

3LTAY04T - Modélisation linguistique pour l'analyse automatique de textes

Devoir maison

- 1 dossier par personne (aucun travail en groupe ne sera accepté)
- 5 à 10 pages maximum, hors annexes
- à rendre au plus tard le 14/11 (envoi par mail en PDF ou impression papier)

À partir du TD 3 (et de l'exercice 14 du TD 2), vous avez dû constituer des automates finis avec le logiciel Unitex, dans le but de repérer les adverbiaux temporels calendaires et spatiaux (voire fréquentiels).

Dans un court dossier (5 à 10 pages maximum), vous présenterez votre travail. Après avoir brièvement rappelé la problématique et la démarche suivie (cf. TD2 et TD3), vous insisterez sur les points suivants :

- les ressources que vous avez créées, dictionnaires (à joindre en annexe) et graphes (capture d'écran). Vous pourrez aussi mentionner les ressources proposées par Unitex que vous avez réutilisées ;
- l'application de vos graphes sur un autre texte que celui sur de Villepin (quelques captures d'écran des concordanciers obtenus seront les bienvenues). À cette étape, vous quantifierez l'efficacité de votre système à l'aide des mesures de rappel, de précision et de la F-mesure ;
- les difficultés, les problèmes que vous avez rencontrés, ce que vous avez mis en œuvre pour les résoudre, les choix que vous avez opérés pour mener à bien l'exercice devront également être explicités. N'oubliez pas de parler des limites de votre travail et des améliorations possibles.

N.B. : sur les dossiers que vous me rendrez, mettez bien votre nom, votre prénom, votre université d'inscription principale et votre numéro d'étudiant.