

Groupe transversal proposant le sujet de post-doctorat

| | |
|---|-------------------------------------|
| Justice, espace, discriminations, inégalités | <input type="checkbox"/> |
| Métabolisme urbain | <input type="checkbox"/> |
| Métiers de la ville | <input type="checkbox"/> |
| Mobilités urbaines pédestres | <input type="checkbox"/> |
| Modélisation et dynamiques urbaines | <input type="checkbox"/> |
| Penser l'urbain par l'image | <input type="checkbox"/> |
| Production urbaine et marchés | <input type="checkbox"/> |
| Réseau international de recherche sur les villes diffuses | <input type="checkbox"/> |
| Usage de l'histoire et devenirs urbains | <input type="checkbox"/> |
| Ville et énergie | <input type="checkbox"/> |
| Ville et numérique | <input checked="" type="checkbox"/> |
| Ville, tourisme, transport et territoire | <input type="checkbox"/> |

1 – Intitulé du sujet

Big data et évolution de la modélisation urbaine

2 – Présentation du projet

Le post doctorat s'inscrit dans les activités du groupe ville et numérique du LABEX futurs urbains et doit contribuer à alimenter une réflexion nodale pour ce groupe qui rassemble des spécialistes de la modélisation urbaine et des chercheurs en sciences humaines intéressés par les effets du numérique sur la ville.

Dans les questions de gestion urbaine, les modèles ont pris une place importante depuis plusieurs décennies. Les modèles allient le traitement de faits mesurés et des hypothèses sur les phénomènes étudiés. Ils permettent d'inférer des données là où il n'y a pas de capteur (dimension spatiale) ou de faire des prédictions (dimension temporelle). Ces modèles ont une histoire ancienne, comme dans le cas des transports où les premières méthodes ont été développées dès les années 1950 aux Etats unis pour traiter des problèmes de congestion (actuelle et anticipée), et déployées en France une décennie plus tard. Ils sont aujourd'hui communément utilisés pour des politiques de transport (politiques tarifaire, interdiction du stationnement, ...) ou la gestion en temps réel du trafic. Ces modèles n'ont depuis cessé d'étendre leurs

domaines d'application et leur sophistication, passant par exemple de modèles déterministes simples aux modèles probabilistes, dans une volonté de mieux appréhender le réel.

Autour de l'usage de ces modèles se sont stabilisées des relations entre les acteurs de la gestion urbaines et les centres d'études et de recherche. Certains résultats de modèles servent à paramétrer l'intervention des opérateurs urbains et toute nouvelle question est renvoyée par eux aux organismes de recherche qui déploient de nouveaux modèles ou adaptent d'anciens. Les opérateurs cherchent dans les modèles des outils de pilotage ou d'aide à la décision, tandis que les chercheurs y voient également des moyens de produire de la connaissance.

Deux phénomènes nouveaux s'imposent cependant. En premier lieu grâce à la diminution des coûts des capteurs et à l'augmentation des capacités de stockage de l'information, les données « classiques » issues de ces capteurs dédiés augmentent en volume. En second lieu des données qui sont produites en masse dans des finalités de service et non dans des visées de connaissance peuvent être dans certaines conditions remobilisées pour la connaissance. C'est ainsi par exemple que les données de localisation des téléphones portables que possèdent les opérateurs téléphoniques ou des données de la télébilletique collectées par les opérateurs de transport peuvent être utilisées pour connaître les déplacements. Le mouvement déjà engagé dans certains domaines d'introduire un caractère stochastique dans la modélisation peut s'amplifier face à ces nouvelles données qui répondent en grande partie aux quatre « V » des big data : volume, variété, vélocité (mises-à-jour) et véracité (par opposition aux données de mauvaise qualité qui ont un coût économique et social).

Sur le papier on peut imaginer d'importantes transformations épistémologiques. Etant donné que les modèles considérés ici servent principalement à produire de nouvelles données, on peut s'interroger sur leur utilité dans une situation où l'observation devient permanente et omniprésente. A l'extrême, on peut s'orienter vers une approche purement descriptive des situations, appuyée par les méthodes de la statistique moderne, rejetant comme obsolètes et inutiles toutes tentatives de poser des hypothèses sur les mécanismes des phénomènes étudiés ou de se référer à des cadres théoriques. Ainsi alors que les panneaux à messages variables donnent des informations à partir de quelques mesures qui alimentent des modèles, Waze estime en temps réel des temps de parcours à partir des positions GPS à un instant donné. Mais on peut anticiper aussi des transformations socioéconomiques : prestation de connaissances par des opérateurs non urbains. De ce point de vue ces évolutions interrogent les politiques publiques de la science et des techniques appliquées à l'urbain.

Ces questions ont été débattues lors du premier séminaire du groupe ville et numérique en relation avec le PUCA : les éléments sont développés dans le compte rendu vidéo.

<http://www.urbanisme-puca.gouv.fr/cycle-villes-intelligentes-2-seance-2-modeles-et-a1150.html>

3 – Mission du post-doctorat

L'objet du post-doc sera d'instruire le dossier de la confrontation entre modèles et big data.

4 – Objectifs détaillés

Pour atteindre cet objectif, trois perspectives sont envisageables.

La première est de confronter la modélisation « traditionnelle » à l'usage de « big data » au travers de cas d'étude, sous l'angle de la performance comparée des approches. L'idée est de « mettre les mains dans le cambouis » en menant simultanément des analyses dans les deux orientations pour qualifier dans la méthode et les résultats les différences entre les approches. Quels apports, quelles informations nouvelles peuvent être tirées de données de masse et qui ne le sont pas de modèles ? Quels types de traitement les données de masse doivent-elles subir pour être utilisées dans des processus d'apprentissage de type machine-learning ? Quels sont ces méthodes d'apprentissage ? Pour quels usages les modèles sont-ils mis en question ? Comment les données de masse, par leur quantité et par leur caractère souvent « atypique »

par rapport aux mesures usuelles, font éventuellement évoluer les modèles initiaux ? Une attention particulière sera portée aux deux catégories d'usage classiques des modèles que sont la gestion opérationnelle et la planification. La gestion opérationnelle suppose l'intégration d'informations et de prévisions à court terme. La planification suppose des horizons moyens à longs, et des modèles reflétant les effets sur le système de telle ou telle orientation ou prise de décision. Ces deux catégories d'usage et de modèles sont susceptibles de réagir très différemment à l'introduction des données de masse. Les choix des domaines d'application (transport, eau, ...) seront fait en lien avec les laboratoires engagés dans le groupe ville et numérique du labex.

La seconde consiste à s'interroger sur les fondements épistémologiques et philosophiques liés à cette confrontation entre modèles et big data. Etant donnée la relation dialectique profonde entre données et mesures, la différence entre modèles et big data est-elle une différence de nature ou simplement de degré ? Comment cette opposition peut-elle être rapportée aux débats qui traversent la modélisation : peut-on modéliser les systèmes anthropiques avec les mêmes approches que les systèmes naturels ? La modélisation numérique et la simulation posent-elles des questions qui leur sont propres, et qu'elles partageraient avec les big data, mais les différencieraient des approches mathématiques traditionnelles ? Les approches probabilistes sont-elles purement descriptives, ou incorporent-elles aussi des présupposés non explicités ? La modélisation par analogie (par exemple entre trafic automobile et fluide partiellement compressible) est-elle dépassée et vouée à disparaître ? Quelles sont les modalités pratiques d'hybridation entre les deux perspectives ?

La troisième est d'appréhender cette question dans une perspective plus sociologique ou socio historique. Qui sont les laboratoires qui promeuvent l'une et l'autre approche ? Quels sont les financeurs de leurs travaux ? Quels sont les acteurs de la convergence et ceux de la confrontation entre ces approches ? Les approches big data permettent-elles à des acteurs économiques nouveaux d'investir les champs balisés de la gestion urbaine ? Comment les pouvoirs publics prennent-ils (ou non) en charge cette question. En somme il s'agirait de dresser le tableau d'une écologie sociale de la confrontation entre ces deux perspectives.

La réponse à l'appel à candidature doit porter sur l'une des trois perspectives ci-dessus. Toutefois, une attention sera tout particulièrement portée à la capacité, à partir de l'une de ces trois perspectives, à intégrer des aspects des autres (par exemple confronter les méthodes en maintenant une réflexion épistémologique, aborder les rapports de force sociologiques en intégrant une analyse des transformations des pratiques de modélisation...)

5 – Résultats attendus

Les résultats attendus varient selon le profil du candidat.

Pour ce qui concerne l'ingénieur modélisateur, un exercice de modélisation sera effectué en mobilisant des méthodes différentes sur des objets liés à l'eau ou au transport. Cet exercice débouchera sur une analyse des avantages comparés des deux méthodes.

Pour ce qui concerne les deux autres profils, le travail devra reposer sur un matériau empirique, en particulier des entretiens menés avec les scientifiques et les amenant à expliciter les enjeux de leurs choix méthodologiques et tout autre document sur le développement de la science. Ce travail empirique débouchera sur un rapport de synthèse instruisant le dossier modèle versus big data.

Dans tous les cas, ces résultats ont vocation à contribuer à l'animation du groupe de travail ville et numérique et seront discutés en séminaire. Le travail de post doctorat doit déboucher sur des publications scientifiques dans les revues adaptées.

6 – Faisabilité

Il s'agit d'un post-doc d'un an potentiellement renouvelable une fois. Le projet de recherche individuel présenté par chaque candidat.e devra donc inclure un phasage avec un projet initial sur une année et une poursuite sur une seconde année conditionnée au renouvellement du contrat. Une attention particulière sera portée sur la faisabilité du travail proposé sur la première année et sur l'articulation projetée des deux phases

Dans le premier cas du profil ingénieur modélisateur, la faisabilité est conditionnée à l'accès à des bases de données pertinentes pour l'exercice. Le candidat devra prendre contact avec Olivier Bonin et José Frédéric Deroubaix pour discuter du choix des exercices de modélisation et des partenariats à établir pour l'accès à des bases de données. olivier.bonin@ifsttar.fr j.deroubaix@enpc.fr

7 – profil du / de la post-doctorant-e

Les trois perspectives proposées correspondent à trois types de profils différents : ingénieur modélisateur / épistémologue / historien des sciences ou sociologue des techniques. Chacun de ces profils est éligible, cependant, un profil hybride sera particulièrement apprécié. Un ingénieur ouvert aux questions de sciences humaine pour la première, un philosophe des sciences ayant une expérience de la pratique scientifique pour la seconde, un sociologue ou historien des techniques capable de discuter avec les chercheurs du contenu de leur travaux pour le troisième.

8 – comité de suivi au sein du groupe transversal

Anne Aguiléra, LVMT
Olivier Bonin, LVMT
Céline Bonhomme, LEESU
José-Frédéric Deroubaix, LEESU
Gilles Jeannot, LATTs

9 – chercheur-e-s à contacter dans le laboratoire d'accueil

Gilles Jeannot, directeur-adjoint du Lattts
E-mail : Jeannot@enpc.fr

10 – Embauche

| | |
|------------------|---|
| Durée du contrat | Un an renouvelable le post doc à compter soit du 1er novembre 2017 ou du 1^{er} janvier 2018 |
| Rémunération | 2 000, 00 € net/m |
| Localisation | Campus de Marne-la-Vallée, Champs sur Marne (77) |

Pièces à joindre au dossier :

- Résumé de la thèse + documents de soutenance (pré-rapports et rapport). Dans le cas où la soutenance n'a pas encore eu lieu, les pré-rapports suffiront.
- Curriculum Vitae.
- Lettre de motivation indiquant à quel projet le/la candidate pourrait prendre part et un exposé de sa compréhension du sujet et de la contribution qu'il/elle pense pouvoir y faire, en deux pages.
- Présentation du projet de recherche pour la 1^{ère} année (3 pages)
- Présentation du projet de recherche exploratoire pour la 2^{ème} année éventuelle (3 pages)
- Date limite de remise du dossier de candidature complet : **vendredi 13 octobre 2017 (à 18 h 00)**
- Le dossier de candidature, accompagné des pièces demandées, doit être renvoyé, par mail ou par l'application « We transfer » à :

Catherine Alcouffe

responsable administrative du LABEX Futurs Urbains

e-mail : catherine.alcouffe@univ-paris-est.fr ou catherine.alcouffe@enpc.fr

P.S. : Tout dossier incomplet ne sera pas recevable