettant à disposition des données sur leurs réseaux, leurs niveaux d'utilisation et leurs projets futurs dans des formats visuels faciles à utiliser, tout le monde peut travailler ensemble pour réaliser la transition énergétique. Par exemple, les acteurs des collectivités locales peuvent voir comment la consommation d'énergie de leur ville se compare à celle des autres, ce qui leur permet de cibler les mesures d'efficacité, tandis que les promoteurs immobiliers peuvent voir quel est le meilleur endroit pour construire de nouvelles infrastructures ou de nouveaux logements.

Exemples de portails open data de référence dans les secteurs de l'énergie et des services publics :

[ENEDIS](#)

[GRDF](#)

[UK Power Networks](#)

[Elia](#)

[Northern Powergrid](#)

[SP Energy Networks](#)

[EDF](#)

[Agence ORE](#)

[EDP \(Energias De Portugal\)](#)

[E-Redes](#)

[Electricity North West](#)

[NIE Networks](#)

Organisations du secteur privé

Les organisations du secteur privé ne sont pas concernées par les mêmes réglementations que celles du secteur public en matière de partage des données . Cependant, les entreprises innovantes voient les avantages que l'ouverture des données peut apporter à leur organisation.



- ✓ **Pour communiquer en toute transparence** sur les initiatives de responsabilité sociale des entreprises (RSE), ce qui contribue à attirer et à fidéliser les clients et les employés,
- ✓ **Partager des offres d'emploi** ou la localisation d'agences physiques sur le territoire pour mieux communiquer,
- ✓ **Donner accès à des informations clés dans l'intérêt du public.** Par exemple, de nombreuses entreprises de télécommunications ont partagé des données d'utilisation anonymes pendant la pandémie, ce qui a permis aux gouvernements de mesurer la fréquentation de certaines zones.

Exemples de portails open data de référence dans les secteurs de l'énergie et des services publics :

[Kering Environmental Profit & Loss Portal](#)

[SAEMES Open Data](#)

[KEOLIS Rennes Open Data](#)

[Infrabel](#)

[SBB](#)

[SP Energy Networks](#)

[MAIF Open Data](#)

[Groupe SNCF](#)

[KAPSARC](#)

[STIB](#)

[Swisscom](#)

5. Comment préparer vos données ?

Les données ouvertes doivent être fiables, structurées et facilement compréhensibles par les humains et les machines. Les gens ne les utilisent pas si elles sont incomplètes, de mauvaise qualité, difficiles à comprendre ou si elles manquent de structure.

Pour s'assurer que les données ouvertes répondent aux besoins des utilisateurs, les organisations doivent préparer leurs données en suivant plusieurs étapes.

La nécessité de disposer de données de qualité

La qualité des données est mesurée en fonction de critères tels que l'exactitude, l'exhaustivité, l'actualité, la cohérence et la fiabilité. Pour garantir la qualité des données, les organisations doivent :



- ✓ Mettre en place une structure de gouvernance des données, contenant des processus pour vérifier et, si nécessaire, améliorer la qualité des sources de données ouvertes,
- ✓ Rendre les employés, tels que les gestionnaires de données, responsables de la qualité de certains jeux de données. Ils doivent savoir d'où viennent les données (la source de données), à quelle fréquence elles sont mises à jour, combien de temps elles doivent être conservées et par quels systèmes elles sont passées ou ont été enrichies,
- ✓ Déployer une plateforme centrale pour garantir la qualité des données et éviter la duplication des efforts.

Le processus d'enrichissement des données

Pour accroître la valeur des données, il est possible de les enrichir avec d'autres informations, provenant de l'intérieur ou de l'extérieur de l'organisation. Ces informations ajoutent un contexte et rendent les datasets plus utiles lorsqu'ils sont partagés.

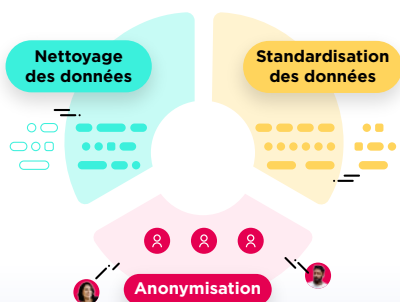


Les types de jeux de données utilisés pour enrichir les données sont les suivants :

- ✓ Les données de géolocalisation, telles que les limites administratives, les adresses, les codes postaux,
- ✓ Les données météorologiques,
- ✓ Les informations démographiques,
- ✓ Les informations économiques, telles que les données relatives à l'emploi ou à la santé.

Le [Data hub d'Opendatasoft](#) propose plus de 30 000 jeux de données librement disponibles qui peuvent être utilisés pour enrichir les données ouvertes.

Le processus de gestion des données (data management)



La gestion des données doit être un processus clair de préparation des données ouvertes à la publication. Les organisations doivent établir une méthodologie de bout en bout qui couvre toutes les étapes de la gestion des données, de la préparation à l'utilisation finale, y compris l'intégration des données pour leurs initiatives en matière de données ouvertes.

Le processus doit comprendre :

- ✓ [Le nettoyage des données](#) (pour supprimer les champs inexacts ou incomplets),
- ✓ [La normalisation des données](#) (pour s'assurer que les données, telles que les dates ou les adresses, sont présentées dans le même format),
- ✓ L'anonymisation pour supprimer toute information personnellement identifiable avant qu'elle ne soit partagée.

Il existe des outils permettant d'automatiser un grand nombre de ces processus afin de gagner du temps tout en garantissant la cohérence.

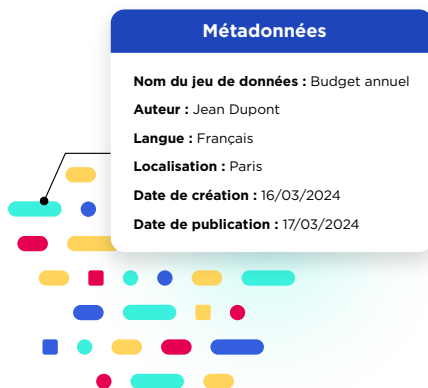
L'importance des métadonnées

Les données sans contexte ni structure n'ont pas de sens. [Les métadonnées](#) contribuent à fournir ce contexte. Les métadonnées sont essentiellement des données qui décrivent d'autres données – elles visent à créer un contexte pour vos données en répondant aux questions de base : qui, quoi, où, quand, comment et pourquoi ?

Les métadonnées permettent aux humains et aux outils technologiques de comprendre facilement ce que vos données couvrent et le format dans

lequel elles se trouvent en fournissant un résumé de leur contenu. **Une bonne gestion des métadonnées est essentielle au succès des données ouvertes.** Les organisations doivent tout d'abord établir un système de classification de leurs données, en créant un vocabulaire commun qui décrit les données. Ce vocabulaire doit être clair et compréhensible par tous, quelles que soient leurs connaissances techniques.

Pour ce faire, il existe un certain nombre de modèles normalisés, tels que la norme ISO, la norme de métadonnées Data Catalog Vocabulary (DCAT-AP) ou la Dublin Core Metadata Initiative (DCIM). Par exemple, la norme ISO 8601 fournit des formats de date universels. Les normes de métadonnées doivent ensuite être partagées et appliquées dans l'ensemble de l'organisation.



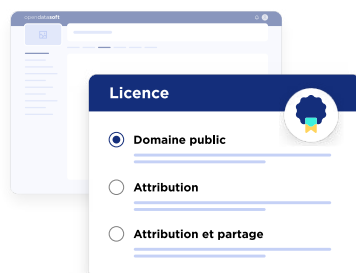
6. Comment publier vos données en open data ?

Une fois que les organisations ont préparé leurs données, elles doivent les publier et les mettre gratuitement à la disposition du monde entier. Pour ce faire, il faut se concentrer sur trois domaines clés :

Choisir la bonne licence

Les données ouvertes peuvent faire l'objet de trois niveaux de licences :

- ✓ **Domaine public.** L'organisation renonce totalement à tous les droits sur les données particulières qu'elle publie. Les gens peuvent réutiliser les données sans avoir besoin de dire d'où elles proviennent, et ils peuvent créer leurs propres applications ou services sur la base de ces données.
- ✓ **Attribution.** L'organisation renonce toujours à tous les droits et les personnes peuvent réutiliser les données de n'importe quelle manière. La seule obligation est d'indiquer que l'organisation qui publie les données en est la source.
- ✓ **Attribution et partage.** Outre l'attribution de l'organisme de publication en tant que source, l'utilisateur des données doit partager tout contenu ou donnée dérivés sous la même licence.



Les organisations peuvent créer leurs propres licences pour leurs données ouvertes. Toutefois, il est plus facile d'utiliser [une licence légale](#) existante et éprouvée. **Il s'agit notamment des licences Creative Commons et de celles de l'Open Data Commons.** Il peut également y avoir des licences spécifiques à un pays ou à un secteur qui peuvent être utilisées – par exemple, le gouvernement britannique exige que les ministères et les organismes gouvernementaux utilisent la licence [Open Government](#). Toutes ces licences offrent les trois mêmes niveaux de protection juridique (domaine public, attribution et attribution/partage).

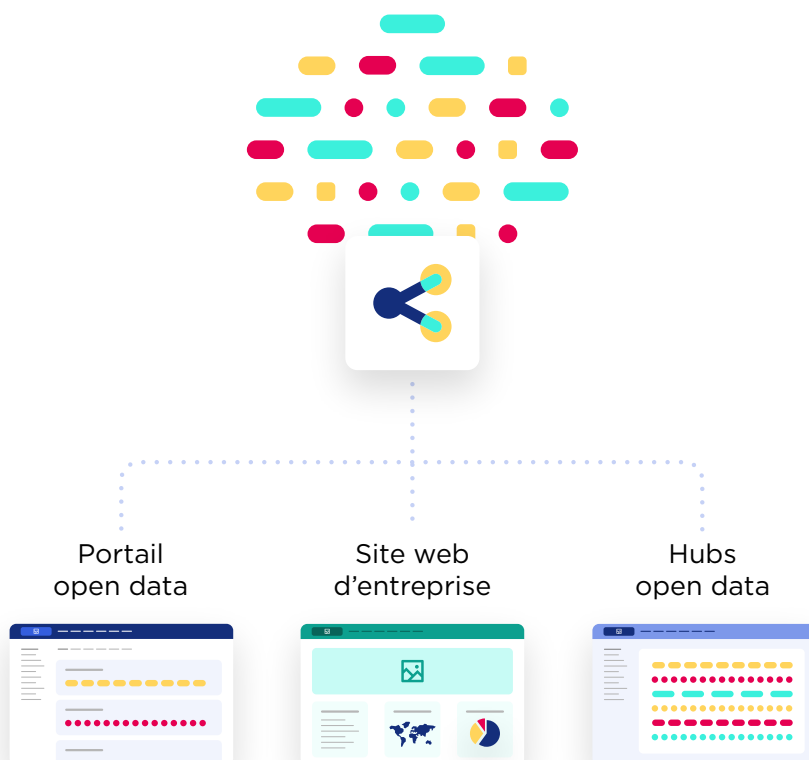
Créer des visualisations de données convaincantes

Alors que les spécialistes des données (comme les data analystes) effectuent des analyses de données ouvertes sur des jeux de données brutes stockées dans une base de données ou une feuille de calcul, les citoyens, les consommateurs et autres non-spécialistes sont susceptibles de trouver plus facile de comprendre et d'agir sur les données si elles sont sous une forme plus visuelle. **Les organisations doivent donc créer des visualisations de données attrayantes à partir de leurs datasets.** Il peut s'agir de cartes ou de graphiques zoomables, ou encore d'histoires de données, qui combinent textes et images pour en expliquer le sens.



Choisir les bons canaux de partage

L'accessibilité est l'un des trois critères clés des données ouvertes. Cela signifie que les données doivent être librement accessibles en ligne. **Cela peut se faire par le biais d'un portail de données ouvertes spécifique, qui collecte et partage toutes les données ouvertes d'une organisation, et/ou par le biais du site web de l'entreprise.** Par exemple, certaines organisations peuvent republier des données ouvertes sur des pages spécifiques de leur site web principal, notamment en ce qui concerne la RSE ou d'autres indicateurs de performance. Pour maximiser le partage, les organisations devraient également s'assurer que leurs données sont disponibles via des hubs de données ouvertes, avec des fonctions de recherche pour faciliter la découverte et la réutilisation.



7. Comment réutiliser des données publiées en open data ?

L'objectif de la publication de données ouvertes est d'aider et d'informer les principales parties prenantes, y compris les citoyens et les consommateurs. Par conséquent, une fois que vous avez assuré la qualité des données et que vous les avez publiées au bon endroit, il est essentiel de faciliter au maximum leur réutilisation. Cela repose sur deux facteurs :

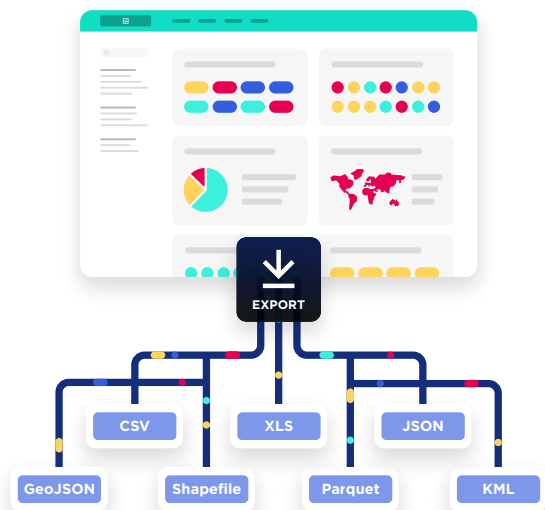
La durabilité

Les données sont-elles fréquemment mises à jour ? Il est clair que la publication de données ouvertes a un coût pour l'organisation, puisqu'elle doit collecter, nettoyer et partager l'information, sans que cela ne coûte rien à l'utilisateur.

Les organisations doivent donc mettre en place des budgets sûrs et stables pour s'assurer que les ensembles de données ouvertes sont toujours à jour. Le fait de disposer d'un historique continu d'un jeu de données particulier (remontant par exemple à plusieurs années) le rend beaucoup plus utile, car il permet de tirer des conclusions essentielles sur les changements intervenus au fil du temps. Des outils sont disponibles pour automatiser les mises à jour afin de maximiser la valeur des données ouvertes, tout en minimisant le temps et les ressources en matière de données ouvertes.



API et options de téléchargement



Bien que tous les types de données puissent être partagés en tant que données ouvertes, elles ont plus de chances d'être réutilisées si elles sont dans des formats qui peuvent être facilement compris et accédés par d'autres. La possibilité d'enregistrer des données dans des formats ouverts, tels qu'un fichier texte ou un fichier de feuille de calcul CSV, signifie qu'elles sont automatiquement accessibles et réutilisables par tous.

Outre la mise à disposition des données par le biais de visualisations et de téléchargements, de nombreux utilisateurs (en particulier les analystes de données) voudront pouvoir automatiser l'accès aux données, par exemple par le biais d'interfaces de programmation (API).

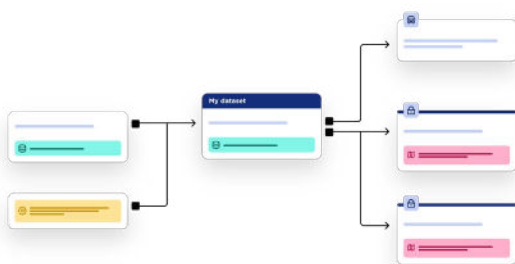
Cela permet aux utilisateurs de lier les jeux de données à leurs propres systèmes et de s'assurer que les mises à jour sont automatiquement intégrées dans leur solution. Cela facilite également

le téléchargement de données en masse pour une analyse détaillée. Proposer des API permet également aux outils d'IA et d'apprentissage automatique d'accéder automatiquement aux données et de les télécharger, sans nécessiter d'intervention humaine.



Comment mesurer la réutilisation des données ouvertes ?

Comprendre comment les données ouvertes sont utilisées – et quels jeux de données sont les plus populaires – est essentiel pour réussir. Pour ce faire, les organisations doivent savoir quelles sont les données que les utilisateurs trouvent les plus utiles, afin de pouvoir les classer par ordre de priorité. [Parler aux utilisateurs et recueillir leurs commentaires](#) est un bon moyen de commencer à établir des priorités.



Il est également possible de demander aux utilisateurs de s'enregistrer gratuitement pour avoir accès à davantage de méthodes de téléchargement ou de manipulation des données. Cela permet de savoir quels sont les ensembles de données qui offrent la plus grande valeur. Toutefois, cela peut décourager les utilisateurs occasionnels qui souhaitent uniquement accéder à des informations de base.

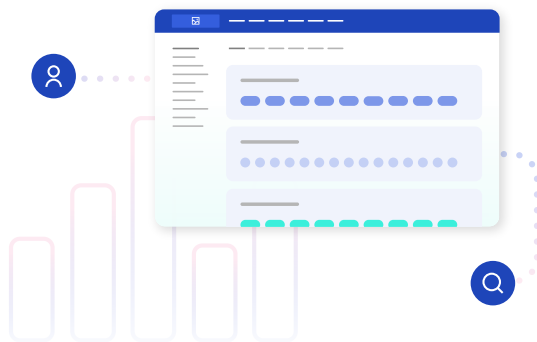
La plupart des plateformes de données ouvertes fournissent désormais des statistiques complètes et des [fonctionnalités de lignage](#) sur le nombre de fois qu'un jeu de données spécifique a été consulté dans un référentiel de données ouvertes et/ou téléchargé par le biais d'API et d'autres méthodes. Cela permet d'évaluer la popularité, même si cela n'aide pas à comprendre exactement ce pour quoi il est utilisé. Le suivi des utilisateurs peut aider à combler cette lacune et à expliquer pourquoi certains jeux de données ouverts sont utilisés, et par qui. Les ressources peuvent alors être concentrées sur des données similaires qui pourraient également être utiles à l'avenir.

8. Gouvernance, outils et culture : les trois leviers pour partager vos données

L'ouverture des données peut sembler simple. Cependant, les données sont souvent stockées dans plusieurs systèmes au sein d'une organisation, peuvent être décrites en utilisant un langage et un vocabulaire différents et être contrôlées par différents départements. **Cela signifie que la collecte et le partage de données, tout en s'assurant de leur fiabilité, de leur qualité et de leur actualité, deviennent rapidement une initiative majeure, qui peut nécessiter des ressources importantes.** Pour réussir, il faut se concentrer sur trois domaines clés :

Élaborer une stratégie de gouvernance des données pour normaliser l'utilisation au sein de l'organisation

La gouvernance des données couvre la manière dont une organisation traite et utilise les données qu'elle collecte. Elle va au-delà de l'ouverture des données, mais compte tenu de l'importance des données pour chaque organisation, une gouvernance solide est vitale. Cependant, ce n'est pas simple, étant donné le nombre d'ensembles de données générés par une organisation, les différents formats dans lesquels ils se trouvent et les centaines d'endroits où ils sont potentiellement stockés. Il est essentiel de gérer toutes ces données de manière à ce qu'elles soient facilement accessibles, de haute qualité et pertinentes.



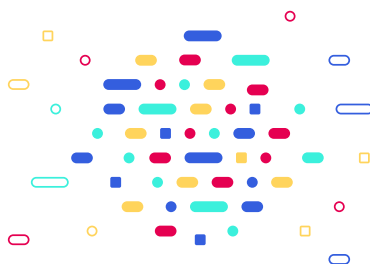
La gouvernance des données permet de s'assurer que les données sont :

- ✓ **Cohérentes** (elles utilisent les mêmes formats pour les champs de données (tels que les noms des clients) dans l'ensemble de l'organisation),
- ✓ **Dignes de confiance** (elles sont exactes et à jour),
- ✓ **Sécurisées** (elles ne peuvent pas être consultées par des personnes mal intentionnées),
- ✓ **Utilisées correctement** (elles ne sont pas utilisées à mauvais escient par le personnel ou les parties prenantes externes),
- ✓ **Conformes** (elles sont conformes aux réglementations et respectent la vie privée des personnes).

La gouvernance des données consiste également à mettre en place l'organisation, [la structure](#) et les processus humains adéquats pour évoluer vers un modèle axé sur les données. Il s'agit notamment de nommer des gestionnaires de données responsables de chaque ensemble de données, de mettre en place des processus clairs que tout le monde suit et d'assurer un suivi suffisant pour que l'ensemble du personnel respecte le cadre de gouvernance. Consultez [notre blog pour une introduction à la gouvernance des données](#).

Créer et diffuser une culture des données pour impliquer toutes les parties prenantes

Les organisations ont l'habitude de travailler en silos et de conserver leurs données au sein des équipes. Par conséquent, de nombreux services, voire des organisations entières, peuvent être réticents à prendre des risques et craindre que les données qu'ils partagent soient mal utilisées ou mal interprétées. Ils peuvent craindre de s'exposer à des critiques potentielles dans le cas des organismes publics. Toutefois, **les avantages l'emportent sur les risques et, dans de nombreux cas, l'ouverture est désormais imposée par la réglementation**. Pour changer cette approche, il faut changer la culture des organisations. Pour que les programmes d'ouverture des données soient couronnés de succès, il faut donc que tous les membres de l'organisation comprennent la valeur et l'importance des données et s'engagent à les partager.



Cela signifie que les organisations doivent créer une culture des données qui s'étend à tous les départements et bureaux. Cela nécessite une plus grande collaboration, à la fois autour de la gouvernance des données, mais aussi pour aider à développer les compétences en matière de données dans l'ensemble de l'organisation et faciliter autant que possible la compréhension et le partage des données. **Des programmes de formation, un enseignement et une communication forte, menés par la direction, doivent être mis en place pour créer des défenseurs des données dans l'ensemble de l'organisation.**

Choisir les bons outils pour garantir l'utilisation des données

La technologie est la troisième clé du lancement d'un projet de données ouvertes. Bien que de nombreuses organisations aient dépensé de grosses sommes d'argent pour leur infrastructure de données, les informations sont encore souvent enfermées dans des silos et ne sont accessibles qu'aux experts data. Pour que les projets de données ouvertes réussissent, il faut que la bonne technologie



soit en place pour collecter, organiser et partager les données. Il doit être simple de connecter toutes les sources de données de l'organisation, d'enrichir et de nettoyer les données, puis de les partager de manière convaincante et dans tous les formats.

Il faut pour cela une solution spécialisée dans la publication et l'exposition de la donnée et qui s'intègre dans la stratégie de gouvernance des données de l'organisation et la soutienne. **Les données ouvertes doivent être disponibles sous différentes formes, qu'il s'agisse de visualisations telles que des cartes, d'histoires de données ou de téléchargements d'APIs.** Les outils de recherche de données doivent aider les visiteurs à trouver facilement des ensembles de données pertinents à utiliser. Enfin, elles doivent automatiser la gestion et la mise à jour des datasets, en supprimant les frais généraux et en rendant l'administration moins gourmande en ressources.

9. Comment trouver des données ouvertes

La possibilité de combiner des ensembles de données provenant de différentes sources de données ouvertes apporte une valeur ajoutée et favorise l'innovation. Le regroupement de jeux de données ouvertes liées entre elles peut apporter de nouvelles informations et permettre aux organisations de visualiser leurs propres données d'une nouvelle manière.

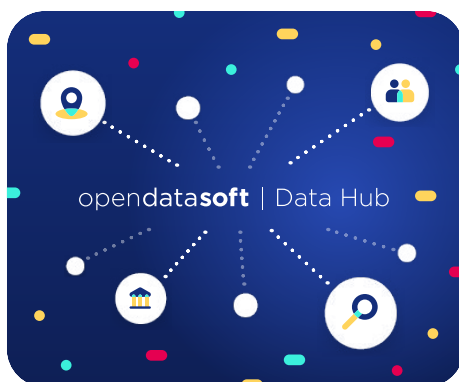
C'est pourquoi un nombre croissant de hubs de données ouvertes ont été créés pour rassembler les catalogues de données de différentes organisations. Le fait d'appartenir à des hubs de données bien connus augmente également les chances que les données ouvertes soient trouvées et réutilisées, ce qui contribue à en accroître l'utilisation.

De nombreux centres de données couvrent des pays ou des secteurs spécifiques – par exemple, la majorité des pays et des régions disposent de leurs propres centres de données publiques :

- ✓ **États-Unis** – le portail [Data.Gov](https://data.gov) contient désormais plus de 250 000 datasets du secteur public.
- ✓ **Europe** – data.europa.eu contient 1 607 000 jeux de données provenant de 36 pays.
- ✓ **Royaume-Uni** – le portail data.gov.uk rassemble des données publiées par le gouvernement central, les autorités locales et les organismes publics du pays.

Le Data hub Opendatasoft

Pour aider les organisations et les individus à maximiser la valeur des données ouvertes, **Opendatasoft a créé son propre portail de données. Le [Data hub Opendatasoft](#) contient plus de 29 000 jeux de données ouvertes provenant d'organisations privées et publiques du monde entier.** Il est conçu pour simplifier au maximum l'accès, la compréhension, la combinaison et l'utilisation des données ouvertes. Nous mettons également à disposition une sélection de 600 jeux de données de référence créés par des producteurs de données officiels, des entreprises de données ouvertes et des autorités, tels que les données ouvertes de l'USAC et les données ouvertes de la Banque mondiale. Ils sont organisés en différents ensembles thématiques, tels que des ensembles de données géographiques, démographiques, économiques, relatives au logement ou au travail. Les exemples incluent une base de données complète des codes postaux américains et des données météorologiques mondiales remontant à 2015.



Le [Data Hub d'Opendatasoft](#) fournit également une gamme complète d'outils pour faciliter la réutilisation, la data visualisation et les téléchargements en masse. Les utilisateurs peuvent simplement cliquer sur un ou plusieurs jeux de données, les combiner et créer des cartes ou des graphiques. La possibilité de rassembler ainsi divers datasets, enrichis de données de référence, permet à tous d'utiliser les données de manière innovante.

10. Maturité des données ouvertes et exemples de bonnes pratiques

Open data maturity

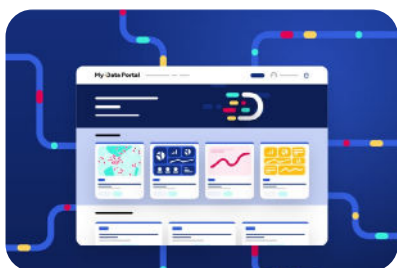
Le partage des données ouvertes est un processus d'amélioration continue. Les organisations comprennent qu'il y a toujours plus de données qu'elles peuvent partager, de nouvelles façons de les rendre disponibles afin de mieux s'engager avec les parties prenantes, et des méthodes innovantes pour les combiner avec d'autres jeux de données afin d'en augmenter la valeur.



C'est pourquoi des études mesurent la maturité des pays et des organisations en matière de données ouvertes. Cela permet de comparer les leaders avec ceux qui sont plus avancés dans leur parcours, et permet à chacun de s'inspirer des meilleures pratiques et de suivre [les dernières tendances](#).

L'étude internationale la plus importante est le rapport sur la maturité des données ouvertes de la Commission européenne. Cette étude annuelle évalue la manière dont 35 pays européens abordent les données ouvertes, en mesurant les progrès réalisés par rapport à quatre critères clés (politique nationale, impact des données ouvertes dans le pays, force de son portail de données ouvertes et qualité des données/métadonnées qu'il contient). [Pour en savoir plus sur l'édition 2022 du rapport, consultez ce blog détaillé.](#)

De même, une série d'organisations à but non lucratif ont développé des outils et des cadres pour aider à accroître la maturité du partage des données ouvertes. Par exemple, l'[Open Data Institute](#) a créé son modèle de maturité des données ouvertes, qui aide les organisations à évaluer dans quelle mesure elles publient et consomment des données ouvertes, et identifie les mesures à prendre pour les améliorer. Ce modèle a été appliqué par des organisations du monde entier. En outre, de nombreux États et organismes gouvernementaux ont créé des modèles de maturité qui peuvent être utilisés par leurs agences affiliées, comme le [DC Data Maturity Model](#) (modèle de maturité des données du district), créé par le district de Columbia.



Contactez nos experts

Créer la data marketplace publique qui correspond à votre besoin

Renforcez vos relations avec les institutions publiques, vos parties prenantes et vos clients en communiquant de manière transparente sur vos performances, vos objectifs et vos offres à l'aide d'une marketplace publique.

opendatasoft

Opendatasoft est le leader de la démocratisation des données. L'entreprise propose une solution de portail de données pour fournir un accès unique en libre service aux actifs data des organisations. Opendatasoft permet aux organisations de mettre en place des expériences data personnalisées à l'échelle et en un temps record.

Opendatasoft est la solution essentielle de data management pour améliorer l'efficacité et réduire les coûts, augmenter et créer de nouvelles sources de revenus, atténuer les risques et gérer les crises.

Opendatasoft sert plus de 400 clients dans 25 pays, alimentant plus de 3 000 portails de données. Grâce à cette expérience, Opendatasoft a développé une expertise unique dans la gestion des données qui lui permet de proposer à ses clients des services premium et réaliser des cas d'usage qui répondent à leurs besoins spécifiques.

En savoir plus

www.opendatasoft.com