

Ecole thématique d'été MISAT

Méthode d'analyse statistique en analyse textuelle

Nice - Villa Clythia, Fréjus (Var) – du 18 au 24 juin 2010

Consulter le site [MISAT](#)

Inscriptions entre le 5 février et le 15 avril

[contacts : D. Mayaffre (mavaffre@unice.fr) ; S. Mellet (mellet@unice.fr)]

PROGRAMME

Vendredi 18 juin :

18h30 : accueil, installation, pot de bienvenue (Sylvie Mellet, Jean-Marie Viprey)

Samedi 19 juin : Cadre et objectifs

- 9h : Introduction générale (Damon Mayaffre)
Ressources numériques, corpus, méthodes
- 10h30 : Pause café
- 11h : Présentation des stagiaires, de leurs corpus et de leurs attentes
- 12h15 : Pause déjeuner
- 14h : Présentation des formateurs et de leurs axes de recherche : tour de table exhaustif
- 15h30 : Pause
- 16h : Présentation plus approfondie de quelques recherches emblématiques
3 interventions de 45 minutes :
- *Etienne Brunet : les corpus littéraires et la classification des textes*
 - *André Salem : ressources textométriques*
 - *Serge Heiden et Bénédicte Pincemin : enjeux techniques actuels des traitements en ADT.*
- 19h : Pause et repas ; soirée libre.

Dimanche 20 juin : Des données brutes aux corpus ; texte et unités textuelles

- 9h : Introduction générale (Jean-Marie Viprey)
Lemmatisation, catégorisation, unités textuelles
- 10h30 : Pause café
- 11h : Ateliers DiaTag / Astartex (JM Viprey, Margareta Kastberg)
- 12h15 : Pause déjeuner
- 14h : Etablissement des données, étiquetage et balisage (Lou Burnard et François Daoust)
- 15h30 : Pause
- 16h : Ateliers Données Sato / Xaira (Lou Burnard et François Daoust)
- 19h : Pause et repas
- 20h30 : Discussion animée par Lou Burnard, François Daoust, Dominique Longrée, Sylvie Mellet, Bénédicte Pincemin, André Salem et Jean-Marie Viprey. Modératrice : Sylvie Mellet
Pourquoi et comment mettre des textes en corpus ? Les corpus numériques. La notion de corpus contrastif. Hétérogénéité vs homogénéité. Clôture et réflexivité des données. Etc. Avec prise en compte et examen des spécificités des corpus des stagiaires.

Lundi 21 juin : Méthodes et outils statistiques en analyse des données textuelles (I)

- 9h : Introduction générale (Ludovic Lebart)
Description et inférence statistique en analyse de textes
- 10h30 : Pause café
- 11h : Cours d'initiation pour les débutants (Margareta Kastberg, Sylvie Mellet)
Richesse lexicale, accroissement du vocabulaire, écart réduit, calcul de distance
Pour les non débutants, temps libre ; mais les formateurs seront à leur disposition pour répondre à toute demande de leur part
- 12h30 : Pause déjeuner

Après-midi libre avec animation proposée par la Villa Clythia (visite archéologique de Fréjus, découverte de l'Estérel).

Mardi 22 juin : Parcours textométriques

- 9h : Introduction générale (André Salem)
Les corpus structurés : séries textuelles chronologiques et corpus alignés
- 10h30 : Pause café
- 11h : Atelier Lexico / Trameur (Serge Fleury, William Martinez, André Salem)
- 12h15 : Pause déjeuner
- 14h : Lexicométrie, Textométrie et sémantique interprétative (Bénédicte Pincemin)
- 15h30 : Pause café
- 16h : Présentation de la plate-forme textométrique (Serge Heiden)
- 18h : Pause et dîner

Mercredi 23 juin : De quelques fonctionnalités spécifiques en ADT

- 9h : Cooccurrences et recherches thématiques (Etienne Brunet)
- 10h : Pause café
- 10h15 : Cooccurrences généralisées (Jean-Marie Viprey)
- 11 h : Pause café
- 11h15 : Poly-cooccurrences (William Martinez)
- 12h15 : Pause déjeuner
- 14h : Classification et partitionnement de textes : l'approche topologique (Dominique Longrée)
- 15 h30 : Pause
- 16h : Ateliers Hyperbase / Cooccs centrés sur les cooccurrences et corrélats (Etienne Brunet, William Martinez, Dominique Longrée, Céline Poudat)
- 19h : Pause et repas ; soirée libre

Jeudi 24 juin : Statistiques multidimensionnelles

- 9h : Initiation à l'AFC (Jean-Marie Viprey)
- 9h45 : Méthodes de bootstrapping (Ludovic Lebart)
- 10h30 : Pause café
- 11h00 : Ateliers Astartex (Jean-Marie Viprey) / DTM (Ludovic Lebart, Céline Poudat)
- 12h15 : Pause déjeuner
- 13h30 : Conférence de clôture (Jules Duchastel)
- 15h : premier bilan de fin de stage (retour oral et écrit *via* des fiches d'évaluation).