

## Atelier Grandes métropoles (axe Information Territoriale Locale)

### COMPTE-RENDU SEANCE 1

Campus Paris Diderot - 17 mai 2016

#### Présents :

Martine Drozd (Latts), Sarah Duché (Pacte, Université Grenoble Alpes), Sylvestre Duroudier (Géographie-cités, Université Paris Diderot), Bernard Elissalde (Idées, Université de Rouen), Julie Fen Chong (Thema, Université de Bourgogne), Claude Grasland (Géographie-cités, Université Paris Diderot), Marianne Guérois (Géographie-cités, Université Paris Diderot), Renaud Le Goix (Géographie-cités, Université Paris Diderot), Françoise Lucchini (Idées, Université de Rouen), Malika Madelin (Prodig, Université Paris Diderot), Florence Nussbaum (Géographie-cités, Université Paris Diderot), Hugues Pecout (CIST), Jean-François Valette (Prodig).

Excusées : M. Carmes (CNAM), C. Carré (Ladyss, Univ. Paris 1), M. Severo (Geriico, Univ. Lille 3).

#### Organisation de la journée :

##### – **Matin : le projet Grandes métropoles, amorce et perspectives**

- Retour rapide sur les **premières analyses** réalisées sur les villes de Paris, Chicago et Mexico et trois thématiques (occupation du sol, démographie, marché immobilier), présentées en mars dernier lors du colloque du CIST (cf. diaporama en ligne sur la page de l'axe ITL).
- Echanges autour des **objectifs** du projet, des sujets possibles des **prochains ateliers**, du **format de la plate-forme** en construction : quelles thématiques ? quelles villes ?

##### – **Après-midi : deux ateliers en salle informatique**

- Le vote EELV aux élections européennes en Ile-de-France : analyse multiscalaire (HyperAtlas) et anamorphose (ScapeToad) (C. Grasland).
- Approches de l'information ponctuelle par méthodes de lissage : les données immobilières AirBnB (M.Madelin, M. Guérois).

Une synthèse de la matinée est proposée ci-dessous. Les diaporamas des ateliers de l'après-midi sont disponibles en ligne.

#### 1. Le projet Grandes métropoles :

---

- **Rappel de l'objectif général du projet :**

- **Dans le cadre du CIST** : initié par l'axe Information territoriale locale (ITL), destiné à être un projet transversal **au croisement des axes** (thématiques) et **des labos** (sollicitation des ingénieurs de l'information géographique, en particulier).

- Construction d'une **plate-forme d'échanges** pour réfléchir aux **problèmes soulevés par le traitement de données locales de plus en plus foisonnantes et hétérogènes** : choix des échelons d'analyse pertinents pour des données élémentaires de résolution très fine, croisement de données thématiques de granulométrie ou de géométries différentes, articulation entre données d'enquête et données contextuelles, traitement et visualisation de phénomènes dont la nature discontinue ressort fortement à l'échelon local, etc.

L'objet des ateliers qui accompagne la construction de cette plate-forme peut être théorique (et/ou) méthodologique, technique ou bien encore lié à l'exploration d'un jeu de données en particulier.

- **Quels terrains pour ce « bac à sable » ?** Les grandes métropoles, affectées par des dynamiques accélérées de recompositions sociales, économiques ou environnementales, sont des terrains d'étude riches car elles appellent à la mise en place de nouveaux dispositifs d'observation, mais aussi de nouvelles techniques de visualisation et de mise en relation de données élémentaires utilisant des géométries ou temporalités différentes.

Trois métropoles ont été choisies dans un premier temps (Paris, Chicago et Mexico), en fonction de la familiarité des membres du projet avec ces terrains et de la disponibilité des données.

- Tout ou partie des données utilisées pour ces « bacs à sable » seront mises en ligne sur la plateforme.

L'objectif n'est pas tant de constituer une base transversale aux axes du CIST que de proposer des jeux de données tests, à partir desquels **soulever des problèmes théoriques ou méthodologiques** et **proposer des démarches génériques (« how to ») pour les résoudre.**

- **Avancement et perspectives :**

- **Juillet 2015-Mars 2016 : premières analyses sur « échantillon d'apprentissage ».** 3 thématiques (utilisation et occupation du sol, démographie, marché immobilier) et 3 métropoles (Paris, Chicago, Mexico) pour amorcer le projet → Communication proposée au colloque du CIST, 17-18 mars 2016 (cf. diaporama)

- **17 Mai 2016 :** 1<sup>er</sup> atelier Grandes métropoles proposé aux membres de l'axe Information Territoriale Locale.

- **Prochaines étapes...** : un nouvel atelier Grandes métropoles à l'automne 2016 ; élargissement des thématiques et des villes + mise en place de la plate-forme, avec l'aide d'Hugues Pecout (ingénieur CIST)

## 2. Les ateliers Grandes métropoles : quelles villes, quelles thématiques pour les prochains séminaires ?

---

- **Nouvelles villes proposées en plus de Paris, Chicago, Mexico**

- **Londres** (M. Drozd) : quelles données disponibles ?

Pas de pb pour accéder aux données démos, notamment depuis 1981.

Problème des découpages locaux : immense mobilité géographique des mailles du recensement. Comparaison diachronique très compliquée. Algorithmes utilisés spécifiquement pour construire des correspondances entre certaines dates (existe pour 2001-2011, sauf ds les quartiers de régénération urbaine). Plus gros pbs pour la période 1991-2001.

Données immobilières : nombreuses mais question de la capture des données en accès libre mais qui ne le seront bientôt plus !

- Autres villes évoquées (sous réserve) : Delhi, Lima, Le Cap ...

- **Exemples de thématiques proposées par les participants**

### 1) Big data et pratiques individuelles de la ville

Intérêt pour les **données issues des médias sociaux** (ex Twitter) ou de la téléphonie mobile, en lien avec la question de la pratique quotidienne des villes (B. Elissalde, F. Lucchini)

1. Question théorique sur la temporalité de ces pratiques : hors des repères intercentraux, aller vers des pas de temps différents. Info temporelle très fine, liberté de découpages temporels et spatiaux (cyclicité, résilience, ... routines ou événement spécifique). Définition d'emboîtements temporels, par analogie avec les contextes spatiaux.
2. Catégorisation de l'espace ou des gens qui pratiquent la ville : enjeux de construction de catégories nouvelles, en croisant espace et temps, nouvelles formes de pratiques de la ville.
3. Dans la comparaison des villes : diffusion et degré d'utilisation de ces nouveaux outils changent énormément d'un pays à un autre.

Autres pistes :

Travaux sur les données Twitter développés dans le cadre du projet ESPON BigData et poursuivis par l'axe media du CIST (M. Severo, H. Pecout)

Médias « traditionnels » et presse : tendances et anomalies détectées à partir des flux RSS (C. Grasland) ; discours sur les conflits d'aménagement dans les médias (M. Drozd mentionne outils d'analyse diachronique des articles de presse développés entre autres par YF Le Lay de l'ENS Lyon).

## 2) Pollution de l'air, mesures et représentations (S. Duché, M. Madelin...)

→ **Disponibilité données** ? En dehors de la métropole parisienne (données d'AirParif), données accessibles à Chicago (11 stations, mais fonctionnement très irrégulier), à Londres (30n), à Mexico.

→ Quelles agrégations spatiales et temporelles des données collectées au niveau des stations ? Peut-on **proposer une mesure globale de pollution** pour comparer les villes entre elles ?

C'est un enjeu important car souvent, dans les études qui visent à comparer le niveau moyen de pollution dans les grandes métropoles, l'agrégation de données hétérogènes (par ex. stations de trafic et stations de fond) pose problème

→ **Comparaison mesures individuelles (capteurs de pollution) / stations « officielles »**

**Stations officielles** : quelle interpolation pour avoir une information sur l'ensemble du territoire ? Ex des cartes lissées de pollution en IDF (AirParif) = utilisation d'une chaîne de modèles (interpolation + modèle de dispersion + cadastre des émissions). Jusqu'à quel échelon d'observation cette information est-elle valable, par exemple pour calculer des indicateurs d'exposition ? Et quid de l'accès aux données ?

**Capteurs** : exemple de la région Rhône-Alpes où se développe un réseau de capteurs citoyens, ou bien des FabLab dans certaines villes qui contribuent au développement de plate-formes de données participatives (Chicago, Londres -enjeux de débat public, données utilisées au niveau métropolitain-, Paris avec le projet Quel est ton air, par exemple).

Soulève des questions de transmission de l'information collectée : quel message on diffuse à partir de ces données issues des sc. participatives ? Comment agréger par rapport au quartier (plus proche des gens) ?

## 3) Trajectoires territoriales : analyse de séquences territoriales (pistes lancées en particulier par M. Drozd, R. Le Goix)

→ Il existe des méthodes d'analyse de séquences appliquée aux trajectoires individuelles, **très utilisées par sociologues** (par ex à Genève, sur insertion professionnelle, carrières militantes...).

→ **Enjeu théorique pour les territoires**, comment conceptualiser la notion de séquence pour les territoires ? Comment réintégrer l'espace dans les analyses sociologiques de trajectoires individuelles types ? En quoi ces méthodes permettraient-elles d'appréhender différemment certaines transitions (notamment gentrification) ? Quels tps d'accélération dans le changement social ?

→ **Enjeu méthodologique** : comment identifier des trajectoires territoriales types ? Comment identifier le comportement particulier de certains quartiers qui changent d'état, par rapport à un « glissement global » ?

Différents types d'outils développés sous R à explorer : *Traminer* (notamment pour la représentation des séquences), *Weighted cluster* (Matthias Studer : classif appliquées aux données séquentielles, changements qualitatifs d'états en très grand nombre).

4) Mobilités à partir des enquêtes sur **trajectoires résidentielles ou enquêtes ménages déplacements** (J. Fen Chong, C. Grasland, JF. Valette...)

5) **Réseaux de transport et accessibilité**, autour de **l'exploitation d'OpenStreetMap**. 2 thématiques envisagées pour l'instant : d'une part accessibilité et calcul de distances réseaux, d'autre part analyse des structures urbaines à partir des caractéristiques de la voirie (C. Grasland évoque les travaux de C. Lagesse inspirés de la « space syntax ») et des types de réseaux viaires (cf article de Wheeler cité ds le diaporama du 17/03/2016).

...

### 3. La plate-forme Grandes métropoles :

---

Au-delà des ateliers... comment assurer la capitalisation des petits travaux qui vont s'accumuler dans le cadre du projet ?

- A court terme : un **portail d'accès aux données** utilisées par les membres du projet
  - Plate-forme de centralisation des BD (données brutes + un important travail de pré-traitement des données a été assuré par H. Pecout pour les premières données mises en ligne)
  - Question de l'accès aux données à préciser
  
- Dans un second temps (à la suite de plusieurs ateliers...), un **blog**
  - Alimenté, entre autres, par (et/ou) des diaporamas, des fiches « How to », des programmes de traitement, mais aussi éventuellement par une veille.
  - Mise en ligne de documents de travail (par ex. sur le modèle du site du GAWC et de ses *working papers* ou du TGE Belgrand)
  - Là encore, questions à résoudre sur l'accès au site.

Prochain atelier : automne 2016