

# AP Mobilité

Clermont-Ferrand 13 novembre 2019 14h à 17h

Salle A1, AgroParisTech

Organisation : Thomas Devogele et Sonia Chardonnel

## Contexte

Après 5 années d'animation de l'AP Mobilités et Trajectoires, nous faisons avant tout le constat que les enjeux scientifiques (théoriques et méthodologiques) soulevés depuis quelques années par l'analyse des mobilités individuelles nécessitent des interactions fortes entre deux domaines :

- L'informatique : base de données, ingénierie des ontologies, fouilles de données
- Les sciences sociales occupées à produire de la connaissance sur la mobilité d'individus ou d'objets : géographie, sociologie, psychologie, démographie, etc.

Alors que les sciences sociales héritent de méthodes de collecte éprouvées pour l'observation et l'analyse des mobilités individuelles (enquêtes transport, enquêtes ménages et déplacements, enquêtes emplois du temps, enquêtes biographiques...), les sciences de l'information permettent aujourd'hui d'acquérir rapidement et de manière (presque) continue des données (très) nombreuses issues de traces numériques décrivant l'activité (c'est à dire *a minima* la présence dans 1 lieu à 1 temps donné) des individus porteurs de capteurs (téléphones géolocalisés, montres connectées...) et/ou utilisant des réseaux sociaux. Les recherches interdisciplinaires informatiques/sciences sociales ont permis, ces dernières années, d'explorer la diversité de ces nouvelles sources en termes de nature des données, de formats, de contenus numériques et d'enrichissement sémantique. En outre, plusieurs revues de littérature existent montrant la portée de l'application de ces données dans des modèles de mobilité ou de dynamiques spatiales urbaines.

Face à ces questions non résolues, il nous semble que l'agenda de recherche que notre communauté interdisciplinaire doit ouvrir aujourd'hui est celui de l'invention de méthodes hybrides qui associent à la fois des données issues de collectes « classiques » (recueil auprès de personnes) souvent très riches mais longues et difficiles à obtenir et des gros volumes de données issues des traces numériques d'activités individuelles. L'enjeu interdisciplinaire est de trouver des modèles et des méthodes pour bien prendre en compte les connaissances sur les comportements de mobilité issues d'enquêtes classiques dans des raisonnements informatiques appliqués à des grands volumes de données et des sources de données variées complémentaires. Cet agenda spécifique à la question des connaissances sur les mobilités individuelles s'inscrit plus largement dans la programmation de recherche sur les humanités numériques.

# Verrous

Si globalement les résultats sont très prometteurs, un certain nombre de verrous persistent. Nous les regroupons en 6 groupes. Nous vous proposons deux séances de travail. Lors de chaque séance, quatre groupes seront définis. L'objectif de chaque groupe est de partager leur connaissance sur cette question et de proposer des pistes de recherche.

## Première séance

### **Groupe 1 : Sources d'information existantes**

Quelles sources de données existantes devons-nous prendre en considération pour mieux comprendre les comportements humains ?

### **Groupe 2 : Gestion de l'hétérogénéité et de la qualité des sources**

Comment gérer l'hétérogénéité des jeux et sources de données qui peuvent être mobilisés pour décrire une même population, un même territoire ?

### **Groupe 3 : Enrichissement sémantique**

Quelles méthodes employées pour enrichir sémantiquement les traces brutes issues des capteurs ?

### **Groupe 4 : méthodes hybrides**

Quelles nouvelles méthodes hybrides (fouille de données, statistiques, analyse spatiale) définir pour prendre en compte la richesse sémantique des données issues d'enquêtes et le volume important des traces issues de capteurs ?

## Deuxième séance

### **Groupe 5 : Etude de mobilité et analyse visuelle**

Comment intégrer une approche d'analyse visuelle de ces données hétérogènes pour faciliter le raisonnement d'experts non informaticiens ?

### **Groupe 6 : Etude de mobilité et IA**

Comment aller vers des processus d'enrichissement sémantique automatique ou semi-automatique des traces à l'aide de techniques d'intelligence artificielle dont l'apprentissage profond ?

### **Groupe 7 : Ontologie et Open data**

Comment intégrer des ontologies ou des sources de données ouvertes (Open data) pour prendre en compte la richesse sémantique des données d'enquêtes dans des processus d'analyse informatique ?

## **Groupe 8 : Ethique**

Quelle éthique adopter dans l'usage de ces données à des fins de recherche alors que leurs producteurs (souvent privés) visent des objectifs commerciaux ?

# **Déroulé de la demi-journée**

**14 h** accueil, présentation rapide des participants et des objectifs de la séance

**14h30** première séance de groupes de réflexion autour verrous scientifiques des 4 premiers groupes : 40 min en petits groupes de travail puis 10 min synthèse partielle

**15h20** pause

**15h40** deuxième séance de groupes de réflexion autour des verrous scientifiques des 4 derniers groupes : 40 min en petits groupes de travail

**16h20** Synthèse globale & conclusion

**17h** fin de la demi-journée