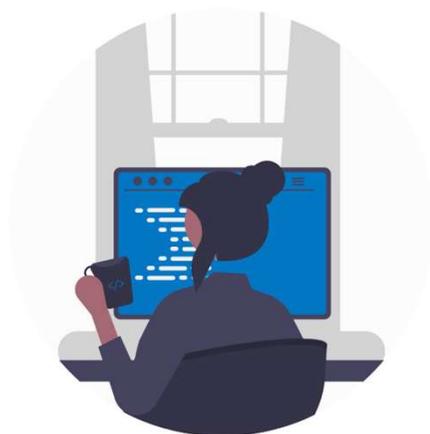




Automatiser des opérations avec Dataverse

Démonstration de l'API



Luc Decker

Service Information Scientifique
et Technique – MCST

Contact : data@ird.fr



A l'aide de l'API, automatiser les échanges d'informations avec le logiciel Dataverse

Exemples d'applications possibles

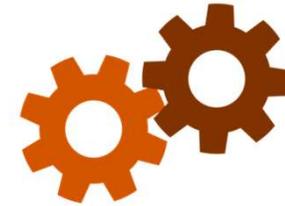
- ❑ Rechercher des fichiers de données afin de les référencer dans un catalogue centralisé, détecter régulièrement les ajouts
- ❑ Extraire le contenu d'une collection, sous forme de liste
- ❑ Télécharger des fichiers de données, comme point d'entrée d'une chaîne de traitement
- ❑ Mettre à jour un jeu de données à partir d'une application ou d'un script lancé périodiquement
- ❑ Transférer de nombreuses données dans un entrepôt à partir d'une base existante

Qu'est-ce qu'une API ? Pourquoi faire ?

Advanced Programming Interface

Dans un logiciel, un **ensemble de méthodes** documentées qui ouvrent la **possibilité à d'autres logiciels d'interagir avec lui** et d'accomplir diverses opérations (lecture/écriture de données...)

→ Passerelles, connexions entre logiciels



→ Permet la conception de chaînes de traitement automatisées

→ S'affranchir d'opérations répétitives réalisées manuellement

En lien avec les principes **F A I R**,

le fait de fournir une API contribue à l'**interopérabilité**

Le logiciel Dataverse offre plusieurs APIs

Native API	toutes les opérations de base
Search API	spécialisée dans les recherches
Data Access API	spécialisée dans les téléchargements
Metrics API	spécialisée dans les mesures

Principe général : toute opération réalisée manuellement au moyen de l'interface web de DataSuds peut être automatisée par l'intermédiaire de l'API.

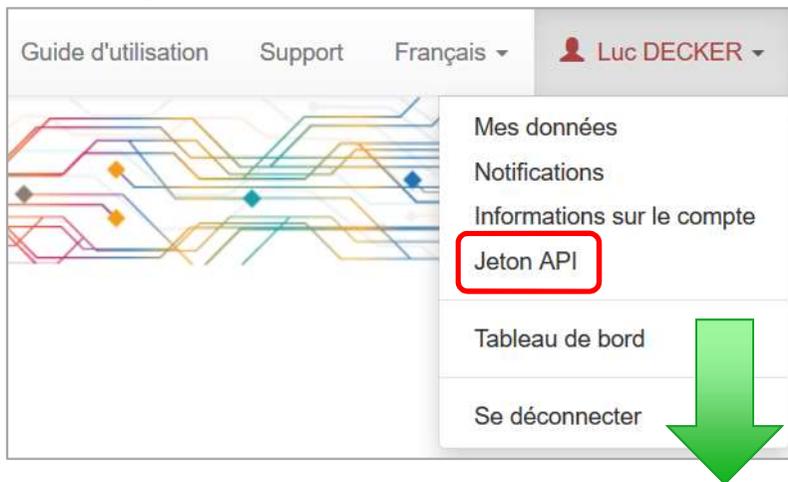


Comment utiliser l'API Dataverse ?

- ▶ Pour tester, mettre au point : commandes lancées directement
 - dans un navigateur web (*Firefox* recommandé)
 - dans un terminal : installer l'outil *CURL* ou équivalent
- ▶ Avec votre langage de programmation préféré
 - Appels à Dataverse dans un logiciel existant
 - Script créé sur mesure pour répondre à un besoin ponctuel, de l'ordre de 10-100 lignes de code
- ▶ A l'aide d'un logiciel dédié à l'utilisation d'APIs

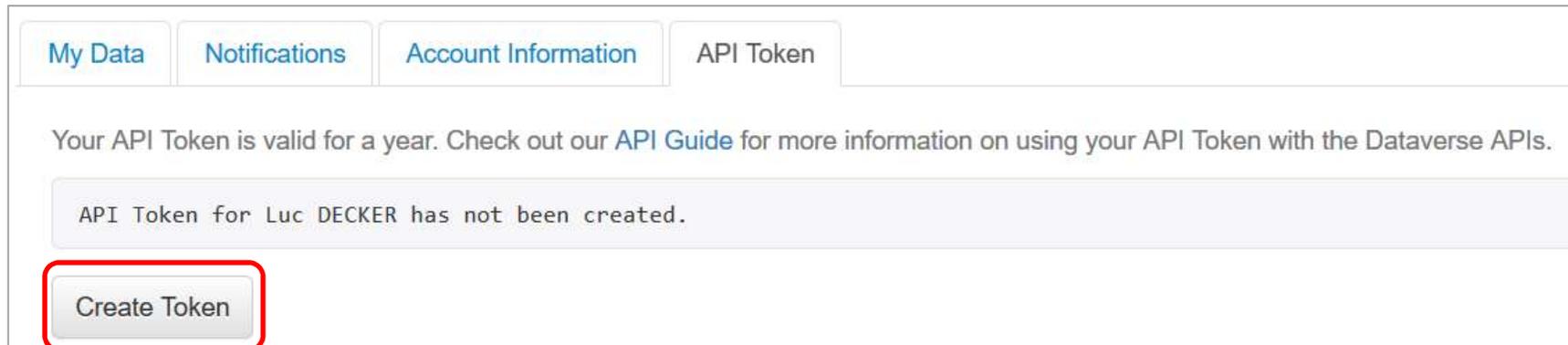
Comment utiliser l'API Dataverse ?

Pour toutes les actions qui nécessitent d'être connecté :
générer et utiliser votre code d'authentification personnel
« *jeton API* » à partir du menu « Utilisateur » de Dataverse
→ opérations automatisées réalisées en votre nom



ⓘ Les jetons API

- doivent être stockés de manière sécurisée
- expirent au bout d'un an
- peuvent être désactivés, re-générés



Démonstration de l'API Dataverse

Point d'entrée (URL) de l'API pour DataSuds

<https://dataverse.ird.fr/api/>

Exemple simple : interroger les métadonnées d'un jeu de données, au format JSON

```
JSON  Données brutes  En-têtes
Enregistrer Copier Tout réduire Tout développer Filtre le JSON
status: "OK"
data:
  id: 36081
  identifier: "SNGQHV"
  persistentUrl: "https://doi.org/10.23708/SNGQHV"
  protocol: "doi"
  authority: "10.23708"
  publisher: "DataSuds"
  publicationDate: "2023-02-14"
  storageIdentifier: "file://10.23708/SNGQHV"
  latestVersion:
    id: 24263
    datasetId: 36081
    datasetPersistentId: "doi:10.23708/SNGQHV"
    storageIdentifier: "file://10.23708/SNGQHV"
    versionNumber: 1
    versionMinorNumber: 0
    versionState: "RELEASED"
    UNF: "UNF:6:ANjAoi1b+71R2DZt5Euwew=="
    lastUpdateTime: "2023-02-14T17:31:30Z"
    releaseTime: "2023-02-14T17:31:30Z"
    createTime: "2022-04-28T10:02:05Z"
    license: "NONE"
  termsOfUse: "<a rel=\\"license\\" href=\\"http://creativecommons.org/licenses/by/4.0/\">This work is licensed under a <a rel="
```

Objet ou action demandée



[https://dataverse.ird.fr/api/datasets/:persistentId/?](https://dataverse.ird.fr/api/datasets/:persistentId/?persistentId=doi:10.23708/SNGQHV)

[persistentId=doi:10.23708/SNGQHV](https://dataverse.ird.fr/api/datasets/:persistentId/?persistentId=doi:10.23708/SNGQHV)

Paramètre transmis à l'API : identifiant DOI du jeu de données

Démonstration de l'API Dataverse

Exemple : script en langage PHP, lecture de métadonnées

```
<?php
// Configuration: Entrepot DataSuds, serveur de production
$SERVER_API_URL = "https://dataverse.ird.fr/api";
// Exemple (fictif) de jeton API, inutile pour lire un jeu de données publié
$apiToken = "xxcff5a3-41xx-4cx0-87x2-b04exx984xxx";
// DOI d'un jeu de données quelconque, pour tester
$persistentId = "doi:10.23708/SNGQHV";

// Initialiser CURL, navigateur internet en ligne de commande et en mode texte
$CurlSession = curl_init();
// curl_setopt($CurlSession, CURLOPT_HTTPHEADER, array("X-Dataverse-key:$apiToken"));
curl_setopt($CurlSession, CURLOPT_RETURNTRANSFER, true);

// Interroger l'API Dataverse - Requete simple : métadonnées d'un jeu de données
$APIQuery = "$SERVER_API_URL/datasets/:persistentId/?persistentId=$persistentId";
curl_setopt($CurlSession, CURLOPT_URL, $APIQuery);
$result = curl_exec($CurlSession);

// Importer les métadonnées au format JSON dans une variable exploitable par le
// programme PHP => traitements ultérieurs possibles
$dataSet = json_decode($result, true);

// Vérifier le contenu de la variable (affichage)
echo "Resultat<pre>\n"; print_r($dataSet); echo "</pre>\n";
?>
```

API Dataverse : pour aller plus loin

Manuel de référence complet, en anglais

<https://guides.dataverse.org/en/latest/api/index.html>

Sur un projet précis, demander conseil à l'équipe support de DataSuds data@ird.fr