**Génération des vocabulaires MIMO au format SKOS**

2 vocabulaires générés :

* La classification Sachs & Hornbostel,
* Le thésaurus des noms d’instruments de musique.

Les versions à jour des répertoires et programmes mentionnés ci-dessous se trouvent sur le serveur de production MIMO (194.250.19.158).

La version de Python à utiliser est la 2.7.

Etape 1

* Générer un export complet des thésaurus depuis IDESIA dans un fichier nommé Export\_Idesia\_initial.xml,
* Copier ce fichier dans le répertoire C:\Program Files (x86)\Common Files\Archimed\VocabulairesMIMO\01\_idesia,
* Lancer le programme clean\_Export\_Idesia\_initial.py : génération du fichier Export\_Idesia\_clean.xml,
* Via un outil tel que oXygen ou tout autre moyen permettant d’exécuter des transformations XSL, appliquer les transformations extract\_Export\_Idesia\_HS.xsl et extract\_Export\_Idesia\_Keywords.xsl pour générer les fichiers hs\_01.xml et keywords\_01.xml.

Etape 2

* Se placer dans le répertoire C:\Program Files (x86)\Common Files\ VocabulairesMIMO\02\_cleaned,
* Lancer les programmes clean\_xml\_hs.py et clean\_xml\_keywords.py : génération des fichiers hs\_02.xml et keywords\_02.xml.

Etape 3

* Se placer dans le répertoire C:\Program Files (x86)\Common Files\03\_skos,
* Via un outil tel que oXygen ou tout autre moyen permettant d’exécuter des transformations XSL, appliquer les transformations hs2skos.xsl et keywords2skos.xsl aux fichiers hs\_02.xml et keywords\_02.xml pour générer les fichiers hs\_skos.xml et keywords\_skos.xml,
* En appliquant les transformations skos2rdf.xsl et skos2jsonld.xsl sur les fichiers hs\_skos.xml et keywords\_skos.xml, on peut générer si besoin est les vocabulaires aux formats RDF et JSON-LD.

Il existe des outils en ligne qui permettent de vérifier la bonne composition des vocabulaires tels que SKOSIFY (<http://demo.seco.tkk.fi/skosify/skosify> - fournir en entrée les fichiers au format SKOS) ou encore SPARNA (<http://labs.sparna.fr/skos-play> - fournir en entrée les fichiers au format RDF).

A noter que SPARNA fournit un outil permettant de générer un PDF (avec navigation interne entre liens) des vocabulaires.

**Chargement des vocabulaires sur Datalift**

Etape 1 : nettoyer les anciennes versions des vocabulaires

* Se connecter à la plateforme Virtuoso (<http://10.0.2.8:8890>) sur laquelle s’appuie Datalift,
* Cliquer sur « Conductor » (identifiant « dba » et mot de passe « dba » demandés),
* Cliquer sur l’onglet « Linked Data »,
* Cliquer sur l’onglet « Graphs »,
* Cliquer sur l’onglet « Graphs » secondaire,
* Cliquer sur le bouton « Delete » en regard des vocabulaires <http://www.mimo-db.eu/HornbostelAndSachs> et <http://www.mimo-db.eu/InstrumentsKeywords> et confirmer la suppression pour chacun.

Etape 2 : charger les vocabulaires sur Datalift

* Se connecter à la plateforme Datalift (<http://10.0.2.8:8180/project#description> – nom d’utilisateur « datalift » et mot de passe « test » demandés),
* Cliquer, si nécessaire, sur la case « Espace projets » puis sur « MIMO » dans la liste des projets,
* Cliquer sur l’onglet « Sources »,
* Supprimer les deux transformations RDF hs\_skos.xml (RDF#1) et keywords\_skos.xml (RDF#1) (une suppression se fait en cliquant sur la source pour la sélectionner puis en cliquant sur l’icône poubelle),
* Supprimer les deux fichiers hs\_skos.xml et keywords\_skos.xml,
* Recharger les nouvelles versions des fichiers hs\_skos.xml et keywords\_skos.xml (cliquer sur le signe « + » et suivre la démarche proposée),
* Une fois les fichiers rechargés, cliquer sur l’onglet « Description »,
* Cliquer sur la case « Transformation XML vers RDF » :
  + Choisir dans la liste proposée la source à convertir (hs\_skos.xml puis keywords\_skos.xml),
  + Veiller que le nom de la source à convertir soit hs\_skos.xml (RDF#1) ou keywords\_skos.xml (RDF#1),
  + Veiller à ce que l’URI du graphe nommé associé soit http://www.mimo-db.eu/HornbostelAndSachs ou http://www.mimo-db.eu/InstrumentsKeywords selon le cas,
  + Veiller à ce que la valeur de l’URI de base soit http://www.mimo-db.eu,
  + Le champ « Modèle de transformation » doit être fixé à « Document entier »,
  + Cliquer sur « Valider ».
* Une fois les transformations effectuées, cliquer sur l’onglet « Description »,
* Cliquer sur la case « Publication vers le RDF store public » :
  + Veiller que le nom de la source à convertir soit hs\_skos.xml (RDF#1) ou keywords\_skos.xml (RDF#1),
  + Veiller à ce que l’URI du graphe nommé associé soit http://www.mimo-db.eu/HornbostelAndSachs ou http://www.mimo-db.eu/InstrumentsKeywords
  + Cliquer sur « Valider ».

**Chargement des vocabulaires sur Skosmos**

Skosmos est un outil Open Source de gestion de vocabulaires. Les vocabulaires sont consultables via une interface web et interrogeables via une API.

L’outil est installé sur une machine virtuelle LINUX (Debian) accessible à l’adresse suivante 10.0.2.30.

Pour accéder à la machine et manipuler les fichiers, il est nécessaire d’installer sur son poste l’utilitaire « puTTY » ainsi que « WinSCP ».

Etape 1 : nettoyer les anciennes versions des vocabulaires

* Se connecter à la machine virtuelle via puTTY (host name : 10.0.2.30 ; port : 22 ; type de connexion : ssh),
* S’identifier (identifiant « ced » et mot de passe « dzalkmz103 » demandés) puis passer en super-utilisateur (su mot de passe « fZ4fghnb »),
* Se positionner dans le répertoire /etc/skosmos,
* La base de données Skosmos se trouve dans le répertoire /etc/apache-jena-fuseki/run/databases/skosmos,

On peut connaître sa taille par la commande :

du /etc/apache-jena-fuseki/run/databases/skosmos.

Les commandes de récupération, suppression, chargement d’un vocabulaire font appel au protocole soh ([Sparql over http](https://jena.apache.org/documentation/fuseki2/soh.html" \l "soh-sparql-http)). Les binaires de ces commandes sont contenus dans le répertoire /etc/apache-jena-fuseki/bin.

Pour récupérer un vocabulaire de la base dans un fichier :

/etc/apache-jena-fuseki/bin/s-get http://localhost:3030/skosmos/data http://www.mimo-db.eu/*vocabulaire*/ > nom\_fichier.ttl (vocabulaire = InstrumentsKeywords ; HornbostelAndSachs ou tout autre vocabulaire qui sera chargé dans Skosmos). Le fichier récupéré est au format Turtle (extension ttl).

Pour supprimer un vocabulaire :

/etc/apache-jena-fuseki/bin/s-delete http://localhost:3030/skosmos/data http://www.mimo-db.eu/*vocabulaire*/

Pour charger un vocabulaire :

/etc/apache-jena-fuseki/bin/s-put http://localhost:3030/skosmos/data http://www.mimo-db.eu/*vocabulaire*/ nom\_fichier

nom\_fichier est le nom du fichier contenant le vocabulaire. Le contenu de ce fichier doit être au format Turtle ou RDF.

A chaque chargement d’un vocabulaire, il faut :

* Arrêter Fuseki (se positionner dans /etc/apache-jena-fuseki et y exécuter la commande ./fuseki stop),
* Arrêter Apache (sudo service apache2 stop),
* Charger le vocabulaire,
* Redémarrer Apache (sudo service apache2 start),
* Redémarrer Fuseki (se positionner dans /etc/apache-jena-fuseki et y exécuter la commande ./fuseki start).

**API Skosmos**