

# Traitement automatique de corpus

## Titulaire

Max DE WILDE (Coordonnateur)

## Mnémonique du cours

STIC-B545

## Crédits ECTS

5 crédits

## Langue(s) d'enseignement

Français

## Période du cours

Premier quadrimestre

## Campus

Solbosch

## Contenu du cours

Le cours comprendra les éléments suivants: méthodes et outils divers pour la création d'un corpus textuel, notamment dans le contexte du Web; introduction à Python et NLTK pour la transformation de données; usage de diverses méthodes de type "machine learning", à la fois supervisées (p. ex. linear regression) et non-supervisées (Word2Vec, Topic Modeling) pour la classification de l'information et l'extraction de concepts.

## Objectifs (et/ou acquis d'apprentissages spécifiques)

Fournir à l'étudiant les connaissances conceptuelles et techniques en matière de gestion de corpus textuels numériques.

## Méthodes d'enseignement et activités d'apprentissages

Combinaison d'exposés "ex cathedra" et de travaux pratiques.

## Contribution au profil d'enseignement

Cet enseignement permet d'appréhender le traitement de corpus structurés et non-structurés de manière automatisée. Il complète les autres cours du Master en STIC en se focalisant plus particulièrement sur la gestion de larges collections de documents textuels via des méthodes programmatiques.

## Références, bibliographie et lectures recommandées

Le support de cours, disponible sur l'UV, propose des références bibliographiques.

## Support(s) de cours

Université virtuelle

## Autres renseignements

### Lieu(x) d'enseignement

Solbosch

### Contact(s)

Max De Wilde (max.de.wilde@ulb.be)

## Méthode(s) d'évaluation

Examen oral, Travail pratique et Rapport écrit

### Construction de la note (en ce compris, la pondération des notes partielles)

50% pour l'examen oral, et 50% pour les travaux pratiques (y compris le rapport final)

### Langue(s) d'évaluation principale(s)

Français

### Autre(s) langue(s) d'évaluation éventuelle(s)

Anglais

## Programmes

### Programmes proposant ce cours à la faculté de Lettres, Traduction et Communication

MA-LING | **Master en linguistique** | finalité Spécialisée/bloc 2, MA-STIC | **Master en sciences et technologies de l'information et de la communication** | finalité Spécialisée/bloc 2 et MA-TRAD | **Master en traduction** | finalité Spécialisée/bloc 2 et finalité Approfondie/bloc 2