

OpenArchaeo : doc d’Administration

|  |  |  |  |
| --- | --- | --- | --- |
| **Historique des versions** | | | |
| **Version** | **Date** | **Auteur** | **Modifications** |
| 1 | 19/02/2019 | Thomas Francart | *Version initiale* |
| 2 | 25/07/2019 | Thomas Francart | Ajout des prérequis Git  Gestion des métadonnées |
| 3 | 10/03/2020 | Thomas Francart | Doc d’administration des Collections |
| 4 | 07/04/2020 | Thomas Francart | Export de données |
| 5 | 13/07/2020 | Thomas Francart | Clarification URI graphe nommé |

Table des matières

[1 Table des matières 2](#_Toc37162440)

[2 Introduction 3](#_Toc37162441)

[3 Serveurs et identifiants 3](#_Toc37162442)

[4 Administrer le tripletore OpenArchaeo 3](#_Toc37162443)

[4.1 Adresse et identifiants 3](#_Toc37162444)

[4.2 Charger un jeu de données dans le triplestore OpenArchaeo 3](#_Toc37162445)

[4.3 Supprimer un jeu de données dans le triplestore OpenArchaeo 7](#_Toc37162446)

[4.4 Recharger une nouvelle version d’un jeu de données 8](#_Toc37162447)

[4.5 Exporter un jeu de données 8](#_Toc37162448)

[5 Configuration de la fédération et de l’explorateur 8](#_Toc37162449)

[5.1 Principe général pour modifier les fichiers de configuration 8](#_Toc37162450)

[5.2 Structure des répertoires de configuration 9](#_Toc37162451)

[6 Administrer la fédération OpenArchaeo 10](#_Toc37162452)

[6.1 Réindexer les données d’une source 10](#_Toc37162453)

[6.2 Ajouter une source dans la fédération 11](#_Toc37162454)

[6.3 Editer les Collections de sources dans la Fédération 12](#_Toc37162455)

[6.4 Modifier les requêtes SPARQL d’exemple dans la fédération 13](#_Toc37162456)

[6.5 Mettre à jour le référentiel Periodo 13](#_Toc37162457)

[7 Administrer l’explorateur OpenArchaeo 14](#_Toc37162458)

[7.1 Modifier le contenu éditorial de la page d’accueil ou des mentions légales 14](#_Toc37162459)

[8 Administrer le serveur 14](#_Toc37162460)

[8.1 Consulter les fichiers de config 14](#_Toc37162461)

[8.2 Consulter les logs 14](#_Toc37162462)

[8.3 Arrêter / redémarrer les services 15](#_Toc37162463)

[8.3.1 Via le git 15](#_Toc37162464)

[8.3.2 En ligne de commande 15](#_Toc37162465)

[8.4 Nettoyer les logs 15](#_Toc37162466)

# Introduction

Ce document décrit les procédures d’administration de la plateforme OpenArchaeo.

# Serveurs et identifiants

* Serveur : **openarchaeo.huma-num.fr**
  + Le même serveur est accessible à **openarchaeo-graphdb.huma-num.fr** pour l’accès à GraphDB (voir ci-dessous) ;
* Les applications sont accessibles à :
  + <http://openarchaeo.huma-num.fr/explorateur>
  + <http://openarchaeo.huma-num.fr/federation>

# Administrer le tripletore OpenArchaeo

## Adresse et identifiants

Le triplestore OpenArchaeo est accessible à <http://openarchaeo-graphdb.huma-num.fr/>.

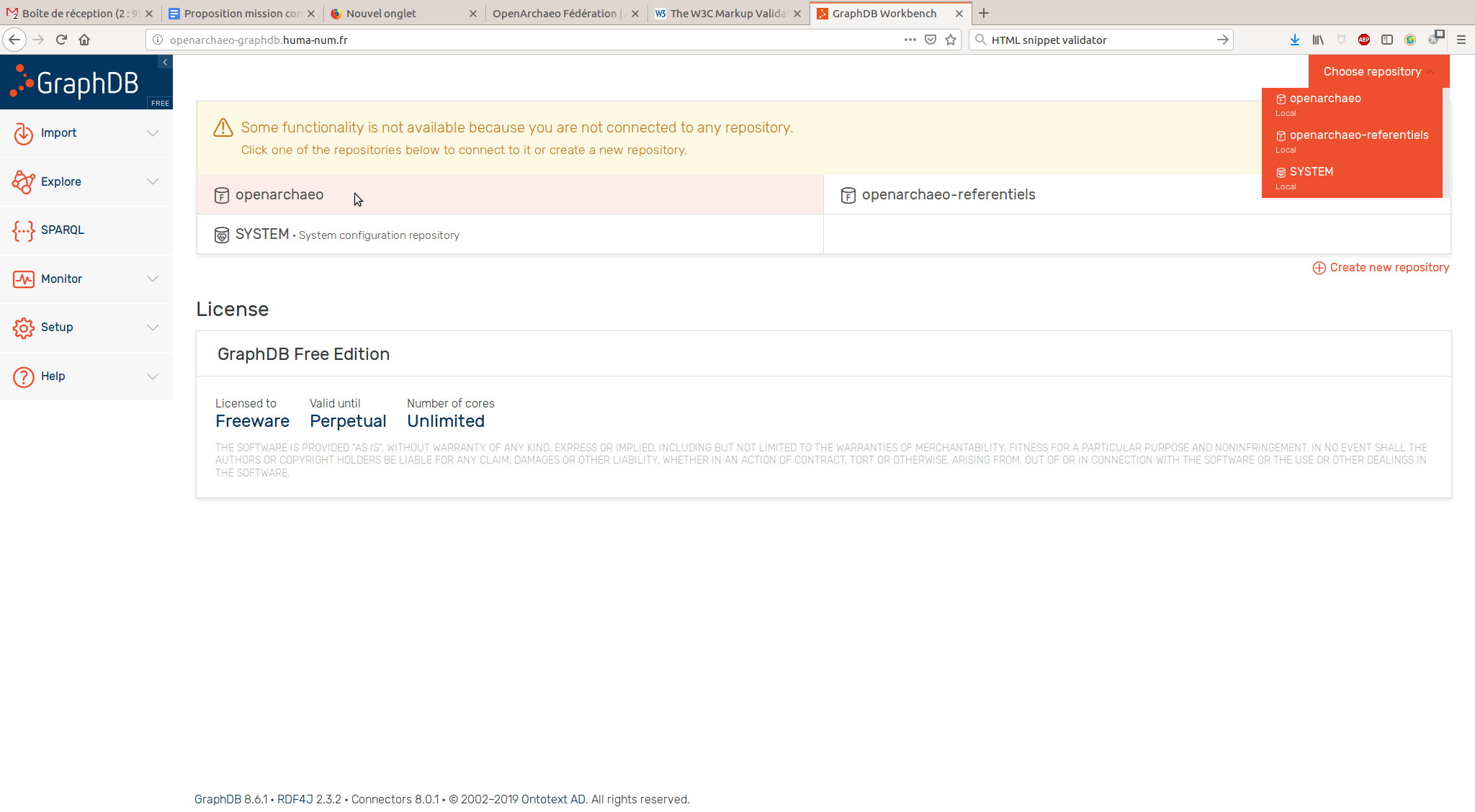
Le triplestore contient 2 repositories (2 bases) :

* **openarchaeo** : repository principal contenant les données ;
* **openarchaeo-referentiels** : repository contenant les données récupérées automatiquement des référentiels (Geonames et Pactols);

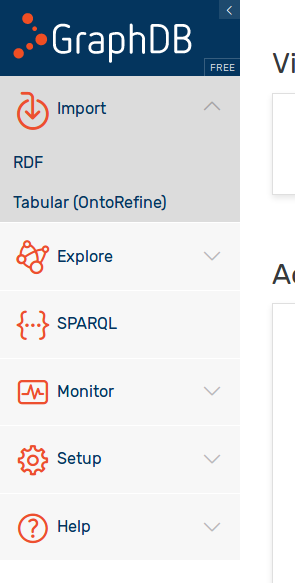
## Charger un jeu de données dans le triplestore OpenArchaeo

Pour charger un nouveau jeu de données :

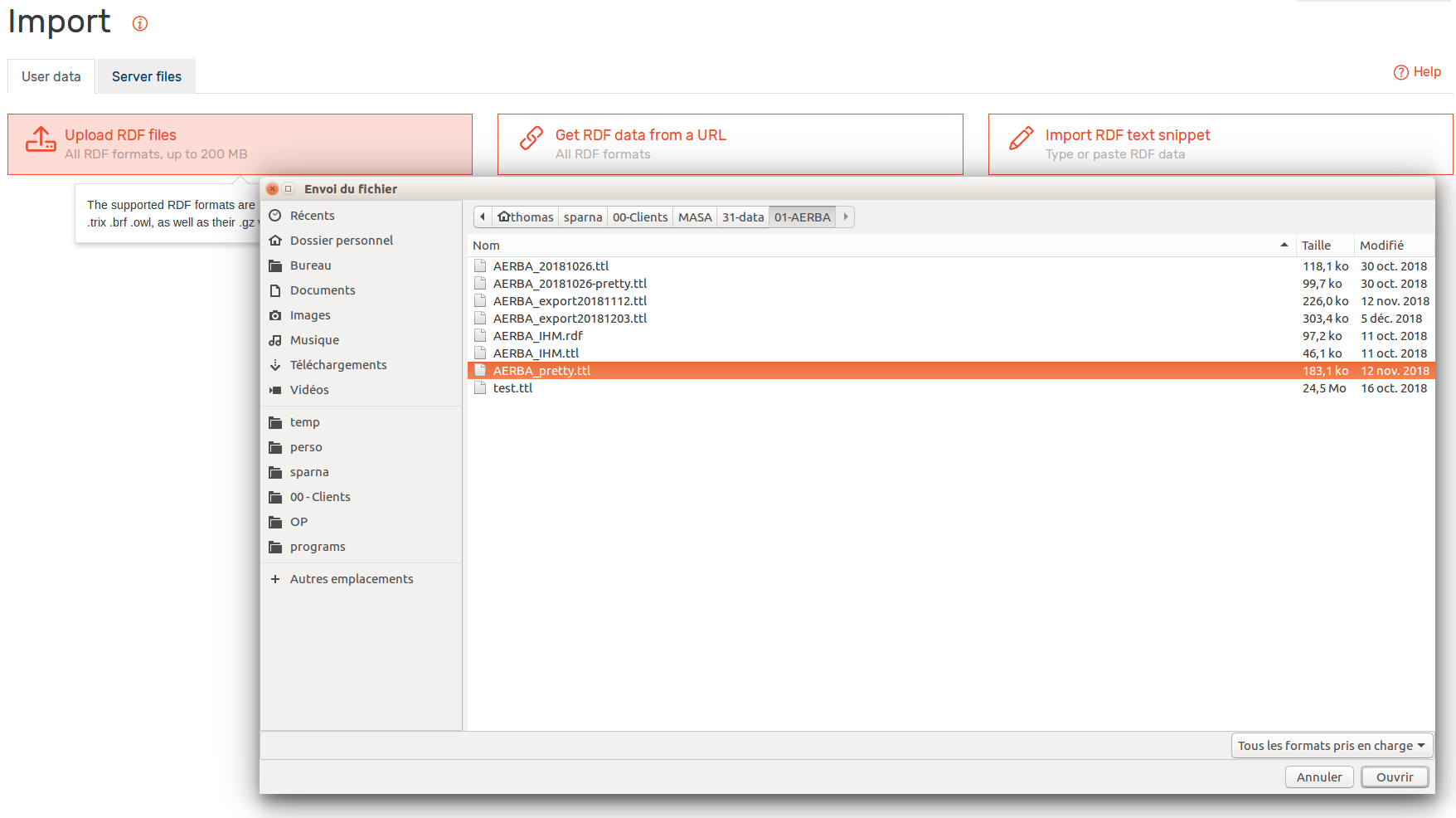
1. Se connecter au repository **openarchaeo** ;



1. Aller dans le **menu** **Import > RDF** ;



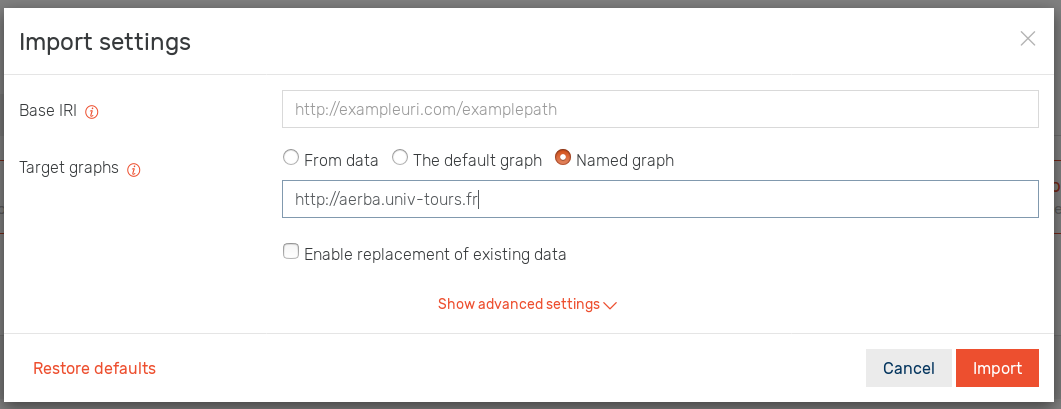
1. Cliquer sur le **bouton « Import RDF files »** et sélectionner le fichier à importer ; Note : si les données se composent de plusieurs fichiers, il est possible de zipper ces fichiers et d’uploader le zip ;



1. Le fichier s’affiche dans la liste au milieu de l’écran ; cliquer sur le **bouton « Import » à droite du nom de fichier** ; un popup de chargement s’affiche ;
2. Dans le popup :
   1. Sélectionner l’option **Target Graphs = Named Graph**
   2. Saisir l’**URI du graphe nommé** dans lequel les données seront enregistrées ; cette URI de graphe nommé devra être répercutée dans le fichier de config federation/sources.ttl (cf procédure correspondante plus bas).

Cette URI est une URI arbitraire, l’important est que la même valeur soit répercutée dans le fichier de configuration des sources ; typiquement, on peut choisir pour cette URI l’adresse de la page d’accueil de la base de données importée, par exemple. A titre d’exemple, les URI pour Arsol et Aerba sont respectivement <http://arsol.univ-tours.fr> et <http://aerba.univ-tours.fr> (l’URI pour ICERAMM pourra donc être <http://iceramm.univ-tours.fr>).

* 1. Cliquer sur **Import** : les données s’importent, sous réserve que les fichiers RDF sont syntaxiquement conformes ;

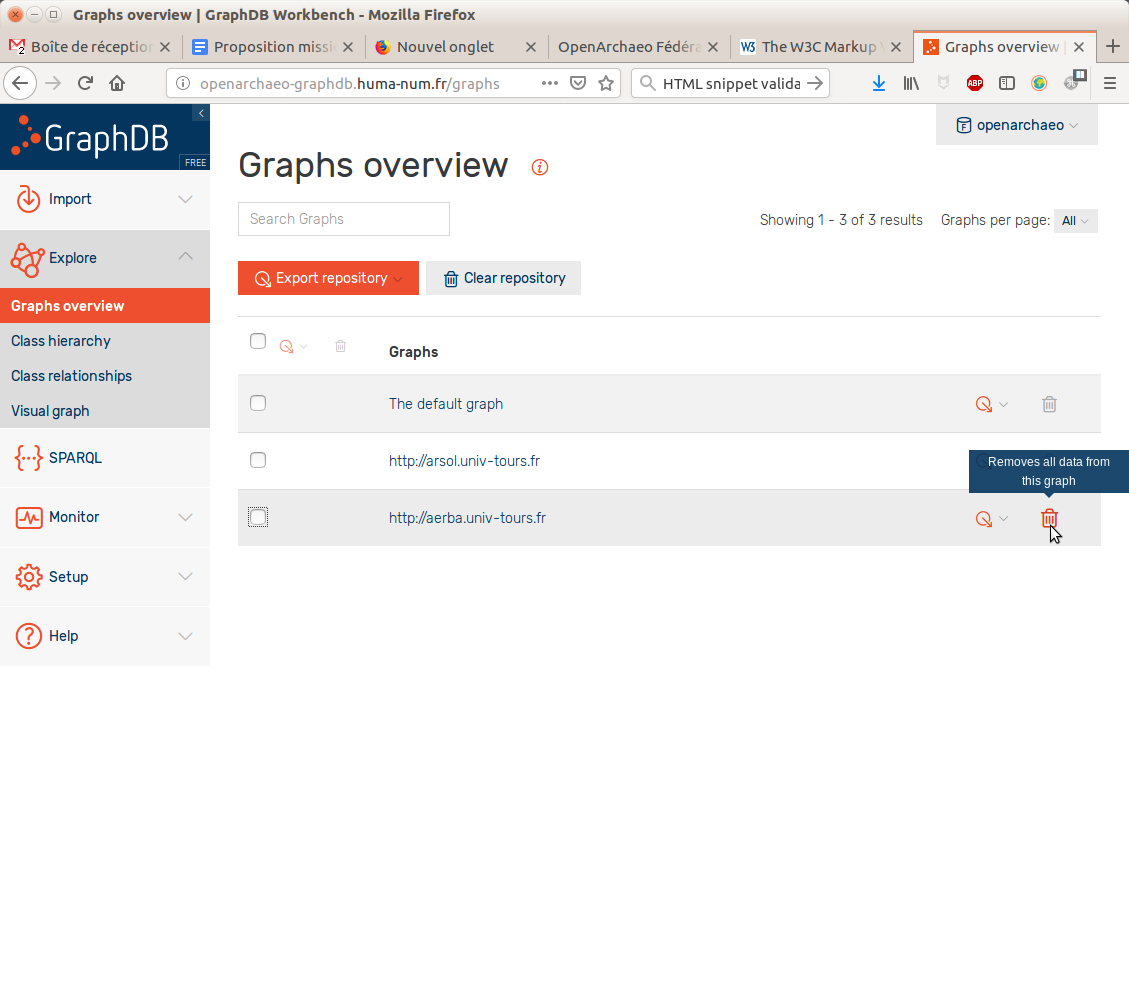


1. Aller dans le menu **Explore > Graphs Overview** :
   1. Vérifier le nouveau graphe nommé apparait bien dans la liste ;
   2. Cliquer dessus pour vérifier son contenu ;
2. Répéter la procédure d’import pour importer le fichier DCAT de description de la source de données. **Attention de bien charger le fichier avec la même URI de graphe nommé que le fichier de données** ;
3. Dans l’administration de la féderation, déclarer une nouvelle source (fichier federation/sources.ttl), puis recalculer les index de recherche de cette fédération (cf. procédures correspondantes) ;

## Supprimer un jeu de données dans le triplestore OpenArchaeo

Pour supprimer un jeu de données du triplestore :

1. Aller dans le menu **Explore > Graphs Overview** ;
2. Cliquer sur la **poubelle en face de l’identifiant du graphe à supprimer** ;



## Recharger une nouvelle version d’un jeu de données

Pour recharger une nouvelle version des données d’une source, il est nécessaire :

* de supprimer l’ancienne version, puis de recharger la nouvelle version dans le même graphe nommé. Cf les 2 autres procédures ci-dessus,
* de recalculer les index de la source (cf. procédure correspondante);

## Exporter un jeu de données

Pour exporter un jeu de données du triplestore :

* Aller dans le menu **Explore > Graphs Overview** ;
* Cocher sur l’icône avec « flèche qui sort du rond » à droite **du graphe à exporter**, à côté de la poubelle ;
* Sélectionner le format d’export voulu, typiquement Turtle ;

# Configuration de la fédération et de l’explorateur

## Principe général pour modifier les fichiers de configuration

Les fichiers de configuration de la fédération ou de l’explorateur OpenArchaeo sont accessibles et modifiables dans un dépôt Git sur le serveur OpenArchaeo. *Toute modification faite sur ces fichiers et transmises au serveur déclenche de façon automatique le déploiement de la nouvelle configuration et le redémarrage automatique des applications pour prendre en compte la nouvelle config*.

Toute la procédure décrite ci-dessous en ligne de commande nécessite l’installation de « git ». Il faut installer pour cela « Git for windows » depuis <https://gitforwindows.org/>. La séquence de lignes de commandes peut aussi être réalisée avec une application Git pour Windows tel que TortoiseGit à <https://tortoisegit.org/>.

Pour modification la configuration, il faut donc :

1. Cloner le dépôt git à l’adresse [git@openarchaeo.huma-num.fr:/home/git/openarchaeo-config.git](mailto:git@openarchaeo.huma-num.fr:/home/git/openarchaeo-config.git). Typiquement avec la ligne de commande suivante :
   1. git clone [git@openarchaeo.huma-num.fr:/home/git/openarchaeo-config.git](mailto:git@openarchaeo.huma-num.fr:/home/git/openarchaeo-config.git)
2. Le mot de passe est **oaGIT!**
3. Faire les modifications dans les fichiers de config (voir ci-dessous) ;
4. « Commiter » les modifications. Typiquement avec la ligne de commande :
   1. git add <nom du fichier modifié>
   2. git commit
5. « Pousser » les modifications sur le serveur. Typiquement avec la ligne de commande :
   1. git push
6. La config est automatiquement redéployée et les serveurs redémarrés ; l’application est inaccessible pendant environ 1 ou 2 minutes, le temps de redémarrer.

Voici la séquence à suivre avec TortoireGit :

1. Faire l’opération « Git clone… » documentée à <https://tortoisegit.org/docs/tortoisegit/tgit-dug-clone.html> en reseignement
   1. l’adresse du dépôt [git@openarchaeo.huma-num.fr:/home/git/openarchaeo-config.git](mailto:git@openarchaeo.huma-num.fr:/home/git/openarchaeo-config.git)
   2. le même mot de passe que ci-dessus ;
2. Faire les modifs dans les fichiers ;
3. « Commiter » les modifications par le menu « TortoiseGit > Commit… » ; voir la doc à <https://tortoisegit.org/docs/tortoisegit/tgit-dug-commit.html>
4. « Pusher » les modifications par le menu « TortoiseGit > Push… » ; voir la doc à <https://tortoisegit.org/docs/tortoisegit/tgit-dug-push.html>;

## Structure des répertoires de configuration

Le dépôt Git contenant la configuration a la structure suivante :

* openarchaeo-config
  + **testgit.md** : un fichier qui ne sert à rien, sinon à tester le fonctionnement du git pour déclencher un redémarrage des serveurs sans toucher à la config (voir plus bas)
  + **explorateur/…**
    - **config.properties** : configuration générale de l’application « explorateur », contenant en particulier les URLs d’appels à la fédération. Il n’est pas nécessaire de modifier ce fichier ;
    - **home\_en.html + home\_fr.html** : contenus HTML anglais et français insérés sur la page d’accueil de l’explorateur ;
    - **legalnotice\_en.html + legalnotice\_fr.htlml** : contenus HTML anglais et français insérés dans les mentions légales de l’explorateur ;
    - **query\_expansion.ttl** : fichier de configuration pour l’extension des requêtes SPARQL. *Attention ce fichier est également présent dans la configuration de la fédération – toute modification dans ce fichier doit être répercutée aux 2 endroits !*
    - **spec-search.json** : specifications du composant graphique du constructeur de requêtes ;
  + **federation/…**
    - **config.properties** : configuration générale de l’application de fédération ;
    - **periodo.jsonld** : fichier contenant les périodes de Periodo (téléchargé directement depuis Periodo)
    - **query\_examples.json** : fichier contenant les exemples de requêtes SPARQL qui s’affichent dans le formulaire de requête SPARQL de la fédération ;
    - **query\_expansion.json** : fichier de configuration pour l’extension des requêtes SPARQL. *Attention ce fichier est également présent dans la configuration de l’explorateur – toute modification dans ce fichier doit être répercutée aux 2 endroits !*
    - **sources.ttl** : fichier contenant la liste et la configuration des sources dans la fédération OpenArchaeo ;

# Administrer la fédération OpenArchaeo

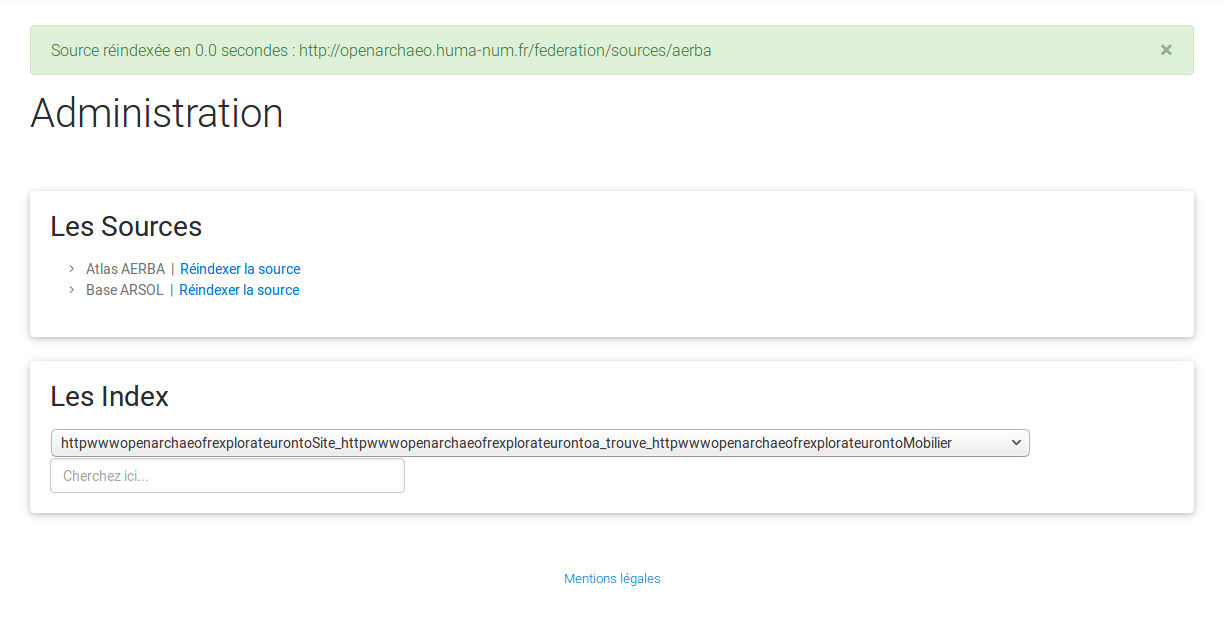
## Réindexer les données d’une source

Les recherche avec autocompletion fonctionnent sur la base d’index plein-texte Lucene qui doivent être précalculés à partir des données RDF.

Il est nécessaire de recalculer ces index si les données de la source changent.

Pour ce faire :

1. Se connecter dans la page d’admin de la fédération à <http://openarchaeo.huma-num.fr/federation/admin>
2. Cliquer sur le lien « Réindexer la source » en face de la source à réindexer ;
3. Attendre ***3 minutes sans recliquer***! un message s’affiche lorsque la réindexation est terminée ;



1. Tester la recherche dans un index en-dessous de l’écran : sélectionner un index, puis recherche quelque chose dans le champ de recherche ;

## Ajouter une source dans la fédération

Pour ajouter ou modifier une source :

1. Modifier le fichier de config federation/sources.ttl (cf la procédure générale de modification de la config) ;
2. Copier-Coller un bloc existant :

sd:namedGraph [

a sd:NamedGraph ;

# source IRI

sd:name <http://openarchaeo.huma-num.fr/federation/sources/aerba> ;

# source labels

dcterms:title "Aerba";

# source endpoint and graph

void:sparqlEndpoint <http://openarchaeo-graphdb.huma-num.fr/repositories/openarchaeo> ;

fed-config:sparqlGraph <http://aerba.univ-tours.fr> ;

] ;

1. Les champs de description d’une source sont les suivants :
   1. sd:name : l’URI de la source telle qu’elle sera interrogeable en SPARQL, et telle qu’elle sera exposée « publiquement » ; cette URI devra être utilisée dans les clauses « FROM » du SPARQL pour interroger cette source ;
   2. dcterms:title et toutes les autres métadonnées dcterms : les métadonnées dcterms présentes dans ce fichier seront ajoutées aux métadonnées lues dans la source pour l’affichage dans l’écran de choix des sources dans l’explorateur ;
   3. void:sparqlEndpoint : l’URL du point d’accès SPARQL des données de cette source ; pour le triplestore OpenArchaeo, l’URL est <http://openarchaeo-graphdb.huma-num.fr/repositories/openarchaeo>
   4. fed-config:sparqlGraph : le graphe nommé auquel restreindre les requêtes dans le point d’accès SPARQL. **Dans le cas du triplestore OpenArchaeo, il faut indiquer ici le graphe nommé dans lequel ont été enregistré les données dans le triplestore** ;
2. Commiter et Pousser les modifications au serveur ;
3. Recalculer les index de recherche sur cette source (cf. procédure correspondante) ;
4. Tester une interrogation SPARQL sur cette nouvelle source ;

## Editer les Collections de sources dans la Fédération

Les sources de la Fédération sont organisées en Collections. Ces Collections sont éditoriales et éditées à la main. L’organisation des sources en Collection se fait dans le fichier config.ttl.

Pour modifier l’organisation des sources en Collections :

1. Modifier le fichier de config federation/sources.ttl (cf la procédure générale de modification de la config) ;
2. Copier-coller ce bloc pour créer une nouvelle Collection :

<http://openarchaeo.huma-num.fr/collections/all> a skos:Concept ;

skos:prefLabel "Jeux de données OpenArchaeo"@fr ;

skos:prefLabel "OpenArchaeo dataset"@en ;

skos:inScheme <http://openarchaeo.huma-num.fr/collections> ;

.

Bien ajuster la fin de l’URI en remplçant « /all » par autre chose, et les libellés français et anglais.

1. Ajuster l’organisation des sources en Collection en indiquant l’URI de la ou des Collections dans lesquelles la source est catégoriser :

sd:namedGraph [

a sd:NamedGraph ;

…

# collection

dcterms:subject <http://openarchaeo.huma-num.fr/collections/all> ;

] ;

Au besoin, dupliquer cette ligne pour indiquer plusieurs collections.

1. Commiter et Pousser les modifications au serveur ;

## Modifier les requêtes SPARQL d’exemple dans la fédération

Pour ajouter ou modifier une requête SPARQL d’exemple dans le formulaire d’interrogation :

1. Modifier le fichier de config federation/query\_examples.json (cf la procédure générale de modification de la config) ;
2. Copier-Coller un bloc existant :

{

"titles": {

"fr": "ARSOL/AERBA : Liste de tous les E27\_Site",

"en": "ARSOL/AERBA : List all E27\_Site"

},

"query": [

"PREFIX wficheweb: <http://arsol.univ-tours.fr/4DACTION/WFICHEWEB/>",

"PREFIX crmsci: <http://www.ics.forth.gr/isl/CRMsci/>",

"PREFIX crmba: <http://www.ics.forth.gr/isl/CRMba/>",

"PREFIX crmarch: <http://www.ics.forth.gr/isl/CRMarchaeo/>",

"PREFIX crm: <http://www.cidoc-crm.org/cidoc-crm/>",

"PREFIX owl: <http://www.w3.org/2002/07/owl#>",

"PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>",

"PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>",

"PREFIX skos: <http://www.w3.org/2004/02/skos/core#>",

"SELECT ?site ?nom",

"FROM NAMED <http://openarchaeo.huma-num.fr/federation/sources/arsol>",

"FROM NAMED <http://openarchaeo.huma-num.fr/federation/sources/aerba>",

"WHERE {",

" ?site a crm:E27\_Site .",

" # AERBA n'a pour l'instant pas les libellés",

" OPTIONAL { ",

" ?site skos:prefLabel ?nom .",

" }",

"}"

]

},

1. Modifier la requête en conséquence ;
   1. ATTENTION aux guillemets et aux virgules en fin de ligne ;
   2. ATTENTION si le bloc est le dernier de la liste, la virgule finale (après la « } » fermante) ne doit pas être incluse ;
   3. EN CAS DE DOUTE tester avec un validateur JSON en ligne tel que <https://jsonlint.com/?code=>
2. Commiter et Pousser les modifs

## Mettre à jour le référentiel Periodo

Pour modifier le référentiel PeriodO :

1. Modifier le fichier de config federation/periodo.jsonld (cf la procédure générale de modification de la config), en téléchargeant un autre fichier de PeriodO ; le fichier est exactement celui téléchargeable depuis PeriodO ;
2. Commiter et Pousser les modifs ;

# Administrer l’explorateur OpenArchaeo

## Modifier le contenu éditorial de la page d’accueil ou des mentions légales

Pour modifier les contenus éditoriaux :

1. Modifier les fichiers de config suivants :
   1. explorateur/home\_en.htm et explorateur/home\_fr.htm pour la page d’accueil en anglais et français ;
   2. explorateur/legalnotice\_en.htm et explorateur/legalnotice\_fr.htm pour les mentions légales en anglais et français ;
2. Tester que le HTML est valide dans un validateur HTML en ligne comme <https://validator.w3.org/#validate_by_input+with_options> en sélectionnant l’option « Validate HTML fragment »
3. Commiter et Pousser les modifs ;

# Administrer le serveur

## Consulter les fichiers de config

Les fichiers de config sont consultables et modifiables via le repository Git.

Sur le serveur :

* le repository Git est dans /home/git/openarchaeo-config.git/
* les fichiers de config effectivement utilisés par l’application sont dans /var/lib/tomcat8/openarchaeo/config-openarchaeo/ (le repository est dupliqué dans ce répertoire à chaque modification).

## Consulter les logs

Les logs de l’application sont dans /var/lib/tomcat8/logs/catalina.out

## Arrêter / redémarrer les services

### Via le git

Pour redémarrer les services, le plus simple est de pousser une modification dans les fichiers de config via le Git qui déclenchera un redémarrage des services.

Pour cela, modifier le fichier **testgit.md** à la racine du répertoire de configuration (ce fichier n’est pas utilisé, on peut introduire n’importe quelle modif dedans), puis commiter et pousser les modifications.

### En ligne de commande

ssh [tfrancart@openarchaeo.huma-num.fr](mailto:tfrancart@openarchaeo.huma-num.fr)

sudo su –

service tomcat8 restart

## Nettoyer les logs

Pour nettoyer les logs :

1. Arrêter le serveur tomcat
2. Supprimer /var/lib/tomcat8/logs/catalina.out
3. Redémarrer le serveur tomcat