

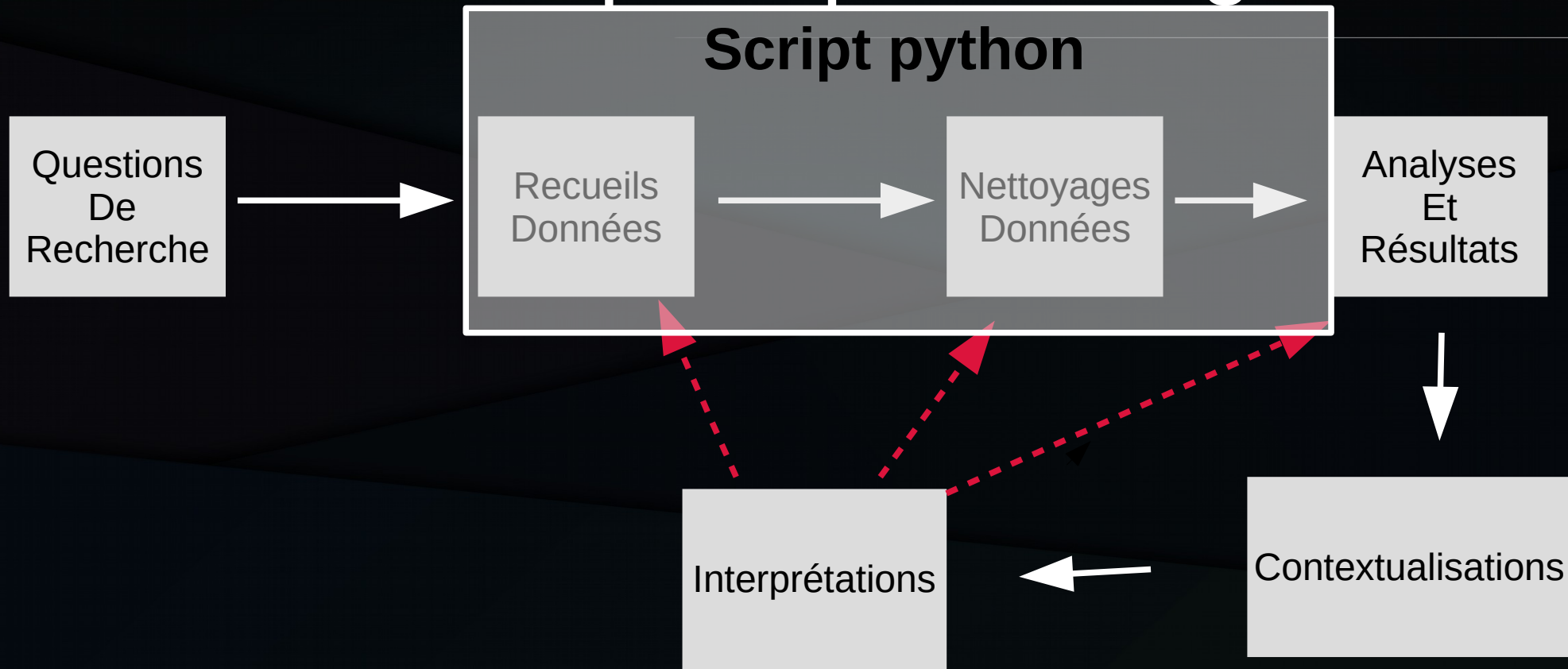
# Phyton en SHS... complémentarités avec la lexico

---

Lucie Loubère, Urfist Lyon

le 12/05/2022

# Focale de mes principaux usages...



# Pourquoi je fais des scripts ???

- Parce que
  - Je peux le faire (gain de temps vs manuel)
  - Les données s'y prêtent et/ou m'y obligent
  - Je ne connais pas d'outil qui ferait exactement ce dont j'ai besoin
  - Les modèles d'analyses me donnent des résultats qu'en fonction de ce que je rentre comme données.

# Pourquoi j'utilise Python ???

- Invitation de mon entourage scientifique
- Langage accessible
- Pertinent pour la gestion des textes
- Nombreuses librairies pré-existantes

# Mes pratiques sur python... pour 80 %

- Ouvrir et lire un fichier .txt

```
16
17 Fichier=open(corpus,'r')
18 Read=Fichier.readlines()
19 Fichier.close()
20
```

- Le lire ligne par ligne et modifier quelque chose

```
10 for line in read :
11     if line.startswith('****') :
12         x=x+1
13         line=line.replace('****','**** *id_'+str(x))
```

- Ouvrir un nouveau fichier et y écrire quelque chose

```
Fichier=open("corpus_id.tx","a")
Fichier.write(line)
Fichier.close
```

# Mes pratiques sur python... pour 15 %





# Quelques exemples d'utilisations

- Ajouter un ID à un corpus pré-existant (10mn)
- Nettoyer une liste d'expression pour la faire ressortir ou pour la neutraliser (10 à 30mn)
- Créer à partir de pdf un corpus en .txt avec ligne étoilée reprenant des variables du .pdf ou d'un fichier (15mn)
- Ajouter une variable dans un corpus (10mn)
- Parser un site bien structuré (1h00)
- Parser Facebook... (temps indéterminé).

# Mes librairies préférés...

- Csv : permet de lire un csv et de le manipuler
- Sélénium : manipule un navigateur web
- BeautifulSoup : permet de lire et prendre du html



- Merci pour votre attention...

