



CorText Plateforme Digitale de l'IFRIS - <http://www.cortext.fr/>

Invitation à une Formation-Atelier le 19 et 20 décembre 2013

Marne La Vallée Bâtiment Bois de l'Etang

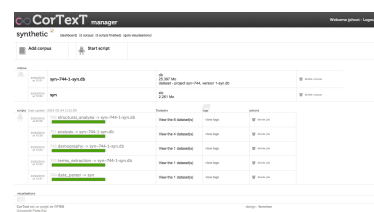
La plate-forme numérique de l'IFRIS combine plusieurs compétences scientifiques et techniques: le traitement du langage naturel, l'extraction d'information, gestion des connaissances, analyse des réseaux complexes, la scientométrie, conception de sites Web et l'informatique. Plateforme numérique de l'IFRIS, CorText a développé une solution logicielle pour répondre aux besoins des chercheurs menant des études empiriques dans les domaines des études sociales de la science. La plate-forme logicielle est une capacité d'appui méthodologique, de construction de jeux de données et d'analyse avec des fonctionnalités de visualisation des données. Outils et logiciels, conçus et développés au sein de la plate-forme, sont mis à la disposition des chercheurs via un service Web.

CorText propose aux chercheurs de l'IFRIS, et notamment aux doctorants et aux post-doctorants, de participer à une Formation-Atelier les 19 et 20 décembre 2013 toute la journée (fin à 17:00 le 20/12). Cette Formation-Atelier est proposée par Andréi Mogoutov qui animera ces journées avec l'appui de collègues de la plateforme. Merci d'exprimer votre intérêt et votre engagement sur 1 jour et demi à l'adresse suivante : equipe@cortext.fr. Le programme est proposé ci-après.

Pour la réussite de cette opération, il est nécessaire de réunir un nombre limité de participants (maximum une dizaine), aussi, en fonction des engagements à participer la Plateforme CorText donnera la priorité aux chercheurs qui n'ont pas encore eu la possibilité de travailler avec elle, et par ordre d'arrivée des candidatures sinon. Bien entendu cette initiative sera renouvelée.

J'en profite pour vous rappeler que tous les Jeudis, sur place à Marne La Vallée, vous êtes les bienvenu(e)s pour présenter vos projets, vos idées, vos envies. Prendre contact à : equipe@cortext.fr

Marc Barbier – Directeur de CorText - Plateforme Digitale de l'IFRIS



Objectif de la Formation-Atelier

La plate-forme logicielle offre un large éventail d'outils et de méthodes: Statistique / analyse exploratoire de données, outils d'analyse textuels/ text mining, la Cartographie du Web, la scientométrie, outils d'analyse de données qualitatives, la cartographie hétérogène et analyse structurale des réseaux sociaux et sociotechniques, l'analyse des séries temporelles. La plate-forme vous propose ainsi : de **découvrir et de vous familiariser avec** certains des modules de collecte et d'analyse de données hétérogènes issues des sources variés, et de les **mettre en œuvre** sur des jeux de données qui sont les vôtres ou qui vous seront apportés.

Données concernées:

- Les bases de données de l'information scientifique et technologique (publications, brevets, projets de recherche)
- Les archives numérique des médias traditionnels (presse / TV / radio) et des sources comme blogs / sites Web et forums
- Des bases de données Spécifiques - des projets de recherche, entreprises, produits etc
- Les données recueillies par des méthodes d'enquête et des entretiens qualitatifs.

Programme

Jour 1 – 19 décembre (9:30 - 18:30)

(repas fourni sur place)

Introduction à la Formation

La Collecte des données

- . Présentation des sources de données, leurs interfaces et les options
- . Stratégies de recherche de données et de délimitation de corpus
- . L'extraction des données, formats de données, transformation de données.

L'analyse des données

- . Exploration de données en utilisant un module de statistiques descriptives. Description des séries temporelles, leur définition et leur utilisation.
- . Analyse des données textuelles. Champs textuels dans les sources de données. Principes et étapes de base de l'extraction de termes. Indexation automatisée de données textuelles, lemmatisation et le filtrage statistique.

La Visualisation de graphes

- . Cartographie des réseaux hétérogènes. Dynamique des réseaux socio-sémantiques. Description mésoscopique de données, clusterisation et détection des communautés. Modèle dynamique des

Retour sur la journée : Discussion

Jour 2 – 20 décembre (9:30 - 17:00)

(repas fourni sur place)

Atelier pratique : reprise des acquis de la veille, mobilisés sur une question et un corpus ou des corpus porté par chaque usager, en trois étapes

Etape 1. La préparation de corpus : mise en oeuvre et partage sur les difficultés, et discussion

Etape 2. Analyse des données : mise en oeuvre et partage sur les difficultés, et discussion

Etape 3. Présentation des résultats : mise en oeuvre et présentation, discussion sur les rendus et leurs sens par rapport au projet de connaissance des usagers

Discussion générale et retour d'expérience sur la journée.

Formation-Atelier performé par Andréi Mogoutov –(IFRIS, CorText)