Ces fichier SPARQL permettent de nettoyer les données du fichier XML http://librairies.ign.fr/geoportail/resources/IGNF.xml une fois celles-ci converties en RDF avec la plateforme Datalift. Le « nettoyage » se fait en plusieurs étapes car la plateforme supporte mal le traitement simultané de trop grandes quantités de données.

Les données sont remises en forme (identifiants convertis en URIs, valeurs converties en littéraux, etc.) et typées à l'aide des ontologies ignf et qudt.

Il faut appliquer ces requêtes dans l'ordre suivant:

- sur le jeu de données RDF brutes:

- GeographicCRS.sparql

- GeocentricCRS.sparql

- ProjectedCRS.sparql

- VerticalCRS.sparql

- CompoundCRS.sparql

- Datum1.sparql

- Conversion1.sparql

- Methode1.sparql

- Parameter.sparql

- Transformation1.sparql

- sur le jeu de données datum1: Datum2.sparql

- sur le jeu de données datum2: Datum3.sparql

- sur le jeu de données conversion1: Conversion2.sparql

- sur le jeu de données conversion2: Conversion3.sparql

- sur le jeu de données methode1: Methode2.sparql

- sur le jeu de données transformation1: Transformation2.sparql

- sur le jeu de données transformation2: Transformation3.sparql

Puis exporter en Turtle les fichiers RDF correspondant aux jeux de données suivants:

- GeographicCRS

- GeocentricCRS

- ProjectedCRS

- VerticalCRS

- CompoundCRS

- Datum3

- Conversion3

- Methode2

- Parameter

- Transformation3

Enfin, coller l'ensemble des données dans un même fichier Turtle qui pourra être chargé directement à l'aide de la plateforme (à créer avec Notepad++ par exemple). Voir chargement en Figure 1.

NB : Attention à bien remettre les bons préfixes (perdus lors de l'export des données depuis la plateforme):

@prefix ignf: <http://data.ign.fr/def/ignf#>.

@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#>.

@prefix qudt: <http://qudt.org/schema/qudt#>.

@prefix unit: <http://qudt.org/vocab/unit#>.

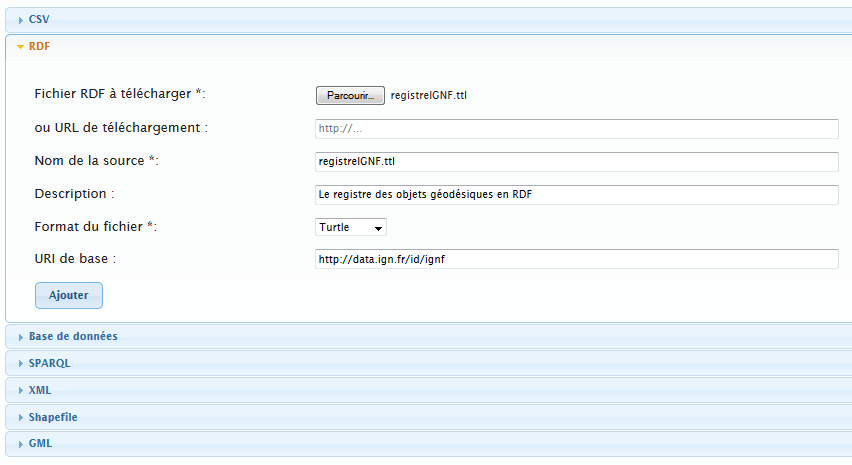


Figure 1: Chargement des données via la plateforme