

Semaine Data SHS

Traiter et analyser des données en sciences humaines et sociales

Du 9 au 13 décembre 2019
à la MSH Lyon St-Étienne, Lyon 7^e

PRÉSENTATIONS ET ATELIERS MÉTHODOLOGIQUES

Contexte et objectifs

Dans le cadre de sa plateforme universitaire de données PANELS, la MSH Lyon St-Étienne organise du 9 au 13 décembre 2019 une semaine « Data SHS », qui aura pour thématique générale « Traiter et analyser des données en sciences humaines et sociales ». Cette semaine, inscrite dans le plan national de formations de la TGIR Progedo, propose une série de présentations et d'ateliers pratiques d'initiation aux méthodes de traitement et d'analyse de données en sciences humaines et sociales.

Ces ateliers ont un double objectif : présenter en détails les fondements théoriques des méthodes d'analyse et de représentation de données SHS et expliciter les moyens de leur mise en œuvre pratique. Le but est de permettre aux participants d'adapter ces méthodes à leurs propres travaux, de façon la plus autonome possible. Ils seront basés, autant que faire se peut, sur des logiciels libres (ex : R, Qgis).

Public : ces présentations et ateliers s'adressent aux chercheurs, enseignants-chercheurs, ingénieurs et techniciens, ainsi qu'aux doctorants des laboratoires SHS du site Lyon St-Etienne.

Thématiques

6 grandes thématiques ont été retenues pour cette semaine de présentations et d'ateliers :

- méthodes et pratiques pour **produire** des données originales en SHS ;
- **plan de gestion** de données en SHS (méthodes, enjeux et RGPD) ;
- la **cartographie** et l'analyse spatiale appliquées aux SHS ;
- grandes **enquêtes internationales** et bases de données quantitatives - TGIR Progedo ;
- des données **images et sons** en SHS : de la production à la valorisation ;
- enjeux et pratiques de l'**analyse textuelle** en SHS.



■

LUNDI 9 DÉCEMBRE 2019

Méthodes et pratiques pour produire des données originales en SHS

Module 1 – Création de données quantitatives à partir de textes : une application à l'étude du comportement des pourvoyeurs d'information

Intervenant : [Sophie Hatte](#), MCF en économie (ENS de Lyon, GATE Lyon St-Étienne - UMR 5824)

L'étude des stratégies de diffusion d'information des médias ou des activistes, ainsi que leurs effets sur le fonctionnement de nos sociétés, se trouve à l'intersection des disciplines qui forment les sciences sociales. Cet atelier propose de faire en premier lieu un tour d'horizon des sources qui peuvent être utilisées afin de créer et exploiter des données quantitatives sur les pourvoyeurs d'informations que sont les médias (traditionnels et sociaux) et les activistes. Dans un second temps, les participants découvriront comment manier ces données sous format « texte » afin de les intégrer dans une démarche de recherche quantitative. Enfin, l'exploration de cas pratiques issus de travaux de recherche récents illustrera comment ces données peuvent être construites, assemblées et intégrées dans des modèles économétriques qui permettent une compréhension fine des phénomènes analysés.

Prérequis, outils/logiciels à télécharger :

Pour les personnes souhaitant participer à la mise en pratique des traitements, merci de télécharger la version d'essai du logiciel Stata (<https://ritme.com/BE-fr/logiciels/stata/>). Pour ces personnes, il sera demandé une familiarité minimale avec le logiciel Stata ou toute autre logiciel orienté lignes de codes (par exemple : R Studio).

Pour les autres participants, cette partie constituera une démonstration des possibilités offertes en termes de maniabilité de traitements des données.

Horaires : 8h30-11h10 (pause 15 min) – **Lieu :** salle André Bollier (rdc) + visioconférence à St-Étienne

Inscription > formulaire en ligne : <http://enquetes.msh-lse.fr/index.php/519354/lang-fr>

Module 2 – Analyser notre paysage informationnel à l'aide d'outils d'intelligence artificielle

Intervenant : [Julien Velcin](#), professeur en informatique (Univ. Lyon 2, ERIC EA 3083)

Analyser toute la complexité du paysage informationnel est devenu une tâche très difficile sans le recours à des techniques avancées de fouille de données et de traitement automatique des langues. L'intelligence artificielle apporte des algorithmes qui peuvent aider à appréhender de manière synthétique les milliers, si ce n'est les millions, de textes publiés dans les médias, sur les réseaux sociaux, et plus généralement sur le Web. Je présenterai un projet qui visait précisément à "digérer" des articles de presse en employant des techniques d'analyse thématique puis en réalisant un post-traitement manuel afin de construire des catégorisations qui font sens pour les chercheurs en SHS. J'esquisserai le périmètre de futures recherches qui viseraient à automatiser l'attribution de ces catégories toujours en considérant une supervision humaine afin de garder un contrôle sur la qualité des résultats obtenus. Ce travail a été réalisé dans le cadre du projet *Journaliste à l'ère du Numérique*, en collaboration avec le Pr. J.C. Soulages. Tout au long de mon exposé, je passerai un peu de temps pour expliciter les mécanismes d'IA à l'oeuvre (leurs hypothèses, leurs possibilités, leurs limites) et les illustrer.

Prérequis : aucun.

Horaires : 15h30-17h30 – **Lieu :** espace Marc Bloch (rdc) + visioconférence à St-Étienne

Inscription > formulaire en ligne : <http://enquetes.msh-lse.fr/index.php/595591/lang-fr>

Module 3 – Récolte automatique des données du web

Intervenant : [Julien Boelaert](#), MCF en science politique (Univ. Lille, CERAPS - UMR 8026)

Cette formation présente l'étape-clé de la collecte automatisée de données du web (webscraping) : l'extraction, à partir de pages HTML ou de bases de données XML, des informations qui intéressent le chercheur. A l'aide du langage de requêtes XPath et d'outils informatiques simples, on transformera une série de pages web en une base de données sous format tableur. La présentation s'articulera en trois temps : d'abord une introduction aux principes des langages à balise, ensuite une exposition des principes du langage XPath, et enfin une mise en pratique de collectes automatisées sous R.

Prérequis : familiarité minimale avec le langage R.

Outils/logiciels à télécharger :

Logiciel R, RStudio : <https://rstudio.com/products/rstudio/download/#download>

Module ChroPath, pour Firefox : <https://addons.mozilla.org/fr/firefox/addon/chro-path-for-firefox/>

ou pour Chrome : <https://chrome.google.com/webstore/detail/chro-path/ljngjbnaijcbnmcnjfhigebomdlkco>

Horaires : 11h20-12h20, 13h20-15h20 (pause déjeuner 1h) –

Lieu : salle André Bollier (rdc) + visioconférence à St-Étienne

Inscription > formulaire en ligne : <http://enquetes.msh-lse.fr/index.php/432974/lang-fr>

Conférence – Méthodes et pratiques des enquêtes internationales

L'importance de la méthodologie pour la qualité de la comparaison dans les enquêtes internationales

Intervenant : **Dominique Joye**, professeur honoraire (Institut des sciences sociales de Lausanne, Centre de recherche sur les parcours de vie et les inégalités)

Les grandes enquêtes internationales sont souvent vues à travers le prisme du duo jeu de données – en format SPSS, STATA, R, etc. – et questionnaire original, le plus souvent en *British English*. De facto cela fait l'impasse sur les différences de terrain dans chaque pays. Cela a des conséquences sur la qualité et la pertinence des analyses subséquentes.

Nous proposerons ici une approche « Erreur d'enquête totale » – *Total Survey Error* – qui permette de parcourir les points particulièrement importants à considérer. Nous proposons d'appliquer ce modèle aussi bien au moments de la construction du questionnaire, de son implémentation – traduction notamment – et de l'analyse, liée aux questions de design national et de pondération. Ceci est d'autant plus important au moment où le format traditionnel de l'enquête en face à face est remis en cause et où de nouvelles méthodologies sont proposées. In fine, l'enjeu n'est pas seulement une question de documentation des enquêtes mais plutôt de questionner nos pratiques de chercheur.

Horaires : 9h-12h – **Lieu** : espace Marc Bloch (rdc), retransmission spéciale depuis la MSH Alpes (Grenoble)

Inscription > formulaire en ligne : <http://enquetes.msh-lse.fr/index.php/784573/lang-fr>

MARDI 10 DÉCEMBRE 2019

Plan de gestion de données en SHS (méthodes, enjeux et RGPD)

Module 1 – Créer et gérer son plan de gestion de données

Intervenant : **Marie Puren**, post-doctorante en histoire et humanités numériques (CNRS, LARHRA - UMR 5190)

La présentation permettra aux participants de comprendre en quoi consiste un plan de gestion de données et à quoi il sert. Les participants commenceront également à rédiger leurs propres plans de gestion de données : qu'est-ce qu'un plan de gestion de données ? à quoi sert un plan de gestion de données ? les différentes parties d'un plan de gestion de données, écrire et gérer son plan de gestion de données avec DMP Opidor.

Prérequis : avoir un ou des jeu(x) de données issu(s) d'un projet de recherche ou envisager de décrire un ou des jeu(x) de données issu(s) d'un projet de recherche. Ces données seront à la base de l'exercice sur DMP Opidor.

Outils/logiciels à télécharger : création préalable d'un compte sur DMP Opidor : <https://dmp.opidor.fr>

Horaires : 10h-13h15 (pause 15 min) – **Lieu** : salle André Bollier (rdc) + visioconférence à St-Étienne

Inscription > formulaire en ligne : <http://enquetes.msh-lse.fr/index.php/484425/lang-fr>

Module 2 – Appliquer le RGPD dans la recherche, c'est (pas) si compliqué ?

Intervenants : **Lucie Le Barreau**, déléguée à la protection des données personnelles (Univ. Lyon 3) et **Céline Faure**, statisticienne (CNRS, MSH Lyon St-Etienne - USR 2005)

Le monde de la recherche, dont celui des sciences humaines et sociales, implique souvent l'utilisation de données et notamment de données personnelles. Les faits de collecter, partager et protéger peuvent sembler contradictoires. Pourtant les acteurs de la recherche peuvent maintenir le niveau d'excellence de leurs recherches, tout en préservant les droits des personnes interrogées ou ayant fait l'objet d'une collecte.

Pour cela, nous exposerons les grands principes du RGPD (Règlement général de la protection des données applicable depuis 25 mai 2018). Nous présenterons les modalités de mise en œuvre en s'appuyant sur les outils disponibles (et validés des tutelles) et les acteurs du site local (DPD/DPO), notamment.

Cette session sera complétée par plusieurs références bibliographiques (Guides, documents, sites web). Enfin, nous proposerons de tester les connaissances en matière de données personnelles, à travers un jeu ! (« Le jeu sérieux pour identifier les données à caractère personnel » une production de La reine Merlin et de Dawizz, CC by SA).

Prérequis : aucun.

Horaires : 14h30-16h30 (1h d'intervention, 30 min de questions, 30 min de jeu) – **Lieu :** salle André Bollier (rdc)

Inscription > formulaire en ligne : <http://enquetes.msh-lse.fr/index.php/745482/lang-fr>

MERCREDI 11 DÉCEMBRE 2019

La cartographie et l'analyse spatiale appliquées aux SHS

Module 1 – Initiation au logiciel SIG QGIS (système d'information géographique)

Intervenant : **Nicolas Ovtracht**, ingénieur d'études (CNRS, LAET - UMR 5593)

Cette initiation découverte de QGIS s'adresse à toute personne impliquée dans la réalisation de cartographie thématique et de traitements de l'information spatiale. Cette formation comportera une partie théorique et une partie pratique, plus importante, de manipulation de l'application par des exercices collectifs. Cette formation vous permettra de voir les principaux modules du logiciel QGIS, aujourd'hui incontournable en géomatique (cartographie, thématique, cartographie des flux et symboles proportionnels, raster, requêtes SQL attributaires et spatiales).

Prérequis : familiarité avec l'outil informatique.

Outils/logiciels à télécharger : logiciel QGIS <https://www.qgis.org/fr/site/forusers/download.html>

Horaires : 9h-11h – **Lieu :** salle André Bollier (rdc) + visioconférence à St-Étienne

Inscription > formulaire en ligne : <http://enquetes.msh-lse.fr/index.php/887217/lang-fr>

Module 2 – Analyser l'espace, entre statistique et cartographie

Intervenant : **Hélène Mathian**, ingénieur de recherche (CNRS, EVS - UMR 5600)

L'analyse spatiale définie comme l'analyse formalisée des configurations spatiales (Pumain, Saint-Julien, 1997) s'est développée dans le sillage de la géographie quantitative et sur la base d'un ensemble de techniques et modèles mettant la localisation au cœur des questions. Dans ces approches l'exploration des données (au sens de Tukey) a toujours eu une grande part et particulièrement l'exploration visuelle, confrontant cartographies et graphiques statistiques, de manière à pouvoir éprouver relations entre variables et liens avec l'espace.

Avec l'accroissement des données localisées, et le développement d'environnements interactifs, ces questions sont aujourd'hui renouvelées : l'exploration est facilitée. Les utilisateurs deviennent actifs et peuvent « toucher » les données. Nous reviendrons dans un exposé généraliste sur des méthodes fondatrices des approches d'analyse spatiale. Nous proposerons ensuite un atelier sous RStudio sur des jeux de données géographiques pour approcher les potentialités de cet environnement dans l'exploration de la dimension spatiale.

Prérequis : bases statistiques et familiarité minimale avec le langage R.

Outils/logiciels à télécharger : logiciel RStudio <https://rstudio.com/products/rstudio/download/#download>

Horaires : 11h15-13h15 – **Lieu :** salle André Bollier (rdc) + visioconférence à St-Étienne

Inscription > formulaire en ligne : <http://enquetes.msh-lse.fr/index.php/887217/lang-fr>

Grandes enquêtes internationales et bases de données quantitatives - TGIR Progedo

Module 1 – L'enquête *European Values Study*, une grande tradition de recherche sociologique et politologique

Intervenant : Pierre Bréchon, professeur émérite en science politique (Science Po Grenoble, PACTE - UMR 5194)

L'intervention commencera par la présentation de la tradition de recherche sur les valeurs des Français et des Européens. On dispose aujourd'hui de cinq vagues d'enquêtes étalées sur 40 ans avec un questionnaire semblable, permettant donc de riches comparaisons dans le temps et dans l'espace. Cette enquête permet de traiter des grandes évolutions culturelles. On présentera de nombreux résultats pour la France ou l'Europe sur une longue période. Ceci devrait permettre de comprendre à quel point cette démarche empirique permet de sortir des propos impressionnistes qui souvent tiennent lieu d'analyses de la part d'observateurs pressés.

Dans un deuxième temps, l'intervention sera centrée sur un point particulier. Les débats sont très nombreux sur l'individualisme, dont beaucoup disent qu'il se développe dans les sociétés développées. Nous montrerons qu'en fait, cette notion est très composite et qu'il convient de distinguer individualisation et individualisme. Après avoir présenté ce qui distingue ces deux concepts, une démarche méthodologique sera exposée : construction d'échelles d'attitudes sur plusieurs vagues, tris croisés et analyse de régression permettent de tirer des conclusions : si l'individualisation est croissante dans notre société, l'individualisme est au contraire plutôt déclinant, contrairement à ce qu'affirment très souvent les observateurs. La montée de l'éducation et la baisse de la religiosité sont deux éléments fondamentaux expliquant les évolutions culturelles vers plus d'individualisation mais plutôt moins d'individualisme.

Prérequis : aucun.

Horaires : 14h15-16h15 – **Lieu :** salle André Bollier (rdc) + visioconférence à St-Étienne

Inscription > formulaire en ligne : <http://enquetes.msh-lse.fr/index.php/132214/lang-fr>

Module 2 – Analyses et usages des données sur les valeurs politiques dans les enquêtes *European Values Study*

Intervenant : Nathalie Dompnier, professeur en science politique (Univ. Lyon 2, TRIANGLE UMR 5206)

L'intervention proposera une présentation des questions des enquêtes EVS portant sur les valeurs politiques, de leur évolution au fil des enquêtes et de perspectives de recherche dans lesquelles elles peuvent être mobilisées. Seront ensuite présentées quelques grandes tendances observées depuis 1981, en France et en Europe, notamment sur les thèmes de la participation, de la politisation et du rapport à la démocratie. Cette présentation, basée sur des questions spécifiques mais aussi sur la construction d'échelles, permettra de se familiariser avec ces données, de donner un aperçu de leur richesse et, en lien avec les spécialités de recherche des participants.e.s, d'en interroger les usages possibles en sciences humaines et sociales.

Prérequis : aucun.

Horaires : 16h30-18h30 – **Lieu :** salle André Bollier (rdc) + visioconférence à St-Étienne

Inscription > formulaire en ligne : <http://enquetes.msh-lse.fr/index.php/132214/lang-fr>

JEUDI 12 DÉCEMBRE 2019

Des données images et sons en SHS : de la production à la valorisation

Module 1 – Exemples de recherches utilisant les données images et sons

9h-9h45 : **Animer des paroles d'enquêtée.s. Retour sur un projet de film de recherche sur les sexualités ordinaires**

Intervenants : Emmanuelle Santelli, directrice de recherche en sociologie (CNRS, Centre Max Weber - UMR 5283) et Jeanne Drouet, anthropologue spécialiste des méthodes visuelles (CNRS, Centre Max Weber - UMR 5283)

Nous présenterons un projet de film de recherche concernant les sexualités ordinaires. La richesse du matériau collecté dans le cadre de la recherche, - à savoir des paroles d'enquêtée.e.s confiées en entretien concernant la sexualité dans le cadre conjugal -, mais aussi le caractère sensible de ces données, nous a poussé à imaginer une forme filmique spécifique, à savoir un court-métrage d'animation allié à une bande son documentaire. Ce format impose une méthodologie de travail adéquate, qui s'invente en collectif, chemin faisant. Nous faisons aujourd'hui le pari que ce « traitement de données » singulier aura des vertus sur le plan heuristique (apport de nouvelles connaissances sur l'objet d'étude), mais aussi et surtout des vertus en terme de publicisation des résultats de recherche, en facilitant l'introduction de questions de recherche dans le débat public.

9h45-10h05 : De la pratique des soins infirmiers à la recherche, le double dialogue du projet CIPSY

Intervenants : **Justine Lascar**, ingénieure d'études en production et analyse de données (CNRS, ICAR UMR 5191), **Louis Maritaud**, doctorant (Univ. Lyon 2, ICAR - UMR 5191) et **Benoît Chalancon**, infirmier de recherche clinique (Centre hospitalier Le Vinatier)

Nous proposons de situer les perspectives d'intervention du chercheur et de l'expert de terrain au sein du projet CIPSY, dont la demande a été faite par une équipe de soignants en psychiatrie.

La spécificité de l'exercice infirmier en santé mentale par le soin relationnel, le souci de continuité des soins et la complexification actuelle des soins, marquée par une augmentation des intervenants auprès des patients, invitent à s'intéresser de près au cheminement de l'oral vers l'écrit et à leur production lors des différents moments de relève infirmière. CIPSY est une étude exploratoire s'intéressant aux compétences d'interaction de l'infirmier actuel en service de psychiatrie adulte. L'objectif principal du projet est de déterminer le répertoire langagier de ces deux productions et leur manipulation, des points de vue de l'analyse conversationnelle, de l'analyse du discours, entre autres. Cette recherche est basée sur un corpus d'une quinzaine d'heures de réunions de relève infirmières, filmées dans un centre hospitalier psychiatrique, ainsi que sur l'extraction des relèves écrites provenant de leur logiciel dédié. L'aspect co-construit de CIPSY amène à repenser les statuts de chacun selon les situations. Ce dialogue, dont le but est de faire avancer la recherche constitue donc un double défi méthodologique en lui-même : comment les chercheurs peuvent aider les pratiques, et comment les soignants peuvent aider la recherche. Ainsi nous décrirons, d'une part, l'implication des experts de terrain dans le processus de recherche et, d'autre part, les différents moyens mis en oeuvre pour concilier les demandes des soignants au regard des outils et objectifs de la recherche. En effet, de nombreux corps de métiers sont invités à s'exprimer pour faire avancer le projet, qu'il s'agisse de chercheur.es en linguistique ou de professionnel.les de la santé. Cette dualité dans la recherche sera représentée dans la communication, l'un des intervenants faisant partie des professionnels de santé du projet, les deux autres intervenants faisant partie du milieu de la recherche.

10h05-10h50 : Une méthode pour étudier les expériences de visite : l'entretien en reviviscence stimulée

Intervenant : **Geoffroy Gawin**, MCF (ENSSIB, ELICO EA 4147)

La méthode d'entretien en reviviscence stimulée sera présentée à travers d'exemples issus d'une recherche menée en 2018. Seront ainsi montrés des enregistrements vidéos réalisés à partir de caméras, lunettes à eye-tracking ou caméra go pro, confiés à des visiteurs de musée. Le regard réflexif proposé amènera à discuter des régimes de souvenance induits par ces deux supports audiovisuels.

11h05-12h35 : Visualiser et classifier des corpus collectés en ligne. Présentation d'un parcours méthodologique

Intervenant : **Oriane Piquer-Louis**, doctorante en sciences de l'information et de la communication (Univ. Lyon 2, ELICO EA 4147)

Collecter des images sur Twitter est une chose. Mais ensuite ? Cet atelier est l'occasion de revenir sur les différents problèmes méthodologiques posés par la manipulation de corpus issus du web, leur étude via des méthodes qualitatives, et débouche sur la présentation d'un outil de visualisation et de classification des données, qui tente de répondre à ces questions (avec la manipulation d'un petit corpus d'exemple).

Prérequis : aucun.

Horaires : 9h-12h35 – **Lieu :** espace Marc Bloch (rdc) + visioconférence à St-Étienne

Inscription > formulaire en ligne : <http://enquetes.msh-lse.fr/index.php/994464/lang-fr>

Module 2 – Ateliers pratiques et démonstrations d'outils et logiciels pour les données numériques images et sons

Atelier 1 – Extraire des images-mots automatisés dans un document

Intervenants : **Véronique Eglin**, MCF en informatique (INSA de Lyon, LIRIS - UMR 5205)

Atelier 2 – Présentation et prise en main du logiciel ELAN

Intervenants : **Isabel Colón**, MCF (ENS de Lyon, ICAR - UMR 5191) et **Laurie Boyer** (CNRS, ICAR - UMR 5191)

Atelier 3 – Ecouter, regarder, transmettre, valoriser avec des images et des sons

Intervenants : **Justine Chapelon**, **Christian Dury**, **Mylène Pardoën** et **Joachim Poutaraud**, pôle image animée audio (CNRS, MSH Lyon St-Étienne USR 2005)

Prérequis : aucun.

Horaires : 14h-17h (30 min par atelier) – **Lieu :** salles Georges Lyvet, Paul Rivière et Fernand Rude (sous-sol)

Inscription > formulaire en ligne : <http://enquetes.msh-lse.fr/index.php/163135/lang-fr>

Conférence – Les jeudis de l'axe Société et humanités numériques

Enjeux éthiques associés à l'introduction de l'intelligence artificielle Exemple de l'application de l'IA au domaine juridique

Intervenant : **Fabrice Muhlenbach**, MCF en informatique (Univ. St-Étienne, lab. Hubert Curien - UMR CNRS 5516)

Les technologies du numérique transforment le monde en facilitant l'échange, le stockage et le traitement des informations. L'informatique et la multiplication des données produites par les hommes ou par les machines amènent à des bouleversements de la société inédits qui procurent une forme de pouvoir à ceux qui savent exploiter ces technologies et ces données (Berry, 2017). Les entreprises du numérique, comme les géants d'Internet, proposent des services dont l'efficacité dépend des *big data* qu'elles collectent auprès de leurs utilisateurs, mais ce faisant elles effectuent des choix qu'elles introduisent dans leurs algorithmes, amenant à des biais, voire à des préjugés, dont les utilisateurs de ces services ne sont pas conscients (Cardon, 2015 ; O'Neil, 2018).

L'intelligence artificielle (IA), et en particulier les dernières technologies associées à l'apprentissage automatique appelées « apprentissage profond » (*deep learning*), se nourrissent de ces *big data* pour améliorer les performances de ces systèmes automatisés. En déléguant à ces systèmes automatisés un certain nombre de tâches, dans le cadre personnel ou professionnel, les êtres humains, utilisateurs de ces technologies, se sentent déresponsabilisés, posant ainsi tout un ensemble de problématiques liées à la gouvernamentalité algorithmique (Rouvroy & Berns, 2013).

Seront ainsi évoquées diverses questions éthiques liées à l'introduction de l'intelligence artificielle dans différents champs de l'activité humaine. Seront notamment présentées certaines conséquences de l'arrivée de l'IA dans le domaine du droit avec l'apparition des « robots juges » et autres applications d'IA juridiques, posant ainsi des questions sur la vie privée, le droit à un procès équitable et la gouvernamentalité algorithmique (Muhlenbach & Sayn, 2019).

Horaires : 17h30-19h30 – **Lieu :** espace Marc Bloch (rdc) + visioconférence à St-Étienne

Inscription > formulaire en ligne : <http://enquetes.msh-lse.fr/index.php/255364/lang-fr>

VENDREDI 13 DÉCEMBRE 2019

Enjeux et pratiques de l'analyse textuelle en SHS

Module 1 – Le Vrai débat - rendre lisibles les données d'un corpus d'intelligence collective, approche méthodologique

Intervenants : **Vincent Ventresque**, ingénieur d'études (TRIANGLE - UMR 5206)
et **François Robert**, ingénieur de recherche (CNRS, TRIANGLE - UMR 5206)

Le mouvement social initié par les gilets jaunes a produit une quantité impressionnante d'écrits que ce soit sous une forme manuscrite (cahiers de doléances) ou numérique via les plateformes du Vrai débat (Gilets jaunes) ou du Grand débat (gouvernementale). Cette verbalisation revendicative est, entre autres, une des caractéristiques de ce mouvement. Face à ce magma de textes, il devient très difficile d'en comprendre le sens. L'opacité générée par l'abondance et la richesse des textes pénalise le mouvement qui peine à afficher ses objectifs programmatiques et facilite aussi toutes les manipulations discursives. Peu de personnes se sont aventurées dans le cœur de ces données.

À la demande des animateurs du Vrai débat qui nous ont donné l'accès à l'ensemble des textes, nous avons essayé d'en dégager les grands axes revendicatifs. Notre contribution vient ainsi compléter celles effectuées par d'autres laboratoires comme le Lerass à Toulouse ou l'Institut des systèmes complexes à Paris.

Nous présenterons nos méthodes d'analyses réalisées collectivement et individuellement au sein du groupe Gilets jaunes de Triangle. Nous présenterons, tout d'abord, le corpus du Vrai débat puis, nous montrerons comment nous avons pu dégager les grandes thématiques pour ensuite, extraire des sous-corpus spécifiques constitués à la demande de chercheurs. Si l'accent sera mis sur les outils utilisés comme les logiciels de textométrie, nous insisterons sur notre capacité de lecture, outil de discernement à la portée de tous.

Prérequis : aucun.

Outils/logiciels à télécharger : logiciel TXM https://groupes.renater.fr/wiki/txm-users/public/preparation_ordinateur
FAQ : https://groupes.renater.fr/wiki/txm-users/public/faq#comment_verifier_que_txm_s_est_installe_correctement

Horaires : 9h-12h15 (pause 15 min) – **Lieu :** salle André Bollier (rdc) + visioconférence à St-Étienne

Inscription > formulaire en ligne : <http://enquetes.msh-lse.fr/index.php/744221/lang-fr>

Module 2 – Lexicométrie et textométrie appliquées à l'analyse de corpus en SHS L'exemple d'une étude en géographie sur la médiatisation des barrages hydroélectriques

Intervenants : **Silvia Flaminio**, post-doctorante (Fonds national suisse, Institut de Géographie et de durabilité, Univ. Lausanne) et chercheuse associée à EVS - UMR 5600

Après une brève introduction sur les enjeux de la lexicométrie et de la textométrie, la présentation détaillera les différentes étapes et les questions soulevées par la construction d'un corpus d'articles de presse (1319 articles publiés en 1945 et 2014 sur des barrages dans le quotidien français *Le Monde*) et de son analyse. Les principales fonctions de deux logiciels libres d'analyse des données textuelles (Iramuteq et TXM) seront présentées.

Prérequis : aucun.

Outils/logiciels à télécharger :

Logiciel R : <https://rstudio.com/products/rstudio/download/#download>

Logiciel Iramuteq : <http://iramuteq.org/telechargement>

Logiciel TXM : https://groupes.renater.fr/wiki/txm-users/public/preparation_ordinateur

Horaires : 13h15-16h – **Lieu :** salle André Bollier (rdc) + visioconférence à St-Étienne

Inscription > formulaire en ligne : <http://enquetes.msh-lse.fr/index.php/666672/lang-fr>

Comité d'organisation

Samuel Bonnand (CNRS, MSH Lyon St-Etienne)
Annie Busson (CNRS, MSH Lyon St-Etienne)
Gilles Pollet (CNRS, MSH Lyon St-Etienne)
Alexandra Dugué (Univ. Lyon 2, MSH Lyon St-Etienne)
Céline Faure (CNRS, MSH Lyon St-Etienne)

Sabine Loudcher (Univ. Lyon 2, ERIC)
Amélie Hugot (CNRS, MSH Lyon St-Etienne)
Christian Dury (CNRS, MSH Lyon St-Etienne)
Mylène Pardoën (CNRS, MSH Lyon St-Etienne)

Informations pratiques

Les personnes intéressées pourront s'inscrire à un ou plusieurs ateliers. La participation est gratuite, avec repas de midi offerts, mais sur inscription obligatoire (nombre de places limité).

Localisation

MSH Lyon St-Étienne, 14 avenue Berthelot, Lyon 7^e

Retransmission de certaines séances en visioconférence à St-Etienne :
Université Jean Monnet, campus Tréfilerie, 33 rue du Onze Novembre
Salle L2.6 (bât. L, 2^e étage), sauf jeudi 12 matin en salle D1.6.11 (bât. D)

Inscriptions

Merci de vous inscrire, avant le 24 novembre 2019, à l'aide des formulaires en ligne spécifiés pour chacun des modules ou conférences (cf programme détaillé ci-dessus).

Pour les visioconférences à St-Étienne : <http://enquetes.msh-lse.fr/index.php/247982/lang-fr>

Contacts

Contenu des ateliers : Samuel Bonnand (MSH Lyon St-Etienne) - samuel.bonnand@msh-lse.fr

Inscriptions, organisation : Annie Busson (MSH Lyon St-Etienne) - annie.busson@msh-lse.fr

Site web : www.msh-lse.fr/evenements/semaine-data-shs

Twitter : [@mshlyon](https://twitter.com/mshlyon) #SemaineDataSHS2019

