


# UE2

## APPROFONDISSEMENTS

### (2 COURS AU CHOIX, AM2 ET AT2)

---

 ECTS  
8 crédits

 Composante  
École  
d'économie  
de la  
Sorbonne  
(EES)

 Période de  
l'année  
Printemps

#### Liste des enseignements

ObligatoireChoixÀ choix11ChoixMatière5.024.0Matière4.019.5Matière18.0Matière4.030.0Matière4.024.0Cette formation vise à doter les participants-es des connaissances et des savoir-faire nécessaires à l'analyse d'un corpus de sources textuelles à l'aide des méthodes de l'analyse des données textuelles. Nous aborderons les techniques d'acquisition et de construction de corpus : reconnaissance optique de caractères manuscrits (HTR) et imprimés (OCR), les méthodes de structuration de corpus, de lemmatisation et d'étiquetage morpho-syntaxique du vocabulaire.Seront également enseignées Les différentes méthodes de mesures lexicales et d'interrogation de corpus étiquetés couramment utilisées en textométrie, les analyses multivariées et les méthodes de classifications ainsi que les méthodes cooccurrenceelles permettant l'étude des relations entre les mots. Une attention sera enfin portée aux techniques venues du domaine du traitement automatique des langues et du machine learning. Pour ce faire différents logiciels pourront être mobilisés : Lexico, TXM, Iramuteq, Hyperbase, des paquets R ou bien Python.Matière3.018.0Matière4.018.0Matière19.5Matière19.5Matière4.018.0Matière4.018.0ChoixÀ choix11Matière18.0ChoixMatière3.018.0Matière5.024.0Matière18.0Matière4.018.0Matière3.018.0Matière3.024.0Matière18.0Matière24.0Matière