ETAPE 01

Traitement de texte sur un fichier .rdf

Principales taches :

- Parcourir les fichiers avec extension .rdf automatiquement et création des fichiers de sortie avec le même nom de fichier + la mention _DB (pour doblon)

```
Dossier courant :/auto_home/nbendjoudi/TER/TER2/mop

liste des fichiers à traiter :
mop-iaml-mimo_valide.rdf
genres_itema3_iaml-valide.rdf
```

- Le programme crée automatiquement un dossier qui contiendra les nouveaux fichiers
- Tous les blocs contenant des relations valides et dont entity1 ou entity2 sont présent au minimum deux fois

```
Traitement de :mop-iaml-mimo_valide.rdf
mop-iaml-mimo_valide_DB.rdf : créé dans /auto_home/nbendjoudi/TER/TER2/mop_DB  <<<<<<<<<<<<<<<

lien présent : http://data.doremus.org/vocabulary/iaml/mop/tha : 2  fois sauvegardé.
lien présent : http://www.mimo-db.eu/InstrumentsKeywords/2281 : 2  fois sauvegardé.
lien présent : http://www.mimo-db.eu/InstrumentsKeywords/2230 : 2  fois sauvegardé.
lien présent : http://data.doremus.org/vocabulary/iaml/mop/mmi : 2  fois sauvegardé.
Veuillez patienter SVP!
ecriture de mop-iaml-mimo_valide_DB.rdf en cours

Liens entity1 présents au min 2 fois

Liens entity2 présents au min 2 fois
```

- Éliminer tous les blocs de balises <map>..</map> qui contiennent la ligne <relation>notvalid</relation>

```
<map>
  <Cell>
    <entity1 rdf:resource='http://www.mimo-db.eu/InstrumentsKeywords/3395'/>
    <entity2 rdf:resource='http://data.doremus.org/vocabulary/diabolo/mop/bandurria'/>
    <relation>notvalid</relation>
    <measure rdf:datatype='http://www.w3.org/2001/XMLSchema#float'>0.86</measure>
  </Cell>
</map>
<map>
  <Cell>
    <entity1 rdf:resource='http://www.mimo-db.eu/InstrumentsKeywords/4126'/>
    <entity2 rdf:resource='http://data.doremus.org/vocabulary/diabolo/mop/cor_naturel'/>
    <relation>http://www.w3.org/2004/02/skos/core#exactMatch</relation>
    <measure rdf:datatype='http://www.w3.org/2001/XMLSchema#float'>1.0</measure>
  </Cell>
</map>
```

- Et sur le résultat de ce premier traitement, on ne garde que les blocs dont entity1|entity2 sont présents dans d'autres blocs aussi, au moins une fois.

```
<map>
  <Cell>
    <entity1 rdf:resource='http://www.mimo-db.eu/InstrumentsKeywords/3730'/>
    <entity2 rdf:resource='http://data.doremus.org/vocabulary/diabolo/mop/aerophone'/>
    <relation>http://www.w3.org/2004/02/skos/core#exactMatch</relation>
    <measure rdf:datatype='http://www.w3.org/2001/XMLSchema#float'>1.0</measure>
  </Cell>
</map>
<map>
  <Cell>
    <entity1 rdf:resource='http://www.mimo-db.eu/InstrumentsKeywords/4310'/>
    <entity2 rdf:resource='http://data.doremus.org/vocabulary/diabolo/mop/aerophone'/>
    <relation>http://www.w3.org/2004/02/skos/core#exactMatch</relation>
    <measure rdf:datatype='http://www.w3.org/2001/XMLSchema#float'>1.0</measure>
  </Cell>
</map>
```

- Ecrire un fichier final_DB.rdf avec les résultats souhaités et en rajoutant l'entete et le pieds de page xml.