

## Lignage des données

### Notice d'information

*Dernière mise à jour : 2024-03-01*

Opendatasoft permet à toutes les équipes utilisatrices de ses services de créer et partager rapidement dans leurs écosystèmes de nouvelles expériences data plus accessibles, plus pertinentes et facilement réutilisables. En démocratisant l'accès à la donnée, vous optimisez vos opérations, accélérez le développement de nouvelles activités, et nourrissez des relations de confiance avec vos parties prenantes.

Dans ce cadre, et pour assurer les objectifs ci-avant, Opendatasoft offre à ses Clients dans le cadre de la fourniture de son Service, une fonctionnalité de Lignage de données permettant de comprendre la manière dont les jeux de données d'un Espace de travail sont utilisés dans l'écosystème Opendatasoft, à la fois au sein de l'Espace de travail concerné mais également vers les autres Espaces de travail, et d'encourager les échanges entre Clients producteurs et utilisateurs de la Plateforme Opendatasoft afin que chaque partie soutienne le rayonnement du réseau de données.

La présente Notice d'information destinée aux sociétés ou entités clientes d'Opendatasoft (ci-après "Client") a pour objectifs de présenter l'intérêt et le fonctionnement du Lignage de données et de décrire les bonnes pratiques de parties intervenantes dans le cadre de l'utilisation de cette fonctionnalité. Elle a un caractère informatif et ne modifie pas le socle contractuel des Services.

Opendatasoft tient également à informer ses Clients que la fonctionnalité de Lignage de données n'implique pas de traitements portant sur les données personnelles au sens du Règlement Général sur la Protection des Données "RGPD". En effet, cette fonctionnalité utilise seulement les données relatives à des personnes morales et des métadonnées qui ne sont pas susceptibles d'identifier une personne physique. Par ailleurs, dans le cadre de la fonctionnalité de Lignage de données, aucune Information confidentielle n'est utilisée.

La présente Notice est accessible sur la page <https://legal.opendatasoft.com/> et pourra faire l'objet de mises à jour dont la version applicable sera accessible sur cette même page.

## 1. Définitions

Tous les termes en majuscules qui ne sont pas définis dans la présente Notice auront la signification qui leur est donnée dans les Conditions Générales des Services et Conditions d'Utilisation d'Opendatasoft.

- **Lignage de données** : désigne une visualisation complète des flux de données en modélisant les dépendances entre les différents Objets, ainsi que la manière et la raison de son évolution au fil du temps ; ce Lignage est documenté en répertoriant l'origine et la destination finale d'une information ainsi que

toutes les transformations qu'elle a subies à chaque étape de son parcours.

**Métadonnées d'usage** : désignent les métadonnées de Lignage disponibles à l'échelle d'un jeu de données ou d'une page et modélisées sous la forme d'un graphe. Cette structure adaptée est une vue d'ensemble des relations directes et indirectes, en amont et en aval des Objets impliqués dans la construction ou l'utilisation d'un jeu de données. Les Métadonnées d'usage sont :

- la dénomination et la nature de l'Objet ODS (nom de l'Espace de travail, titre, identifiant et champs liés de l'Objet ODS)
- la dénomination et la nature de l'Objet tiers (titre et type de source)
- la relation, c'est-à-dire le type d'utilisation.
- **Objet ODS** : désigne un jeu de données, une page ou un éditeur de cartes ou de graphes.
- **Objet tiers** : désigne une source externe (fichier, URL ou service distant) d'un jeu de données.
- **Objet** : désigne un Objet ODS ou un Objet tiers.
- **Relation** : désigne une dépendance entre deux Objets qui établit un **lien orienté** entre un Objet **origine** et un Objet ODS **destination**.

## 2. Principes du Lignage de données

Le Lignage des données d'Opendatasoft propose des outils pour comprendre les dépendances entre les Objets et mettre en valeur le réseau de données.

Sur l'exemple ci-dessous, A est l'Objet origine et B l'objet destination. B dépend de A.



Il existe deux types de Relations :

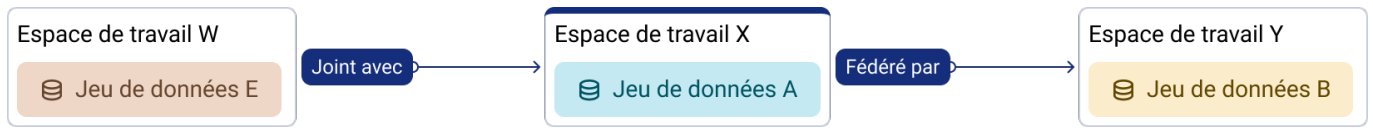
**directe** : quand deux Objets sont liés sans aucun intermédiaire. L'un est identifié comme **origine** et l'autre comme **destination**.

Sur l'exemple ci-dessous, le jeu de données A est l'Objet ODS origine source et le jeu de données B est l'Objet ODS destination.



**indirecte** : quand un Objet ODS **destination** dépend d'une autre relation directe ou indirecte avec d'autres Objets.

Sur l'exemple ci-dessous, le jeu de données E est l'Objet ODS origine du jeu de données A qui est lui-même l'Objet ODS origine du jeu de données B. Les jeux de données E et B ont une relation indirecte.



### 3. Partage des données

#### 3.1. Partage entre le Client et Opendatasoft

Opendatasoft a intégralement développé en interne l'ensemble des fonctionnalités relatives au Lignage des données. Cette fonctionnalité est maintenue et hébergée au même titre que la Plateforme Opendatasoft, et en application des conditions contractuelles d'Opendatasoft.

Afin d'offrir une fonctionnalité de Lignage de données à haute valeur ajoutée, la visualisation des Relations entre les Objets implique de partager les Métadonnées d'usage avec Opendatasoft et les autres Clients possédant des Espaces de travail Opendatasoft.

Dans ce cadre, et selon les conditions décrites dans la présente Notice, Opendatasoft réunit toutes les Métadonnées d'usage collectées pour le Lignage de données dans le but d'améliorer la qualité de la Plateforme Opendatasoft et de mettre en valeur le réseau de données en résultant et proposer des fonctionnalités pertinentes pour la gestion et la diffusion des catalogues au profit des utilisateurs.

Le contenu des Données Client n'est ni modifié ni lu par la fonctionnalité. La solution de Lignage des données scanne la Plateforme Opendatasoft pour analyser la configuration des jeux de données et leurs tâches de traitement ou le code physique des pages/éditeurs pour extraire uniquement les Métadonnées d'usage nécessaires. La fonctionnalité ne transforme pas les Données Client pendant ses analyses.

Les Métadonnées d'usage entre Objets sont extraites du système et mises à jour après la publication. Des données agrégées anonymes peuvent par ailleurs être utilisées par Opendatasoft aux fins d'évaluation, d'amélioration et de maintenance de la fonctionnalité, à des fins statistiques et pour la promotion du partage de données.

#### 3.2. Partage entre Clients

Il est important pour Opendatasoft de garantir à chaque Client une information minimum exploitable (*quoi et comment*) sur l'usage de ses données et de permettre de choisir ou non de partager la dénomination de leurs Espaces de travail. Des modes de partage sont disponibles pour empêcher l'identification de l'Objet ODS destination aux producteurs des Objets ODS origines.

Les Métadonnées d'usage sont réparties en quatre niveaux de restitution :

Par défaut et pour chaque relation, son type et une information sur la nature de l'Objet ODS destination sont transmis au Client producteur de données.

La dénomination de l'Espace de travail utilisateur est soumise à un mode de partage choisi par le Client (mode déclaré/mode incognito).

La dénomination des Objets ODS est soumise aux conditions d'accès définies sur le portail propriétaire.

La dénomination et le type de source des Objets tiers sont soumis aux conditions d'accès définies sur le portail propriétaire et ne sont pas transmises en dehors de celui-ci.

**Opendatasoft recommande aux Clients d'être en mode « déclaré » pour augmenter la qualité des Métadonnées d'usage en faveur des producteurs de données. Connaître l'usage de ses jeux de données permet de créer une dynamique communautaire entre plusieurs acteurs.**

### **3.2.1 Application du mode de partage de la dénomination de l'Espace de travail**

Ce choix se fait à l'échelle d'un Espace de travail et s'appliquera aux autres Espaces de travail qui consomment directement ou indirectement ses données. Son activation sera faite par les équipes d'Opendatasoft via nos outils internes.

### **3.2.2 Changement de mode de partage de la dénomination de l'Espace de travail**

Pour effectuer un changement entre les modes « déclaré » et « incognito », la demande devra être formulée auprès du service client Opendatasoft qui effectuera les démarches opérationnelles nécessaires et adaptera les conditions tarifaires applicables. Après validation par Opendatasoft, le changement de mode de partage sera effectif et les données existantes seront mises à jour sur les différents supports.

### **3.2.3 Utilisation des Métadonnées d'usage**

La fonctionnalité de Lignage des données et les Métadonnées d'usage est destinée aux seuls Clients et leur utilisation est limitée à des usages strictement internes. Tout usage à des fins commerciales des données partagées est fait aux risques exclusifs de l'organisme utilisateur.