



Programme Semaine DATA SHS

La semaine Data-SHS est une idée originale de la PUD de Lille (Héloïse Petit, Marie Cros et Sofia Papastamkou), qui a vu le jour en 2017 et a ensuite été reprise et élargie à partir de 2019 par l'IR* PROGEDO. Il s'agit un événement organisé localement par chacune des PUD que l'IR* PROGEDO fait résonner au niveau national. Il se déroule durant la deuxième semaine de décembre.

Cette semaine vise à mettre en avant l'utilisation et la réutilisation des données quantitatives en SHS, en particulier les grandes enquêtes et les données issues de la statistique publique. Elle vise aussi à promouvoir les bonnes pratiques autour de leur production (questionnaire, échantillonnage, collecte, documentation, diffusion) et de leur usage (méthodes, outils, valorisation).

➔ **Inscription : [cliquez ici](#)**

Lundi 5 décembre 2022

9h30 à 11h

Présentation de la PUD : Une plateforme au service de la communauté scientifique en SHS dans la recherche et l'analyse de données

Axel DELARUE, Ingénieur d'étude à la Plateforme Universitaire de Données de Clermont-Ferrand

Créée depuis plus d'un an, appartenant à l'infrastructure de Recherche PROGEDO, la PUD est au service de la communauté scientifique. La plateforme accompagne les chercheurs dans la recherche de données quantitatives en SHS, apporte son soutien à de nombreux projets ou encore organise des événements scientifiques (formations, ateliers, journées d'études). Le rôle et le fonctionnement de la plateforme seront expliqués et un premier bilan sera dressé. PROGEDO sera également présentée ainsi que son catalogue de données.

Lundi 5 décembre de 9h30 à 11h - [amphi 220 de la MSH de Clermont-Ferrand](#)

13h30 à 15h

Comment interpréter les questions d'une enquête : les pièges à éviter

Simone BERTOLI, Professeurs des universités à l'école d'économie de Clermont-Ferrand et directeur adjoint du CERDI

L'utilisation de données d'enquête doit s'appuyer sur un travail préalable d'étude du questionnaire, ainsi que du manuel de l'enquêteur. Cela est indispensable car, « le diable se loge souvent dans les détails » et l'interprétation d'une question pourrait être ambiguë. De plus, les données d'enquête font souvent l'objet d'un traitement avant leur diffusion, qui peut avoir un impact sur l'utilisation des données de la part d'un chercheur. Nous allons voir un certain nombre d'exemples issus du travail de recherche de Simone BERTOLI.

Lundi 5 décembre 2022 de 13h30 à 15h - salle 230 de la MSH de Clermont-Ferrand

Mardi 6 décembre 2022

9h à 11h

Nouvelles technologies et nouveaux comportements du consommateur, deux retours d'expériences

- Environnement socio-technique et comportements : comment nos habitudes sont bousculées par l'hyperconnexion, Pascal BRASSIER, Négociation, Développement commercial, Management commercial, Relation client, Vente, Marketing B2B, Réseaux et médias sociaux.

Cette session a pour objectif d'évoquer l'impact de la technologie sur nos comportements dans un contexte professionnel ou de consommation. Un premier travail montrera comment la gamification d'une activité pédagogique, au moyen d'un jeu massif en ligne, permet aux étudiants de s'engager dans leur apprentissage. Dans un second temps, Pascal BRASSIER expliquera comment comprendre les besoins des consommateurs pour s'exprimer au cours ou à la suite de leurs expériences de consommation de produits ou services. L'exemple d'un site touristique sera choisi.

- Réseaux sociaux numériques et empowerment des consommateurs, Emna CHERIF, Maître de Conférences e-Marketing, gestion de la relation client, interaction homme-machine.

Les réseaux sociaux numériques ont fortement contribué à l'empowerment (processus sociopolitique) des consommateurs dans divers domaines. Les recherches présentées portent principalement sur le rôle joué par les influenceurs en ligne sur les plateformes comme Instagram ou YouTube. Deux projets en cours seront présentés, l'un porte sur le type d'influenceur (« topdog » versus « underdog ») dont l'empowerment de sa communauté. Le second s'intéresse à l'expertise de l'influenceur (médecin versus pairs) dans l'adoption de comportements alimentaires sains.

Mardi 6 décembre 2022 de 9h à 11h - Salle 201 IADT à Clermont-Ferrand

13h30 à 15h30

Apprendre à utiliser et interpréter des statistiques descriptives fondamentales

Frédéric DELARUE, Professeur de mathématiques agrégé en 2016.

Dans l'ensemble des domaines de la recherche en SHS, il est aujourd'hui nécessaire de s'appuyer sur des statistiques, qu'elles soient complexes ou non, pour justifier, démontrer son travail. Frédéric DELARUE, professeur de Mathématiques agrégé va vous expliquer les statistiques élémentaires pour vous familiariser avec elles. Au programme : la définition mathématique et les interprétations qui s'en-suivent, les variances, les médianes, les moyennes, les écarts types, les intervalles de confiance ou même les « boîtes à moustaches », et bien d'autres notions encore.

Mardi 6 décembre 2022 de 13h30 à 15h30 - Salle 332 à la MSH de Clermont-Ferrand

Mercredi 7 décembre 2022

9h à 12h

Atelier Excel : Se familiariser avec le logiciel

Axel DELARUE, Ingénieur d'étude à la Plateforme Universitaire de Données de Clermont-Ferrand

Excel est aujourd'hui utilisé dans tous secteurs de la recherche et propose une multitude de services. L'objectif de cette session est d'apprendre à utiliser le logiciel de manière standard. L'interface du logiciel sera présentée, des exemples de manipulations (insérer/supprimer des colonnes/lignes) seront expliqués et quelques fonctions standards (somme ; moyenne etc) seront montrées. Cet atelier s'adresse à toutes les personnes qui utilisent Excel sans réellement exploiter toutes les possibilités qu'offre le logiciel.

Mercredi 7 décembre 2022 de 9h à 12h - Salle 332 à la MSH de Clermont-Ferrand.

13h30 à 16h30

Atelier EXCEL : Fonctions et Tableaux croisés dynamiques

Axel DELARUE, Ingénieur d'étude à la Plateforme Universitaire de Données de Clermont-Ferrand

Le logiciel offre une multitude fonctions méconnues des utilisateurs. Certaines sont souvent la solution à beaucoup de problème lorsque l'on souhaite manipuler un jeu de données pour ensuite l'exploiter. L'après-midi sera donc consacrée à la découverte de ces fonctions à travers plusieurs exemples. Les Tableaux Croisés Dynamiques seront également au cœur de l'atelier. Souvent peu utilisés car méconnus, ils permettent l'exploration des données de façon rapide et soignée.

Mercredi 7 décembre 2022 de 13h30 à 16h30 - Salle 332 de la MSH de Clermont-Ferrand

Jeudi 8 décembre

9h à 12h

Utiliser des données quantitatives dans les recherches sur le langage : l'exemple de trois approches disciplinaires

- **La méthode d'analyse combinée des représentations sociales des langues (MAC) : limites liées à la collecte de données en ligne - Marie SARRAUTE ARMENTIA**, doctorante CIFRE en sociolinguistique (Office public de la langue occitane et UMR 5478 IKER, Université Bordeaux-Montaigne)

La méthode d'analyse combinée des représentations sociales des langues, ou méthode MAC, est une méthode d'enquête par questionnaire quanti-quali qui a été créée par Bruno Maurer, avec l'aide de Pierre-Antoine Desrousseaux, en 2013. Le but est de pallier plusieurs problèmes méthodologiques identifiés dans les méthodes qualitatives "classiques" de l'étude des représentations sociales, à savoir, les entretiens semi-directifs, qui ne permettent pas de décrire ou de modéliser la structure de la représentation. Cependant, la méthode MAC, de par les solutions mêmes qu'elle propose, a tendance à souffrir d'autres problèmes, comme la méfiance des répondant.es vis-à-vis d'une méthode d'enquête qu'ils/elles perçoivent comme biaisée et non scientifique. Le problème se pose d'autant plus dans le cas d'une diffusion en ligne, au cours de laquelle le/la chercheur.e ne peut pas expliquer ses choix méthodologiques afin de rassurer les répondant.es.

Dans le cadre d'une thèse qui porte sur les représentations sociales de l'occitan en milieu scolaire, Marie Sarraute Armentia a diffusé un questionnaire MAC auprès de parents et d'enseignant.es via des chaînes de mails au sein de l'Education Nationale et d'associations de parents d'élèves. Dans cette communication, elle revient sur la réalisation de ce projet, et plus précisément, sur la réflexion "technique" visant à récolter le moins de "déchets" possibles.

- **Dictionnaires vs. Données orales : Le cas de la gémination de /r/ en anglais - Quentin DABOUIS**, Maître de Conférences en linguistique anglaise & Responsable du N3 LLCER anglais

Les dictionnaires de prononciation de l'anglais rapportent régulièrement que des mots comme *irrational* pourraient voir leur /r/ réalisé comme un /r/ « double » (dit « géminé ») en anglais américain, mais non en anglais britannique. Ce phénomène de gémination a été étudié pour d'autres consonnes, comme /n/ ou /l/, et peut nous informer sur le traitement des mots complexes par les locuteurs. L'équipe en charge de l'étude a donc extrait des occurrences orales de 25 mots en <irr-> sur YouTube via le site *Youglish* et a étudié la durée du /r/ dans ces mots. Les données confirment l'existence d'une différence entre les deux variétés et apportent de nouveaux éléments sur les paramètres conditionnant la durée de cette consonne. La PUD a apporté son soutien au chercheur dans le cadre de ses analyses de données quantitatives.

- **L'approche bayésienne, une approche statistique alternative en psycholinguistique. Un exemple d'application : l'étude du rôle des syllabes dans la production écrite - Cyril PERRET**, maître de conférences, CeRCA - CNRS UMR 7295, Université de Poitiers

Y a-t-il des représentations stockées en mémoire dont le format correspondrait aux syllabes ? Et, si tel est le cas, est-ce que ces représentations sont mobilisées lors de l'écriture ? Le chercheur en psycholinguistique va construire une hypothèse empiriquement testable pour

aborder ce type de question. Dans le cas présent, celle-ci pourrait être : si de telles représentations existent en mémoire, alors leur accès devrait être en partie modulé par la fréquence d'occurrences. La démarche du chercheur consistera ensuite à collecter des données. Il peut, par exemple, faire produire par écrit le label d'un groupe d'images afin de mesurer le temps de réponse, c'est-à-dire le temps s'écoulant entre la présentation d'un stimulus et le début de l'écriture. Enfin, la tâche du chercheur consiste à décider, au vu des données dont il dispose, si son hypothèse doit être conservée ou non. Le critère utilisé en sciences expérimentales pour prendre ce type de décision scientifique repose sur les statistiques. D'aucuns diront que les statistiques permettent de statuer sur la présence ou non d'un effet. Or, les choses sont plus complexes que cette vision naïve peut le laisser penser. Dans cette présentation, Cyril Perret revient sur ces critères de décision. Il discutera notamment de l'approche dite "fréquentiste", c'est-à-dire l'approche de Fisher (1935), fondée sur l'obtention d'une valeur de probabilité appelée p-value. L'objectif est de montrer les limites de ce système de référence et de décrire l'alternative que constitue une approche bayésienne.

Jeudi 8 décembre 2022 de 9h à 12h - salle 332 de la MSH de Clermont-Ferrand

Vendredi 9 décembre

9h à 12h

Initiation au logiciel R

Axel DELARUE, Ingénieur d'étude à la Plateforme Universitaire de Données de Clermont-Ferrand

Le logiciel R est un logiciel d'analyse de données qui ne cesse de se répandre dans le monde de la recherche. Accessible gratuitement, il offre de nombreuses opportunités d'analyses de données, comme des analyses factorielles, des méthodes de clustering, croisement de variables, tests statistiques et bien d'autres possibilités. Durant cette initiation nous apprendrons comment fonctionne le langage R, nous importerons des données dans le logiciel, nous effectuerons quelques traitements standard et enfin nous verrons comment visualiser et exporter des résultats. Il s'agit d'une initiation, par conséquent aucun prérequis n'est attendu !

Pour cela veuillez télécharger R puis RStudio :

Télécharger R : <https://pbil.univ-lyon1.fr/CRAN/>

Télécharger RStudio : <https://rstudio.com/products/rstudio/download/#download>

Vendredi 9 décembre 2022 de 9h à 12h - Salle 332 de la MSH de Clermont-Ferrand