

ANNEXE : DOCUMENTATION D'INSTALLATION

Cette section a pour but de guider à travers l'installation des différents logiciels requis pour pouvoir manipuler le code source.

1) Installation de l'IDE

La première étape est de faire en sorte d'installer l'IDE de votre choix. Par pure expérience, je vous conseille d'installer :

- **Visual Studio Code** sur les distributions linux (ou bien son clone, CODE OSS si le premier n'est pas disponible)
- **PyCharm** sous Windows 10 (Visual Studio Code provoque un bug lors de l'importation du module Numpy à la date du 17/05/2020 sous W10. Si ce dernier est corrigé, préférez VS Code).

Privilégiez l'installation par le terminal sous Linux si possible (sudo apt install sous Ubuntu, sudo apt-get sous Debian, sudo pacman -S sous Arch, etc...)

2) Installation de GIT

- Sous Windows : se rendre sur le site officiel de Git :
<https://git-scm.com/>
Et télécharger la dernière version, puis l'installer.
- Sous Linux : installer Git via le terminal :
sudo apt install git (Ubuntu)

3) Clonage du dépôt NLPvs

- Sous Windows : taper "git" dans la barre de recherche de Windows 10, et démarrer le terminal de git.
- Sous Linux : démarrez simplement le terminal système

Ensuite, pour les deux OS :

- entrer la commande suivante pour choisir le répertoire d'accueil du dossier "NLPvs" qui sera cloné depuis GitHub :
cd dossier_choisi (remplacer "dossier choisi" par le chemin d'accès du dossier d'accueil souhaité. Par exemple, pour le bureau, la commande sera **cd Desktop** ou **cd Bureau**)
- entrer la commande suivante :
git pull https://github.com/VincDub/NLPvs

Un dossier nommé NLPvs apparaîtra alors, et contient les fichiers sources du dépôt (il sera possible de mettre à jour le clone en réitérant la commande dès qu'une nouvelle version sort).

4) Installation de Miniconda

Miniconda est une version sans interface graphique d'Anaconda, et qui est donc bien plus légère, et s'utilise uniquement par le biais d'un terminal.

Sous Windows :

- Se rendre sur la page web de téléchargement de miniconda:
<https://docs.conda.io/en/latest/miniconda.html>
et télécharger la dernière version en .exe

Sous Linux :

- Installer miniconda via le terminal

OU
- Se rendre sur la page web de téléchargement de miniconda:
<https://docs.conda.io/en/latest/miniconda.html>
et télécharger la dernière version en .sh
- une fois le .sh téléchargé, se rendre dans le dossier où il se trouve via le terminal (avec **cd dossier**)
- lancer la commande suivante :
sudo sh nom_du_sh.sh (remplacer nom_du_sh par le nom de l'archive)

5) Installation des modules

Sous Windows :

- Taper "Anaconda" dans la barre de recherche du bureau, et ouvrir le terminal d'Anaconda

Sous Linux :

- Ouvrir simplement le terminal système

Pour les deux OS , les commandes d'installation sont les mêmes :

- PDFMiner
pip install pdfminer.six
- NLTK
conda install NLTK
- French Leff Lemmatizer
pip install git+https://github.com/ClaudeCoulombe/FrenchLefffLemmatizer.git
- Pandas
conda install pandas
- Scikit Learn
conda install scikit-learn
- Numpy
conda install -c anaconda numpy
- CuML (!\ Ne fonctionne pas sous Windows)
conda install -c rapidsai -c nvidia -c conda-forge -c defaults cuml=0.13
- Matplotlib
conda install matplotlib
- Networkx
conda install networkx
- Pillow
conda install -c anaconda pillow

6) Configuration de l'IDE

Lancez votre IDE :

- Ouvrez le script **import_nltk_stopwords.py**
- Sous **Visual Studio Code**, changez la version de python en bas à gauche de l'interface, et choisir celle d'Ananconda (3.7.6 à la date du 17/05/2020).
- Sous **PyCharm**, le changement est automatique
- (Optionnel) : Sous VS Code, installer l'aide à la rédaction du langage python
- Exécutez le script : Si tout se passe bien, il devrait vous indiquer que le téléchargement s'est bien effectué.
- Fermez le script
- Ouvrez **source.py**
- Lancez-le : Si l'interface apparaît, le lancement a réussi. Vous pouvez désormais l'essayer comme bon vous semble.