

Web

http://ceditec.u-pec.fr
http://textopol.u-pec.fr

Contact

jean-marc.leblanc@u-pec.fr

Intervenants

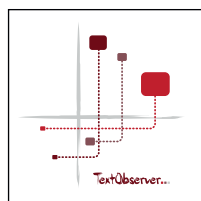
Philippe Gambette, Claude Martineau, Emilie Née, Marie Pérès

Localisation

Salle multimédia, I2-317 / i2 306 bâtiment i - Campus centre (CMC) 61 avenue du Général de Gaulle 94010 Créteil Cedex

Horaire

samedi 10h-13h
6 séances - 20 heures



Tutoriels, manuel, téléchargement :
http://textopol.u-pec.fr/textobserver/

Stages de formation à TextObserver

Préparation de corpus, prise en main sur les corpus des participants.

En semaine, deux séances (dates à préciser).

17 novembre 2018 : Introduction - Approches textométriques des discours.

J.M. Leblanc :
Présentation et typologie pratique de logiciels standard et d'analyseurs de données textuelles. Options théoriques, principes méthodologiques, limites interprétatives.

Initiation à **Lexico 3 / 5** :
- Fonctions documentaires, décomptes statistiques, modèles probabilistes.
- Distributions statistiques, distributions linguistiques.
- Analyse factorielle des correspondances, spécificités, fréquences...

15 décembre 2018 : Présentation et prise en main de TextObserver [10h-16h]

J.M. Leblanc :
- Expliciter l'analyse factorielle des correspondances
- Analyser la variation lexicométrique
- Introduction aux opérations de catégorisation
- Recension de corpus et balisage semi-automatisé : présentation de la base Textopol

19 janvier 2019 : Des corpus textuels aux corpus multimodaux (annoter, catégoriser, étiqueter, visualiser, interpréter).

J.M. Leblanc, M. Pérès :
- Transformer des textes pour les soumettre à des traitements automatisés. Repérer les régularités d'un document, extraire des motifs textuels, concaténer des fichiers.
- Prise en main de quelques catégoriseurs, évaluateurs, étiqueteurs (**cordial, treetager, tropes**)
- Base de données textuelles et outils de constitution et de balisage de corpus.
- Outils de visualisation, de caractérisation de corpus : **Gephi, R, Xlstat, Textstat.**

16 février 2019 : De la lexicométrie au traitement automatique des langues (TAL) [10h-16h]

P. Gambette (LIGM-MLV) : Les nuages arborés dans **TextObserver** et **Treecloud**.
C. Martineau (LIGM-MLV) : Présentation et prise en main du logiciel **UNITEX**.

PROGRAMME 2018-2019

16 mars 2019 : Exploration textométrique sur une base annotée - Prise en main du Trameur

E. Née (Céditec-UPEC) :
- **i-trameur**
- La logique **Trameur**
- Création d'une base étiquetée et intervention sur l'annotation (annotation manuelle, ajout de niveaux d'annotation).
- Présentation de quelques fonctionnalités textométriques classiques sur une base annotée : cooccurrences spécifiques, patrons, cartes des sections

13 avril 2019 : De la textométrie à l'analyse des données, quels outils pour quels usages?

J.M. Leblanc :
Cooccurrences généralisées et mondes lexicaux : analyses comparées **Alceste** et **Iramuteq**.
Ontologies et mondes sémantiques (**Tropes, Alceste, Astartex**)
Quantifier les données en sciences sociales : **Prospero, Nvivo** (sous réserve)

Corrélations et causalités interprétatives. Expérimentations, distance intertextuelle et voisinages.

J.M. Leblanc :
Distances, cooccurrences, voisinage
Présentation et prise en main d'**Hyperbase** : de la lexicométrie à la stylométrie.
Présentation d'**Hyperbase** en ligne.

Séances spécifiques

Certaines séances spécifiques sont organisées sur la journée entière (10h-16h).
Le samedi après-midi sera consacré aux questions des participants (sur rendez-vous).

S'inscrire à la liste de diffusion de Textopol

https://listes.u-pec.fr/www/subscribe/textopol

PROGRAMME 2018-2019