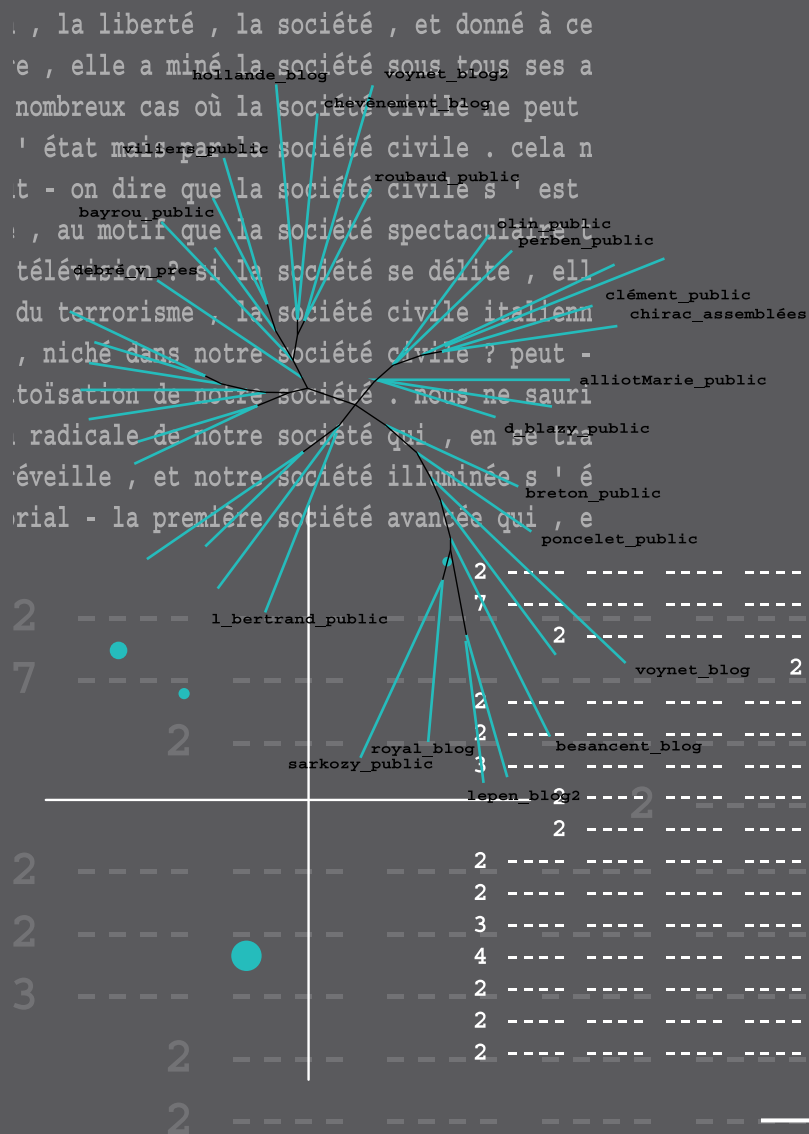


Cette formation doctorale constitue une introduction à l'analyse quantitative et qualitative des corpus textuels et à la sémantique discursive, qui s'adresse à tous les doctorants de SHS dont la recherche porte sur l'analyse de textes littéraires, le dépouillement d'enquêtes et de formulaires, l'analyse des discours sociaux et politiques. Chaque séance, outre la prise en main d'un logiciel et des problématiques théoriques qui s'y rattachent comprend des exercices pratiques en linguistique de corpus et en statistique textuelle qui permettent de tester les fonctionnalités, de répondre aux questions pratiques. Les doctorants et les étudiants en Master 2 de l'UPEC, ou d'autres universités dans la limite des places disponibles, peuvent y trouver des réponses concrètes dans leurs recherches individuelles ou collectives.

Les résultats produits par les outils textométriques intègrent désormais divers modules de mises en forme des données et des résultats demandant aux utilisateurs une maîtrise des outils de codage et de visualisation et une réflexion ergonomique. La formation comprend la mise en oeuvre d'outils infographiques s'appuyant sur les compétences bureautiques universitaires nécessaires à la constitution et gestion des corpus (recueil et documentation des données, exploitation et construction de bases de données) ainsi qu'une initiation à l'infographie. Ces notions sont présentées en fonction des demandes et du niveau des participants.



Formation Doctorale

INFORMATIQUE POUR LA RECHERCHE

Approches textométriques,
ergonomies numériques



Web

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

Contact

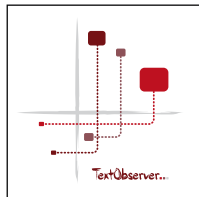
jean-marc.leblanc@u-pec.fr

Localisation

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

Horaire

samedi 10h-17h
6 séances - 35 heures



19 octobre 2013 : Introduction - Approches textométriques des discours.

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

Initiation à Lexico 3:

- Fonctions documentaires, décomptes statistiques, modèles probabilistes.
- Distributions statistiques, distributions linguistiques.
- Analyse factorielle des correspondances, spécificités, fréquences.....

23 novembre 2013: Construire et analyser des corpus textométriques.

Transformer des textes pour les soumettre à des traitements automatisés. Repérer les régularités d'un document, extraire de motifs textuels.

Approfondissement : Opérations sur le tableau lexical
Analyse factorielle des correspondances
Lexico3/TextObserver

14 décembre 2013: Des corpus textuels aux corpus multimodaux (annoter, catégoriser, étiqueter).

Forme graphique, catégories morphosyntaxiques, univers sémantiques
Prise en main de quelques catégoriseurs, évaluateurs, étiqueteurs (cordial, treetager, tropes)
Etude de cas : Analyser des corpus multimodaux.
Application de TextObserver à l'analyse automatisée de pages web.

Comprendre le web : Outils graphiques, développement Web, ergonomie des interfaces.

18 janvier 2014 : Visualisation et interprétations des résultats. Analyse de contenu et analyse de discours.

Ontologies et mondes sémantiques (Tropes, Alceste, Astartex)

Présentation et prise en main de deux outils open source: UNITEX et TreeCloud. (Claude Martineau, Philippe Gambette)

PROGRAMME 2013-2014

08 mars 2014: Corrélations et causalités interprétatives. Expérimentations, distance intertextuelle et voisinages.

Distances, cooccurrences, voisinages

Prise en main d'Hyperbase
Prise en main d'Alceste (2/2)

05 avril 2014 : De la textométrie à l'analyse des données, quels outils pour quels usages?

TXM, plate-forme ouverte pour le traitement des données textuelles.

Cooccurrences dans les textes, Weblex : Lexicogrammes récursifs, extraction de motifs textuels.

Prise en main d'un logiciel open source : Iramuteq
Outils de dépouillement d'enquêtes et de formulaires: Sphinx, Modalisa.

Ateliers infographie

En fonction des demandes des participants on proposera les ateliers suivants. (Voir détail sur le site Textopol)

Atelier 1 : Analyser et utiliser le web: installation, gestion, édition de dispositifs en ligne. Trois ateliers web.

Atelier 2 : Outils pour la recherche: Charte graphique et identité visuelle des supports de communication. Posters, diaporama, flyers...

Atelier 3 : Outils multimédia pour la publication. Edition numérique pour la recherche, intégrer le multimédia aux articles scientifiques.

mai 2014 : journée de présentation des outils de traitements de corpus textuels

Gramlab/Unitext, Treecloud, TextObserver/webObserver...
Echanges, présentations, ateliers. Outils développés par les équipes de recherche de Paris-Est.

Programme à définir. Organisation: Claude Martineau, Philippe Gambette, JM Leblanc

Le programme détaillé sera mis en ligne sur Textopol

PROGRAMME 2013-2014