



# Semaine DATA-SHS 2025 de Clermont-Ferrand

Chaque année, la Semaine Data-SHS offre à la communauté des sciences humaines et sociales une occasion privilégiée de plonger dans les enjeux actuels de l'analyse quantitative en science humaine et sociale : de la collecte à l'analyse, de la gestion à la diffusion des données. Sous l'égide de PROGEDO, cet événement se décline dans le réseau des Plateformes Universitaires de Données (PUD) et mobilise chercheurs·ses, ingénieur·e·s, doctorant·e·s et étudiant·e·s autour d'ateliers pratiques, conférences, présentations de jeux de données et retours d'expérience. Une belle occasion d'apprendre, d'échanger et de (re)découvrir tout le potentiel des données SHS dans une ambiance conviviale !

**Public :** Les sessions sont indépendantes et s'adressent à toute la communauté scientifique du site clermontois et d'ailleurs : chercheurs et chercheuses, doctorantes et doctorants, enseignantes et enseignants, étudiantes et étudiants (uniquement pour les séminaires du lundi et jeudi après-midi), ainsi que les personnels d'appui à la recherche, documentalistes, ingénieures et ingénieurs, etc.

Les sessions se dérouleront dans la maison des sciences humaines de Clermont-Ferrand.

**Distanciel :** Possibilité de suivre les conférences du lundi et du jeudi en distanciel, veuillez-vous inscrire et le lien vous sera communiqué par mail. Veuillez faire une demande par mail si vous souhaitez participer en distanciel aux formations à [renaud.hass@uca.fr](mailto:renaud.hass@uca.fr)

**Inscription (obligatoire uniquement pour les formations) :**

<https://enquetes.uca.fr/index.php?r=survey/index&sid=122185&lang=fr>

## Lundi 8 décembre

**10h30 - 12h :** Présentation de la PUD et de ses missions ainsi que de l'actualité nationale du réseau PROGEDO et Huma-Num.

**14h - 16h :** Usage des données d'enquête en SHS : exemple concret

## Mardi 9 décembre

**9h - 12h et 14h - 17h :** Journée d'initiation à R (de l'interface à la création de graphique)

## Mercredi 10 décembre

**9h – 11h :** Utiliser Excel efficacement, trucs et astuces

**14h – 16h :** Initiation à la cartographie avec magrit.cnrs.fr

## Jeudi 11 décembre

**10h – 12h :** Initiation à la gestion des données de la recherche avec GIT

**14h – 16h :** IA et sciences humaines : démystifier, interroger, expérimenter

## Vendredi 12 décembre

**14h – 17h :** Créer son questionnaire avec Lime Survey



# Lundi 8 décembre

## **Ce que la Plateforme Universitaire des Données (PUD) de la MSH peut faire pour vous.**

**10h – 11h en salle 332**

La Plateforme Universitaire de Données (PUD) de Clermont-Ferrand fait partie des quatre plateformes de la MSH et constitue le relais local de l'infrastructure nationale Progedo. Elle accompagne et forme la communauté de recherche en sciences humaines et sociales (SHS) à l'usage des données et des méthodes quantitatives, dans une démarche de science ouverte.

La PUD valorise également les données issues de la statistique publique, disponibles via l'entrepôt Quetelet-Progedo, ainsi que de nombreuses enquêtes internationales, afin d'encourager leur utilisation pour contextualiser ou mener des travaux de recherche. Venez échanger avec notre ingénieur PUD pour découvrir l'ensemble des services proposés par la plateforme.

## **Actualités des infrastructures nationales de recherche en SHS avec Progedo et Huma-Num** (Visio-conférence nationale retransmise en direct)

**11h – 12h en salle 332**

Cette séance commune des deux infrastructures de recherche (IR) en sciences humaines et sociales (SHS) a l'objectif d'exposer les actualités respectives d'une manière conjointe. La mission principale de l'infrastructure de recherche Hum-Num est de mettre en œuvre une infrastructure numérique permettant aux communautés SHS de développer, de réaliser et de préserver sur le long terme les programmes de recherche, alors que l'infrastructure de recherche PROGEDO a pour but de développer la culture des données quantitatives, d'impulser et structurer une politique des données d'enquêtes pour la recherche en sciences sociales.

*Olivier Baude (IR\* Huma-Num) et Nicolas Sauger (IR\* PROGEDO)*

## **Usage des données d'enquêtes en Sciences Humaines et Sociales**

**14h – 16h dans l'amphi 125**

**L'esthétique des prairies : enquête autour d'un service écosystémique culturel sur la Chaîne des Puys.**

La Chaîne des Puys est une petite région naturelle au patrimoine géologique et écologique exceptionnel où se croisent une diversité d'usagers des espaces naturels. Si les prairies sont renommées localement pour faciliter la lecture des reliefs dans un paysage, leur dimension esthétique intrinsèque n'est pas connue. Quels sont les éléments visuels qui permettent d'évaluer l'esthétique d'une prairie ? Quelles sont les variations des préférences d'esthétique des prairies selon le profil des usagers ? Une enquête réalisée en 2022 et 2023 auprès des

randonneurs sur la Chaîne des Puys a permis de recueillir les préférences de 202 usagers à l'aide d'ensembles d'illustrations photoréalistes à la composition contrôlée. La réalisation d'analyses multivariées permet de dégager des préférences générales et par profil de répondant-es pour la perception de l'esthétique des prairies.

*Benjamin Allard, doctorant Geolab*

### Construire son enquête et récolter ses propres données

Au cours de ma thèse, j'ai travaillé sur les publics du théâtre de Shakespeare. Seulement, aucunes données (quantitatives ou qualitatives) n'étaient disponibles à ce sujet. Les enquêtes sur les pratiques culturelles étaient trop larges, proposant des données sur la culture en générale et rarement sur le théâtre en particulier. J'ai donc décidé de récolter mes propres données, à la fois quantitatives et qualitatives. Pendant cette présentation, je souhaite exposer les étapes de la création du questionnaire, la récolte et l'analyse des données au moyen d'outils statistiques comme R, ainsi que l'analyse des données quantitatives au moyen d'outils d'encodage et de textométrie.

*Méline Dumot, doctorante IHRIM*

### Le catalogue des enquêtes de l'Ined, un outil au service de la science ouverte

L'Institut National d'Études Démographiques (Ined) mène des enquêtes quantitatives ambitieuses et novatrices. Dans ce cadre la diffusion de ces enquêtes est un point crucial. La plateforme NADA, s'appuyant sur le standard DDI (Data Documentation Initiative) constitue la solution technique idéale pour leur partage, dans le cadre de la Science Ouverte. Elle est la fondation technique du catalogue de métadonnées DataIned, permettant ainsi la découvrabilité, l'accès et la réutilisation des productions de l'Ined.

*Thibaud Ritzentaller, Ingénieur d'études INED*



## Mardi 9 décembre

### Journée d'initiation à R tous niveaux

**9h - 12h et 14h - 17h en salle 232**

Cette formation s'adresse aux personnes n'ayant jamais utilisé RStudio et souhaitant découvrir ce logiciel pour le traitement de données statistiques. Aucun prérequis en programmation n'est nécessaire, mais il est recommandé d'être à l'aise avec le vocabulaire statistique de base (variable, moyenne, distribution, etc.) et d'avoir déjà manipulé des bases de données, notamment sous Excel. Au cours de cette journée, les participant-es apprendront à utiliser RStudio à partir de données d'enquête réelles.

Penser à ramener votre ordinateur personnel pour pouvoir participer.

*Renaud Hass, IE MSH, plateforme universitaire des données*



# Mercredi 10 décembre

## Utiliser Excel efficacement, trucs et astuces

**9h – 11h en salle 28**

Vous utilisez Excel tous les jours, mais avez-vous vraiment exploré tout son potentiel ? Cette séance de formation tous niveau vous guidera dans la découverte d'astuces et de méthodes efficaces pour exploiter vos données. Au menu : raccourcis clavier, astuce et fonctionnalité pour pouvoir réaliser vos tableaux, analyses et graphiques. Cette formation vous propose une approche conviviale et concrète, basée sur des exemples tirés du quotidien professionnel, pour vous permettre de gagner en autonomie, en rapidité et en précision.

*Renaud Hass, IE MSH, plateforme universitaire des données*

## Initiation à la cartographie avec [magrit.cnrs.fr](http://magrit.cnrs.fr)

**14h - 16h en salle 230**

Les cartes et représentations géographiques font désormais partie de notre quotidien et leur réalisation n'a jamais été aussi simple. Localiser des objets, des personnes ou des phénomènes, ou encore représenter des données statistiques, sont devenus des exercices incontournables.

Nous verrons, à l'aide d'un outil gratuit développé par le CNRS et disponible directement sur navigateur, comment réaliser des cartes thématiques permettant de représenter des indicateurs sociaux, économiques et démographiques.

Penser à ramener votre ordinateur personnel pour pouvoir participer.

*Renaud Hass, IE MSH, plateforme universitaire des données*





# Jeudi 11 décembre

## Initiation à la gestion des données de la recherche avec GIT

10h - 12h en salle 230

Cette formation propose une introduction pratique aux principes de la gestion de version appliquée aux données et aux scripts de recherche. Les participant-es y découvrent comment utiliser Git pour suivre l'évolution de leurs projets, collaborer efficacement et garantir la traçabilité de leurs travaux. À travers des exercices concrets, la formation aborde les notions de dépôt, de commit, de branche et de fusion. Elle met également l'accent sur les bonnes pratiques pour documenter et partager ses données dans une démarche de science ouverte.

*Loris Croce, Ingénieur d'études MSH*

## IA et sciences humaines : démystifier, interroger, expérimenter

14h – 16h dans l'amphi 219

IA, machine learning, de quoi parle-t-on exactement ?

L'intelligence artificielle, terme désormais omniprésent dans le discours scientifique et médiatique, suscite autant d'enthousiasme que de questionnements. Cette introduction ne cherchera pas à débattre de la supposée intelligence de l'IA, mais à en clarifier les fondements en retraçant brièvement l'histoire de ses concepts, que nous ferons débiter en 1942. En explorant les grandes étapes du développement de l'apprentissage machine, cette présentation vise à démystifier ce qu'est réellement l'IA et à offrir des repères pour mieux comprendre les enjeux contemporains.

*Renaud Hass, IE MSH*

*Ce que l'IA change : Promesses, faits, doutes.*

À partir d'exemples tirés de mes recherches et d'études SHS récentes, j'examine ce que l'IA change vraiment.

**Ce qu'elle apporte** : passage à l'échelle (textes, images, rasters), accélération des analyses et standardisation de pipelines reproductibles.

**Ce qu'elle ne remplace pas** : le terrain, l'identification causale et les diagnostics (pré-tendances, robustesses), la délibération scientifique.

**Zones d'ombre** : biais linguistiques/géographiques, boîtes noires des LLM, confidentialité spatiale/RGPD, et risques d'« IA-washing ».

*Pierre Beaucoral, doctorant CERDI*

### L'IA peut-elle aider à (re)lire l'*Encyclopédie* de Diderot et D'Alembert ?

Ouvrage monumental de la seconde moitié du 18ème siècle, l'*Encyclopédie* de Diderot et D'Alembert reste un ouvrage dont la lecture ne pourra jamais être entièrement achevée. Pour cause, les 17 volumes de textes et 11 volumes de planches constituent une somme d'articles que la lecture humaine est incapable de saisir. Pourtant, les catégories du dictionnaire et les branches du savoir, constituées par les éditeurs eux-mêmes, devraient suffire aux chercheur-euses pour s'orienter dans l'ouvrage, dégager des items et constituer des bases de données thématiques (géographie, philosophie, médecine, chimie, etc.). Or, ces données textuelles restent insatisfaisantes : Diderot lui-même s'en plaignait à la fin de l'aventure encyclopédique. L'IA avec son plongement lexical et ses requêtes sémantiques sur les gros corpus de textes pourrait-elle éventuellement être appelée à la rescousse ?

*Nai's Sabatier, doctorante PHIER*



Vendredi 12 décembre

### Créer son questionnaire avec Lime Survey

14h - 17h en salle 230

Limesurvey est un des outils les plus populaires pour réaliser des questionnaires. Nous verrons, après une introduction sur le questionnaire comme technique d'enquête, comment utiliser Limesurvey pour créer un questionnaire (types de questions, natures de variables) mais aussi comment exploiter les données avec excel une fois l'enquête terminée.

Facultatif : Demander l'accès au service Lime Survey de l'université pour avoir accès à la version Premium (<https://dsi.uca.fr/catalogue-de-services/enquetes-uca>)

*Renaud Hass, IE MSH, plateforme universitaire des données*