



Analyser les données biographiques Quelles méthodes ? Quelles approches ? Quels défis ?

28 novembre 2024

INED salle Sauvy, Campus Condorcet, 9 cours des Humanités, Aubervilliers

Cette journée s'inscrit dans le cadre du projet exploratoire du CIST *Parcours de vie et territoires. Collecte et analyse* qui associe l'Ined, le Laboratoire informatique de Grenoble (LIG – Université Grenoble Alpes) et le laboratoire Espaces et sociétés de Rennes (ESO – Université Rennes 2) sur la collecte et l'analyse des parcours de vie.

[sur inscription gratuite mais obligatoire <https://page.hn/1ps6vx>]

- 9h30 *Accueil café*
- 10h00 Introduction
France GUÉRIN-PACE (UR12 / Ined) et **Pascal SEBILLE** (UR12 / Ined & ESO / U. Rennes 2)
- 10h15-11h00 **Nicolas ROBETTE** (UVSQ) – De l'étude des trajectoires à l'analyse de séquences
- 11h00-11h45 **Mathias STUDER** (U. Genève) – Défis méthodologiques et développements récents en analyse de séquences
- 11h45-12h15 Questions
- 12h15 *Pause déjeuner libre*
- 13h45-14h30 **Valérie GOLAZ** (UR15 / Ined & LPED / IRD) et **Arnaud BRINGÉ** (SMS / Ined) – L'analyse des données biographiques et l'approche multi-niveau
- 14h30-15h15 **Thomas GUYET** (AlstroSight / Inria) – Modélisation et analyse des parcours de soins (approche web sémantique, *pattern mining*)
- 15h15-15h45 Questions
- 15h45-16h00 *Pause*
- 16h00-16h45 Table ronde : quels défis pour l'analyse des données biographiques en sciences numériques et sciences humaines ? Regards croisés
- 17h00 *Fin de la journée*

L'objectif de cette journée est d'apporter une réflexion sur les méthodes d'analyse mobilisées à partir des données collectées en sciences humaines et sociales sur les parcours de vie. Elle s'articulera autour de deux axes. Dans un premier temps, la présentation des différentes approches d'analyse des données sur les parcours de vie (analyse de séquences, modèles de durée, *machine learning*, etc.) permettra de préciser les objectifs et les apports de ces méthodes pour l'étude des trajectoires de vie (étude de la survenue d'événements, analyse de trajectoires, etc.). Dans un deuxième temps, un espace de discussion sera proposé pour échanger sur les difficultés rencontrées dans l'utilisation de ces méthodes ainsi que pour mettre en lumière les écueils et les limites qu'elles recouvrent.

Les méthodes de collectes des données des parcours de vie ne seront pas ici abordées. En revanche, la richesse des données biographiques collectées et leur traitement en fonction des objectifs d'analyse recherchés feront l'objet d'une attention toute particulière. L'exploitation de ces données s'avère complexe. On peut citer parmi les obstacles rencontrés : i) la difficulté de prise en compte du temps en raison de durées d'observation inégales selon les individus, ii) l'hétérogénéité des découpages territoriaux au fil du temps (évolution des délimitations géographiques), iii) l'évolution des contextes de réalisation des parcours de vie (transformations sociales, spatiales, économiques, etc.).

Cette journée sera l'occasion de confronter les méthodes d'analyse des trajectoires de vie utilisées en sciences humaines et sociales avec celles expérimentées en informatique. Une attention sera portée à l'intérêt que peuvent susciter les méthodes issues du *machine learning* pour l'analyse des données sur les parcours de vie et la compréhension des problématiques propres aux sciences humaines et sociales.