



HAL
open science

Belgrand : un grand équipement pour l'utilisation des bases de données dans le domaine de la ville et de la mobilité

Jean-Paul Hubert, Olivier Bonin, Martin Henriquez

► **To cite this version:**

Jean-Paul Hubert, Olivier Bonin, Martin Henriquez. Belgrand : un grand équipement pour l'utilisation des bases de données dans le domaine de la ville et de la mobilité. colloque GIS - Modélisation urbaine " Modélisation de la ville : du modèle au projet urbain ", Feb 2011, Champs-sur-Marne, France. pp.296-298. halshs-00737394

HAL Id: halshs-00737394

<https://shs.hal.science/halshs-00737394>

Submitted on 1 Oct 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Belgrand : un grand équipement pour l'utilisation des bases de données dans le domaine de la ville et de la mobilité

Jean-Paul Hubert, Olivier Bonin, Martin Henriquez
UPE-IFSTTAR

Les équipes de recherche ont intérêt à mutualiser les acquis des travaux d'ingénierie pour l'intégration de données sur la ville, qu'elles soient statistiques, géographiques, environnementales, techniques ou résultant de simulation, mais aussi les démarches pour l'acquisition de ces données. Le projet de centre de ressources « Belgrand » vise à répondre à cet enjeu majeur en s'appuyant sur : (1) un réseau scientifique multidisciplinaire ; (2) un portail web de référencement des projets, données, méthodes et outils, simulation et modèles, le cas échéant ; (3) des travaux méthodologiques mettant l'accent sur la référence spatiale, clé naturelle de croisement des données portant sur la ville ; (4) un séminaire régulier. L'infrastructure web Belgrand mettra en lien projets, équipes et terrains, offrant une vue synoptique des travaux de modélisation urbaine pour les chercheurs et les décideurs.

Capitaliser les savoir-faire sur les données pour la ville

Le besoin de rassembler les compétences sur les données utiles à la recherche sur la ville, la mobilité et l'environnement a été maintes fois souligné. Les données utilisées dans les projets de recherche sont rarement produites dans ce but. Leur acquisition requiert des démarches auprès des producteurs ou diffuseurs et leur utilisation, des travaux d'ingénierie allant du simple reformatage à la reconstruction de bases de données. Ce travail est coûteux, *ad hoc*, mais réutilisable en grande partie dans d'autres projets, à condition d'être bien référencé et documenté. La recherche pourrait y gagner en efficacité, et élargir le spectre des données utilisables.

Le contexte est porteur. Nombreuses sont les impulsions politiques, scientifiques ou commerciales qui poussent à faciliter l'échange des données et leur valorisation : la directive européenne Inspire⁸⁶, le livre blanc du consortium PARADE⁸⁷, les travaux sur l'interopérabilité des systèmes d'information... ou encore les produits de Google qui entrent massivement dans notre quotidien. La France réalise depuis quelques années un effort important :

- pour la mise à disposition des données (Réseau Quételet, Geoportail et Geocatalogue, politique de l'IGN à l'égard de la recherche, initiatives de collectivités telles que Rennes ou Paris),
- dans l'organisation des observatoires statistiques, régionaux ou nationaux, intégrant des données souvent d'origine administrative,

dans les domaines de l'environnement, de l'urbanisme, de la santé...

- dans la technologie sur les données à des fins pratiques ou d'observation : création de l'Agence française pour l'information multimodale et la billétique (MEDDTL-DGITM), enquête pilote GPS intégrée à l'enquête nationale transports et déplacements (SOeS-Insee-Inrets), financement des EQUIPEX DIME-SHS (panel équipé de « Smartphones » ScPo, Ined) et CASD (accès sécurisé aux données ENSAE Réseau Quételet), recherches en sécurité routière utilisant des enregistreurs embarqués (INRETS-DSCR : projets S-VRAI, LAVIA).

Mais la multiplication et l'informatisation des données ne vont pas sans de nouvelles contraintes. Pour protéger la vie privée, la CNIL⁸⁸ impose une rigueur croissante dans les procédures d'anonymisation, qui a pour effet pratique de supprimer des fichiers beaucoup de références spatiales. Le comité du secret statistique du CNIS contrôle l'accès à ces données indirectement confidentielles [1]. Parce qu'elles ont un caractère commercial, bien des données ne sont pas rendues publiques ou ne le sont plus suite à la privatisation d'opérateurs de transport ou de télécommunication. Enfin, si les innovations technologiques semblent permettre aux instituts statistiques de faire des économies substantielles, les méthodologues rencontrent de grandes difficultés pour garantir la représentativité des

⁸⁶ <http://inspire.jrc.ec.europa.eu/>

⁸⁷ <http://www.csc.fi/english/pages/parade>

⁸⁸ Commission nationale de l'informatique et des libertés

données collectées grâce à ces nouvelles méthodes.

Pour mobiliser ce corpus d'informations potentiellement utiles chaque jour plus large et intégrer les données, les équipes de recherche ont intérêt à mutualiser les acquis du travail d'ingénierie et des démarches pour l'acquisition

L'objectif est de créer un cercle vertueux d'accumulation de compétences techniques et pratiques sur l'ingénierie des données, à partir des projets de recherche sur la ville durable, en s'appuyant sur :

- un réseau scientifique sur les données ouvert aux équipes de toutes disciplines ;
- un portail web de référencement des projets, données, méthodes et outils ;
- des travaux méthodologiques sur l'intégration de données, ressource scientifique ;
- un cycle de séminaire de deux ans, avec publication d'actes, sur l'ingénierie des données (au sens large) pour la ville.

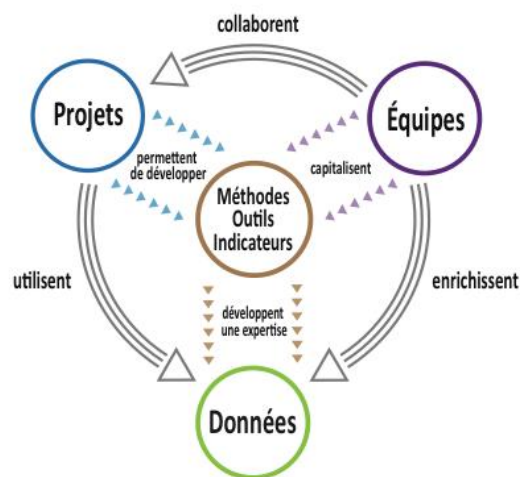
Les points forts du projet

Quatre points peuvent être soulignés :

- le projet innove par son organisation distribuée et adossée à une communauté d'intérêt scientifique ; il opérera en toute neutralité à l'égard de la conduite et de la thématique des projets de recherche ;
- son action pour clarifier au maximum les conditions d'utilisation et de partage des données aidera au montage de partenariats élargis, impliquant notamment des industriels ;
- l'équipe du projet rassemble des compétences à l'interface du monde de la statistique, de l'observation environnementale, de l'information géographique et de la modélisation ; il bénéficie d'une forte dynamique de pôle à Marne la Vallée (Labex « Futurs Urbains », IEED...) ;
- d'envergure nationale, il est voué par nature à fonctionner en réseau avec d'autres équipements analogues dans d'autres pays.

Le centre Belgrand sera ouvert à l'ensemble des acteurs : chercheurs, producteurs de données, collectivités, bureaux d'étude, etc. Il appliquera le principe de « l'équipe la plus favorisée » pour mettre à disposition de tous les modalités pratiques et juridiques d'accès aux données (e.g. conventions pour des partenariats

des données [2]. Le projet de centre de ressources « Belgrand » vise à répondre à cet enjeu majeur. La configuration générale du projet a été définie dans une réponse à l'appel à projet EQUIPEX en septembre 2010 et une première phase de travaux est lancée, qui associe l'IGN, l'FSTTAR et l'IRSTV grâce à une subvention du MEDDTL.



particuliers protégeant contre le pillage des données etc.) ainsi que toutes les informations techniques et métadonnées.

Ses ressources web accéderont directement aux catalogues des producteurs pour aider à traduire une problématique en modèle d'intégration de bases de données. Elles rendront compte, par des moyens bibliométriques, des utilisations précises des fichiers de données dans les publications et les études. Producteurs et chercheurs pourront ainsi être informés des utilisations réelles et potentielles des données.

Le centre Belgrand archivera et diffusera des méthodes et des résultats du travail d'ingénierie des données. Il mutualisera au besoin la production et la maintenance d'outils d'intérêt général ; par exemple, pour la manipulation de référentiels spatiaux évolutifs, le changement de maille ou l'imputation spatiale. Un accent particulier est mis sur la référence spatiale qui est la clé naturelle de croisement des données portant sur la ville. Le SIG *open source* est privilégié comme vecteur de diffusion des méthodes, à l'exemple de l'IRSTV à Nantes [3].

Les opérations de ressource scientifique seront guidées par la volonté de rapprocher les données statistiques, les données

géographiques et celles issues des modèles. Elles visent aussi à apporter une meilleure reconnaissance académique du travail sur les données. Il est en effet indispensable de mieux

valoriser ce travail qu'il ne l'est actuellement pour inciter les équipes à le capitaliser et à le rendre réutilisable.

Pourquoi Belgrand ?

Eugène Belgrand participa à la rénovation de Paris dirigée par le Baron Haussmann. On lui doit les égouts de Paris, l'aqueduc de la Vanne, celui de la Dhuis, ainsi que le réservoir de Montsouris. Le travail d'ingénierie des données partage avec les réseaux d'adduction d'eau et d'assainissement de nombreux points communs : caché, il part de la source, partage une ressource précieuse, l'épure, la recycle, irrigue et relie les individus au-delà des cloisonnements.

Un grand équipement pour l'utilisation des bases de données au service de la modélisation urbaine

Du fait de la complexité de l'objet urbain dans le prisme du développement durable, de la densité informationnelle de ce milieu et du développement des capacités de calcul, la modélisation urbaine est une consommatrice compulsive de données. Le problème de l'intégration de ces données requiert une accessibilité continue des informations, et l'interopérabilité spatiale et sémantique des données.

Classiquement, les données obtenues par observations sont des entrées pour les modèles ou encore servent à calibrer des paramètres. Elles peuvent également être mobilisées pour de la vérification *a posteriori* des résultats de modèles. Mais lorsque des modèles urbains font appel à de la simulation, leurs résultats ont un statut comparable aux produits d'ingénieries créés à partir des bases de données brutes : la valeur ajoutée par la création et transformation de données doit être rendue visible. Il est nécessaire de documenter et d'évaluer les processus qui ont créé les données simulées, et utile de les capitaliser lorsqu'elles sont réutilisables.

Le référencement des modèles conceptuels, des modèles de simulation ainsi que de leurs résultats est aussi complexe et utile que celui des données d'observations. L'infrastructure web Belgrand permettra de les mettre en lien avec les projets, équipes et terrains, offrant une vue synoptique des travaux de modélisation urbaine pour les chercheurs et les décideurs. Les enjeux sont d'autant plus importants dans un contexte de dématérialisation des traitements sous forme de services web ou de *living labs* urbains associant collectivités, industriels, chercheurs et habitants modifiant les rapports aux données et aux territoires.

Modèles et observations doivent donc bénéficier d'une dynamique commune, et non être opposés. Un équipement intégrant avec les mêmes outils : données brutes, modèles conceptuels, modèles de simulation, données intégrées et données simulées permettra une synergie plus importante entre les différentes approches scientifiques. ■ ■ ■

Bibliographie

- [1] Gaeremynck, J.
Le comité du secret statistique
Courrier des statistiques, n° 128, septembre-décembre 2009, pp.15-18
- [2] Cassier, M.
Le partage des connaissances dans les réseaux scientifiques : l'invention des règles de "bonne conduite" par les chercheurs
Revue française de sociologie, vol. 39 n°4, 1998, pp. 701-720
- [3] Rouzet, C., Labbe, S
SIG en «multi-partenariat»: mobiliser les acteurs et partager les données
Revue Internationale de Géomatique, vol. 7, N°3-4, 1997 pp. 279-295