

Semaine Data SHS

Traiter et analyser des données en sciences humaines et sociales

9-13 décembre 2019

Misha | Université de Strasbourg

La Plateforme Universitaire de Données de Strasbourg organise 5 jours de formation aux **méthodes de traitement et d'analyse de données en sciences humaines et sociales**. Cette semaine est organisée au niveau national par la très grande infrastructure de recherche **Progedo** et le Réseau National des Maisons des Sciences de l'Homme.

🎯 Objectifs

- **Acquérir** des notions théoriques sur des méthodes d'analyse et de représentation de données numériques afin de les adapter à ses propres travaux ;
- **Initier** les participant.e.s via des **travaux pratiques** au traitement, à la structuration et à l'analyse de données.

👥 Public

Ces ateliers s'adressent aux Master 2, doctorant.e.s, aux enseignant.e.s-chercheur.e.s, aux chercheur.e.s et ingénieur.e.s.

📅 Programme

📅 **Lundi 9 décembre** ⌚ 9h-17h

Comment utiliser des données de l'internet dans le cadre d'une recherche en sciences sociales ?

Etienne Ollion, *chercheur CNRS - Professeur associé à l'ENSAE*

L'objectif de cette formation est de **proposer une introduction à la collecte automatisée et à la curation de ces données numériques**. Il s'agira d'apprendre à les localiser, à mettre en place une stratégie pour les collecter, les nettoyer et les préparer en vue d'un traitement. Cet apprentissage ne sera pas dissocié d'une réflexion sur les enjeux que pose leur utilisation. L'accent sera mis sur la diversité des types de données, leurs avantages comme leurs limites.

📅 **Mardi 10 décembre** ⌚ 9h-17h

Analyse textuelle

Pablo Ruiz Fabo, *maître de conférences en Informatique (LiLPa)*

Cette formation consiste à une introduction à l'analyse textuelle selon une approche lexicométrique ou textométrique. Les outils prévus pour illustrer cette approche seront principalement les suivants :

- TXM : <http://textometrie.ens-lyon.fr/spip.php?article60>
- le Trameur : <http://www.tal.univ-paris3.fr/trameur/>

Nous aborderons les fonctions principales d'exploration de corpus de ces outils. Une introduction aux langages de requêtes pertinents sera présentée, notamment les expressions régulières (regexp), et Corpus Query Language (CQL). Les corpus à analyser seront fournis lors de la formation.

 **Mercredi 11 décembre**  10h-12h30

Préserver, partager et réutiliser des enquêtes qualitatives : présentation de la plateforme « beQuali »

Jeremie Vandebunder, *ingénieur de recherche (CDSP - Sciences Po, CNRS)*

« beQuali » est une banque d'enquêtes qualitatives en sciences sociales à disposition de la communauté scientifique. L'objectif de cet atelier est de présenter les possibilités d'usage de cet outil pour les chercheur.e.s et les enseignant.e.s, et de proposer un retour d'expérience en matière de préparation de données qualitatives dans un objectif de préservation pérenne et de partage scientifique. Une démonstration de l'outil, des discussions collectives et une prise en main rapide de l'outil seront proposées aux participant.e.s.

 **Mercredi 11 décembre**  9h-17h

Cartographie en anamorphose

Anne Christine Bronner, *Cartographe, ingénieure de recherche CNRS (SAGE)*

- 9h-12h : Quelques bases en cartographie

Quelles sont les étapes pour produire une carte thématique ? De la donnée à la représentation en passant par les traitements statistiques et graphiques. Prise en main du logiciel Magrit avec un jeu de données autour de la cartographie électorale.

- 14h – 17h : Représenter des données en SHS : l'intérêt du cartogramme

Le fond de carte classique offre une vision discutable de certaines réalités sociales, notamment les résultats électoraux, le chômage, les données démographiques, etc. Il s'agira d'aborder la question du fond de carte en lien avec les données et le message de la carte en mobilisant la technique du cartogramme.

 **Jeudi 12 décembre**  9h-17h

Les premiers traitements statistiques en SHS avec l'aide d'Excel

Floriane Varieras, *ingénieure d'étude de la Plateforme universitaire de Données de Strasbourg*

La journée sera consacrée à la prise en main du logiciel Excel, afin de pouvoir réaliser des traitements de statistiques descriptives à partir des données de recensement : construction d'indicateur, réalisations de graphiques, tableaux croisés dynamiques, khi-2, etc.

 **Vendredi 13 décembre**  9h-12h

La réutilisation des données publiques à des fins de recherche

Sarah Piquette, *déléguée à la protection des données de l'Université de Strasbourg*

Cette présentation reviendra sur le nouveau cadre de l'ouverture des données publiques et leur possible réutilisation à des fins de recherche (RGPD). Elle abordera également le cas du webscraping.

Informations et inscription

La participation est libre dans la limite de 12 places par atelier.

Inscription en ligne obligatoire sur le site web de la MISHA : misha.fr/pud_actu.htm

 **Lieu** Salle informatique et salle Océanie | 1^{er} étage | Misha | Université de Strasbourg | 5 allée Rouvillois
