Mettre à disposition des données ouvertes avec le Kit de démarrage OGDI DataLab

Solution Open Data Clé en main

Publication : Juin 2013

Version: 6.0

Auteurs : Philippe Beraud (Microsoft France), Sébastien Brasseur (Microsoft France), Rémi Olivier (Microsoft France), Frédéric Romanczuk (Microsoft France), Alexandre Fournier (Microsoft France)

Pour les dernières informations, consultez

aka.ms/OpenDataFrance

Sommaire

[Licence du Kit de démarrage OGDI DataLab 3](#_Toc358205225)

[Résumé 4](#_Toc358205226)

[Commentaires 4](#_Toc358205227)

[Stockage 5](#_Toc358205228)

[Prérequis 5](#_Toc358205229)

[Stratégies de téléchargement montant des données 6](#_Toc358205230)

[Téléchargement des données au format texte 8](#_Toc358205231)

[Préparation des données pour le téléchargement au format CSV 8](#_Toc358205232)

[Données de géolocalisation 9](#_Toc358205233)

[Segmentation des données 10](#_Toc358205234)

[Préparation des données pour le téléchargement au format KML 10](#_Toc358205235)

[Importation d’un ensemble de données sous forme de points 10](#_Toc358205236)

[Importation d’un ensemble de données sous forme de polygones 11](#_Toc358205237)

[Segmentation des données 12](#_Toc358205238)

[Envoi des données 13](#_Toc358205239)

[Téléchargement des données vers votre instance de stockage 13](#_Toc358205240)

[Installation de l'outil de téléchargement 13](#_Toc358205241)

[Configuration de l'application de téléchargement des données 13](#_Toc358205242)

[Préparation du téléchargement montant 14](#_Toc358205243)

[Téléchargement des données 18](#_Toc358205244)

[Téléchargement des données Blob 19](#_Toc358205245)

[Exemple de données 19](#_Toc358205246)

[Téléchargement de données binaires 19](#_Toc358205247)

[Préparation des données 23](#_Toc358205248)

# Licence du Kit de démarrage OGDI DataLab



Le Kit de démarrage **OGDI** (Open Government Data Initiative) **DataLab** est publié sous le contrat de licence de logiciel libre Microsoft Public License (Ms-PL) tel que décrit à l’adresse Internet <http://opensource.org/licenses/ms-pl.html>.

# Résumé

Le kit de démarrage **OGDI** (Open Government Data Initiative) **DataLab** sous license libre est une initiative menée par les équipes Microsoft en charge du [Secteur Public](http://www.microsoft.com/france/entreprises/secteur-public/solutions-pour-secteur-public.aspx)[[1]](#footnote-1). OGDI DataLab utilise la [plateforme Windows Azure](http://www.microsoft.com/france/windows-azure/)[[2]](#footnote-2) dans le Cloud pour faciliter la publication et l'utilisation d’informations publiques ou ouvertes (*open data*) variées provenant d'entités publiques (gouvernement, ministères, collectivités locales et territoriales, mairies, services déconcentrés de l'Etat, etc.).

Ce document vous guide tout au long du processus de chargement des données pour OGDI France (<http://ogdifrance.cloudapp.net>). Le site OGDI France est fondé sur le kit de démarrage OGDI DataLab, en offrant la possibilité de publier des données ouvertes pour une réutilisation étendue en France.

Dans ce guide, nous allons aborder les différentes stratégies de mise en ligne par téléchargement (montant), et vous aider dans le processus de téléchargement. Ce document décrit également la méthode permettant d’ajouter des données binaires liées aux enregistrements (essentiellement des fichiers tels que les images).

Toutes ces procédures sont aisément transposables à une toute solution personnalisée fondée sur le kit de démarrage OGDI DataLab.

# Commentaires

Votre opinion nous intéresse. N’hésitez pas à nous transmettre vos commentaires via la boîte aux lettres Open Data France (<mailto:ogdifrance@live.fr>) afin que nous puissions améliorer le kit de démarrage OGDI DataLab, l’un des accélérateurs de la [solution Open Data Clé en main](https://mspartner.microsoft.com/fr/fr/Pages/Solutions/Open-Data-Cle-en-main.aspx)[[3]](#footnote-3) disponible sur le site Microsoft Partner Network.

Le [blog MSDN Open Data France](http://blogs.msdn.com/b/ogdifrance/)[[4]](#footnote-4) vous tient informé dans le même temps des dernières évolutions de la solution Open Data Clé en main, des différents accélérateurs proposés dans ce contexte, de l’actualité et des nouveautés relatives à ceux-ci

# Stockage

Dans ce guide, nous abordons le thème du stockage dans Windows Azure. Nous partons du principe que vous disposez d’un compte de stockage Windows Azure configuré pour une utilisation avec OGDI DataLab, si vous souhaitez fournir les éléments suivants :

* Jeux de données de grande taille au format texte (ou besoin d'avoir un contrôle accru sur les données) ;
* Données binaires.

Dans ces cas, vous avez un accès complet à cette ressource.

Remarque importante : Une telle ressource requiert un compte et une souscription active avec Windows Azure. Vous pouvez accéder à Windows Azure avec un paiement à l’usage, sans engagement, ou bien des forfaits comme décrit à l’adresse <http://www.microsoft.com/france/windows-azure/Offres.aspx>.

# Prérequis

Les procédures décrites dans la suite de ce document s’appuient sur les utilitaires de chargement de données du Kit de démarrage OGDI DataLab.

Elles supposent que ces derniers ont été compilés conformément à ce qui est décrit dans le document Démarrer avec le Kit de démarrage OGDI DataLab. Vous pouvez également télécharger ces mêmes utilitaires (en version anglaise) sur le site GovData.eu : (<http://frogdidata.blob.core.windows.net/public/OGDI_DataLoader.zip>).

* La solution Cloudberry explorer for Azure Blob Storage de téléchargement des fichiers binaires. Elle est disponible gratuitement à l'adresse <http://www.cloudberrylab.com/?page=explorer-azure>.

Remarque : Vous pouvez utiliser le programme que vous souhaitez pour effectuer le téléchargement des fichiers vers votre compte de stockage Windows Azure. Dans ce guide, nous utilisons cet outil spécifique à des fins d'illustration.

Nous partons du principe que vous disposez déjà d'un compte de stockage Windows Azure, configuré pour une utilisation avec OGDI DataLab. Pour télécharger les données, vous avez besoin des informations d'identification suivantes :

* Intitulé du compte de stockage ;
* Clé d'accès au stockage.

# Stratégies de téléchargement montant des données

Actuellement, nous proposons deux stratégies de téléchargement pour OGDI France (<http://ogdifrance.cloudapp.net>). La méthode la plus simple pour fournir des données au format texte consiste simplement à nous envoyer vos fichiers. Cette stratégie est parfaitement adaptée aux jeux de données au format texte de petite taille, lorsque la fréquence de mise à jour n'est pas importante (par exemple, plus d'une fois par an).

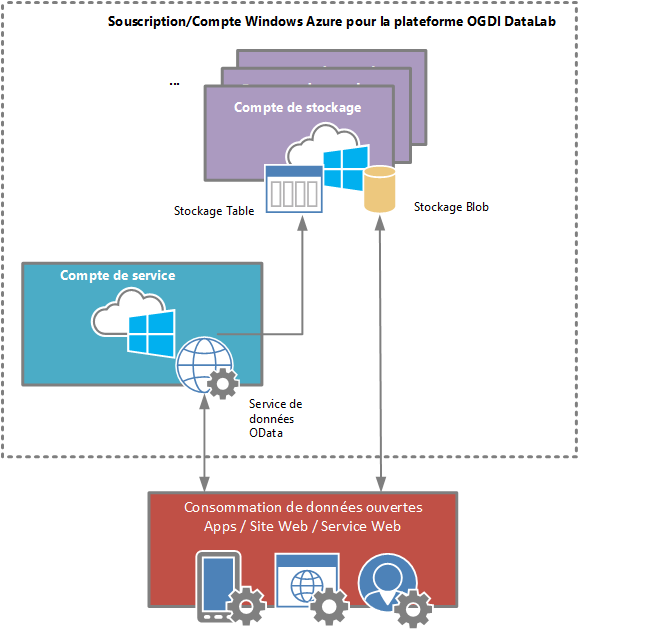
Remarque importante : Ces jeux de données sont destinés uniquement à des fins de tests, par exemple vis-à-vis d’applications en cours de développement qui les consommeraient.

Pour les utilisateurs qui recherchent des fréquences de mise à jour élevées, la fourniture de jeux de données de grande taille, ou un contrôle accru sur leurs données, etc. nous proposons notre assistance pour ajouter des instances supplémentaires de service de stockage Windows Azure à notre catalogue de données. Cette option est la solution qui fournit le meilleur contrôle quant au téléchargement et à la mise à jour des données.



Le [protocole ouvert de données OData (*Open Data Protocol*)](http://www.odata.org/developers/protocols)[[5]](#footnote-5) ne gère pas nativement les objets Blob ou les données binaires. Pour permettre aux utilisateurs d'accéder aux données binaires, vous devez télécharger vos fichiers vers le stockage Blob, puis lier ces fichiers depuis les enregistrements au format texte.

Lorsqu'il effectue ses requêtes auprès d’OGDI DataLab ou de votre instance du kit de démarrage OGDI DataLab (Cf. guide Déployer le kit de démarrage OGDI DataLab dans Windows Azure), le service consommateur va appeler les enregistrements auprès de Open Data France selon la méthode normale. Les enregistrements retournés comprennent une référence aux fichiers qui pourront être récupérés directement depuis le stockage Blob, comme illustré dans la figure ci-dessous :

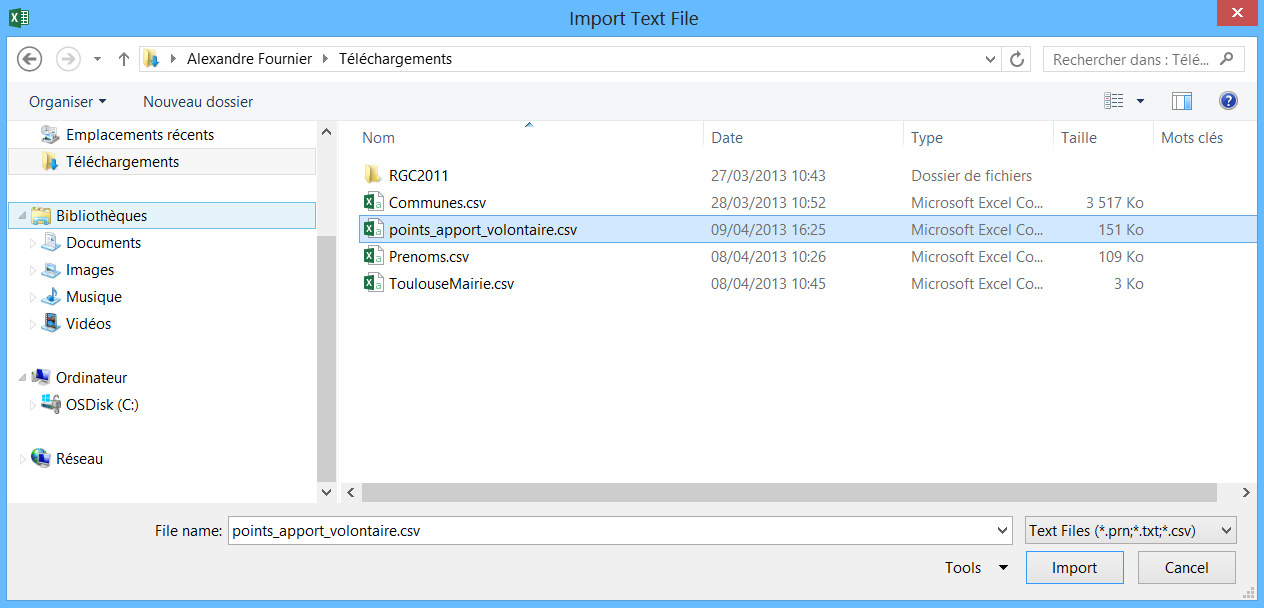


# Téléchargement des données au format texte

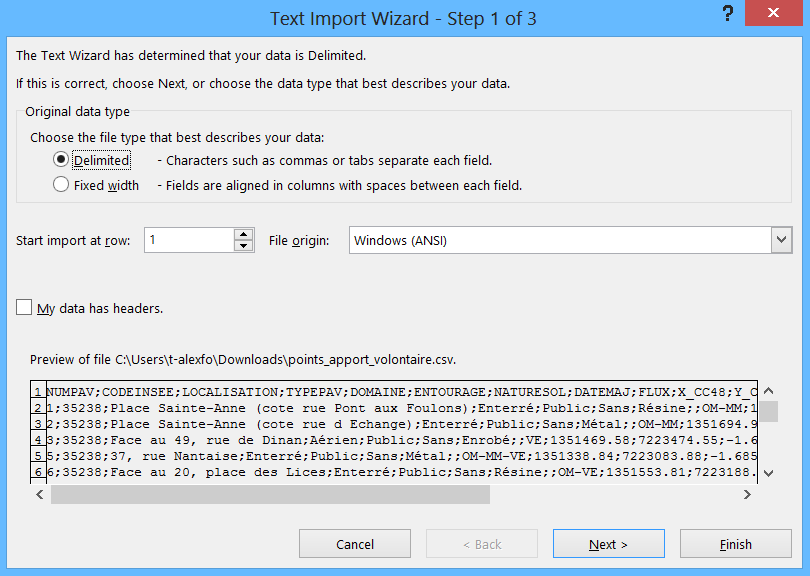
## Préparation des données pour le téléchargement au format CSV

Avant d'effectuer le téléchargement, vérifiez les données et agencez-les si nécessaire en fonction des besoins des utilisateurs. Pour mener à bien l'opération, vous devez convertir les données en un fichier CSV contenant uniquement du texte.

Pour cela, par exemple, vous pouvez utiliser en premier lieu l’assistant d’importation de texte de Microsoft Excel. Cela suppose simplement que les données sont contenues dans un fichier portant l’extension .*txt* et que vous ouvrez ce fichier en spécifiant Fichier texte.



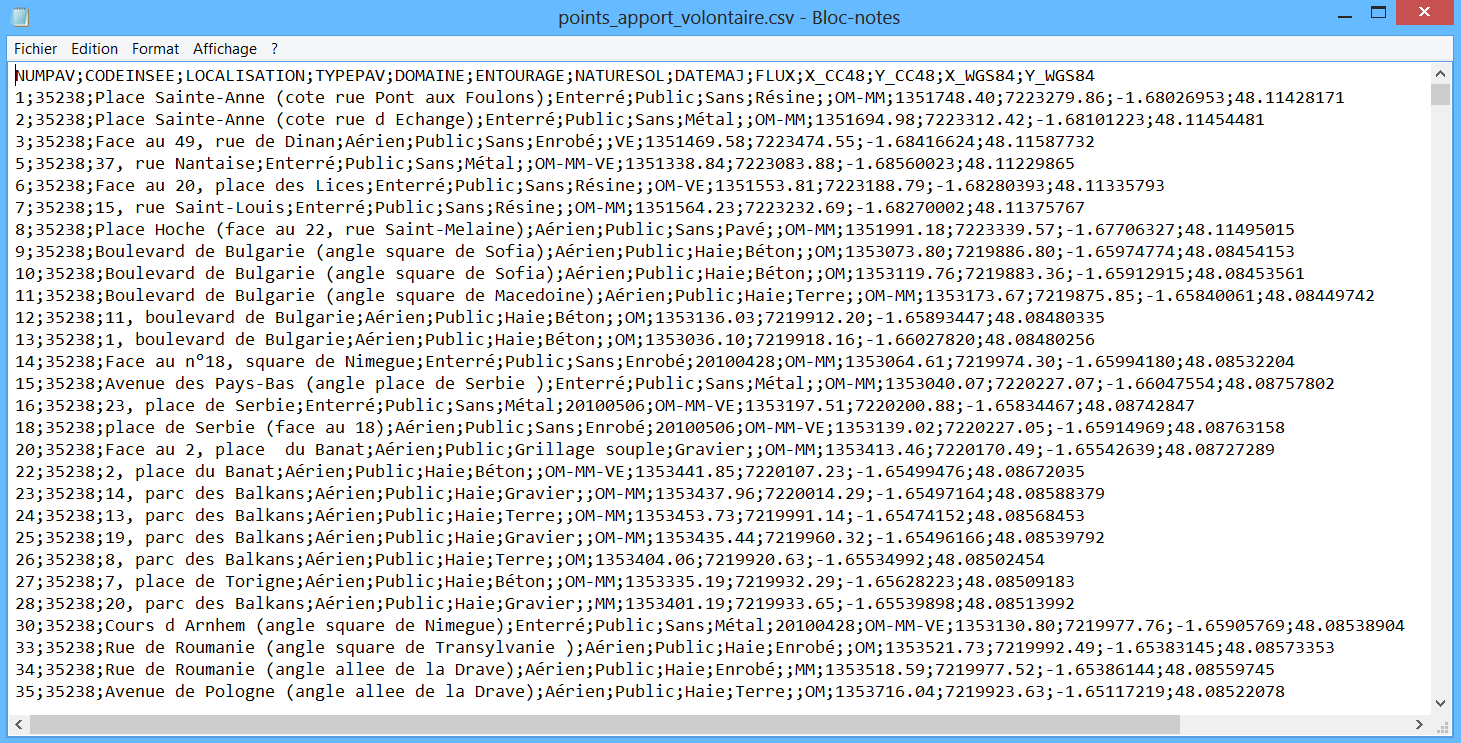
L’assistant d’importation de texte s’exécute alors. Vous êtes alors guidés dans l’importation la plus adaptée du fichier.



Vous pouvez alors sauvegarder des données au format Excel sous la forme d'un fichier CSV. Pour cela, pensez à respecter les points suivants :

* Chaque colonne doit être surmontée d'un en-tête. Cet élément sera utilisé ultérieurement ou lors de la création du fichier de configuration ;
* Pour toutes les valeurs numériques, le séparateur décimal doit être représenté par un point (.) et non une virgule (,). Cela s'applique également aux données de géolocalisation ;
* Toutes les lignes et colonnes contenant des données doivent être adjacentes, de manière à ne pas laisser de lignes ou de colonnes vides.

Enfin, le fichier CSV doit utiliser des virgules comme séparateur. La fonction de remplacement de texte proposé par le bloc-notes permet de réaliser l’opération. Lorsque le fichier CSV est préparé et enregistré, vous pouvez vérifier son contenu en l'ouvrant dans le bloc-notes. Le fichier doit ressembler à l'exemple suivant :



Remarque : Nous nous appuyons ici sur un ensemble de données pour les points d'apport volontaire (PAV) des déchets ménagers mis à disposition par le [Service SIG Rennes Métropole](http://www.data.rennes-metropole.fr/les-donnees/catalogue/?tx_icsopendatastore_pi1%5buid%5d=47)[[6]](#footnote-6). Cet ensemble contient une liste de PAV, chacun ayant des propriétés telles que le code INSEE de la commune sur laquelle est situé le PAV, le type de PAV, le ou les flux de déchets pris en charge (verre, journaux-magazine, ordures ménagères ou multi-matériaux), si le PAV est implanté sur le domaine public ou pas, une longitude et une latitude en termes de coordonnées géo spatiales, etc.

### Données de géolocalisation

OGDI DataLab fournit son assistance pour ajouter des coordonnées au jeu de données. Pour utiliser cette fonction, vous devez fournir ces coordonnées sur deux colonnes, avec un en-tête de colonne explicite :

|  |  |  |
| --- | --- | --- |
| Nom | Latitude | Longitude |
| monEmplacement | 3.5112959870828035 | 54.3750878143349 |

### Segmentation des données

Avec OGDI DataLab, les données sont présentées en faisant appel au protocole OData ; les utilisateurs peuvent ainsi effectuer des requêtes sur les données, et en extraire les éléments qui les intéressent.

Vous devrez télécharger les données dans des colonnes séparées pour permettre aux utilisateurs d'effectuer ces opérations :

#### Exemple erroné

|  |
| --- |
| **Adresse** |
| 22 Rue de Brest, 35000 Rennes |

#### Exemple correct

| **Rue** | **Numéro** | **Code Postal** | **Ville** |
| --- | --- | --- | --- |
| Rue de Brest | 22 | 35000 | Rennes |

Dans le deuxième exemple, les utilisateurs ont la possibilité de faire des requêtes sur les données, rechercher toutes les adresses d'une ville donnée, ou encore rechercher toutes les adresses pour une gamme de codes postaux définis. Ainsi, les utilisateurs ont la possibilité d'exploiter et de présenter les données dans leurs propres applications, selon la méthode qui leur convient le mieux.

## Préparation des données pour le téléchargement au format KML

Avant d'effectuer le téléchargement, vérifiez les données et agencez-les si nécessaire en fonction des besoins des utilisateurs. Pour mener à bien l'opération, vous devez convertir les données en un fichier KML contenant uniquement du texte manuellement ou via un SIG (Système d’Information Géographique).

Plusieurs options d’importation sont possibles :

* Importation d’un ensemble de données sous forme points ;
* Importation d’un ensemble de données sous forme de polygones.

Pour vérifier la conformité des données vous pouvez ouvrir votre fichier KML avec un éditeur de texte classique. Nous allons voir dans les parties suivantes la mise en forme attendue.

### Importation d’un ensemble de données sous forme de points

Pour importer un ensemble de points, aussi appelé marqueurs, au format KML le document doit présenter la syntaxe suivante :

<kml xmlns="http://www.opengis.net/kml/2.2">

<Document>

<name></name>

<Placemark>

<name>00009df5-7a1f-439b-bfc1-caede7182208</name>

<Point><coordinates>-1.61892019094906,48.0832344136664</coordinates></Point>

<description><![CDATA[Texte de description]]></description>

</Placemark>

<Placemark>

<name>00314d3b-3cb1-4d1f-b858-25d2637bd890</name>

<Point><coordinates>-1.67963229485677,48.0906598077195</coordinates></Point>

<description><![CDATA[Texte de description 2]]></description>

</Placemark>

<Placemark>

<name>0054697d-72bd-4aee-9f60-24fffae5e05b</name>

<Point><coordinates>-1.64350134517347,48.1008868659243</coordinates></Point>

<description><![CDATA[Texte de description 3]]></description>

</Placemark>

…

</Document>

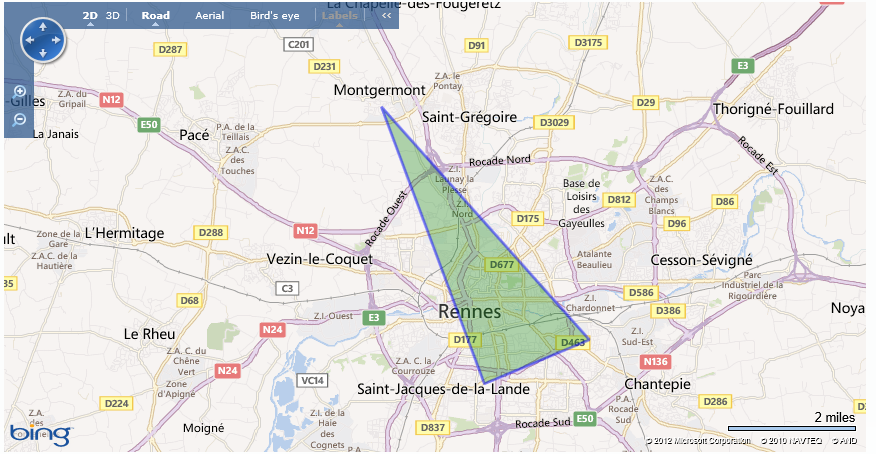
</kml>

#### Description des éléments

| **Nœud** | **Description** |
| --- | --- |
| Placemark | Ce nœud représente un élément (Marqueur) que vous voulez importer. Il contiendra toutes les informations que vous souhaitez télécharger et qui décriront cet élément. |
| Point | L’élément *Point* permet de spécifier les coordonnées du marqueur qui sera affiché sur la carte. |
| Coordinates | *Coordinates* est l’élément contenant les coordonnées. Il se compose de la façon suivante : <coordinates>*latitude,longitude*</coordinates>. Il est possible d’ajouter une information supplémentaire à savoir l’altitude ce qui donnera : <coordinates>*latitude,longitude,altitude*</coordinates> |
| Name | Le nom du marqueur. |
| Description | La description du marqueur qui pourra être affiché sur la carte. |

Remarque : Vous pouvez rajouter d’autres éléments dans la balise *Placemark* ceux-ci seront importés et pourront être réutilisés via le service de données.

### Importation d’un ensemble de données sous forme de polygones



Pour importer un ensemble de données au format KML représentées par des polygones le document doit présenter la syntaxe suivante :

<kml xmlns="http://www.opengis.net/kml/2.2">

<Document>

<name></name>

<Placemark xmlns="http://www.opengis.net/kml/2.2">

<name>00009df5-7a1f-439b-bfc1-caede7182208</name>

<Polygon>

<outerBoundaryIs>

<LinearRing>

<coordinates>48.0832344136664 -1.67963229485677,48.0906598077195 -1.64350134517347,48.1008868659243 -1.71492356265051,48.1542711149675 -1.61892019094906</coordinates>

</LinearRing>

</outerBoundaryIs>

</Polygon>

<description><![CDATA[Texte de description]]></description>

</Placemark>

…

</Document>

</kml>

#### Description des éléments

| **Nœud** | **Description** |
| --- | --- |
| Placemark | Ce nœud représente un élément (Marqueur) que vous voulez importer. Il contiendra toutes les informations que vous souhaitez télécharger et qui décriront cet élément. |
| Polygon | L’élément *Polygon* est le conteneur permettant de dessiner un polygone sur la carte. La balise *Polygone* ainsi que tous ses enfants présentés dans le code ci-dessus sont obligatoires. |
| outerBoundaryIs | Représente la balise qui définit le type de tracé du polygone. Cette balise est obligatoire et ne doit pas être modifiée. |
| LinearRing | Décrit les frontières du polygone comme étant un trait continu. Cette balise est obligatoire et ne doit pas être modifiée. |
| Coordinates | *Coordinates* est l’élément contenant l’ensemble des coordonnées composant le polygone. Afin d’afficher un polygone de façon optimale les coordonnées doivent respecter les caractéristiques suivantes :   * La latitude et la longitude doivent être séparées par un **espace** * Chaque ensemble latitude/longitude doit être séparé par une **virgule** * La première **latitude** doit être placée à la fin de l’ensemble des coordonnées **et** séparée de son prédécesseur par une **virgule**. Ceci permet de fermer le polygone   Exemple :  <coordinates>**longitude1**,latitude2 longitude2,latitude3 longitude3,**latitude1**</coordinates> |
| Name | Le nom du polygone. |
| Description | La description du polygone. |

Remarque : Vous pouvez rajouter d’autres éléments dans la balise *Placemark* ceux-ci seront importés et pourront être réutilisés via le service de données.

### Segmentation des données

Avec OGDI DataLab, les données sont présentées en faisant appel au protocole OData ; les utilisateurs peuvent ainsi effectuer des requêtes sur les données, et en extraire les éléments qui les intéressent.

Vous devrez télécharger les données dans des balises séparées pour permettre aux utilisateurs d'effectuer ces opérations :

#### Exemple erroné

<adresse>22 Rue de Brest, 35000 Rennes</adresse>

#### Exemple correct

<rue>Rue de Brest</rue>

<numero>22</numero>

<codepostal>35000</codepostal>

<ville>Rennes</ville>

Dans le deuxième exemple, les utilisateurs ont la possibilité de faire des requêtes sur les données, rechercher toutes les adresses d'une ville donnée, ou encore rechercher toutes les adresses pour une gamme de codes postaux définis. Ainsi, les utilisateurs ont la possibilité d'exploiter et de présenter les données dans leurs propres applications, selon la méthode qui leur convient le mieux.

## Envoi des données

Lorsque vous êtes prêts à nous envoyer vos données, compressez les fichiers au format CSV ou utilisez le fichier KML généré et faites-les nous parvenir à l'adresse [ogdifrance@live.fr](mailto:ogdifrance@live.fr). Nous les vérifierons et les mettrons en ligne. Nous vous rappelons qu'il s'agit d'un processus manuel. Nous acceptons uniquement les fichiers ne faisant pas l'objet de modifications trop fréquentes, et dont la taille est limitée.

## Téléchargement des données vers votre instance de stockage

Pour bénéficier du téléchargement de vos données vers votre propre compte de stockage de données, vous devez disposer d'un abonnement Windows Azure et configurer votre compte Windows Azure. Veuillez-vous reporter au guide Déployer le Kit de démarrage OGDI DataLab dans Windows Azure pour plus d'informations sur les abonnements Azure et la configuration de votre compte de stockage de données.

Une fois l'instance prête à l'emploi, vous pouvez commencer à télécharger les données. Les prérequis sont :

* Nom de votre compte de stockage de données ;
* Clé d'accès à votre compte de stockage de données ;
* Instance de stockage configurée pour OGDI DataLab ;
* Données préparées sous forme de fichier CSV ou KML, tel que décrit dans ce guide.

### Installation de l'outil de téléchargement

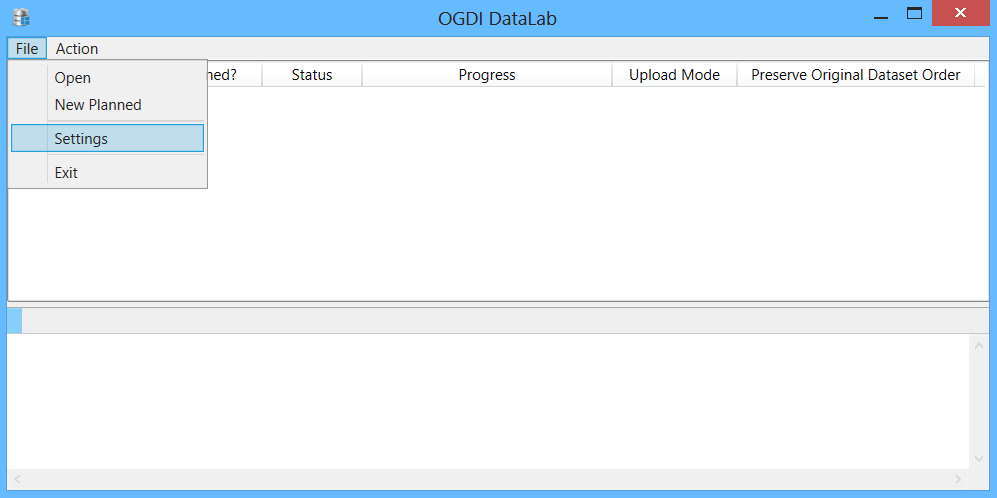
Vous pouvez vous procurer l'outil de téléchargement sur le site Web GovData.eu (<http://frogdidata.blob.core.windows.net/public/OGDI_DataLoader.zip>), ou compiler le code source des utilitaires de chargement fournis avec le Kit de démarrage OGDI DataLab.

### Configuration de l'application de téléchargement des données

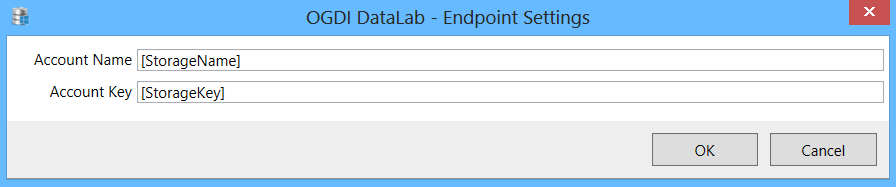
Au lancement de l'application, vous devez configurer les informations d'identification pour le moteur de stockage.

Procédez de la manière suivante :

1. Cliquez sur le menu **File**, puis sur **Settings**.



1. Entrez les informations d'identification du compte de stockage de données Windows Azure.

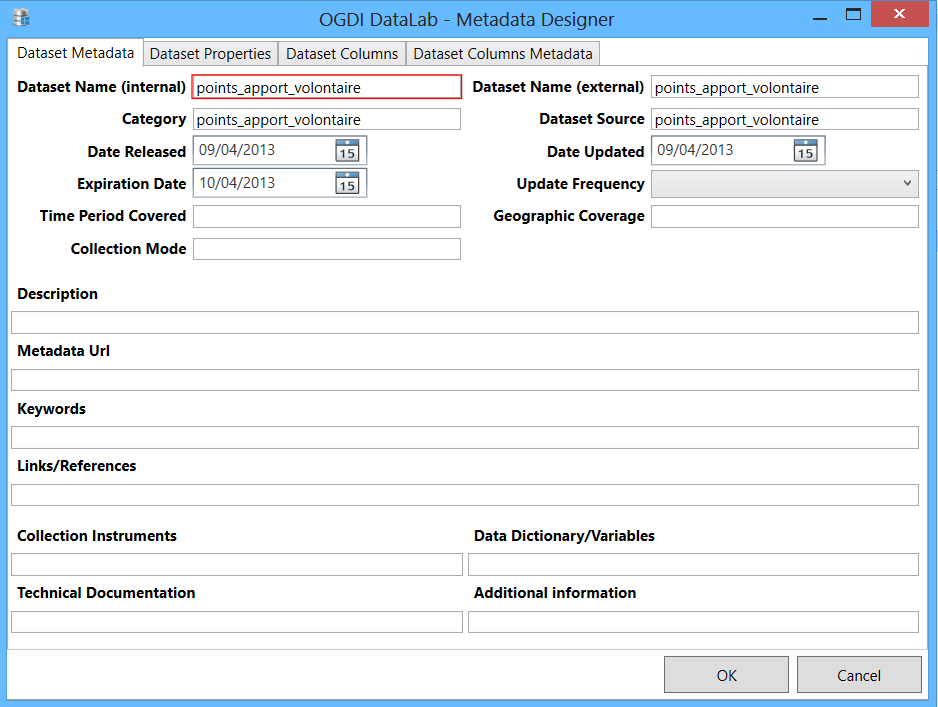


1. Cliquez sur **OK**.

### Préparation du téléchargement montant

Pour préparer le téléchargement, procédez de la manière suivante :

1. Ouvrez le fichier CSV préparé précédemment en sélectionnant **File** | **Open**.
2. Recherchez puis ouvrez le jeu de données que vous voulez télécharger. À l'ouverture du fichier CSV/KML, la fenêtre *Metadata Designer* s'ouvre.



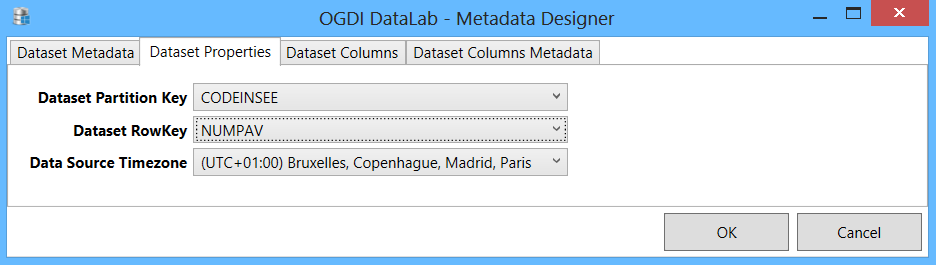
#### Onglet Dataset Metadata

L'interface de l'application contient les éléments suivants :

| **Élément** | **Description** |
| --- | --- |
| Dataset name (internal) | Nom du jeu de données dans le service Web OData. Ce nom est exposé en externe dans le service Web. |
| Dataset name (external) | Nom explicite sur le portail du développeur de l'instance OGDI DataLab. |
| Category | Catégorie affichée sur le portail du développeur de l'instance OGDI DataLab. |
| Dataset Source | Source des données |
| Date released | Date de publication |
| Date updated | Date de la dernière mise à jour |
| Expiration date | Date d'expiration des données |
| Update frequency | Fréquence de mise à jour |
| Time Period Covered | Période de temps couverte |
| Geographic Coverage | Aire géographique couverte |
| Collection mode | ? |
| Description | Description du jeu de données |
| Metadata URL | Adresse URL permettant d'obtenir plus d'informations sur le jeu de données |
| Keywords | Mots-clés pour la recherche |
| Links/References | Adresse(s) URL contenant plus d'informations sur les jeux de données |
| Collection Instruments | Instruments utilisés pour la collecte des données |
| Data dictionary / Variables | ? |
| Technical Documentation | Informations complémentaires sur les parties techniques des jeux de données |
| Additional information | Informations complémentaires sur les jeux de données |

Remarque importante : Nous vous remercions de fournir les informations sur la source originale et la licence des données dans l'Éditeur des métadonnées. Vous pouvez vous référer pour cela à la page <http://ogdifrance.cloudapp.net/Home/TermsOfUse>.

#### Onglet Dataset Properties

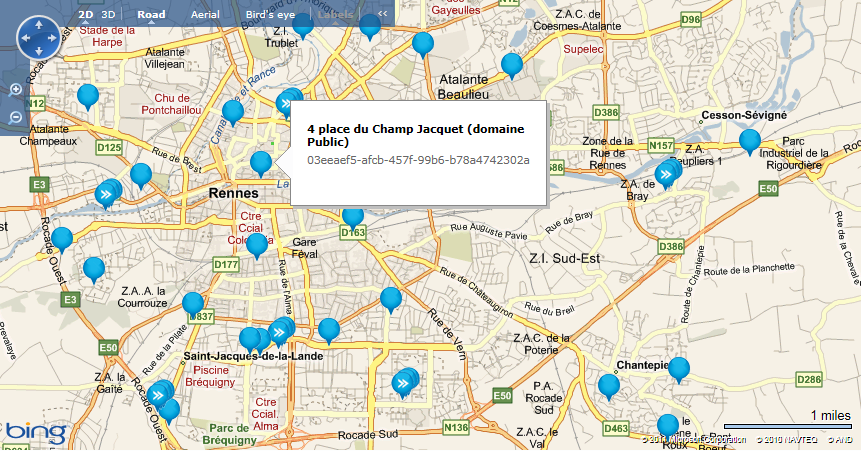


Cette boîte de dialogue permet de définir le partitionnement des jeux de données, ainsi que le fuseau horaire de la source des données.

Remarque : Pour plus d'informations sur la sélection de la partition et la propriété RowKey, veuillez-vous reporter à la page Understanding the Table Service Data Model (<http://msdn.microsoft.com/en-us/library/dd179338.aspx>).

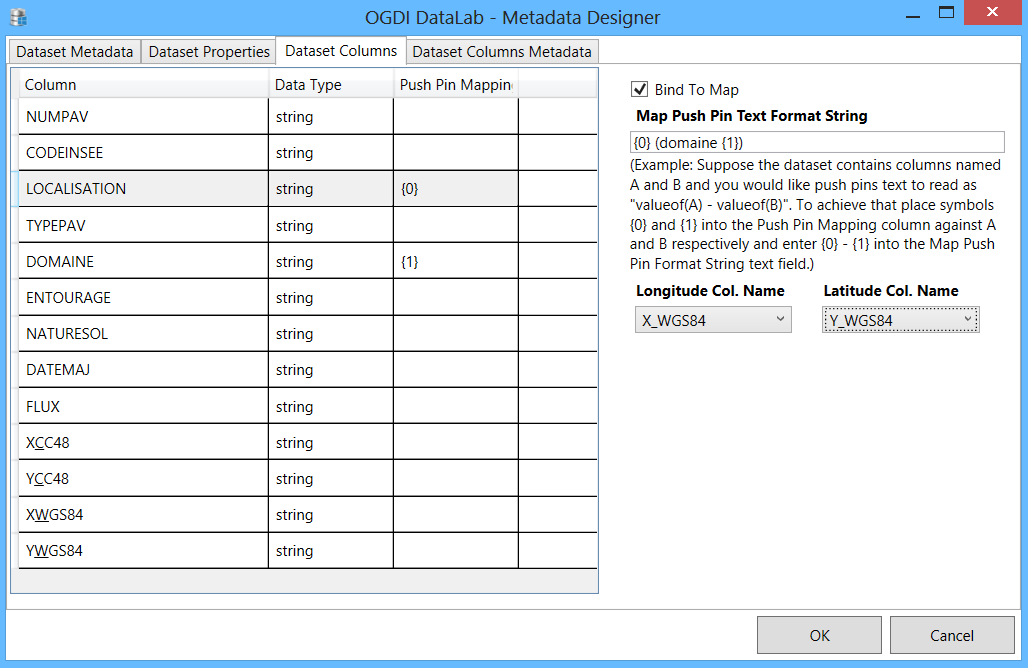
#### Onglet Dataset Columns

Dans l'onglet **Dataset Columns**, vous pouvez définir la chaîne de caractères qui sera affichée sur la carte d'un jeu de données.

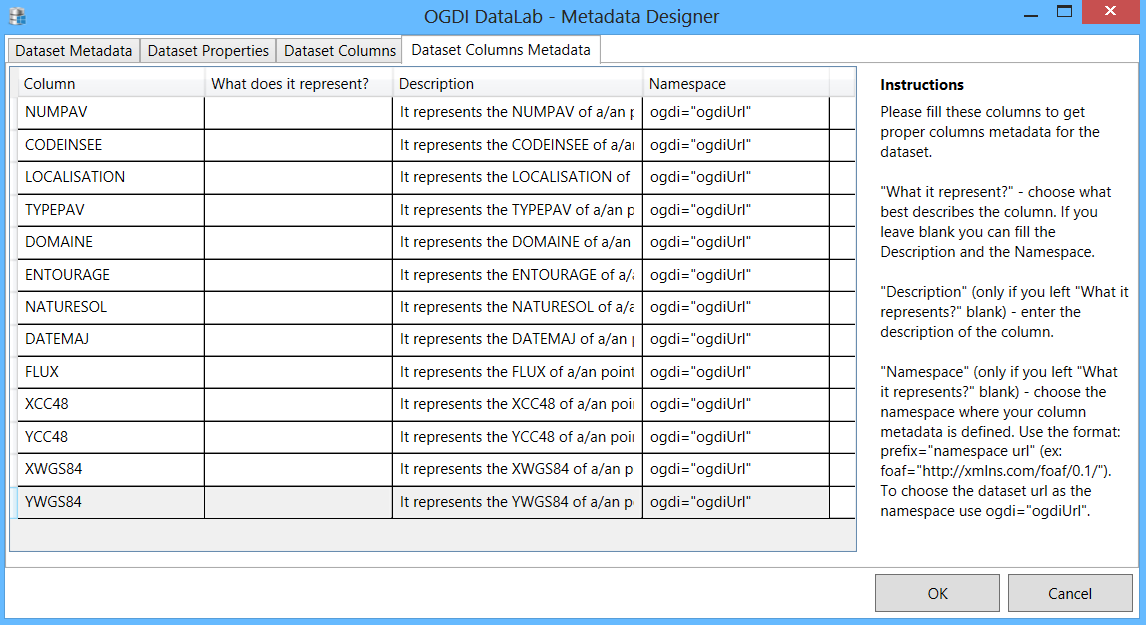


Dans cet exemple, à savoir les points d’apports volontaires (PAV) des déchets ménagers, le titre contient à la fois la localisation et le domaine du point d’apport volontaire. Ces éléments peuvent être configurés dans **Map Push Pin Text Format String**. Définissez d'abord un symbole pour chaque champ qui sera utilisé dans la chaîne de format.

Chaque champ doit avoir son propre symbole de positionnement sur la carte (*{0}*, *{1}*, *{2}*, etc.). Utilisez ensuite ces symboles dans la zone de texte **Format String**.



#### Onglet Dataset Columns Metadata



Cet onglet permet de définir les métadonnées RDF (*Resource Description Framework*) pour les colonnes du jeu de données. Dans cet onglet, deux approches s’offrent à vous pour associer des métadonnées aux colonnes :

1. Vous pouvez choisir une option de ce que représente la colonne ;
2. Vous pouvez éditer manuellement, pour chaque colonne, la description et l’espace de noms de la description.

Pour ce qui est de la première approche, le champ **What does it represents?** est une liste déroulante qui vous permet de choisir depuis celle-ci un attribut décrivant au mieux la colonne.

La liste des attributs disponibles peut être éditée manuellement depuis le fichier de configuration *RDFNamespaces.xml* qui doit se trouver dans le même dossier que l’exécutable du chargeur de données.

Plusieurs attributs peuvent être ajoutés à ce champ simplement en ajoutant le nom de l'option (propriété), celui de l'espace de noms (préfixe) et enfin l'URL de l'espace de noms.

Une structure typique du fichier *RDFNamespaces.xml* est la suivante :

<?xml version="1.0" encoding="utf-8" ?>

<metadata>

<description>It represents the {0} of a/an {1}.</description>

<namespaces>

<namespace>

<property>Name</property>

<prefix>foaf</prefix>

<url>http://xmlns.com/foaf/0.1/</url>

</namespace>

<namespace>

<property>Address</property>

<prefix>foaf</prefix>

<url>http://xmlns.com/foaf/0.1/</url>

</namespace>

</namespaces>

</metadata>

L’élément **description** est un champ obligatoire et représente la description par défaut qui sera associée aux métadonnées. Sa valeur peut être changée, mais les deux paramètres *{0}* et *{1}* doivent être impérativement utilisés et renseignés :

1. Le paramètre *{0}* représente le nom de la colonne ;
2. Le paramètre *{1}* représente le nom du jeu de données.

Les éléments relatifs aux espaces de noms sont les nouveaux attributs qui peuvent être modifiés ou ajoutés pour remplir le champ **What does it represent?** Ainsi, pour ajouter un nouvel attribut, vous devez ajouter la structure suivante après le dernier nœud :

<namespace>

<property>[name of the attribute]</property>

<prefix>[name of the namespace (no spaces)]</prefix>

<url>[namespace URL]</url>

</namespace>

S’il n’y a pas d’attribut choisi dans la liste déroulante, alors les champs **Description** et **Namespace** doivent être renseignés afin de générer correctement les métadonnées. Ceci correspond à la seconde approche.

Le champ **Description** est automatiquement rempli avec la phrase par défaut configurée dans le fichier *RDFNamespaces.xml* et représente la métadonnée associée à chaque colonne.

Le champ **Namespace** définit l’espace de noms de la métadonnée associée à la colonne. Cet espace de noms doit respecter la syntaxe *[prefix]=”[URL namespace]”* ou *[prefix]=[URL namespace]*.

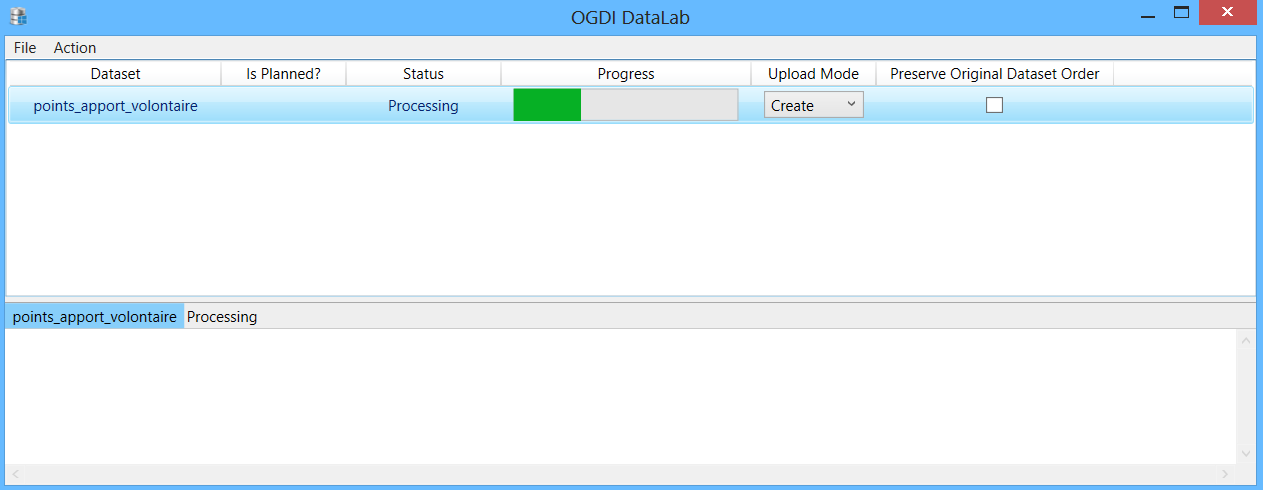
Lorsque le jeu de données est chargé, un nouveau champ **rdfsnippet** est créé dans la table qui contient l’information RDF associée à l’entité en question.

Une table **TableColumnsMetadata** supplémentaire est aussi créée pour stockée l’information précisée dans l’onglet *Dataset Columns Metadata*.

### Téléchargement des données

Après le paramétrage de la connexion et la configuration des métadonnées, l'utilitaire de téléchargement de données est prêt pour la transmission des données. Vous pouvez définir le mode de téléchargement (création, ajout ou modification des données) et conserver l'ordre d'origine du jeu de données (case à cocher).

Pour télécharger vos données, cliquez sur **Action |** **Start Transfer** pour démarrer le téléchargement. La barre de progression montre l'avancement de l'opération.



# Téléchargement des données Blob

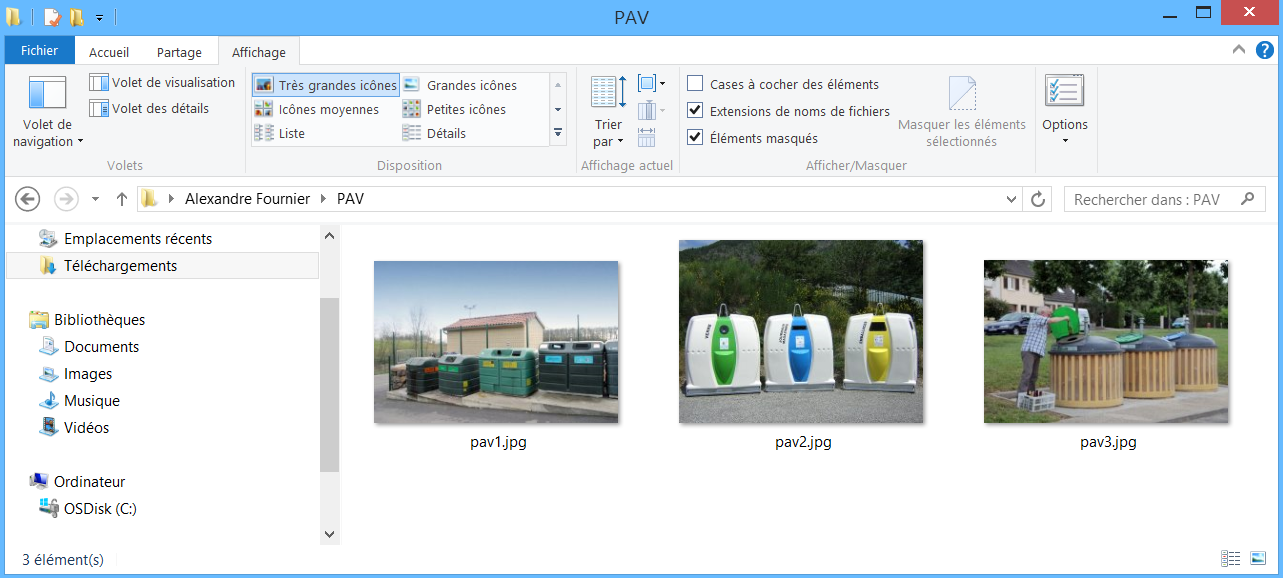
## Exemple de données

Dans cette section, nous allons télécharger des données factices à des fins d'illustration. Supposons que nous ayons des fichiers image avec une description et des données de géolocalisation (latitude/longitude).

Voici la table contenant les informations que nous allons télécharger :

| **Nom** | **Description** | **LAT.** | **LONG.** | **Nom de fichier** |
| --- | --- | --- | --- | --- |
| PAV 1 | PAV Avenue de Chèvre | 6.366823 | 99.667107 | pav1.jpg |
| PAV 2 | PAV Rue du clos des vignes | 1.234567 | 99.876543 | pav2.jpg |
| PAV 3 | PAV Rue de Cornouailles | 1.234567 | 99.876543 | pav3.jpg |

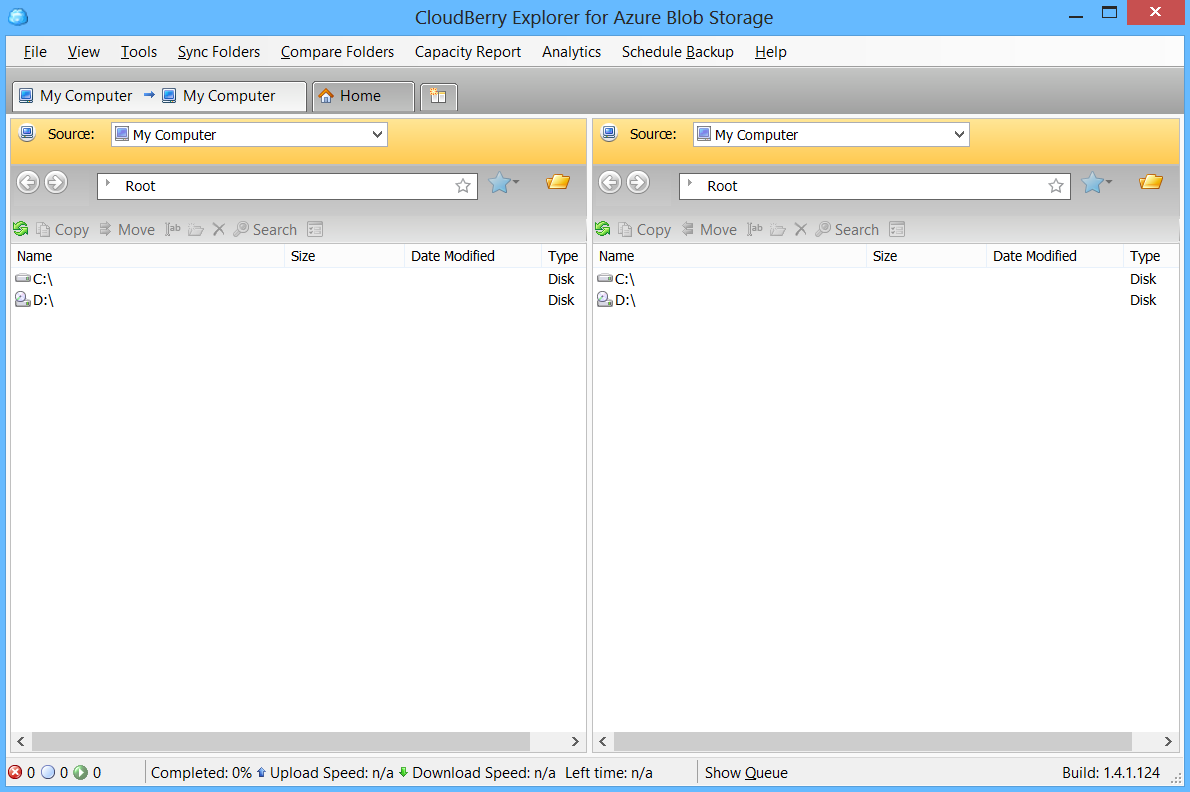
Nous voulons ajouter les images au stockage Blob Windows Azure (comme s'il s'agissait d'un partage de fichiers). La première étape constitue le téléchargement des images.



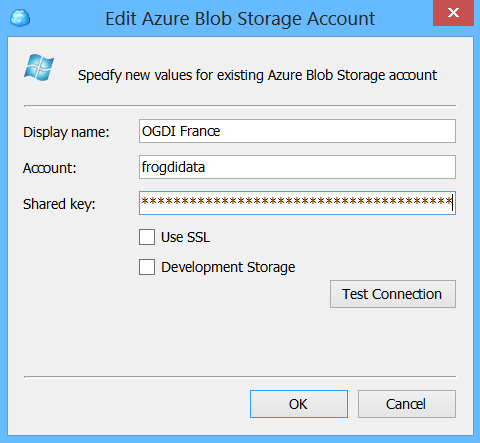
## Téléchargement de données binaires

Procédez de la manière suivante :

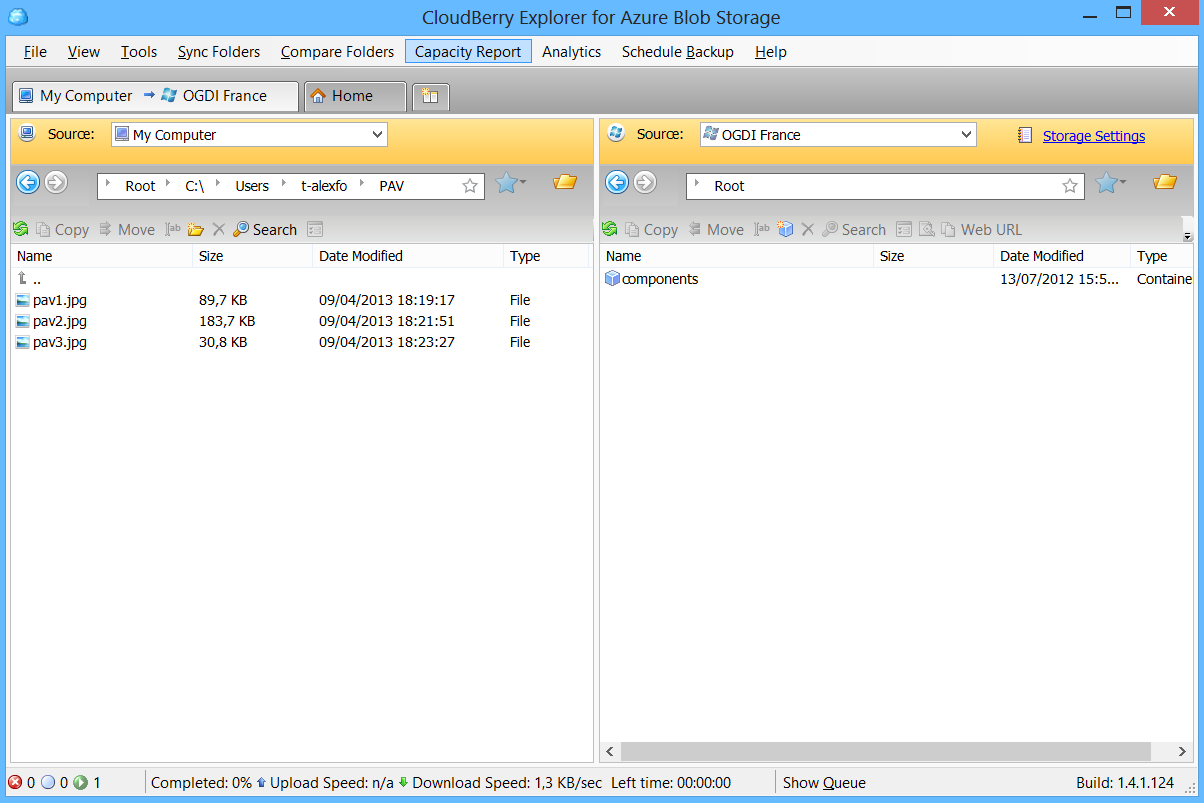
1. Pour le téléchargement des données binaires, nous utilisons ici l'application CloudBerry Explorer for Azure Blob Storage comme indiqué précédemment. Lancez l'application.



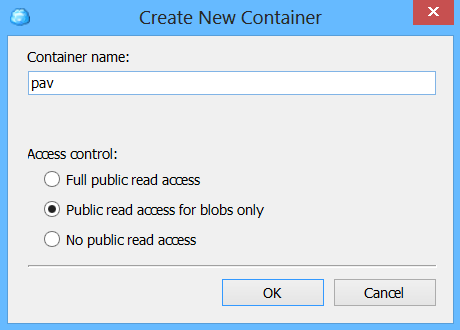
1. Une fois l'application ouverte, ajoutez votre compte de stockage de données Windows Azure au moyen de la commande **File** | **Azure Blob Storage Accounts** | **New Account**.



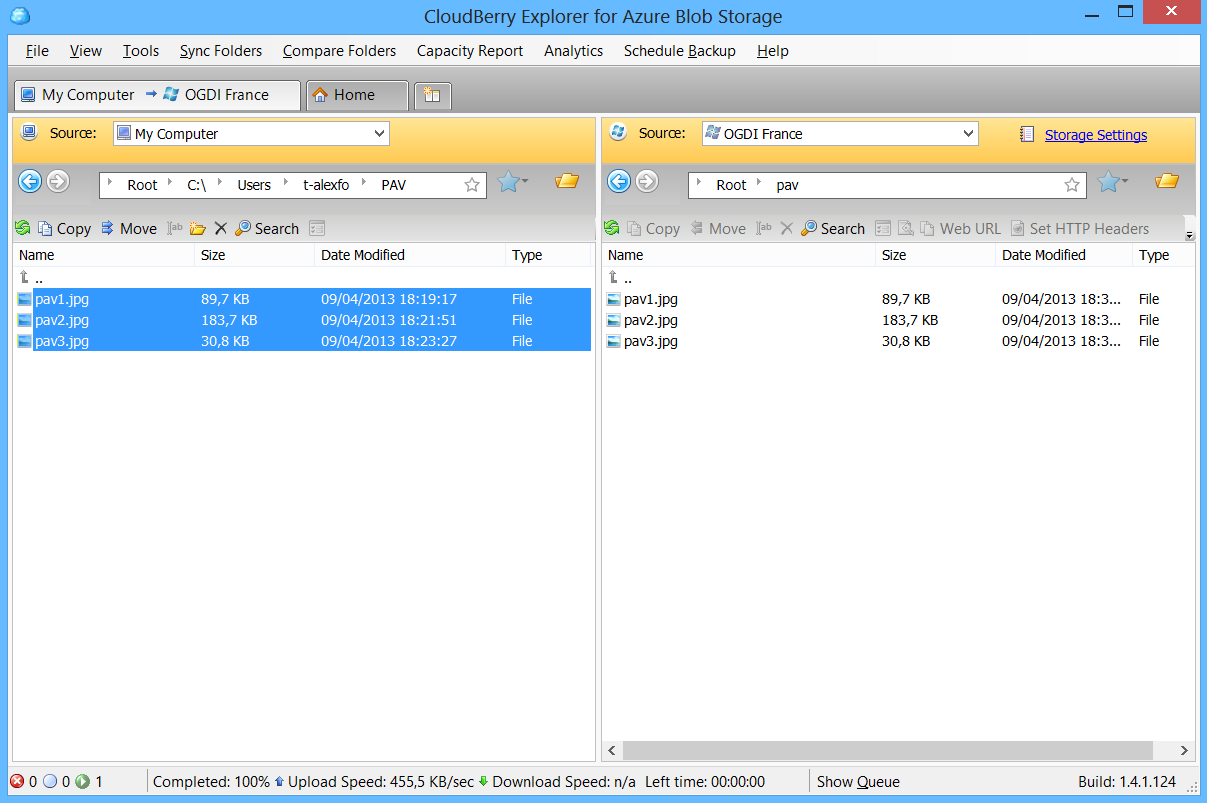
1. Paramétrez le volet de gauche pour qu'il pointe vers le dossier local qui contient les données à télécharger, le volet de droite vers votre compte de stockage de données, dans le cas présent le compte *OGDI France*.



1. Créez un nom au conteneur au niveau de votre compte de stockage de données, *OGDI France* dans le cas présent, et définissez le contrôle d'accès associé (en cochant par exemple l'option **Public read access for blobs only** dans le cas présent).

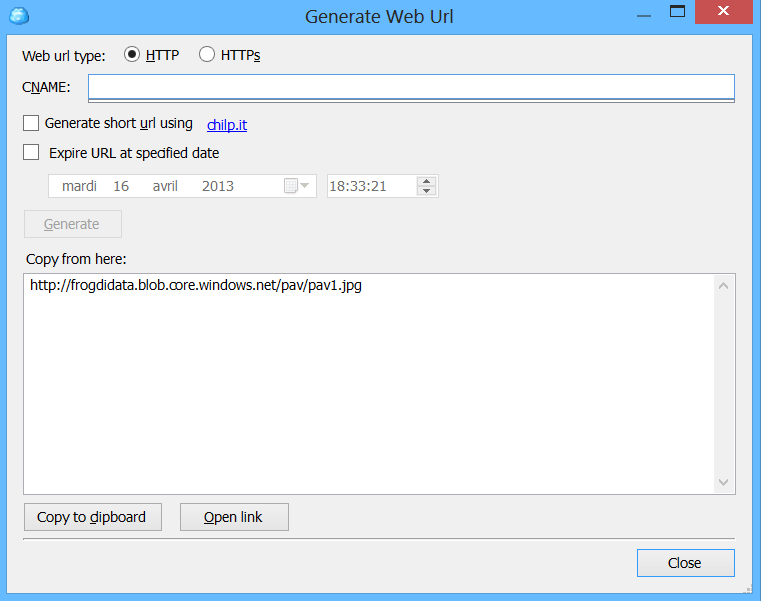


1. Ouvrez le conteneur que vous venez de créer par un double-clic.
2. Faites glisser les fichiers depuis le volet de gauche vers le volet de droite pour commencer leur téléchargement.

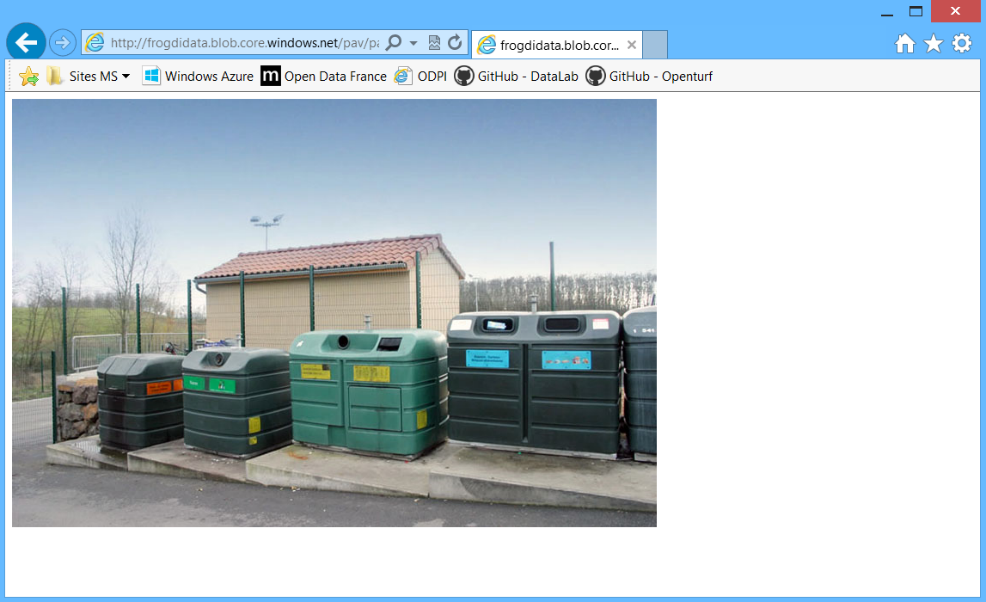


Vos fichiers sont ajoutés à la file d'attente et téléchargés vers votre compte de stockage de donnée Windows Azure dans le Cloud.

1. Lorsque le téléchargement est terminé, sélectionnez un des fichiers dans le volet de droite et cliquez sur **Web URL**. Cela vous permet de visualiser (et de copier-coller) l'adresse URL de votre fichier.



1. Pour tester l'accès au fichier, copiez l'adresse URL et collez-la dans un navigateur pour afficher votre image ou fichier.



## Préparation des données

Maintenant que les fichiers sont disponibles dans le stockage Blob de votre compte de stockage de données Windows Azure, nous devons ajouter l'adresse URL complète de l'image au jeu de données. Au lieu du nom de fichier, nous ajoutons le chemin complet au fichier.

Vous pouvez le faire en ajoutant une colonne et en la renseignant en fonction du nom de fichier :

http://[YOUR STORAGE ACCOUNT NAME].blob.core.windows.net/pav/<fichier>

Où <*fichier*> se rapporte au nom de fichier. Ensuite, par la commande **Paste special**, collez les valeurs dans une nouvelle colonne. Après cela, vous pouvez supprimer les colonnes obsolètes.

| **Nom** | **Description** | **LAT.** | **LONG.** | **Nom de fichier** | |
| --- | --- | --- | --- | --- | --- |
| PAV 1 | PAV Avenue de Chèvre | 6.366823 | 99.667107 | | http://frogdidata.blob.core.windows.net/pav/pav1.jpg |
| PAV 2 | PAV Rue du clos des vignes | 1.234567 | 99.876543 | | http://frogdidata.blob.core.windows.net/pav/pav2.jpg |
| PAV 3 | PAV Rue de Cornouailles | 1.234567 | 99.876543 | | http://frogdidata.blob.core.windows.net/pav/pav3.jpg |

Désormais, vous pouvez utiliser le processus de téléchargement normal pour ajouter ces données à OGDI DataLab, en vous assurant que les liens pointent vers les « bons » fichiers.

1. Site portail Web Microsoft Secteur Public : http://www.microsoft.com/france/entreprises/secteur-public/solutions-pour-secteur-public.aspx [↑](#footnote-ref-1)
2. Microsoft Windows Azure : http://www.microsoft.com/france/windows-azure/ [↑](#footnote-ref-2)
3. Solution Open Data clé en main: https://mspartner.microsoft.com/fr/fr/Pages/Solutions/Open-Data-Cle-en-main.aspx [↑](#footnote-ref-3)
4. Blog MSDN Open Data France : http://aka.ms/OpenDataFrance/ [↑](#footnote-ref-4)
5. Documentation du protocole OData : http://www.odata.org/developers/protocols [↑](#footnote-ref-5)
6. Catalogue de données Rennes Métropole : http://www.data.rennes-metropole.fr/les-donnees/catalogue/?tx\_icsopendatastore\_pi1%5buid%5d=47 [↑](#footnote-ref-6)