



Potentiels et limites des traces (géo)numériques dans l'analyse des mobilités : l'exemple des données de la plateforme de covoiturage BlaBlaCar

Boris Mericskay

► To cite this version:

Boris Mericskay. Potentiels et limites des traces (géo)numériques dans l'analyse des mobilités : l'exemple des données de la plateforme de covoiturage BlaBlaCar. *Cybergeographie : Revue européenne de géographie / European journal of geography*, 2019, 10.4000/cybergeographie.31990 . halshs-02091115

HAL Id: halshs-02091115

<https://shs.hal.science/halshs-02091115>

Submitted on 5 Apr 2019

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.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

Potentiels et limites des traces (géo)numériques dans l'analyse des mobilités : l'exemple des données de la plateforme de covoiturage BlaBlaCar

Boris Mericskay



Édition électronique

URL : <http://journals.openedition.org/cybergeo/31990>

DOI : 10.4000/cybergeo.31990

ISSN : 1278-3366

Éditeur

UMR 8504 Géographie-cités

Ce document vous est offert par Université Rennes 2



Référence électronique

Boris Mericskay, « Potentiels et limites des traces (géo)numériques dans l'analyse des mobilités : l'exemple des données de la plateforme de covoiturage BlaBlaCar », *Cybergeo : European Journal of Geography* [En ligne], Cartographie, Imagerie, SIG, document 889, mis en ligne le 03 avril 2019, consulté le 05 avril 2019. URL : <http://journals.openedition.org/cybergeo/31990> ; DOI : 10.4000/cybergeo.31990

Ce document a été généré automatiquement le 5 avril 2019.



La revue *Cybergeo* est mise à disposition selon les termes de la Licence Creative Commons Attribution - Pas d'Utilisation Commerciale - Pas de Modification 3.0 non transposé.

Potentiels et limites des traces (géo)numériques dans l'analyse des mobilités : l'exemple des données de la plateforme de covoiturage BlaBlaCar

Boris Mericskay

Introduction

- 1 Avec l'avènement du numérique, les données massives alimentent de plus en plus de réflexions sur le fonctionnement des sociétés contemporaines et notamment sur les modalités de production de savoirs et de connaissances à leur propos (Bastin et Tubaro, 2018). La multitude de traces numériques¹ produites par les individus et les objets connectés offre de multiples perspectives de recherche en permettant notamment de saisir et de donner à voir des pratiques spatiales avec une granularité et des échelles spatio-temporelles difficilement accessibles avec d'autres méthodes (Beaude, 2015 ; Mericskay *et al.*, 2018). Qu'il s'agisse de données de billettique, de téléphonie mobile ou de réservation d'un trajet, toutes ces empreintes (géo)numériques viennent renouveler les approches de compréhension des mobilités individuelles comme collectives.
- 2 Au-delà de perspectives riches de promesses, la mobilisation de traces (géo)numériques, émanant d'acteurs publics et privés dans la production de connaissances, soulève de multiples interrogations, tant au niveau de la nature des données, des modalités de mise en circulation que des outils et méthodes de traitement associés. D'un côté, l'accessibilité à ces nouveaux gisements de données se caractérise par une série de contraintes limitant leurs utilisations (techniques, légales, éthiques, économiques, etc.). De l'autre, les méthodes et outils d'analyse sont encore en phase de consolidation et nécessitent

d'explorer et d'expérimenter les potentialités de ces données d'un type particulier (volumineuses, hétérogènes et souvent en temps réel).

- 3 Cet article a pour ambition d'explorer les potentialités des traces numériques dans la compréhension du covoiturage par le spectre des données de la plateforme BlaBlaCar. Par l'analyse temporelle et spatiale d'un corpus de 150 000 trajets, l'objectif est à la fois de brosser un portrait du covoiturage organisé au sein de la ville de Rennes et de faire un tour d'horizon des avantages et des limites de ces données dans la compréhension de cette nouvelle forme de mobilité.
- 4 La première partie de l'article s'attache à faire un retour sur les données massives comme objet et matériau au service de l'analyse des mobilités. Les principales formes de données disponibles relatives aux déplacements sont présentées (téléphonie mobile, billettique, vélos en libre-service, réseaux sociaux, etc.) ainsi que leurs modalités d'accessibilité (*Web scraping*, *API*). Les enjeux techniques et méthodologiques à l'analyse de ces données particulières sont également explicités.
- 5 La seconde partie présente une étude exploratoire des données de la plateforme BlaBlaCar portant sur les trajets au départ et à destination de Rennes pendant 5 mois. Elle revient sur les différentes étapes de la chaîne de traitement mise en place et présente les principaux résultats. La méthodologie de collecte et de constitution de la base de données est préalablement explicitée. Les grandes logiques de l'offre de trajets d'un point de vue temporel sont ensuite exposées selon différentes caractéristiques. Enfin, la dimension spatiale et territoriale des données est abordée à deux échelles d'analyse : l'une nationale, qui envisage les trajets selon leurs origines, leurs destinations et leurs distances ; l'autre locale, qui revient sur la configuration spatiale des lieux de covoiturage au sein de la ville de Rennes.

Les traces numériques comme marqueurs spatiaux de mobilités

Analyser les mobilités spatiales par les traces

- 6 Les chercheurs ont longtemps compté sur des enquêtes par comptages ou des enquêtes rétrospectives (comme les enquêtes ménages-déplacements) pour tenter de qualifier, quantifier et modéliser les mobilités spatiales. En complément, le recours à des dispositifs de collecte de données sur les pratiques spatio-temporelles individuelles (enregistreurs de données ou applications mobiles dédiées) a introduit de nouvelles modalités dans la compréhension des formes de déplacement (Depeau et Quesseveur, 2014 ; Gwiazdzinski et Klein, 2014). Enfin, plus récemment, l'arrivée de grandes masses de données produites par des opérateurs publics et privés à l'image des données de billettique, de téléphonie mobile ou d'activité sur les réseaux sociaux a ouvert un champ d'investigation important dans l'étude des déplacements (Chen *et al.*, 2016 ; Schwanen, 2017). Ces nouvelles sources de données, de plus en plus explorées par les chercheurs, participent aujourd'hui à un renouvellement des méthodes de collecte, de traitement et d'analyse des mobilités individuelles et collectives (Vidal, 2015).
- 7 Les données de téléphonie mobile, qui permettent d'étudier de gros volumes d'individus sur des temporalités longues, constituent des marqueurs intéressants à explorer dans une optique de compréhension des mobilités. Plusieurs travaux mettent bien en avant

l'intérêt de ces données pour appréhender et venir quantifier les modalités spatio-temporelles de déplacement individuel au sein d'une ville ou d'un quartier (Toole *et al.*, 2014 ; Senaratne *et al.*, 2018). D'un point de vue méthodologique, les études mobilisant ces données s'articulent pour beaucoup autour de l'identification et la formalisation de logiques récurrentes de déplacement (Calabres *et al.*, 2012 ; Jiang *et al.*, 2017), l'exploration de logiques origine-destination (Iqbal *et al.*, 2014 ; Alexander *et al.*, 2015) ou l'intégration de ces nouvelles sources au sein de processus de modélisation existants.

- 8 Les données relatives à la billettique des réseaux de transport en commun permettent quant à elles de mettre en lumière des dynamiques spatio-temporelles de déplacement (Barry *et al.*, 2002 ; Zhong *et al.*, 2015 ; Mohamed *et al.*, 2017), des évolutions de pratiques dans le temps (Briand *et al.*, 2017) ou encore des logiques d'intermodalité (Richer *et al.*, 2018). Dans le même esprit, les données concernant les systèmes de vélos en libre-service sont également étudiées par des chercheurs pour en extraire des informations sur les dynamiques de déplacement urbain. Les informations compilées, agrégées et traitées (horaires, occupation des stations, localisation des stations de départ et d'arrivée, durée du trajet, etc.) permettent de visualiser sous forme de cartes et de graphiques le comportement des usagers dans le temps et dans l'espace. Au-delà de la simple visualisation, de nombreuses études explorent davantage la dimension analytique à des fins d'optimisation de gestion (Vogel *et al.*, 2011 ; Wergin et Buehler, 2017). Les travaux d'Etienne Côme et Latifa Oukhellou (2014) autour de l'utilisation des Vélib's, proposent par exemple d'appréhender les pratiques de ce service sous des formes complexes basées sur une palette d'analyses (réurrences d'utilisation des stations, matrices origine-destination dynamiques, typologie des stations).
- 9 Les données issues des réseaux sociaux à l'image de Twitter ou de Foursquare sont également de plus en plus explorées pour comprendre et appréhender sous des formes renouvelées, des dynamiques de déplacement ou de fréquentation, notamment urbaines. Les travaux sur le sujet portent essentiellement sur l'analyse des mobilités de touristes ou d'immigrants à des échelles nationales et internationales (Hawelka *et al.*, 2014 ; Jurdak *et al.*, 2015 ; Rashidi *et al.*, 2017). Enfin, le rôle grandissant des géants du numérique qui collectent et produisent une multitude de données relatives aux mobilités individuelles ouvre de nombreuses perspectives. D'un côté, les plateformes spécialisées dans la mise en relation de conducteurs et de passagers comme Uber ou BlaBlaCar, constituent des bases de données volumineuses aux contenus diversifiés (pratiques des usagers, matrices origine-destination, conditions de circulation, etc.)². De l'autre, les applications mobiles d'aide à la navigation à l'image de Waze ou Google Maps, collectent, archivent et traitent en temps réel des flux d'informations volumineux sur la vitesse de leurs utilisateurs, les temps de trajet, les congestions en présence ou la localisation des accidents³.

Accessibilité et traitement des données massives : des enjeux économiques et techniques

- 10 Tenter de faire parler toutes ces traces numériques hétérogènes, c'est se confronter aux techniques multiples de transformation des données en informations et en savoirs, en cherchant à ouvrir les boîtes noires algorithmiques dont elles sont issues. A l'heure des données massives, la prolifération des données ne rime pas forcément avec mise à disposition. Même si le *big data* se développe au côté du mouvement *open data*, l'accessibilité aux données est de plus en plus contrôlée par les acteurs publics et privés

qui les produisent. Il existe trois grandes méthodes d'accessibilité à ces nouveaux gisements informationnels : le recours au *Web scraping*, la mise en place de partenariats avec les producteurs de données et la mobilisation des API (Bastin et Tubaro, 2018).

- 11 Premièrement, le *Web scraping*, qui permet de récupérer de manière automatisée des éléments de pages Web, est de plus en plus utilisé pour la constitution de jeux de données. Le principe est assez simple : transformer des informations non structurées présentes au sein de pages Web en données structurées plus facilement exploitables. Ensuite, la mise en place d'un partenariat avec une plateforme ou l'achat sont des options qui permettent d'accéder à certaines bases de données (billettique, réseaux sociaux, téléphonie mobile). Enfin, l'utilisation des API⁴, permettant de récupérer certains contenus sous des formes structurées, se généralise au sein des pratiques des professionnels et des chercheurs. Le recours grandissant à ces services en ligne s'explique en partie par le fait que de plus en plus d'entreprises et d'organisations publiques mettent à disposition certaines de leurs données via des API. Toutefois, beaucoup de données possiblement récupérables sont, dans les faits, de plus en plus verrouillées par les acteurs privés, désireux de contrôler, à des fins commerciales ces données éminemment stratégiques (Quesnot, 2016). La généralisation des API participe ainsi à la fois à l'ouverture des données par leur mise à disposition sous la forme de protocoles standardisés et, dans le même temps, à leur fermeture à travers la généralisation de restrictions d'usages (licences d'utilisation, monétisation des données, limitations des requêtes). Le recours grandissant aux API accentue également l'opacité des processus de collecte, d'agrégation et de mise en circulation des données. Une API n'est finalement qu'un automate qui exécute une requête chez un prestataire externe sans qu'on puisse connaître sa façon d'organiser et de traiter l'information (Masure, 2014).
- 12 Une fois les données collectées, récupérées ou achetées, leur traitement relève bien souvent de la gageure. D'une part, les contenus produits sont rarement structurés pour être croisés et analysés avec d'autres types de données plus classiques (statistiques publiques, référentiels administratifs, données métiers, etc.). Les données issues du *big data* sont bien souvent peu documentées et l'interopérabilité technique (formats d'échange) et sémantique (informations associées) demeure encore limitée avec les systèmes existants (Rashidi *et al.*, 2017 ; Yang *et al.*, 2017). D'autre part, il n'existe pas encore de véritables solutions logicielles dédiées aux traitements de ces données d'un type particulier (volumineuses, hétérogènes et en temps réel). Quelques outils permettent des opérations basiques, mais les fonctionnalités de transformation et d'analyse spatiale par exemple, sont assez limitées par rapport aux systèmes d'information géographique (SIG). Quant aux nouvelles approches mobilisant la fouille de données, l'apprentissage supervisé ou l'analyse prédictive, elles restent encore limitées et éloignées des pratiques opérationnelles⁵.
- 13 Même si la question de l'analyse des données est au centre du développement du *big data*, il est important de souligner que les usages autour de ces données restent exploratoires et s'inscrivent essentiellement dans des approches inductives, ou comme le soulignent justement Plantin et Russo (2016), la multiplication des bases de données disponibles marque clairement le retour à des conceptions empiristes et déterministes. En d'autres termes, la formulation des hypothèses et des théories est de plus en plus dérivée de l'exploration de nouvelles masses de données, avec comme leitmotiv « d'abord les données, ensuite la méthode ». C'est dans cette approche que s'inscrit la seconde partie de l'article, laquelle propose de jeter un regard compréhensif sur les pratiques de

covoiturage par le biais de l'analyse de données de la plateforme BlaBlaCar. L'objectif n'est pas de venir affirmer ou infirmer des hypothèses existantes sur le sujet mais bien, à travers l'exploration de traces (géo)numériques originales, de tenter de faire ressortir des grandes tendances, des logiques et des dynamiques relatives à ce nouveau mode de déplacement.

Appréhender l'offre de covoiturage à travers l'analyse des données de la plateforme BlaBlaCar

Le covoiturage, un mode de déplacement en développement encore peu étudié

- 14 Le covoiturage connaît depuis une dizaine d'années un développement important, pour des raisons à la fois économiques, politiques, idéologiques et technologiques. Popularisé dans les années 1980 par le milieu associatif, le covoiturage s'organise aujourd'hui essentiellement sur le Web avec le développement de plateformes publiques et privées de mise en relation (ADEME, 2015). Il est important de préciser qu'il n'existe pas un seul type de covoiturage tant les pratiques autour de ce mode de déplacement sont diversifiées (Jacquot, 2018). Il coexiste ainsi plusieurs formes de covoiturage (organisé, dynamique, pendulaire, occasionnel, courte distance, longue distance, etc.) renvoyant à différentes modalités de mise en relation, de répartition des frais, de fréquence ou encore de types de trajets (CEREMA, 2016 ; ADEME, 2017).
- 15 Malgré un intérêt grandissant pour ce nouveau mode de déplacement, le covoiturage reste encore mal qualifié car peu étudié. Hormis quelques enquêtes essentiellement qualitatives sur les profils et les motivations des covoitureurs, peu de travaux viennent véritablement qualifier et surtout quantifier les pratiques de cette forme de mobilité en développement (ADEME, 2015 ; Castex, 2015 ; CGDD, 2016). Le faible intérêt pour cette question peut, en partie, s'expliquer par la faible part modale du covoiturage, lequel pèse environ 2 % des déplacements longue distance (CGDD, 2016), ou par le manque de données disponibles sur le sujet. Toutefois, avec le développement des services en ligne spécialisés dans le covoiturage à l'image d'iDVROOM ou de BlaBlaCar, de nouvelles perspectives se dessinent en matière de compréhension des pratiques de covoiturage. Ces plateformes, pour qui les données sont au cœur des services, du marketing et de la relation client, ont compris l'importance de mettre à disposition certaines de leurs données pour la création de services innovants et la production de nouvelles connaissances.

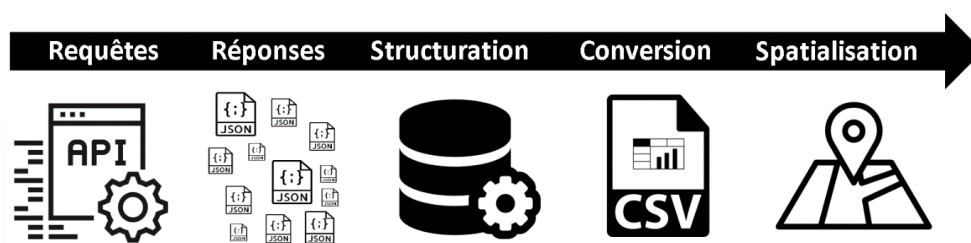
L'exemple des données de la plateforme BlaBlaCar

- 16 Lancée en 2006 en France, la plateforme de covoiturage BlaBlaCar s'est en quelques années largement imposée comme le principal service de mise en relation de conducteurs et de passagers pour des trajets en covoiturage organisé⁶. En France, BlaBlaCar détient plus de 95 % du marché du covoiturage organisé, laissant très peu de place à ses concurrents. En quelques chiffres, la plateforme rassemble plus de 40 millions de membres, lesquels proposent plus de 10 millions de trajets annuellement dans toute l'Europe, avec une offre majoritaire sur le territoire français. Même si la plateforme a lancé il y a plusieurs mois un service dédié aux déplacements quotidiens intitulé

BlaBlaLines, l'essentiel des trajets de covoiturage publiés sur BlaBlaCar concerne des trajets occasionnels de moyenne et longue distance.

- 17 En 2016, l'entreprise BlaBlaCar a mis en ligne une API⁷ permettant d'accéder gratuitement à certaines données relatives aux trajets publiés sur la plateforme. Il est important de préciser que cette API est une API de découverte des trajets, laquelle a vocation à être mobilisée par des sites Web proposant des comparateurs d'itinéraires intégrant le train, le bus et le covoiturage. Côté requête, il est ainsi possible d'interroger uniquement les trajets qui sont en vente sur la plateforme, les trajets effectués comme les trajets complets⁸ ne sont donc pas disponibles. Concernant les critères de recherche de trajets, il est possible de sélectionner les trajets publiés en fonction d'une ville de départ et/ou d'arrivée, la date, l'heure, le prix ou encore le nombre de places disponibles. Du côté des réponses renvoyées par l'API, autrement dit les attributs qui caractérisent les trajets recherchés, le service propose une trentaine de champs relatifs à la localisation des lieux de départ et d'arrivée (coordonnées géographiques, adresse, nom de la ville, pays, etc.), au temps (date et heure du départ, durée du trajet), au prix, aux voitures ou encore aux modalités de réservation.
- 18 Les données disponibles via l'API de BlaBlaCar sont proposées sous des formats structurés (JSON et XML) et rapidement réutilisables pour des analyses relatives à de courtes temporalités. En effet, il n'est pas possible de récupérer des corpus de trajets portant sur de longues temporalités (semaines, mois ou années) dans la mesure où l'API impose des limites et des quotas dans l'interrogation et l'extraction des données. En réponse à ce verrou, une série de requêtes quotidiennes ont permis de collecter les trajets publiés au départ et à destination de Rennes sur une période de cinq mois⁹. La chaîne de traitement concernant l'extraction et la structuration des données s'articule autour de quatre étapes principales (figure 1) :

Figure 1. Chaîne de traitement des étapes de préparation du jeu de données (source : auteur)



- Collecte des trajets au départ et à destination de Rennes par journée sous forme de jeux de données individuels.
 - Agrégation et structuration des jeux de données quotidiens sous forme de deux bases de données : trajets-origines et trajets-destinations.
 - Conversion du formalisme JSON en CSV pour un usage plus simple au sein des logiciels de traitement statistique, de visualisation de données et de SIG.
 - Spatialisation des données à travers le géoréférencement des coordonnées géographiques des trajets (lieux de départ et lieux d'arrivée) sous un logiciel SIG.
- 19 Le jeu de données constitué pour l'étude recense ainsi 151 125 trajets, dont 75 390 trajets au départ et 75 735 trajets à destination de Rennes. Concernant les variables mobilisées, le choix s'est porté sur certains attributs qui paraissent les plus pertinents dans une optique d'analyse des trajets selon des entrées temporelles, spatiales et territoriales (figure 2).

Figure 2. Récapitulatif des attributs récupérés pour chaque trajet étudié (source : auteur)

Date départ	Ville Départ	Adresse départ	Lat départ	Long départ	Ville arrivée	Adresse arrivée	Lat arrivée	Long arrivée	Dist trajet
01/08/2018 00:00	Champrond-en-C	28240 Champrond-en-C	48.401476	1.073473	Rennes	Rennes, France	48.117266	-1.677793	236
01/08/2018 02:00	Châteaubriant	Choisel	47.732876	-1.386532	Mordelles	22 Rue de la Libé	48.075248	-1.847642	70
01/08/2018 02:00	Bordeaux	Gare de Bordeaux St-Jez	44.826431	-0.556205	Rennes	Gare de Rennes,	48.103497	-1.672278	453
01/08/2018 03:00	Vienne	Place Saint-Louis, 3820	45.528368	4.87457	Saint-Grégoire	L'Auge de Pierre,	48.135233	-1.69371	795
01/08/2018 03:50	Saumur	49400 Saumur, France	47.260135	-0.080893	Rennes	Rennes, France	48.117266	-1.677793	190
01/08/2018 04:30	Irodouër	35850 Irodouër, France	48.249738	-1.949785	Rennes	Villejean-Univers	48.121265	-1.704041	30
01/08/2018 05:00	Lannion	Gare de Lannion, Avenu	48.727432	-3.460248	Saint-Grégoire	KFC, Rue Chesnay	48.136407	-1.69168	162
01/08/2018 05:00	Chartres	Rue Blaise Pascal, Chart	48.451754	1.539477	Rennes	Rennes	48.08347	-1.682726	265
01/08/2018 05:30	Lanvallay	13 Rue de Rennes, 2210	48.455714	-2.027952	Rennes	2 Rue d'Alsace, 3	48.122176	-1.703885	49
01/08/2018 05:30	Rouen	Rue Emile Duployé, Rou	49.453574	1.096445	Rennes	Station Service T	48.107168	-1.725825	314
01/08/2018 05:50	St-Malo	Saint-Malo, France	48.649338	-2.025674	Rennes	50 Ctre Commerç	48.121273	-1.711575	69

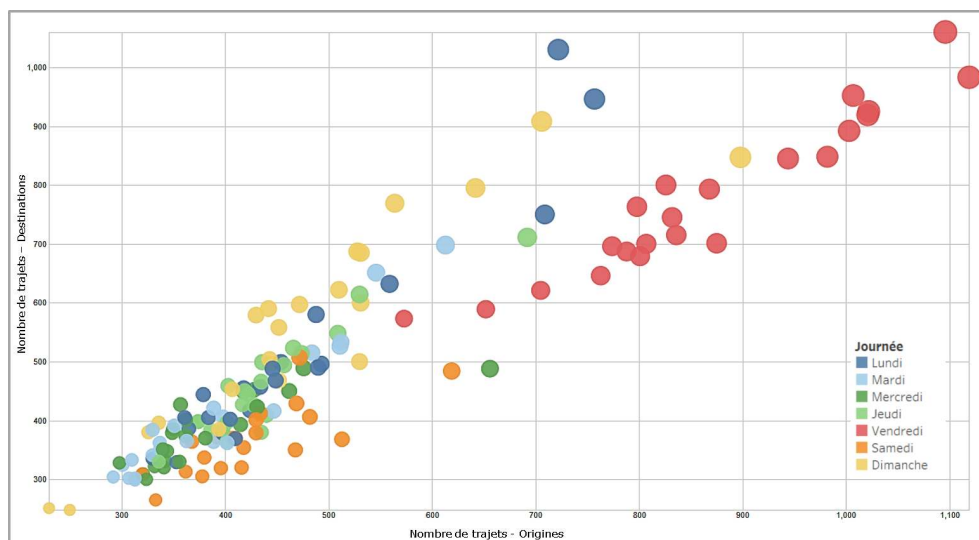
Analyse temporelle des trajets au départ et à destination de Rennes

- 20 En analysant la fréquence quotidienne de l'offre de trajets au départ et à destination de Rennes sur les cinq mois de l'étude, il ressort une logique cyclique fortement en lien avec les déplacements relatifs aux fins de semaines (figure 3). Les vendredis apparaissent clairement comme les journées se caractérisant par le plus grand nombre de trajets, qu'ils soient au départ ou à destination de Rennes. Au-delà des pics des week-ends, certaines périodes touristiques font également l'objet d'une offre importante de trajets comme les ponts du mois de mai, la semaine du 15 août ou la dernière semaine des vacances d'été. De manière plus générale, le nombre de trajets au départ apparaît comme fortement corrélé à la fois aux nombres de trajets d'arrivées et surtout au jour de la semaine (figure 4). Cette forme de covoiturage aux temporalités spécifiques est essentiellement pratiquée par des jeunes actifs, des étudiants, des personnes non-motorisées, des salariés rentrant chez eux le week-end ou encore des vacanciers (ADEME, 2015).

Figure 3. Nombre de trajets quotidiens sur les cinq mois de l'analyse (source : auteur d'après les données de BlaBlaCar)

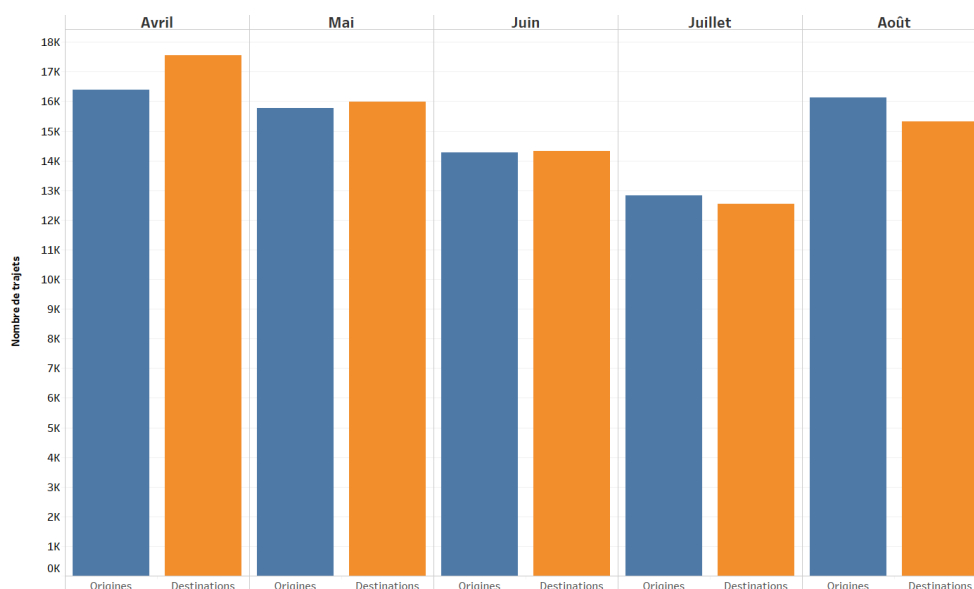


Figure 4. Distribution des trajets selon le nombre de départs, le nombre d'arrivées et le jour de la semaine (source : auteur d'après les données de BlaBlaCar)



- 21 Les fréquences mensuelles de trajets, qui oscillent entre 25 000 trajets (en juillet) et 34 000 trajets (en avril), mettent en évidence un certain usage saisonnier de la plateforme. L'offre de trajets diminue avec l'arrivée de l'été, et notamment au mois de juillet, pour reprendre au mois d'août avec la fin des grandes vacances (figure 5).

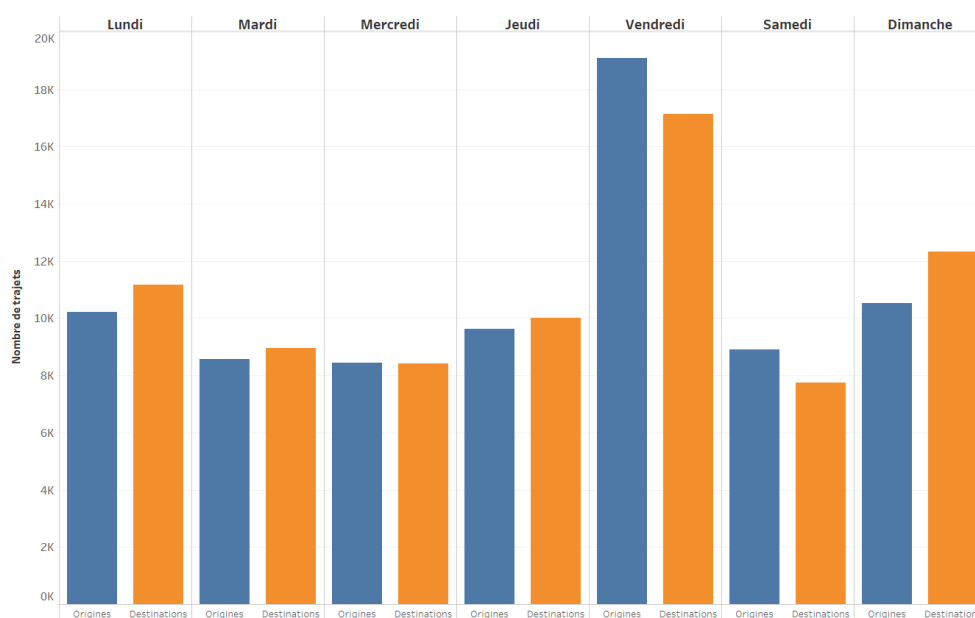
Figure 5. Nombre de trajets au départ et à l'arrivée de Rennes par jour de la semaine (source : auteur d'après les données de BlaBlaCar)



- 22 Concernant le rythme hebdomadaire de l'offre de covoiturage, l'agrégation des trajets par jour de la semaine fait ressortir un usage cyclique de la plateforme sur une semaine type (figure 6). Comme évoqué précédemment, l'offre de trajets se concentre essentiellement en fin de semaine avec un pic principal le vendredi (24 % des trajets) et deux autres pics moins importants, le dimanche (15 %) et le lundi (14 %). Les autres jours de la semaine, l'offre de covoiturage est plus modeste mais il existe un réel usage de la plateforme pour des déplacements en milieu de semaine (35 % des trajets pour les mardis, mercredis et

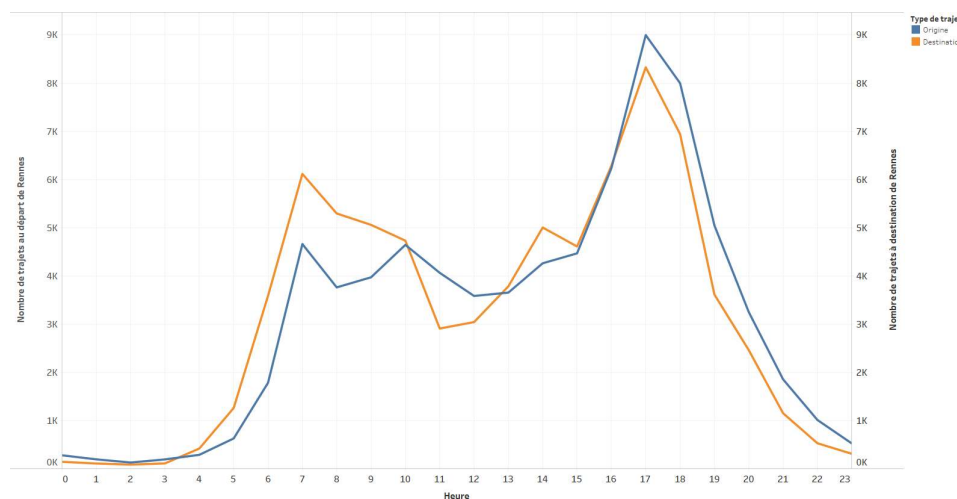
jeudis). Concernant la dynamique origine-destination, il est intéressant d'observer une certaine logique d'émission et de réception de Rennes le week-end. On constate ainsi une plus grande offre de trajets au départ de Rennes les vendredis, et davantage de trajets de retour le dimanche et le lundi. Concernant la relation entre jour de la semaine et distance des trajets, les trajets des fins de semaine ont des distances plus élevées en moyenne (220 km le vendredi contre 177 km le mardi).

Figure 6. Nombre de trajets au départ et à l'arrivée de Rennes par jour de la semaine (source : auteur d'après les données de BlaBlaCar)



- 23 Au niveau des horaires, il existe une certaine différence dans les rythmes quotidiens entre les trajets au départ et ceux à destination de Rennes (figure 7). D'un côté, la fréquence des heures de départ des trajets partant de Rennes suit une logique d'offre qui débute tôt le matin et se stabilise en milieu d'après-midi pour ensuite connaître un pic de 16h à 19h qui regroupe plus de 40 % des trajets. De l'autre, le rythme quotidien des heures de départ des trajets à destination de Rennes se caractérise par deux pics journaliers, un le matin de 6h à 10h (32 % des trajets) et un autre en fin d'après-midi de 16h à 20h (28 % des trajets). Concernant la relation entre horaire de départ et distance, il ressort que les trajets les plus longs partent généralement tôt le matin ou en fin de soirée.

Figure 7. Nombre de trajets au départ et à l'arrivée de Rennes selon l'heure de la journée (source : auteur d'après les données de BlaBlaCar)



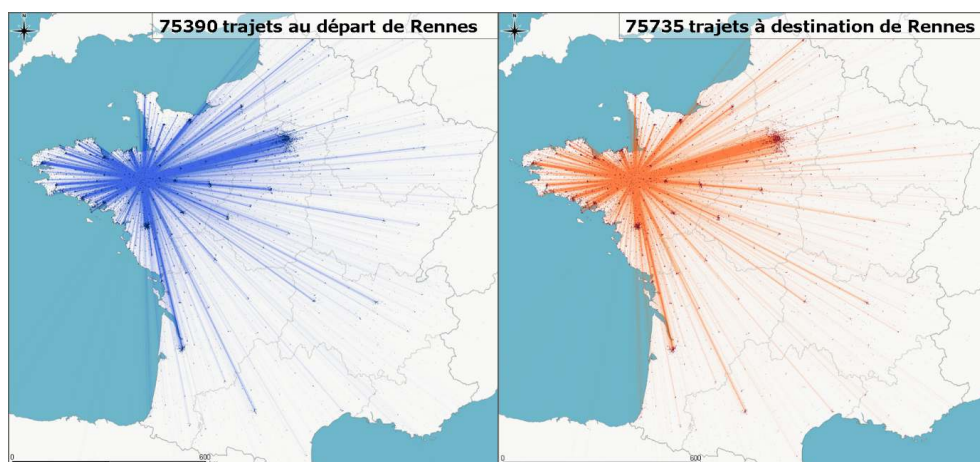
Analyse spatiale des trajets au départ et à destination de Rennes

- 24 La structuration des données de BlaBlaCar, laquelle se base sur le principe d'une matrice origine-destination, permet de venir quantifier les trajets entre des secteurs d'analyse. Le renseignement des coordonnées géographiques des lieux de départ et de destination de chaque trajet constitue une piste intéressante à explorer dans une recherche de mise en visibilité de possibles dynamiques territoriales de covoiturage. Afin de montrer les potentialités heuristiques de ces traces (géo)numériques, la dimension spatiale des trajets est envisagée à deux échelles d'analyse, l'une nationale et l'autre locale.

Analyse des trajets au départ et destination de Rennes à l'échelle nationale

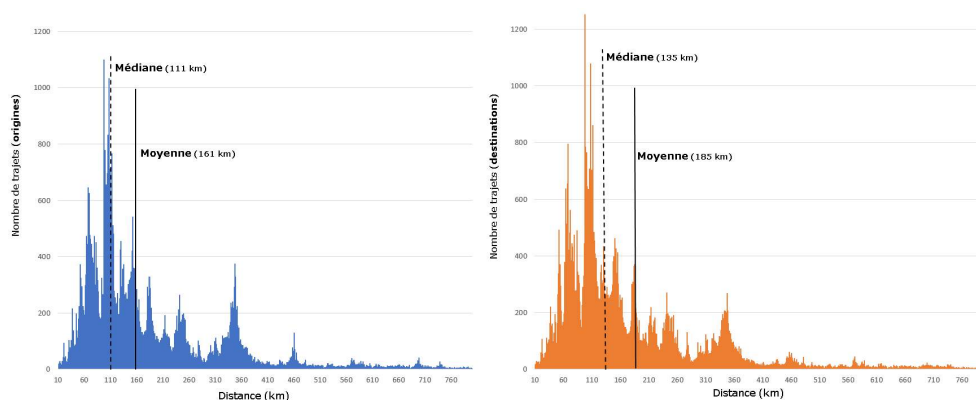
- 25 Pour comprendre les origines et les destinations des trajets publiés, une première analyse à l'échelle nationale permet de mettre en lumière les dynamiques spatiales du covoiturage rennais. Après spatialisation des lieux de départ et d'arrivée (par géoréférencement des coordonnées géographiques associées), la création de cartes de flux (ici des cartes en oursin¹⁰) permet de visualiser de manière générique les lieux d'attraction et d'appréhender l'aire d'influence de Rennes en matière de covoiturage occasionnel (figure 8).

Figure 8. Trajets au départ et à destination de Rennes (source : auteur d'après les données de BlaBlaCar et OpenStreetMap)



- 26 À la lecture de ces cartes, il ressort que l'offre de covoiturage au départ et à destination de Rennes couvre une grande partie du territoire national mais selon des concentrations plus ou moins fortes. On constate que la majorité des trajets partent et arrivent au sein de la même zone d'influence, à savoir en Bretagne, dans les Pays de la Loire, en Normandie et en Ile-de-France. On peut ainsi avancer l'idée que la proximité des villes joue un rôle prépondérant dans l'offre de trajets.
- 27 Cette polarité régionale est largement mise en évidence par l'analyse des distances des trajets (figure 9). La distance moyenne des trajets au départ de Rennes est de 161 km et la médiane de 111 km. Concernant les distances des trajets à destination de Rennes, les valeurs sont plus élevées avec une moyenne de 185 km et une médiane de 135 km. Au niveau de la part des distances des trajets, quelques indicateurs permettent de mettre en avant que les usages de la plateforme se destinent essentiellement à des déplacements de moyenne distance. Par exemple la proportion des trajets entre 10 et 100 km est de 28 %, celle des trajets entre 100 et 300 km monte à 53 %, quant aux trajets de plus de 500 km, leur part est de seulement 5 %.

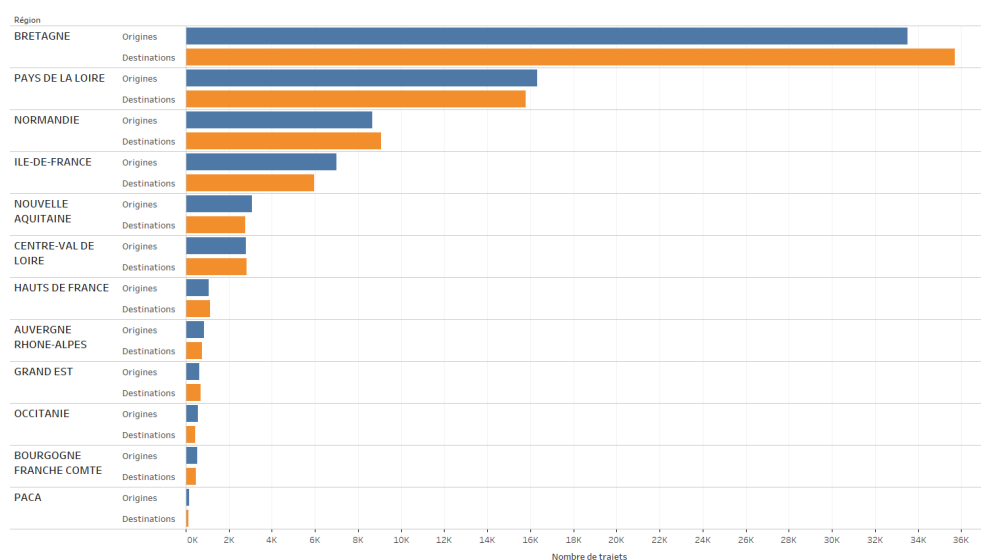
Figure 9. Distances des trajets entre 10 km et 800 km (source : auteur d'après les données de BlaBlaCar)



- 28 Dans une optique de qualification et de quantification des relations interurbaines, l'agrégation spatiale des lieux de départ et d'arrivée à l'échelle des régions fait ressortir des logiques territoriales bien marquées (figure 10). La région Bretagne concentre par

exemple à elle seule plus de 46 % des villes de départ et 47 % des villes d'arrivée des trajets, démontrant ainsi une très forte dynamique régionale dans les pratiques de covoiturage. En élargissant aux régions connexes à la Bretagne, 80 % des trajets au départ et à destination de Rennes ont lieu entre la Bretagne, les Pays de la Loire et la Normandie. On retrouve ensuite les trajets vers l'Ile-de-France et, fait intéressant, plus les autres régions sont éloignées et plus l'offre de covoiturage est faible.

Figure 10. Répartition des trajets au départ et à l'arrivée de Rennes selon la région d'origine et de destination (source : auteur d'après les données de BlaBlaCar)



- 29 En complément, le dénombrement des trajets au départ et à destination de Rennes à l'échelle des EPCI permet de mettre en avant des dynamiques territoriales plus fines¹¹. Il apparaît que l'offre de covoiturage se concentre essentiellement aux principaux pôles urbains régionaux autour de Rennes (figure 11). Seule exception, l'Ile-de-France et notamment Paris qui polarise un nombre important de trajets en raison de son attractivité. Fait intéressant, il ressort clairement un équilibre dans les dynamiques d'origine et de destination des trajets.
- 30 En se concentrant sur les quinze principaux EPCI en connexion avec Rennes (lesquels concentrent 50 % des trajets publiés), on retrouve des dynamiques très localisées (figure 12). Avec plus de 8 % de tous les trajets transitant par Rennes, la métropole nantaise est la zone qui concentre le plus d'offre de trajets. On trouve ensuite le Grand Paris suivi de Saint-Brieuc, Vannes, Saint-Malo et Lorient. A côté de cette offre vers les grandes villes de l'Ouest et la région parisienne, il existe une offre de trajets à destination des grands pôles urbains comme Bordeaux, Lyon ou Lille, mais très limitée.

Figure 11. Nombre de trajets par EPCI avec comme origine Rennes – nombre de trajets par EPCI avec comme destination Rennes (source : auteur d'après les données de BlaBlaCar et OpenStreetMap)

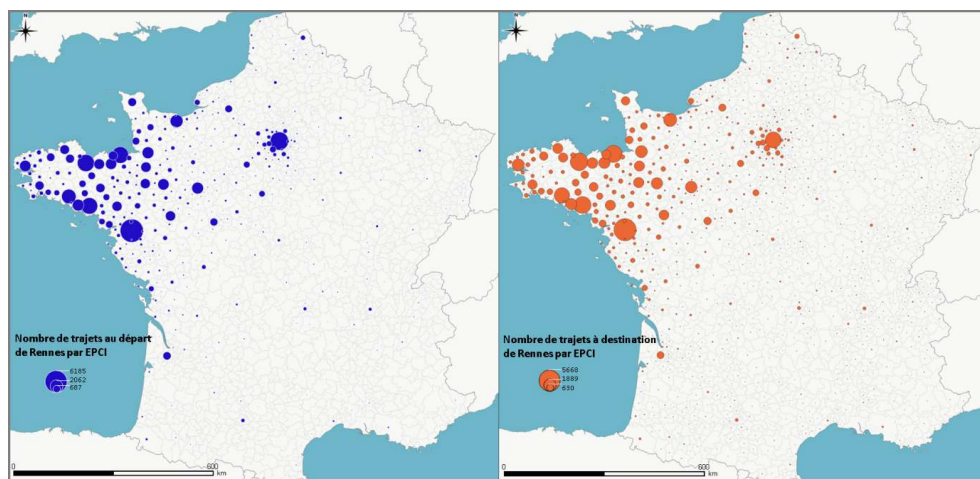
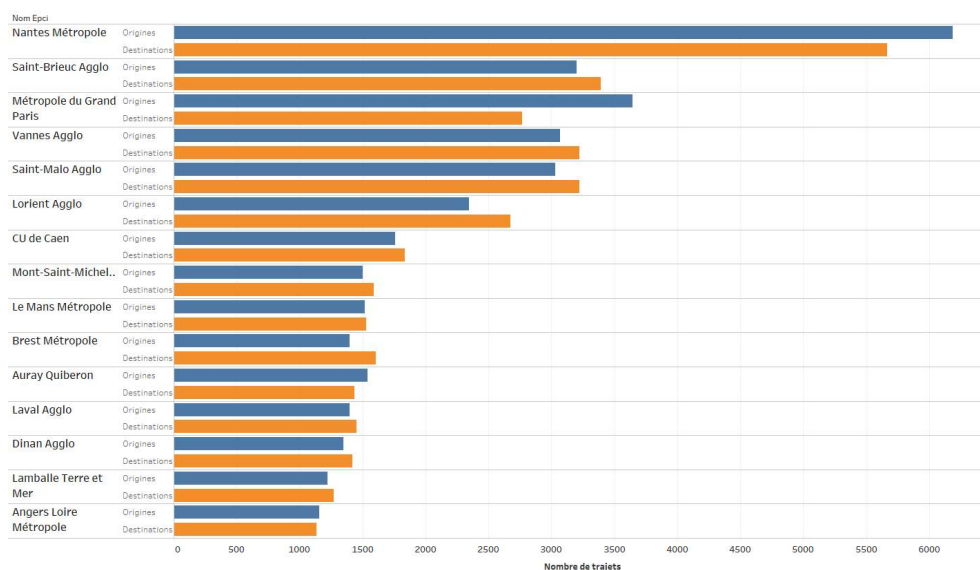


Figure 12. Nombre de trajets au départ et à destination des quinze principaux EPCI « connectés » à Rennes (source : auteur d'après les données de BlaBlaCar)



- 31 Ces quelques résultats illustrent bien que la plupart des trajets proposés sur la plateforme BlaBlaCar viennent plutôt renforcer les dynamiques existantes autour des zones urbaines et n'offrent dans le même temps qu'un accès limité aux espaces ruraux. Ainsi, une personne qui souhaite covoiturer *via* la plateforme aura plus de chance de trouver une opportunité de déplacement si elle veut se rendre vers une grande ville ou une ville moyenne que si elle souhaite aller vers des destinations situées dans les zones peu denses (Castex, 2015). A titre d'exemple, plus de 90 % des trajets arrivent ou partent d'une commune d'une aire urbaine (zonage INSEE). Autre chiffre intéressant, 57 % des trajets partent ou arrivent à moins de 5 km d'une commune de plus de 20000 habitants et ce chiffre monte à 72 % pour les villes de plus de 10 000 habitants.
- 32 Cette polarisation de l'offre de covoiturage aux villes moyennes et aux grandes villes renvoie également pour beaucoup à la proximité spatiale entre ces villes. Selon les

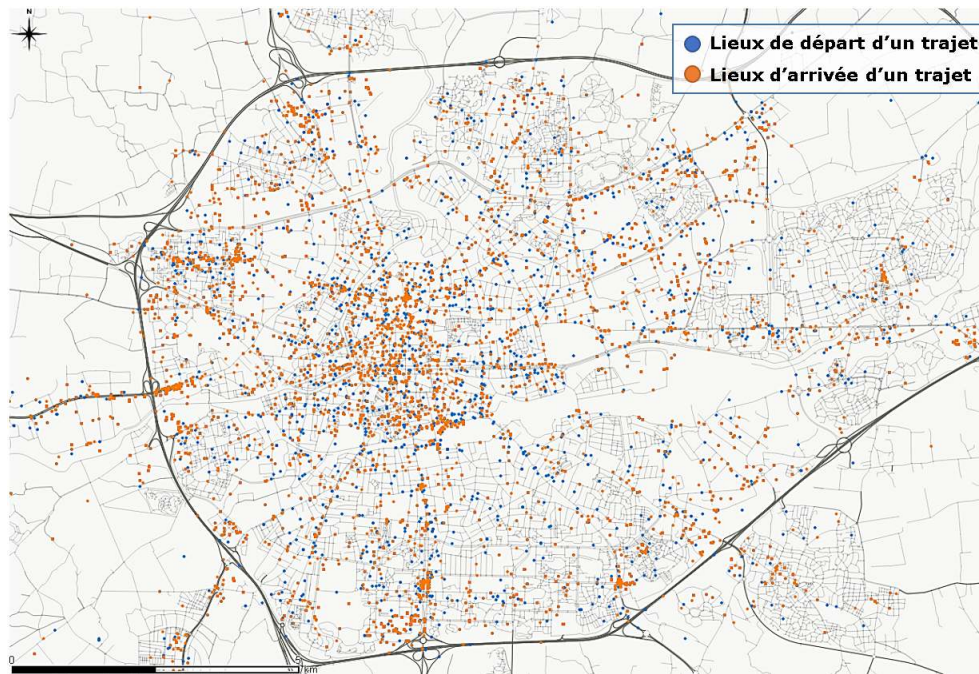
principes d'un modèle gravitaire, les échanges entre deux villes seront d'autant plus importants à la fois si le poids des villes est grand et si la distance qui les sépare est faible. Les exemples de Nantes, Saint-Brieuc ou Vannes illustrent très bien cette logique. A la vue de l'offre de trajets (fréquences, destinations et distances), le covoiturage occasionnel organisé par l'intermédiaire de la plateforme BlaBlaCar, au départ comme à destination de Rennes, se destine ainsi essentiellement à des déplacements au sein d'une aire d'influence régionale avec une exception, Paris.

- 33 Cette tendance à une offre de covoiturage « régionale » est à mettre en perspective avec la question de la complémentarité avec les autres modes de transport, notamment le train. Pour un passager au départ de Rennes et désirant se rendre à Nantes par exemple, il existe certes une offre en train mais celle-ci apparaît peu compétitive en termes de temps, de coût et de fréquence par rapport aux trajets proposés sur BlaBlaCar¹². Cette logique est similaire dans le cas de Saint-Brieuc, Saint-Malo, Vannes ou Lorient où le covoiturage organisé se positionne pour certaines catégories (étudiants, personnes non-motorisées) comme une alternative aux transports collectifs (train, bus).

Analyse des lieux de covoiturage à l'échelle de Rennes

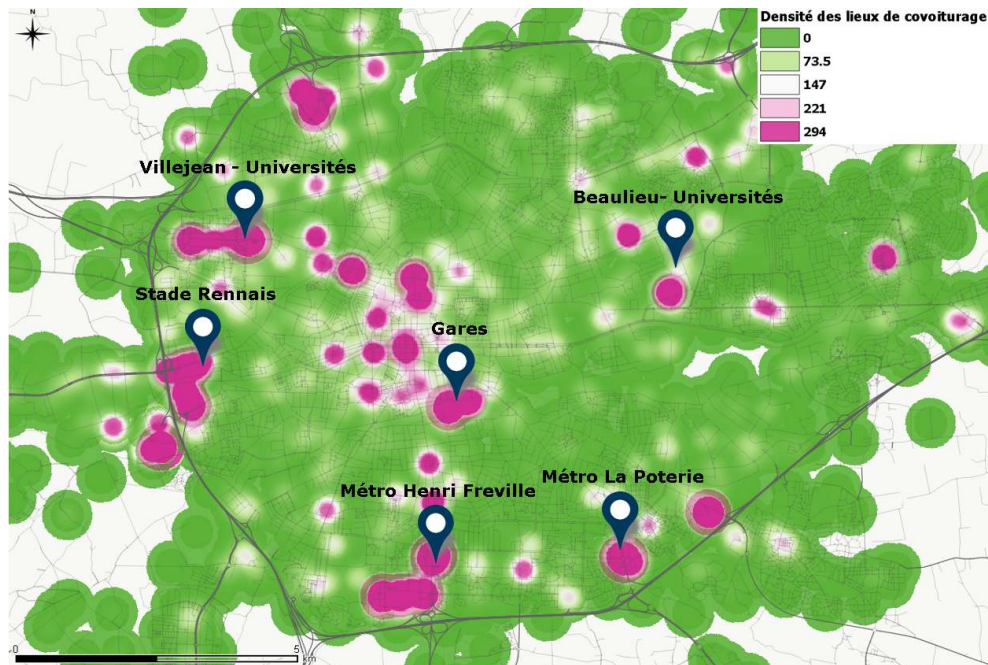
- 34 La nature des données de BlaBlaCar, à travers le renseignement systématique des coordonnées géographiques des lieux de départ et d'arrivée, autorise également d'envisager les pratiques de covoiturage à un niveau plus fin. L'analyse des lieux de covoiturage à l'échelle locale est une piste pertinente à explorer dans une optique de compréhension de pratiques de covoiturage au sein de la ville de Rennes. Toutefois, la simple spatialisation des lieux de covoiturage au sein d'un référentiel cartographique ne suffit pas à comprendre la configuration spatiale de ces lieux stratégiques comme l'illustre la figure 13. En effet, à l'échelle de la ville, la précision spatiale des données est assez hétérogène. Plusieurs trajets peuvent posséder les mêmes coordonnées géographiques d'origine ou de destination dans la mesure où leur renseignement par les conducteurs s'opère de différentes manières (saisie d'une adresse, partage de la géolocalisation, choix dans une liste de lieux prédéfinis). Visuellement, une grande partie des lieux de covoiturage sont ainsi superposés, limitant par là même l'analyse de leur distribution spatiale.

Figure 13. Localisation des lieux de covoiturage à Rennes sur cinq mois (source : auteur d'après les données de BlaBlaCar et OpenStreetMap)



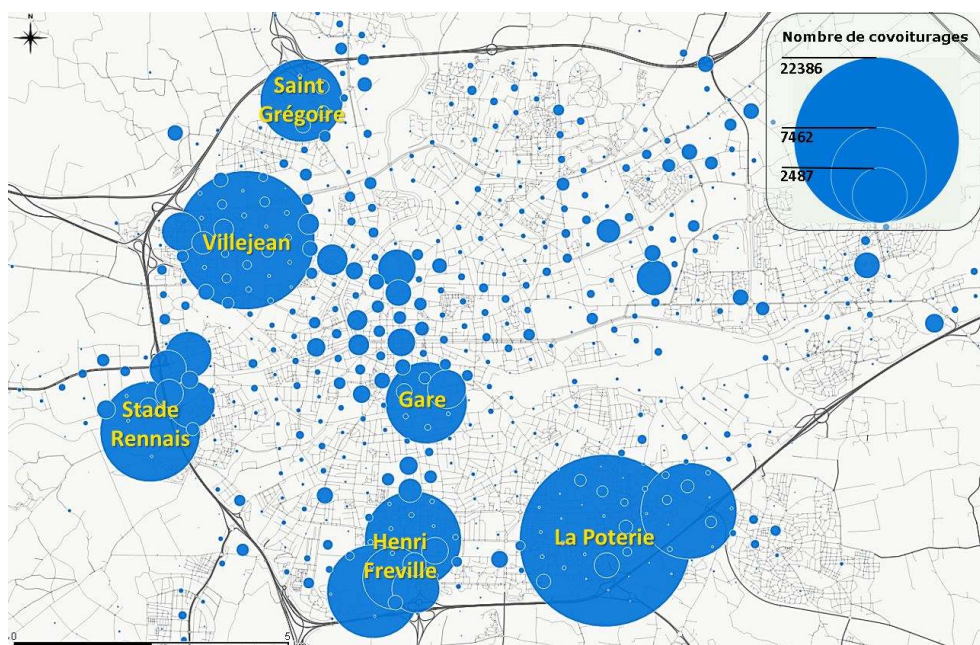
- 35 En réponse à ce biais, deux modes de représentation basés sur la transformation des données ponctuelles sont ici présentés. Premièrement, le passage à une modélisation matricielle à travers une carte de chaleur¹³ permet de traiter et surtout de visualiser la concentration des lieux de covoiturage sous une forme simplifiée. Cette forme de représentation offre à ces données multiples un aspect visuel plus facile à saisir et permet de faire ressortir les « zones chaudes » du covoiturage à Rennes (figure 14). Les principales zones de covoiturage mises en avant par la carte de chaleur se concentrent autour de lieux d'intermodalité (stations de métro, gare) et d'enseignement supérieur (universités).

Figure 14. Concentration des lieux de covoiturage à Rennes sur cinq mois (source : auteur d'après les données de BlaBlaCar et OpenStreetMap)



- 36 Autre méthode, l'agrégation spatiale des lieux de départ et de destination au sein d'un carroyage vectoriel qui permet une lecture plus fine de la question par la discrétisation spatiale de la zone d'étude. Le dénombrement des lieux de covoiturage au sein d'un maillage régulier (ici des hexagones de 300 m¹⁴) permet de mettre en lumière les logiques de distribution spatiale des lieux de covoiturage. En complément de l'identification des zones de concentration, cette approche permet surtout une quantification du nombre de trajets transitant par ces lieux sur la période de l'étude (figure 15).

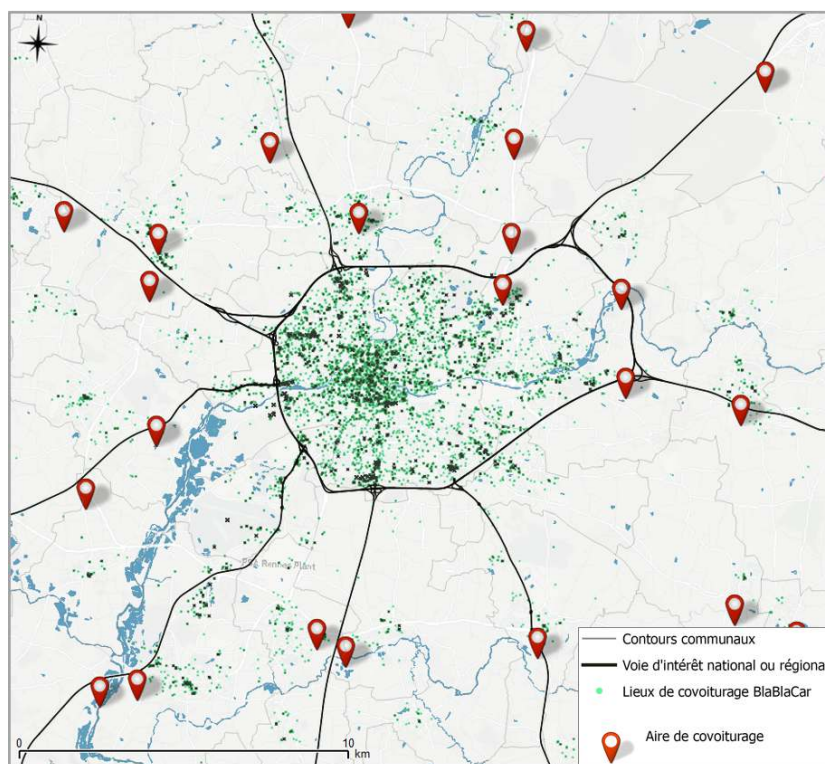
Figure 15. Nombre de covoiturages ayant transité au sein de mailles hexagonales de 300 m entre le 1^{er} avril et le 31 août 2018 (source : auteur d'après les données de BlaBlaCar et OpenStreetMap)



- 37 Deux principales zones de concentration des lieux de covoiturage ressortent de cette analyse : la Poterie au sud-est de Rennes (terminus de la ligne de métro) qui rassemble plus de 20 % des lieux et Villejean au nord-est (campus universitaire) qui regroupe plus de 15 % des lieux. On trouve ensuite la zone commerciale d'Henri Fréville, le Stade Rennais, la gare et l'entrée nord de Rennes. Les zones de rencontres privilégiées par les utilisateurs de BlaBlaCar renvoient à certaines caractéristiques communes, à savoir la présence d'infrastructures de transport en commun (stations de métro, gare, arrêts de bus), la présence de lieux générateurs de déplacements (université, centre commercial, etc.) et la proximité à la rocade¹⁵. Le critère de proximité à une infrastructure de transport en commun apparaît particulièrement important dans la mesure où une grande partie des passagers se rendent ou repartent des lieux de covoiturage en utilisant les transports collectifs. Par exemple, une étude de l'ADEME de 2015 sur le covoiturage longue distance met en avant que 44 % des passagers rejoignent leur covoiturage en utilisant les transports en commun.
- 38 Il est important de préciser que le choix de ces lieux par les conducteurs qui publient un trajet ne se fait pas de manière anodine. Même s'ils peuvent spécifier n'importe quel endroit de la commune pour récupérer ou déposer les passagers, la sélection s'opère dans la majorité des cas via une liste prédéfinie de lieux considérés comme « propices » par la plateforme. On retrouve ainsi dans cette liste, les lieux mis en avant dans les précédentes analyses, à savoir la gare, les universités et les stations de métro. Il est intéressant d'observer que la liste ne suggère pas les aires de covoiturage existantes aux conducteurs publiant un trajet. Ces équipements mis en place par les pouvoirs publics ne sont dans les faits que très peu utilisés par les covoitureurs de la plateforme (figure 16). Situés en périphérie des villes et mal desservis par les transports en commun, ces aménagements, qui se destinent essentiellement au covoiturage pendulaire, ne sont pas adaptés au covoiturage occasionnel. Ils permettent essentiellement aux travailleurs de stationner

leur voiture pour la journée mais ne sont pas conçus pour favoriser la rupture de charge entre conducteurs et passagers bien souvent non véhiculés.

Figure 16. Localisation des lieux de covoiturage BlaBlaCar et des aires de covoiturage (source : auteur d'après les données de BlaBlaCar et OpenStreetMap)



- 39 Dans la pratique, les membres de la plateforme BlaBlaCar se rencontrent majoritairement dans la ville de Rennes au sein d'aires informelles (stationnements public et privés, dépose-minute, voies de bus, etc.) où il n'existe pas forcément d'aménagements dédiés pour le covoiturage. Cette situation peut engendrer des problèmes de stationnement ou de congestion au sein de zones déjà très fréquentées par les piétons, les voitures ou les bus¹⁶. Plusieurs lieux sont ainsi quotidiennement le théâtre de conflits d'usages à Rennes ; comme aux pôles multimodaux de la Poterie ou de l'Université Villejean où la cohabitation entre bus et covoitureurs est compliquée à certaines heures de la journée (stationnement au sein de voies de bus, blocage de voies d'accès, occupation des dépose-minute, etc.).
- 40 Face à cette situation, Rennes Métropole a engagé depuis quelques mois des réflexions et des actions en faveur du covoiturage occasionnel. La ville a par exemple récemment aménagé deux aires de covoiturage pour le stationnement de courte durée à proximité de stations de métro. Comme autre initiative, on peut citer l'installation d'un capteur au pôle multimodal de la Poterie pour mesurer l'utilisation de l'aire de covoiturage. La prise en compte du covoiturage occasionnel par les pouvoirs publics est encore très récente et les premiers aménagements spécifiques commencent seulement à émerger. Il est fort probable qu'avec le développement de cette forme de mobilité, d'autres actions prendront forme au sein des villes pour à la fois favoriser la rencontre entre passagers et conducteurs et minimiser les problèmes de circulation et de stationnement relatifs au covoiturage à proximité de lieux d'intermodalité stratégiques.

Perspectives

- 41 Cette analyse exploratoire des données de la plateforme de BlaBlaCar a permis de jeter un regard compréhensif sur l'offre de covoiturage organisé et occasionnel tout en mettant en évidence une série de limites inhérentes à la nature et à la structure de ces données produites et mises à disposition par un acteur privé. Même si les données étudiées ne sont pas exhaustives (pas de prise en compte des trajets complets), pas forcément très bien renseignées (notamment en ce qui concerne les taux de remplissage des trajets) et, finalement assez peu accessibles (limites de l'API), elles ouvrent néanmoins de nombreuses perspectives de recherche dans la compréhension du covoiturage.
- 42 L'étude d'un corpus de plus de 150 000 trajets sur une période de cinq mois démontre bien les potentialités de ces traces à mettre en exergue de manière quantitative les dynamiques de covoiturage d'une métropole française. Les différentes modalités temporelles et spatiales des trajets mises en lumière dans l'étude participent à venir qualifier et quantifier le covoiturage occasionnel organisé. De plus, les données fournissent la possibilité de changer de niveau de zoom pour des analyses fines de phénomènes plus localisés, à l'image de l'organisation spatiale des lieux de covoiturage à l'échelle de la ville de Rennes.
- 43 Certes le tour d'horizon du covoiturage proposé dans l'article ne reflète que les pratiques des membres de BlaBlaCar, cependant la mobilisation de ces données non institutionnelles permet d'élargir les cadres traditionnels de l'analyse du covoiturage grâce à la précision et la quantité de données disponibles. Des analyses similaires au sein d'autres métropoles permettraient de voir s'il existe des logiques semblables et pourquoi pas des spécialisations locales. En complément, la mise en place d'études prenant en considération un ensemble de villes à des échelles régionales voire nationale participerait à mettre en évidence de potentielles logiques de connexions territoriales relatives au covoiturage comme l'illustre la figure 16.

Figure 17. Trajets au départ et à l'arrivée de Nantes, Paris, Bordeaux, Toulouse, Montpellier, Grenoble et Lyon le 24 août 2018 (source : auteur d'après les données de BlaBlaCar et OpenStreetMap)



- 44 Comme nous l'avons démontré tout au long de l'article, la mobilisation de ces nouvelles données d'un type particulier nécessite à la fois un travail de préparation en amont et surtout la mise en place de méthodes spécifiques pour les rendre intelligibles et leur apporter une réelle valeur ajoutée analytique. La multiplication des nouvelles sources de données issues de traces (géo)numériques s'accompagnent de nouveaux questionnements relatifs à leur utilisabilité dans des contextes de recherche. Les questions de fiabilité, de représentativité, d'exhaustivité, d'accessibilité ou encore de pérennité des données sont au cœur de nouvelles interrogations pour les chercheurs. Le passage des données brutes en connaissances requiert plus que jamais le recours à des méthodes et des techniques adaptées et toujours plus sophistiquées, lesquelles doivent être explorées, renouvelées mais surtout constamment questionnées.
- 45 Enfin, malgré un réel potentiel analytique, ces données ne doivent pas être envisagées comme des substituts aux méthodes plus classiques, à l'image des enquêtes, mais davantage comme des sources complémentaires. Il ne faut pas perdre de vue que les données massives, issues d'acteurs publics et privés, sont produites de manières automatisées et existent avant tout pour détecter et formaliser des comportements et des catégories et non mettre en lumière des spécificités. La mobilité comme fait social ne peut uniquement être appréhendée par le traitement de gros volumes de données standardisées, lesquelles ne permettent pas de venir enrichir les analyses d'informations plus qualitatives à l'image des motifs de déplacement ou des caractéristiques des individus. De nouvelles perspectives méthodologiques plus hybrides, combinant approches et données quantitatives et qualitatives, sont ainsi à explorer et à construire pour comprendre et analyser autrement les mobilités d'aujourd'hui et de demain.

BIBLIOGRAPHIE

- ADEME, 2015, "Enquête auprès des utilisateurs du covoiturage longue distance", Étude réalisée pour le compte de l'ADEME par 6t-bureau de recherche.
- ADEME, 2017, "Développement du covoiturage régulier de courte et moyenne distance : Guide méthodologique et Synthèse", Collection Clés pour agir.
- Alexander L., Jiang S., Murga M., & González M. C., 2015, "Origin-destination trips by purpose and time of day inferred from mobile phone data", *Transportation research part c: emerging technologies*, Vol.58, 240-250.
- Barry J.J., Newhouser R., Rhabee A., Sayeda S., 2002, "Origin and destination estimation in New York City with automated fare system data", *Transportation Research Record*, No.1817, 183-187.
- Bastin G. & Tubaro P., 2018, "Le moment big data des sciences sociales", *Revue française de sociologie*, Vol.59, No.3, 375-394.
- Beaude B., 2015, "Spatialités algorithmiques", In : Severo M. et Romele A. (dir.), *Traces numériques et territoires*, Paris, Presses des Mines, 135-162.
- Briand A. S., Côme E., Trépanier M., & Oukhellou L., 2017, "Analyzing year-to-year changes in public transport passenger behaviour using smart card data", *Transportation Research Part C: Emerging Technologies*, Vol.79, 274-289.
- Bonnel P., Hombourger É., Olteanu-Raimond A.M., & Smoreda Z., 2017, "Apports et limites des données passives de la téléphonie mobile pour la construction de matrices origine-destination", *Revue d'Économie Régionale & Urbaine*, No.4, 647-672.
- Calabrese F., Diao M., Di Lorenzo G., Ferreira J., & Ratti C., 2012, "Understanding individual mobility patterns from urban sensing data: A mobile phone trace example", *Transportation research part C: Emerging Technologies*, Vol.26, 301-313.
- Castex E., 2015, "Organisation des pratiques de covoiturage entre inconnus dans les territoires", *Netcom*, Vol.29, No.1/2, 153-176. URL : <https://journals.openedition.org/netcom/1905>
- CGDD, 2016, "Covoiturage longue distance : état des lieux et potentiel de croissance", Collection Études et documents du Commissariat Général au Développement Durable (CGDD), No.146.
- Chen C., Ma J., Susilo Y., Liu Y., & Wang M., 2016, "The promises of big data and small data for travel behavior (aka human mobility) analysis", *Transportation research part C: Emerging Technologies*, Vol.68, 285-299.
- CEREMA, 2016, "Analyse économique exploratoire du covoiturage longue distance", Rapport d'étude.
- Côme E., Oukhellou L., 2014, "Model-based count series clustering for bike sharing system usage mining: a case study with the Vélib' system of Paris", *ACM Transactions on Intelligent Systems and Technology*, Vol.5, No.3, 39-66.
- Courmont A., 2018, "Plateforme, big data et recomposition du gouvernement urbain", *Revue française de sociologie*, Vol.59, No.3, 423-449.

Depeau S., Quesseveur E., "A la recherche d'espaces invisibles de la mobilité", *Netcom*, Vol.28, No. 1-2, 35-54. URL : <https://journals.openedition.org/netcom/1539>

Gwiazdzinski L., 2014, "Du suivi GPS des individus à une approche chronotopique", *Netcom*, Vol.28, No.1-2, 77-106. URL: <https://journals.openedition.org/netcom/1604>

Hawelka B., Sitko I., Beinat E., Sobolevsky S., Kazakopoulos P., Ratti C., 2014, "Geo-located Twitter as the proxy for global mobility patterns", *Cartography and Geographic Information Science*, Vol.41, No.3, 260-271.

Iqbal M. S., Choudhury C. F., Wang P., & González M.C., 2014, "Development of origin–destination matrices using mobile phone call data", *Transportation Research Part C: Emerging Technologies*, Vol.40, 63-74.

Jacquot M., 2018, "Quel futur pour le covoiturage ? Comment surmonter les obstacles ?", *Annales des Mines-Réalités industrielles*, No.2, 52-55.

Jiang S., Ferreira J., & González M. C., 2017, "Activity-based human mobility patterns inferred from mobile phone data: A case study of Singapore", *IEEE Transactions on Big Data*, Vol.3, No.2, 208-219.

Jurdak, R., Zhao, K., Liu, J., AbouJaoude, M., Cameron, M., & Newth, D. (2015). Understanding human mobility from Twitter, *PloS one*, Vol.10, No.7, URL: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0131469>

Measure A., 2014, "Le design des programmes : des façons de faire du numérique", Thèse de doctorat, Université Paris 1, URL : <http://www.softphd.com>

Mericskay B., Noucher M., & Roche S., 2018, "Usages des traces numériques en géographie : potentiels heuristiques et enjeux de recherche", *L'Information Géographique*, Vol.82, No.2, 39-61.

Mohamed K., Côme E., Oukhellou L., & Verleysen M., 2017, "Clustering smart card data for urban mobility analysis", *IEEE Transactions on Intelligent Transportation Systems*, Vol.18, No.3, 712-728.

Plantin J.C., Russo F., 2016, "D'abord les données, ensuite la méthode ?", *Socio*, No.6, 97-115, URL : <https://journals.openedition.org/socio/2328>

Quesnot T., 2016, "L'involution géographique : des données géosociales aux algorithmes", *Netcom*, Vol.30, No.3/4, URL : <http://netcom.revues.org/2545>

Rashidi T. H., Abbasi A., Maghrebi M., Hasan S. & Waller T.S., 2017, "Exploring the capacity of social media data for modelling travel behaviour: Opportunities and challenges", *Transportation Research Part C: Emerging Technologies*, Vol.75, 197-211.

Richer C., Côme E., El Mahrssi M. K., & Oukhellou L., 2018, "La mobilité intermodale par les données billettiques. Analyses spatio-temporelles du réseau bus-métro de Rennes Métropole", *Cybergeo : European Journal of Geography*, Systèmes, Modélisation, Géostatistiques, article No.854 URL : <https://journals.openedition.org/cybergeo/29132>

Schwanen T., 2017, "Geographies of transport II: Reconciling the general and the particular", *Progress in Human Geography*, Vol.41, No.3, 355-364.

Senaratne H., Mueller M., Behrisch M., Lalanne, F., Bustos-Jiménez, J., Schneidewind, J., Schreck, T. 2018, "Urban Mobility Analysis With Mobile Network Data: A Visual Analytics Approach", *IEEE Transactions on Intelligent Transportation Systems*, Vol.19, No.5, 1537-1546.

Severo M., Romele A., 2015, *Traces numériques et territoires*, Paris, Presses des Mines, Territoires numériques.

- Toole, J. L., Colak, S., Sturt, B., Alexander, L. P., Evsukoff, A., & González, M. C., 2015, "The path most traveled: Travel demand estimation using big data resources", *Transportation Research Part C: Emerging Technologies*, Vol.58, No.1, 62-177.
- Vidal P., 2015, "Tracer sa route, en toute intransparence numérique ? ", *Netcom*, Vol.29, No.1/2, 5-12. URL: <http://journals.openedition.org/netcom/1841>
- Vogel P., Greiser T., & Mattfeld D. C., 2011, "Understanding bike-sharing systems using data mining: Exploring activity patterns", *Procedia-Social and Behavioral Sciences*, Vol.20, 514-523.
- Yang C., Huang Q., Li Z., Liu K., & Hu F., 2017, "Big Data and cloud computing: innovation opportunities and challenges", *International Journal of Digital Earth*, Vol.10, No.1, 13-53.
- Wergin J., & Buehler R., 2017, "Where Do Bikeshare Bikes Actually Go? Analysis of Capital Bikeshare Trips with GPS Data", *Transportation Research Record: Journal of the Transportation Research Board*, No.2662, 12-21.
- Zhong C., Manley E., Arisona S. M., Batty M., & Schmitt G., 2015, "Measuring variability of mobility patterns from multiday smart-card data", *Journal of Computational Science*, Vol.9, 125-130.

NOTES

1. On peut définir les traces numériques comme des empreintes laissées derrière soi lorsque l'on navigue sur Internet ou lorsque l'on recourt à un service à partir d'un dispositif numérique (Severo et Romele, 2015). Elles sont dites (géo)numériques, si des données (ou métadonnées) géographiques autorisant leur spatialisation, leur sont associées. Il peut s'agir de coordonnées géographiques, d'un toponyme, d'un numéro de borne Wifi, d'une cellule de réseau de téléphonie mobile ou encore d'un nom de lieu (Mericskay *et al.*, 2018).
2. Le service *Uber Movement* par exemple, propose un accès gratuit aux données anonymisées de plus de deux milliards de courses dans le monde sous forme de matrices origine-destination temporelles.
3. L'application Waze (propriété de Google) proposent aux autorités publiques un programme de partenariat intitulé *Connected Citizens* qui leur permet d'accéder à des données trafic issues de la plateforme en contrepartie d'informations sur les travaux, les modifications de voirie ou les événements à venir (Courmont, 2018).
4. Une API (*Application Programming Interface*) est une interface de programmation qui permet de mettre en relation un programme avec un autre. Ces outils permettent de réutiliser des données et des fonctions provenant de différents sites dans un autre projet (Masure, 2014).
5. L'exemple des données de téléphonie mobile est symptomatique des verrous en présence. Au-delà de leur caractère dispendieux, ces données se caractérisent aussi par de nombreuses contraintes et limites d'usage à l'image de leur disponibilité, leurs difficultés d'enrichissement (modes, motifs, caractéristiques des individus), leurs difficultés de traitement (mauvaise précision, non-interopérables avec les systèmes existants) ou encore les effets de bord géographiques liés à la couverture des antennes (Bonnell *et al.*, 2017).
6. Le covoiturage organisé sous-entend une planification organisée entre le conducteur et le(s) passager(s) qui passe par un service de mise en relation (gratuit ou payant). Dans le cas de BlaBlaCar, la plateforme apporte aux usagers des services et des outils facilitant les échanges et l'organisation des trajets (moteurs de recherche, messagerie, paiements sécurisés, systèmes d'évaluation, etc.) en contrepartie d'une commission sur chaque réservation.
7. URL de l'API : <https://dev.blablacar.com/docs/versions/1.0>

8. Le nombre de trajets complets au départ et/ou à destination d'une ville est affiché dans chaque réponse renvoyée par l'API. Le taux de trajets complets quotidiens oscille selon les jours entre 20 % et 40 %.
 9. Lors de la phase de collecte des données, il était encore possible de récupérer l'historique des trajets sur trois mois. Depuis octobre 2018, l'API ne propose plus d'accéder aux trajets passés.
 10. La carte en oursin crée une ligne entre le lieu de départ et le lieu d'arrivée de chaque trajet.
 11. Le choix de l'EPCI comme échelle d'analyse permet de prendre en compte les aires d'influence des métropoles et des villes moyennes, à la différence de l'échelon communal qui introduirait de nombreux biais, à l'image de zones de covoiturage en périphérie des villes. Au sein de la métropole nantaise par exemple, 40 % des trajets s'effectuent au départ et à destination d'Orvault, une commune située au nord de Nantes disposant d'un accès à la rocade et d'un arrêt de tramway. Idem pour Paris, où une grande partie des trajets partent et arrivent dans des communes de la petite et de la grande couronnes.
 12. Un covoiturage entre Rennes et Nantes coûte en moyenne sept euros et dure environ une heure alors qu'un trajet en train revient environ à vingt euros et dure entre 1h15 et 2h. La fréquence quotidienne est en moyenne d'une quinzaine de trains directs chaque jour contre une centaine de trajets en covoiturage avec des départs pouvant s'effectuer toutes les dix minutes.
 13. Grille raster pour laquelle on attribue à chaque pixel une valeur de mesure correspondant à la densité de points dans et autour du carreau. La carte présentée se base sur un rayon de densité de points de 300 m pour chacun des lieux de covoiturage dans une grille raster de 10 m par 10 m. A l'échelle de la ville et à la vue de la distribution spatiale des lieux de covoiturage, un rayon de densité de 300 m permet de bien mettre en évidence des dynamiques très localisées et dans le même temps de conserver une vision globale synthétique.
 14. De manière similaire à la carte de chaleur, le choix de mailles vectorielles de 300 m sur 300 m offre une échelle d'agrégation pertinente pour appréhender les dynamiques de distribution spatiales des lieux de covoiturage.
 15. A titre d'exemple, 50 % des lieux de covoiturage se situent à moins de 300 m d'une station de métro et ce chiffre monte à 70 % pour une distance de 500 m. Concernant la proximité à la rocade de Rennes, plus de 50 % des lieux de covoiturage se trouvent à moins de 500 m de la rocade et 70 % à moins d'un kilomètre.
 16. Le stationnement de la station de métro de la Poterie peut accueillir plus de 700 trajets lors de certains vendredis avec des pics d'une centaine de véhicules par heure en fin d'après-midi.
-

RÉSUMÉS

Les données massives représentent un champ d'investigation riche de promesses mais encore complexe pour renouveler l'analyse des mobilités spatiales. Toutefois, penser les pratiques de déplacements par le prisme des traces (géo)numériques soulève de multiples interrogations tant au niveau de la nature des données, des modalités d'accessibilité que des méthodes et des techniques de traitement associées. Cet article a pour ambition d'explorer les potentialités des traces numériques dans la compréhension du covoiturage à travers l'exploration des données de la plateforme BlaBlaCar. Par l'analyse (temporelle et spatiale) des trajets publiés au départ et à destination de la métropole rennaise pendant cinq mois, l'objectif est à la fois de dresser un portrait du covoiturage au sein de la capitale bretonne et de faire un tour d'horizon des

avantages et des limites de ces données dans la compréhension de cette nouvelle forme de mobilité.

Big data is a field of investigation rich in promises but still complex in renewing spatial mobility analysis. In fact, thinking about mobility through the prism of digital footprints raises many questions regarding the data nature, the accessibility procedures and the methods and techniques of treatment. This paper aims to explore the potential of digital footprints in the analysis of mobility through the example of data from the Carpooling platform BlaBlaCar. By analyzing (temporal and spatial) trips to and from city of Rennes over the course of 5 months, the goal is both to draw a portrait of Carpooling in the capital of Brittany and to give an overview of advantages and limitations of these data in understanding this new form of mobility.

INDEX

Keywords : exploratory data analysis (eda), spatial analysis, data, mobility, GIS

geographyun 908, 926, 250

Mots-clés : exploration de données, analyse spatiale, mobilité, SIG

AUTEUR

BORIS MERICKSKAY

Université Rennes 2 – UMR Espaces et Sociétés (ESO-Rennes), France

Maître de conférences en géographie

Boris.merickskey@univ-rennes2.fr

<https://perso.univ-rennes2.fr/boris.merickskey>