

**TP encadré n°2 :**  
“Manipuler des données à partir  
de bases de données relationnelles, de fichiers XML et de base SQLite”

Exercice 1 : Base de données MySQL -> fichier .csv

a) Ecrivez un programme Java qui vous permette de vous connecter à la base de données MySQL `gmd` dont les paramètres de connexion sont donnés ci-dessous. Récupérez ensuite les données contenues dans la base et écrivez les dans un fichier texte en format `.csv` (Comma Separated Value) en suivant le format précisé ci-après.

*Les paramètres de connexion à la base de données sont :*

- *host* : `neptune.telecomnancy.univ-lorraine.fr`
- *database* : `gmd`
- *login* : `gmd-read`
- *pwd* : `esial`

Vous pouvez vous connecter depuis une machine de l'école où depuis n'importe quelle machine connectée en vpn au réseau de l'Université de Lorraine.

*Format de la ligne demandé pour le fichier CSV* : Id, 'Nom', 'Prénom', Age, Tel, Ville, Email  
*Exemple* : 1, 'Perinova', 'Raphael', 37, 047564738, Moscou, raphael.perinova@rinolav.ru

b) Ecrivez un programme Java qui vous permette de vous connecter à la base de données `gmd` et de récupérer toutes les données relatives aux personnes qui vivent à Firenze, Madrid et London. Une fois ces données obtenues, vous les écrivez dans un fichier de texte au format `.csv`. dont le format est précisé ci-dessous.

*Format de la ligne demandé pour le fichier CSV* : Ville;Id;'Nom Prénom';Age;Tel;Email

*Exemple* : Firenze;15;'Rossi Manuela';25;044556732;manurossi@libero.it

Exercice 2 : Fichiers XML -> fichier .csv

Ecrivez un parseur de fichiers XML qui, en utilisant la *Simple API for XML (SAX)*, produira un fichier texte `annuaire.csv` contenant tous les données des trois fichiers XML suivants :

- `italie.xml`
- `espagne.xml`
- `angleterre.xml`

Ces trois fichiers ont la même DTD qui est « `annuaire.dtd` ».

*Format de la ligne demandé pour le fichier CSV* : Ville, Id, Age, 'Nom Prénom', Email, Tel

*Exemple* : London, 15, 33, 'Smith John', jsmith1967@btelecom.co.uk, 022569741

Exercice 3 : Fichiers .csv -> base de donnée SQLite

Le script Python ci-dessous permet d'extraire le contenu d'un fichier csv (que vous pouvez voir à /depot/2A/gmd/tp2/clinicalAnnotations.csv) et de le charger dans une base de données SQLite.

```
#!/usr/bin/python

import csv, sqlite3

conn = sqlite3.connect("./base/pgx.db")
curs = conn.cursor()
curs.execute("DROP TABLE ClinicalAnnotation;")
curs.execute("CREATE TABLE ClinicalAnnotation (variant TEXT, gene
TEXT, type TEXT, strength_of_evidence_level TEXT, drugs TEXT,
diseases TEXT);")
list_of_tuple = []

with open('./clinicalAnnotations.csv','r') as infile:
    # 'r' because the file is in read only
    # csv.DictReader uses first line in file for column headings by
default
    dr = csv.DictReader(infile, delimiter=',')
    for i in dr:
        list_of_tuple.append( (i['variant'], i['gene'],i['type'],
i['strength_of_evidence_level'],i['drugs'], i['diseases']) )

curs.executemany("INSERT INTO ClinicalAnnotation (variant, gene,
type, strength_of_evidence_level, drugs, diseases) VALUES
(?,?,?,?,?,?);", list_of_tuple)

conn.commit()
curs.close()
```

- a) Lisez et comprenez le script précédent.
- b) Utilisez le pour écrire un nouveau script Python qui effectue la même opération à partir du fichier `annuaire.csv` généré lors de l'exercice 2.
- c) Ecrivez un script Python qui retourne, comme dans la question 1.b), la liste les habitants de Firenze, Madrid et London.