



De la production à la maintenance : faire exister collectivement des données géographiques ouvertes

Jérôme Denis, Clement Marquet

► To cite this version:

Jérôme Denis, Clement Marquet. De la production à la maintenance : faire exister collectivement des données géographiques ouvertes. Questions de communication, 2019, 36, pp.63-83. 10.4000/questionsdecommunication.20914 . hal-02557885

HAL Id: hal-02557885

<https://minesparis-psl.hal.science/hal-02557885>

Submitted on 29 Apr 2020

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.

De la production à la maintenance : faire exister collectivement des données géographiques ouvertes

Jérôme Denis

Centre de sociologie de l'innovation
i3 (UMR 9217), MINES ParisTech, PSL Research University
jerome.denis[at]mines-paristech.fr

Clément Marquet

Département Sciences Économiques et sociales
i3 (UMR 9217), Télécom Paris, Institut Polytechnique de Paris
clement.marquet[at]meandres.me

Denis, J. & Marquet, C. 2019. De la production à la maintenance : faire exister collectivement des données géographiques ouvertes, *Questions de communication*, (36), p. 63-83.

Résumé

L'article décrit les différentes étapes qui ont ponctué la collaboration entre les équipes de Transilien et celles d'OpenStreetMap France dans l'élaboration d'une base de données géographique ouverte cartographiant les gares du réseau d'Île-de-France. Il montre que les partenaires de ce projet se sont confrontés à une série d'épreuves qui les a progressivement conduits à reconsidérer la place des données qu'ils souhaitaient rendre disponibles, à trouver les moyens d'équiper le travail des données lui-même, mais aussi à se réorganiser au fil du temps. De la redéfinition des conditions de production des données jusqu'à la prise en compte des modalités de leur maintenance, ces épreuves mettent en lumière l'épaisseur sociotechnique des données géographiques destinées à être « ouvertes » et du travail sur lequel repose leur existence et leur disponibilité. Ce cas invite à prêter attention aux formes de coopération qui s'inventent lors de la fabrique des données et qui font émerger des possibles alternatifs pour la gouvernance numérique des territoires.

Mots clés

OpenStreetMap, open data, données géographiques, cartographie, accessibilité

De l'apparition des premiers systèmes d'information géographiques (Schuurman, 2009) jusqu'à l'avènement des outils et services dits « participatifs » (Dodge, Kitchin, 2013 ; Goodchild, 2007), les mutations qu'ont connues les pratiques de cartographie ces dernières années ont eu d'importantes répercussions, à commencer par la reconfiguration de la géographie elle-même (Turner, 2006). Parmi ces mutations, nombreuses sont celles qui touchent les villes, participant d'un mouvement qui dépasse le seul cadre de la fabrique des cartes en tant que telle, et qui se sont cristallisées autour du terme *villes intelligentes* (Picon, 2018 ; Townsend, 2014). La cartographie est aujourd'hui un élément central non seulement dans le renouvellement des formes de représentation des villes, mais aussi, plus généralement, dans la transformation des modalités mêmes de leur gouvernement et de leurs modes d'existence politique (Burns, Meek, 2015). Ainsi certains voient-ils dans la production de cartes – activité désormais accessible à un nombre exceptionnel de contributeurs – un instrument permettant de faire exister de nouvelles réalités urbaines (Zook, Graham, 2007) et de renforcer les possibilités critiques de la cartographie (Crampton, Krygier, 2005 ; Plantin, 2011). À l'inverse, d'autres insistent sur les désillusions accompagnant ces promesses de démocratisation (Haklay, 2013) et inscrivent les transformations de la cartographie dans une tendance « néolibérale » qui voit l'État se désengager de cette part cruciale de la fabrique urbaine au profit d'entreprises privées dont les intérêts ne garantiraient plus le bien commun (Leszczynski, 2012 ; Plantin, 2018).

Toutefois, une grande partie des travaux académiques qui se sont intéressés à cette « néogéographie », et en particulier aux dispositifs de cartographie participative, oriente les débats sur des questions très générales, qui laissent de côté les activités concrètes et les modes d'organisation que mettent en œuvre celles et ceux qui s'engagent dans la production de ces nouvelles cartes numériques. Dans cet article, plutôt que de chercher à qualifier le régime politique dont ses pratiques témoigneraient ou à évaluer leur caractère plus ou moins démocratique, nous nous inscrivons dans la lignée des recherches qui défendent une approche « processuelle » de la cartographie (Kitchin, Dodge, 2007) afin de comprendre ce qui se joue dans la dynamique même de la fabrique contemporaine des données géographiques et ce que cette dynamique donne à voir des articulations concrètes entre les outils et infrastructures dits « numériques » et les processus territoriaux.

Pour cela, nous analyserons un cas français : la constitution d'une base de données géographique des gares d'Île-de-France. Au départ destinée à équiper de futurs services d'aide à l'accessibilité, cette base de données a donné lieu à une collaboration entre Transilien, alors branche de la SNCF chargée du réseau de trains francilien, et OpenStreetMap France, association régie par la Loi du 1^{er} juillet 1901 relative au contrat d'association chargée d'animer au niveau local les initiatives attachées à OpenStreetMap (OSM), projet international de base de données géographique libre et ouverte, alimentée par des contributeurs volontaires bénévoles.

Une telle collaboration n'est pas anodine. Pour l'équipe d'OSM France, elle est l'occasion d'apporter la preuve de sa capacité à produire une base de données suffisamment fiable et précise pour satisfaire les exigences d'une grande entreprise. En 2013, au moment où

commence notre observation, la base de données a une petite dizaine d'années et commence à avoir quelques succès de collaboration aux niveaux international et français¹.

Si les outils du logiciel libre et, plus rarement, de la cartographie ouverte ont pu être intégrés par de nombreuses entreprises, les cas de partenariat entre une communauté bénévole et un grand groupe sont exceptionnels. Du côté de Transilien, le rapprochement avec OSM intervient, comme nous le verrons, dans le cadre d'une politique d'ouverture des données et d'une stratégie d'innovation ouverte, mais aussi dans un contexte de position hégémonique de Google Maps, dont certaines personnes au sein de l'organisation se méfient. Pour elles, l'enjeu est d'éviter toute dépendance cartographique à Google qui, dans de nombreux pays, a pris la main sur les données de transport, devenant ainsi incontournable aussi bien pour les entreprises concernées que pour leurs usagers². En s'associant dans un projet de coproduction de données cartographiques ouvertes, les deux organisations s'inscrivent dans une dynamique politique et économique contemporaine qui voit s'articuler les enjeux autour de la production et de la propriété des données cartographiques à la question de la souveraineté d'une entreprise publique sur son périmètre d'action territoriale.

Pour comprendre ce qui s'est joué dans cette relation partenariale, nous nous appuyons sur l'enquête conduite par l'un des auteurs (Marquet, 2016, à paraître) durant plus de quatre ans. Inspirée des perspectives de l'ethnographie institutionnelle (Smith, 2005), celle-ci a démarré par la participation de Clément Marquet au Hackathon organisé par Transilien (voir *infra*). Parmi les pistes suivies pour explorer les enjeux de la politique de mise en accessibilité numérique du territoire, la question des données géographiques et de leur place dans cette politique est peu à peu devenue centrale. À partir de cette première situation d'observation privilégiée, l'enquête a donc consisté à reconstituer la genèse de ces données, puis à accompagner les étapes qui rythmaient leur production et leur entretien. Elle a donné lieu à une série d'entretiens semi-directifs menés auprès des principales personnes impliquées dans le projet aussi bien du côté de la SNCF que de celui d'OSM France, à la récolte et l'analyse des nombreux documents produits pour l'occasion et à des observations directes à l'occasion de rassemblements³ et de campagnes de cartographie participative.

Mobiliser ce matériau permettra d'opérer un double mouvement analytique. D'abord, nous reconstituerons le parcours des données (Bates, Lin, Goodale, 2016), mais en nous focalisant sur le temps long de leur production, c'est-à-dire en amont de leurs usages et de leurs traitements. Ce faisant, nous rompons avec la posture de recherche qui s'est stabilisée à la croisée des problématiques de la ville « intelligente » ou « numérique », et qui questionne le rôle

¹ Parmi les plus marquants, on peut noter en France la mise en œuvre de la « Base d'Adresses Nationale » issue d'une convention signée entre l'IGN, le Groupe La Poste, l'État et OpenStreetMap France. À l'échelle internationale, OpenStreetMap a mené de nombreuses interventions dans les domaines du développement et de l'aide humanitaire, qui ont donné notamment lieu à des partenariats avec la Banque Mondiale et l'ONU.

² Les augmentations tarifaires de l'utilisation de Google Maps à partir du 11 juin 2018 ne sont pas sans donner raison aux inquiétudes des promoteurs de l'usage d'OSM chez Transilien. L'ancien président d'OSM France, Christian Quest (2018), a explicité sa position sur le sujet.

³ Les « State of the Map » sont des réunions annuelles de la communauté OSM au cours desquelles des ateliers et conférences sont organisés par les bénévoles pendant deux jours et demi. C. Marquet s'est rendu à ces rassemblements en 2015, 2016 et 2017.

des données en se préoccupant presque exclusivement des conséquences de leur utilisation sans interroger les opérations qui président à leur existence et à leur disponibilité. Nous montrerons que le projet s'organise précisément, non pas à partir d'une volonté de traiter des données déjà-là, mais, au contraire, autour du constat de l'absence de données. Comment faire exister des données pertinentes pour cartographier l'accessibilité des gares ? Qui peut et doit se charger de les produire ? Quelles formes doivent-elles prendre ? Comment s'assurer de leur validité ? Dans la lignée d'une sociologie du travail des données (Denis, 2018), nous montrerons que les questions que posent les différentes personnes concernées par ce projet et les manières qu'elles ont d'y répondre inventent progressivement une place pour ces données et une place pour le travail nécessaire à leur existence. Ces choix qui peuvent sembler « purement » techniques sont toujours organisationnels et, *in fine*, politiques.

Donner à voir les différentes étapes de ce projet permettra, dans le même temps, d'analyser les formes de collaboration qui s'inventent à l'interface des services « numériques » et de l'organisation des territoires. Ici aussi, nous ferons un pas de côté. En effet, le projet que nous étudions invite à dépasser deux des grandes oppositions binaires qui sont largement utilisées pour rendre compte de la transformation récente des équilibres au sein des acteurs de la ville. Avec comme point de départ le partenariat entre une association dont les outils sont en grande partie conçus et utilisés pour concurrencer les acteurs privés de la cartographie numérique mondiale, et une entreprise publique de transport, le projet en question ne peut d'abord pas être compris au seul prisme de l'opposition public/privé. Nous le verrons au fil du texte, les déplacements des personnes au sein des deux organisations et l'intégration progressive de nouveaux acteurs invitent à dépasser cette distinction trop rigide qui plaquerait l'analyse sur des identités stabilisées et *a priori* explicatives. Être attentifs à ces déplacements et hybridations empêche par ailleurs de marquer une séparation nette entre les initiatives cartographiques qui feraient intervenir des contributeurs « amateurs » d'un côté et celles qui impliqueraient des « professionnels » ou des « experts » de l'autre.

Des données pour l'accessibilité : une récolte à peu de frais ?

Le projet de construire une base de données géographique des gares du réseau francilien, d'abord consacrée à l'accessibilité, a émergé dans le cadre de la politique « d'innovation ouverte » mise en place au sein de Transilien au début des années 2010. Cette politique a dessiné deux principales lignes stratégiques dans l'entreprise : l'investissement dans des projets de diffusion publique de données (programme Open Data) et l'organisation de hackathons, événements rassemblant des développeurs en informatique, des représentants d'utilisateurs et des personnels de l'institution organisatrice dans le but de faire naître dans un temps resserré des initiatives innovantes qui se traduisent dans des prototypes d'applications et de services. C'est à l'issue d'un premier hackathon en 2012 consacré à l'information du voyageur qu'a d'abord été identifié au sein de Transilien l'intérêt de travailler la question de l'accessibilité et d'y consacrer le hackathon suivant, qui serait intitulé « Hackcess » (Marquet, 2016). Le choix de cette thématique est le résultat d'échanges, notamment entre l'agence de consultants qui accompagnait à l'époque la SNCF dans sa démarche, les responsables des services concernés dans l'entreprise (le pôle « Innovation et partenariats » de Transilien, la personne responsable du programme Open Data), et le porte-parole d'OSM France qui avait lui-même participé au premier hackathon. Au démarrage du projet, l'accessibilité représentait

un point d'articulation entre les préoccupations d'OSM France et des différents services impliqués à la SNCF (Marquet, à paraître). Toutefois, si la thématique pouvait initialement paraître familière à tous et donc offrir un terrain commun assez facile à explorer, elle s'est rapidement avérée plus problématique qu'attendu. Comme celui qui l'avait précédé, le hackathon sur l'accessibilité était d'abord pensé comme un moyen de mettre en avant l'utilité des données ouvertes par l'entreprise et ainsi de poursuivre la démonstration de la pertinence de sa politique d'*open data*. Or, dès qu'il a commencé à se pencher sur les données qui étaient effectivement disponibles dans les départements de la SNCF et qui auraient pu servir de matériau aux développeurs invités à concevoir des services d'accessibilité innovants, le responsable du programme Open Data au sein de Transilien n'a pu que constater une absence flagrante. Si les gares étaient bien représentées dans les systèmes d'information géographique de l'entreprise, et si une partie des objets architecturaux qui pouvaient constituer à première vue des ressources ou des obstacles pour y circuler en situation de handicap figuraient bien dans ses bases de données, aucun de ces éléments n'était suffisant pour considérer que les participants au futur hackathon auraient à leur disposition des données géographiques de qualité pour imaginer leurs services. Les plans des gares étaient très succincts et n'étaient pas accompagnés de données géocodées.

Quant aux équipements liés à l'accessibilité, les données informatiques les concernant n'indiquaient que des comptages globaux pour chaque gare, sans référence à leur emplacement précis (figure 1).

C'est à la suite de ce constat que le projet qui nous intéresse a vu le jour. En effet, en juin 2013, Transilien a signé un partenariat avec OSM France qui prévoyait de fournir aux développeurs un matériau utile sous la forme d'un jeu de données géographiques répertoriant et localisant l'ensemble des 124 gares des lignes C et L du RER. L'enjeu était de taille. Sans ces données, qui devaient être rassemblées par les contributeurs d'OSM d'ici la fin de l'été, le hackathon consacré à l'accessibilité prévu en novembre ne pouvait avoir lieu.

Figure 1. Extrait du fichier de données indiquant la seule présence (1) ou l'absence (0) des équipements de l'accessibilité par gare (source : <https://ressources.data.sncf.com/explore/dataset/equipements-accessibilite-gares-transilien/table/?disjunctive.gare&sort=gare> consultée le 26 novembre 2019)

gare	guichet_adapte_pmr	boucle_magnetique	portes_automatiques_acces
LES CLAIRIERES DE VERNEUIL	1	1	0
GAGNY	1	1	1
COUILLY SAINT-GERMAIN QUINCY	0	1	0
LA FERTE SOUS JOUARRE	0	1	0
AUVERS SUR OISE	1	1	0
MASSY VERRIERES	0	0	0

Ce premier geste n'est pas ordinaire au regard des projets d'*open data* qui avaient cours dans cette même période dans les administrations françaises (Denis, Goëta, 2017). En effet, il a consisté à ajouter à un projet de « libération » de données (généralement supposées présentes et disponibles dans les systèmes d'information existants), une étape de génération à part entière, reconnue comme telle. Toutefois, si les participants de ce projet affirmaient alors la nécessité de produire de nouvelles données, ils considéraient que celles-ci pouvaient émerger de l'activité bénévole de la communauté des cartographes d'OSM, au même titre que les

données que celle-ci avait l'habitude de produire. L'occasion, qui représentait pour l'association et l'entreprise publique une opportunité de collaboration officielle, entérinait l'idée que la production des données adéquates pouvait se faire à un coût presque nul, sans soulever de défis techniques autre que ceux que l'on imaginait traitables par les contributeurs attentionnés d'OSM.

Cette première étape s'est traduite par un appel lancé par les responsables d'OSM France et de Transilien sur les listes de diffusion françaises d'OSM⁴, invitant les participants à un « *challenge* » cartographique. Elle a aussi donné lieu à une première collaboration concrète entre les porteurs de l'opération chez Transilien et les représentants d'OSM France. Pour produire les données géographiques attendues, il fallait, comme pour n'importe quelle base de données, équiper les contributeurs d'un guide – un « référentiel » – désignant précisément ce qui devait être identifié et localisé et comment le faire. Ce guide a été construit à partir d'un cahier des charges d'une dizaine de pages dans lequel étaient listés les éléments architecturaux à faire apparaître dans la base.

Toutefois, ces informations ne pouvaient suffire, ne disant rien des modalités concrètes de l'intégration de ces entités au sein de la base de données d'OSM. Alimentée par l'intermédiaire d'une interface de programmation applicative (API), la base en question est organisée à partir de trois éléments distincts : des *nœuds* ; des *chemins* (qui peuvent être « fermés », « ouverts », ou constituer une « zone ») ; des *relations*. À chaque élément correspondent, outre sa géolocalisation, des attributs descriptifs qui sont appelés *tags* et qui sont mobilisés pour qualifier l'objet recensé. C'est l'utilisation systématique de ces tags par les contributeurs qui assure à la base sa cohérence et permet à des applications tierces de s'y greffer pour produire automatiquement des cartes ou des services localisés. Par exemple, pour entrer une poubelle de recyclage dans la base, on identifie un nœud avec le tag *amenity =recycling* auquel on peut éventuellement ajouter les matériaux concernés (*i.e. recycling :paper =yes*)⁵. Ces tags sont rassemblés en ligne dans des pages de « wiki ». Mais ils ne sont pas imposés par une organisation hiérarchique ; au contraire, ils sont proposés, puis débattus par les contributeurs eux-mêmes au sein de forums de discussion⁶. Pour les tags voués à être utilisés à l'échelle nationale, voire internationale, un vote est organisé à l'issue duquel le jeu de tags est officiellement inscrit dans la liste normalisée des outils de cartographie.

Pour qu'il puisse donner lieu à la cartographie à proprement parler, le référentiel initial de description des gares a donc été traduit dans le système de tags d'OSM. Ce moment d'intégration représente un premier temps fort de la collaboration entre les équipes de Transilien et d'OSM. Il montre que les échanges n'ont pas simplement consisté à trouver un consensus entre deux organisations aux positions arrêtées à l'avance. Dans la construction du référentiel et dans sa traduction, s'est non seulement jouée une certaine définition de l'accessibilité (Marquet,

⁴ La coordination des contributeurs passe pas divers outils en ligne (wiki, forums) ainsi que par un ensemble de listes de diffusion thématiques (par exemple : transport, accessibilité), nationales ou internationales. Ainsi une partie du travail cartographique réalisé par les contributeurs est-il dirigé par des projets ciblés, sur un ou deux mois, tels : cartographier toutes les bornes incendies, toutes les sorties de secours, ou cet été 2013, tous les équipements d'accessibilité sur les lignes d'Île-de-France.

⁵ Pour en savoir plus, voir : <https://wiki.openstreetmap.org/wiki/Database> (consulté le 13/08/2018).

⁶ On trouvera un exemple d'analyse de ce type de discussions à propos de la cartographie des infrastructures cyclables en France dans J. Denis et D. Pontille (2013).

à paraître), mais aussi une série de premiers ajustements et déplacements qui ont vu le personnel de Transilien produire des représentations inédites du territoire dont l'entreprise a la charge, et OSM ajouter à sa base de données de nouveaux attributs⁷. Ces ajustements auraient pu être les seuls en jeu dans ce projet qui devait déboucher, une fois les tags stabilisés, sur la production rapide d'un jeu de données géographiques exhaustif pour décrire l'accessibilité des gares concernées. Mais les choses se sont vite avérées plus compliquées, et les partenaires ont fait face à une série d'épreuves à travers lesquelles ils ont progressivement découvert les difficultés qu'impliquait la production des données qu'ils souhaitaient fournir aux participants du hackathon consacré à l'accessibilité. Ce faisant, les modalités de leur collaboration et l'objet du projet cartographique se sont transformés.

La première épreuve a été radicale : tandis que l'été avançait, le rythme des contributions a posé problème, bien trop éloigné de ce qui avait été envisagé au départ. Le défi que représentait la cartographie des gares n'était pas relevé, en tout cas pas complètement, et le projet s'est trouvé en danger. Ce premier échec a été l'occasion pour les acteurs impliqués d'explicitier un aspect rarement mis en avant dans les débats autour de la cartographie participative : le coût qu'engendre la production de données géographiques, coût que l'on minimise largement, voire que l'on ignore, dès lors que, pour invoquer la ville « numérique » ou encore les vertus de l'*open data*, on se focalise sur les *start-up* et les prototypes d'applications en oubliant les conditions concrètes de la génération des données.

De ce point de vue, le projet en question s'est avéré particulièrement ardu. Bien plus, finiront par comprendre les représentants mêmes d'OSM France, que la plupart des initiatives qui avaient nourri la base de données collective depuis des années déjà. Loin des métaphores de la fluidité et du déluge de données qui abondent dans les descriptions de la ville dite intelligente, les territoires, qu'il s'agissait ici de cartographier pour qu'ils deviennent à la fois le matériau et le théâtre de services innovants, semblaient résister à leur mise en données. Deux dimensions sont apparues particulièrement délicates : l'appréciation de l'accessibilité elle-même qui nécessitait que les contributeurs localisent des objets qu'ils n'avaient pas l'habitude de traiter, et la pratique de la cartographie dans les espaces intérieurs de la gare (cartographie dite *indoor*), une première pour OSM. Deux difficultés qui ont eu raison de la motivation des contributeurs bénévoles.

Le constat d'échec était sans appel. Trois mois après le démarrage de l'opération, à peine un quart des gares visées avaient fait l'objet d'une cartographie à peu près complète. La solution trouvée a consisté à internaliser d'urgence l'activité de production des données et à abandonner l'une des lignes, afin que le hackathon puisse être maintenu. C'est le responsable du programme Open Data au sein de Transilien qui a finalement pris en charge l'alimentation de la base et achevé de réunir le jeu de données jugé nécessaire à l'exercice. La collaboration entre OSM et l'entreprise de transport a pris à cette occasion une tout autre tournure puisque, pour s'atteler à cette lourde tâche, ce nouveau contributeur a été formé pendant son temps personnel par le porte-parole d'OSM France lui-même, qui l'a initié aux pratiques du relevé sur site et à l'utilisation des logiciels libres qui permettent de renseigner la base.

⁷ Ces déplacements de part et d'autre sont rendus explicites par une première participation commune des équipes de Transilien et des responsables d'OSM France à la conférence *State of the Map* dès avril 2014, conférence annuelle durant laquelle les projets les plus importants de la communauté d'OSM sont présentés. D'autres interventions communes suivront à l'occasion des conférences suivantes, nous y reviendrons.

Le relevé de la ligne C (84 gares) a été accompli en quelques semaines, et le hackathon a bien eu lieu. Il a été considéré par tous comme un succès, en particulier parce que les données mises à disposition ont bel et bien servi. Plusieurs équipes de développeurs ont inscrit dans les prototypes d'applications et dans les services numériques qu'ils ont imaginés des formes d'accessibilité étroitement liées à cette infrastructure informationnelle (Marquet, à paraître).

Si l'événement visait en partie à démontrer la capacité de la SNCF à mettre en œuvre une politique d'innovation ouverte qui dépassait la déclaration d'intention, ce succès a également eu des répercussions du côté de la relation entre l'entreprise de transport et OSM France. Aux yeux de certains, c'est en effet l'importance même des données géographiques et l'utilité des services cartographiques associés aux gares qui ont été confirmées, au-delà de la seule question de l'accessibilité, et bien au-delà de la seule ligne C du RER. Pour suivre cette piste, il allait falloir cartographier l'ensemble des 383 gares du réseau régional en relevant tous les *Points of interest* (POI) de chaque gare.

L'opération de mise en place d'un équipement informationnel censé simplement outiller les expérimentations des participants au hackathon s'est donc trouvée elle-même requalifiée d'expérimentation à part entière. L'élaboration d'une infrastructure de données s'est trouvée au centre des préoccupations, son succès plaçant pour qu'elle soit répliquée sur l'ensemble du territoire concerné. Renouvelé, le partenariat entre la SNCF et OSM France s'est trouvé profondément reconfiguré. Il était désormais clair que l'exercice ne pouvait rentrer dans les prérogatives de la communauté des contributeurs volontaires d'OSM, ni même d'une seule personne, fût-elle bien formée. Du côté de Transilien, les responsables en étaient convaincus : il fallait rémunérer le travail d'alimentation de la base. Du côté du conseil d'administration d'OSM France, le bénévolat n'apparaissait pas non plus comme une bonne modalité de collaboration :

« On a été échaudé par les premiers travaux sur les gares à Paris, il fallait faire la carte, besoin de résultat, on s'est confronté à la limite de la mobilisation des bénévoles et on a fonctionné comme un bureau d'étude à cartographier à trois ou quatre » (Louis-Julien de la Bouère, contributeur, directeur de l'association Tiriad et membre du conseil d'administration d'OSM France, intervention lors de la conférence *State of the Map*, Avignon, 3 juin 2017).

Pour poursuivre la cartographie, les responsables de Transilien confient un contrat de prestation à une « Junior-Entreprise »⁸ aux tarifs relativement bas. 16 étudiants sous contrats à durée déterminée, payés « à la gare », allaient devoir produire les données géographiques qui permettraient de fournir une infrastructure informationnelle exhaustive sur laquelle des services innovants allaient pouvoir s'appuyer. Ici encore, on le voit, cela n'est pas seulement les relations entre l'entreprise de transport et l'association de cartographie participative qui se modifient, mais les organisations elles-mêmes. On voit ici la fluidité des modes d'organisation de la

⁸ Les Junior-Entreprises sont des associations dites « à but non lucratif » à vocation économique et pédagogique. Elles doivent permettre aux étudiants de réaliser des expériences professionnelles en continuité avec leur parcours dans l'enseignement supérieur.

communauté d'OSM avec l'apparition d'une situation plutôt rare dans son fonctionnement : l'alimentation de la base de données par des personnes salariées⁹.

Cela dit, cette première incursion de la collaboration entre OSM France et Transilien dans un modèle de production de données rémunérée n'a, à son tour, pas été jugée concluante. Brièvement formés par les responsables d'OSM France et outillés des mêmes instruments que n'importe quel contributeur volontaire, les étudiants ont rapidement été confrontés à des problèmes comparables à ceux rencontrés par les premières personnes qui s'étaient attelées à la cartographie des gares de la ligne C. Les situations de relevé étaient difficiles, les objets à inscrire dans la base particulièrement délicats à appréhender. Et l'encadrement même de l'équipe n'a pas été une tâche facile pour le responsable de l'*open data* au sein de Transilien.

À ces difficultés, s'est ajouté le caractère précaire de l'emploi et du mode de rémunération. Certaines personnes n'ont réalisé le relevé que de quelques gares, sans aller au bout de ce qu'elles étaient censées faire, se contentant des premiers paiements avant de quitter le projet. L'intégration même à OSM était délicate, et le fait que certains contributeurs volontaires se chargeaient de corriger, voire d'effacer, des données enregistrées par les étudiants n'était pas sans poser problème. L'initiative n'a pas été renouvelée, bien que l'alimentation de la base de données ne soit pas terminée, tandis que, en parallèle, le responsable de l'*open data* côté Transilien quittait l'entreprise pour des raisons personnelles.

Le projet de cartographie n'en a pas été abandonné pour autant. De fait, l'épreuve de cet échec a donné lieu à une nouvelle reconfiguration sociotechnique de la production des données, qui a, encore une fois, transformé les deux organisations. Définitivement plus une opération bénévole, ni même une tâche réalisable à bas coût, la génération des données est alors devenue un travail à part entière.

Co-produire des données de qualité : hybrider, affûter, innover

En mars 2015, pour faire face au départ de la personne qui portait l'initiative de la cartographie depuis son démarrage, Transilien a recruté un chargé de projet dont la mission principale était de mener à bien l'alimentation de la base de données. Avec ce recrutement, le lien avec OSM s'est encore renforcé, puisque la personne en question était à l'époque un contributeur habituel de la base très engagé dans la communauté. C'est à l'occasion de l'arrivée de ce nouveau responsable qu'a été définitivement acté l'échec de l'intervention des personnes recrutées précédemment. Non seulement l'entreprise n'avait pas été menée à son terme, mais le chargé de projet a diagnostiqué des problèmes importants dans la qualité des données : les relevés géographiques n'étaient pas assez précis dans le positionnement des objets et le référentiel devant orienter les relevés des contributeurs n'était pas assez cohérent.

Les choses ont alors radicalement changé. Sur le plan organisationnel d'abord, on vient de le voir, les deux partenaires se sont presque hybridés puisque la personne engagée au sein de Transilien était directement issue de la communauté d'OSM. À partir de cet épisode, il est devenu impossible d'attribuer de manière claire des stratégies autonomes distinguant nettement les organisations. Du côté de la production concrète des données, les choses ont

⁹ Si le bénévolat est la principale forme d'engagement des personnes alimentant les bases de données d'OSM, il faut noter que certaines situations ont déjà donné lieu à des formes salariales, notamment autour des projets humanitaires développés par la Humanitarian OpenStreetMap Team.

elles-mêmes profondément évolué, y compris du point de vue des procédures habituelles d'OSM.

C'est d'abord l'objet même du projet qui s'est déplacé. En effet, à son installation, le nouveau chargé de projet a décidé de ne plus s'en tenir aux aménagements liés à l'accessibilité pour cartographier l'intérieur des gares dans leur totalité. Cela s'est traduit par l'élaboration d'un nouveau référentiel en collaboration avec des membres d'OSM France et ses collègues du service Innovation et partenariat de Transilien et de la stabilisation des tags permettant d'identifier 56 objets dans la base de données¹⁰. Loin de se résumer à un changement d'échelle, cette nouvelle ambition, qui consistait à faire exister l'objet « gare » du point de vue des données, a eu des répercussions sur les pratiques cartographiques au sein de la communauté d'OSM :

« Les données sont enregistrées dans une forme “relation de gare”, ce qui n'est pas terrible vis-à-vis d'OSM car c'est très difficile d'ajouter des données dans une relation, et cela pose problème si on tague un nouveau banc, il ne sera pas dedans. Cette solution temporaire a été adoptée pour pouvoir saisir toutes les données extérieures à la gare (par exemple, le banc à l'extérieur) ainsi que ce qui est dans la gare et ne peut être tagué “operator =snct”, comme le Relais » (chargé de projet open data Transilien, entretien, janv. 2017).

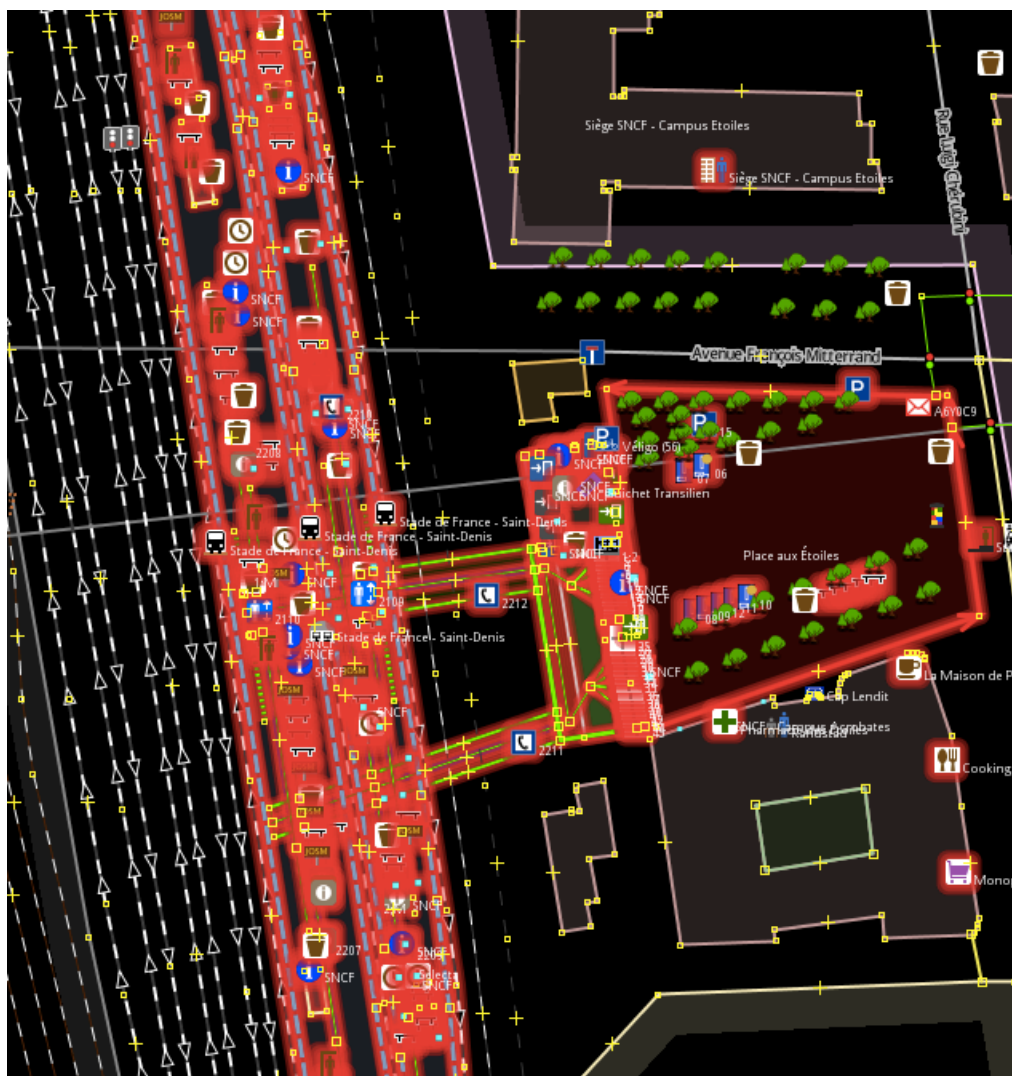
Sans entrer dans les détails, il faut insister sur la complexité de ce qui s'est joué ici, qui montre à quel point produire des données géographiques est une opération délicate. Le problème peut s'expliquer relativement simplement : ce qu'est une gare pour le chargé de mission *open data* de Transilien et ce qu'est une gare dans la structure traditionnelle des données d'OSM sont deux choses sensiblement différentes. Dans le deuxième cas, il suffit, pour faire exister chaque gare, d'associer un attribut spécifique aux objets recensés (par exemple avec l'identifiant de son « opérateur ») qui permet de les assembler par l'intermédiaire d'une requête dans la base. Cette situation ne convenait pas au responsable du projet du côté de Transilien qui considérait que, en s'en tenant à ce système, il perdait en partie la possibilité de maîtriser les modalités d'existence de ce qui compte comme une gare du côté de l'entreprise de transport, notamment au regard de ses propres infrastructures informationnelles et des applications futures que la cartographie en cours était vouée à équiper. C'est à ce titre qu'a été proposée l'utilisation des « relations », une méthode qui consiste à grouper un ensemble d'objets hétérogènes pour les faire appartenir à une même entité, ici la gare (figure 3). Or, l'utilisation de cette méthode est délicate. Elle complique la pratique de la cartographie et limite les opportunités de contribution pour les nouveaux entrants dans OSM, alors même qu'engager davantage de contributeurs est un souci important pour la communauté. L'emploi des relations dans l'organisation des données est l'objet de controverses régulières dans OSM¹¹, et, en cours de projet, un contributeur étranger a reproché à l'équipe française de ne pas avoir

¹⁰ Accès : https://wiki.openstreetmap.org/wiki/FR:Gares_ferroviaires. Consulté le 31/08/2018.

¹¹ Voir notamment la page expliquant les bonnes pratiques de l'usage des relations : https://wiki.openstreetmap.org/wiki/FR:Relations/Les_relations_ne_sont_pas_des_cat%C3%A9gories. Consulté le 31/08/2018.

respecté « les principes d'OSM »¹² dans la cartographie des gares. L'usage des relations a fini par être accepté dans le cadre de ce projet, et montre que les ajustements et compromis se sont joués jusque dans les aspects les plus techniques de la fabrique des données.

Figure 3. Interface de cartographie JOSM¹³ : vue de la gare Stade de France Saint-Denis¹⁴



La précision des relevés du bâti des gares constituait un autre problème. En effet, le responsable du projet a considéré que, si les étudiants de la Junior-Entreprise avaient peiné à localiser précisément les équipements de l'accessibilité, c'était en partie parce que les positions

¹² Une partie de la discussion, qui montre à quel point l'objet gare constitue en tant que tel un problème au regard de certaines pratiques cartographiques, est accessible à cette adresse : <https://www.openstreetmap.org/changeset/53267689> (consulté le 31/08/2018).

¹³ JOSM (éditeur Java pour OpenStreetMap) est un logiciel qui permet de contribuer à la base de données d'OpenStreetMap à partir d'un poste informatique standard.

¹⁴ Tous les éléments en clair (en rouge dans la version originale) sont pris dans la relation « gare ». Sur la partie droite de l'image, la relation gare incorpore la place aux étoiles, les distributeurs de billets, les bancs et places de parking, mais pas les arbres, les poubelles, ni la boîte postale ou la colonne publicitaire. Plus étonnant et difficile à expliquer, l'inclusion de la pharmacie, du Monop' de la Maison de Perle et du restaurant Cooking Cuisine Maison dans la relation gare.

et orientations géographiques des gares n'avaient pas toujours été réalisées correctement par les premiers contributeurs d'OSM. Pour corriger ces défauts, il a décidé de s'appuyer sur des documents internes, qui ne pouvaient être en possession des premiers cartographes sollicités. Ainsi un prestataire a-t-il été sollicité pour numériser et orienter les plans d'architecte des gares. Ensuite, le responsable du projet s'est appuyé sur un ensemble de cartes disponibles dans les différents services de l'entreprise pour recenser et localiser les équipements de chaque gare.

Il est ici aussi remarquable que la méthode employée par le chargé de projet – on le rappelle, contributeur de longue date à OSM – apparaisse à ce point en contradiction avec les principes méthodologiques défendus au sein de la communauté. En effet, la cartographie *in situ* est au cœur de l'initiative qui est régulièrement présentée comme devant s'appuyer sur une activité extérieure et non sur des opérations de simples saisies informatiques sur ordinateur. Comme l'a expliqué le responsable du projet au fil de discussions informelles, l'utilisation des plans d'architecte et autres cartes existantes a facilité ce qui se serait autrement révélé un « travail de titan ». La collaboration des deux organisations a rendu possible cette étape préalable (qu'on retrouve aussi dans certains projets humanitaires s'appuyant en partie sur des activités cartographiques menées à distance) qui a permis de faire ensuite à nouveau appel aux cartographes volontaires d'OSM pour compléter au mieux les gares restantes, au fil de leurs passages, après que référentiels et orientations du bâti ont été clarifiés et précisés.

Malgré ces ajustements non négligeables, le projet ne s'est pas clos pour autant. Huit gares ont résisté à cette méthode. Avec notamment plusieurs niveaux, elles présentaient un degré de complexité qui dépassait largement les compétences cartographiques du responsable du projet au sein de Transilien. Pour les traiter, il a été convenu de solliciter une agence de géomatique¹⁵. Ce nouvel épisode a fait à nouveau bifurquer les relations entre OSM France et Transilien puisqu'il a été l'occasion de greffer un nouvel acteur, sous la forme d'une entreprise prestataire plus traditionnelle. Mais, de fait, loin de distendre les liens entre les deux organisations, l'apparition de celui-ci a encore participé à les renforcer. En effet, l'agence en question, Carto'Cité, a été créée par un expert diplômé en géomatique étroitement lié à OSM, contributeur bien connu et respecté au sein de la communauté française. En signant le contrat de prestation, celui-ci s'est trouvé dans une situation inédite, rémunéré pour ce qu'il faisait bénévolement depuis des années : alimenter la base de données d'OSM. Toutefois, la position de Carto'Cité n'était pas complètement comparable à celle des contributeurs habituels. Son statut de prestataire lui a notamment donné accès aux précieux plans d'architecte, documents internes confidentiels qui avaient déjà servi de base au chargé de projet de la SNCF, un matériau qui n'aurait pas pu circuler tel quel au sein de la communauté des contributeurs bénévoles :

« Au début, je me posais des questions parce qu'eux ils peuvent être bénévoles et nous on est payés. Ça peut mettre mal à l'aise, on fait la même chose que [...] certains bénévoles, et on a accès à des plans d'architecte pour cartographier, ce qui serait un rêve pour certains contributeurs. Après, pour les plans d'architecte, il y a des accords de confidentialité, et ça ne peut pas se régler avec des individus seuls, il faut une structure

¹⁵ La géomatique est une discipline à la croisée de la géographie et de l'informatique qui regroupe les méthodes et technologies liées à la production, au traitement et à la diffusion des données géographiques.

pour pouvoir signer quelque chose comme ça » (cofondateur de Carto'Cité, entretien, mars 2017).

Ainsi, dans le cadre de ce montage contractuel, l'intervention de Carto'Cité a-t-elle élargi encore le périmètre de la production des données géographiques tel qu'il était envisagé au démarrage du projet. Et cet élargissement ne s'est pas arrêté à l'addition de compétences nouvelles en géomatique, elles-mêmes complétées par les documents internes de la SNCF. Produire des données géographiques pour les gares dites « complexes » représentait une véritable gageure technique pour les experts désormais concernés, qui a donné lieu à deux innovations qui se sont avérées cruciales pour tous les acteurs impliqués : la création d'un dispositif de prises de vue à 360 degrés permettant le relevé géographique « *indoor multilevel* », et un outil de visualisation pour les données géographiques multiniveaux sans lequel aucune carte n'aurait pu être produite à partir de ces données d'un nouveau genre. Portées par des membres actifs de la communauté d'OSM, ces innovations ont représenté à de nombreux égards le point d'orgue de l'initiative. D'une part, elles ont permis de conduire le projet de cartographie à son terme, fournissant à la SNCF un jeu de données géographiques ouvertes à grande échelle, et, d'autre part, elles ont renforcé l'expertise cartographique de l'entreprise prestataire (par la suite, le créateur de l'outil de relevé a été recruté par Carto'Cité) et élargi les possibilités cartographiques d'OSM dans son ensemble.

Le projet a été considéré par tous comme un succès. À l'interface d'OSM et des données ouvertes de Transilien, une véritable infrastructure de données était désormais disponible, ouvrant la voie à une multitude d'usages. Toutefois, comme une série de travaux l'a souligné depuis quelques années (Denis, Pontille, 2015 ; Graham, Thrift, 2007 ; Jackson, 2014), le temps de l'innovation ne suffit pas pour qu'une infrastructure technique existe et surtout fonctionne dans la durée (Ribes, Finholt, 2009). Pour jouer ce rôle, la base de données géographiques finalisée a dû passer par une dernière épreuve : celle de l'organisation de sa maintenance.

Des données, et après ? Assurer la pérennité, organiser la maintenance

Du côté de Transilien, les données créées avaient un double horizon. D'un côté, elles visaient à alimenter un ambitieux programme d'*open data* et représentaient un matériau important mis à la disposition du public ; de l'autre, elles étaient destinées à équiper la politique d'innovation ouverte revendiquée par l'entreprise, en particulier pour renforcer ses efforts dans le domaine de l'information voyageur, d'abord sur la question de l'accessibilité, puis sur un projet beaucoup plus généraliste. En tant qu'infrastructure informationnelle, la base de données désormais disponible sur OSM devait permettre à la SNCF de fournir à ses usagers franciliens des services d'aide aux déplacements plus précis que ce qui existait jusqu'alors.

Néanmoins, fournir ces services suppose non seulement l'existence d'une base de données complète, mais aussi que l'intégrité et la validité de ces données soient assurées. Cela signifie deux choses. D'abord, les données doivent être à jour. Comme tout environnement urbain, les espaces de la gare sont en constante transformation et sujets à des modifications temporaires plus ou moins importantes et plus ou moins durables. Ensuite, il faut éviter qu'elles ne soient dégradées, volontairement ou non. Cette question de la maintenance de la donnée est d'autant

plus importante que Transilien fait partie d'un écosystème de transport en Île-de-France, sous l'autorité du Syndicat des transports d'Île-de-France (Stif), dont les choix en termes d'infrastructures techniques doivent être suffisamment robustes pour pouvoir être partagés et/ou articulés aux pratiques et aux systèmes d'information de 70 organisations hétérogènes, voire concurrentes¹⁶.

Or, c'est le propre du partenariat entre OSM France et Transilien, alors même que les données des gares sont mobilisées dans les applications métiers et clientes de l'entreprise, elles restent ouvertes et donc modifiables comme n'importe quelles données d'OSM. La question de la mise à jour des données renvoie donc à l'une des caractéristiques centrales de la base de données qui est précisément organisée comme une infrastructure informationnelle jamais complètement stabilisée, et capable d'intégrer les transformations des territoires avec une grande précision et bien plus vite que la plupart des institutions traditionnellement chargées de produire des cartes. C'est notamment à ce titre qu'OSM est devenu en quelques années un outil incontournable dans l'accompagnement des crises humanitaires pour de nombreux acteurs, y compris l'Organisation des Nations unies et la Banque mondiale (Soden, Palen, 2014). Ce qui est habituellement présenté comme l'une des forces d'OSM s'est donc trouvé stigmatisé comme un défaut, voire un risque, lorsque la collaboration avec Transilien s'est installée dans le temps et que l'infrastructure de données a été confrontée, pour parachever cette collaboration, à son intégration dans le « feuilleté » (Bowker et Star, 1999) d'autres systèmes d'information et au respect des standards de l'institution de tutelle. Une fois le plus gros du travail effectué, à qui confier cet autre versant de la génération des données ? Autrement dit, comment mettre en œuvre ce que Karen S. Baker et Helena Karasti (2018) appellent le « soin des données » ?

Pour ce qui est de la mise à jour des données, au sein de Transilien, il a été décidé de ne pas seulement compter sur l'intervention des cartographes volontaires d'OSM, mais de mobiliser les agents en gare. Ainsi, équipés d'une application *smartphone* dédiée, ceux-ci ont-ils été chargés de mettre à jour les informations concernant leur gare à l'occasion des tournées quotidiennes d'inspection, ce qui permet une grande réactivité.

Quant à la question de l'intégrité des données, elle s'est avérée plus délicate. Au moment où nous écrivons cet article, la solution mise en place est encore à l'état d'expérimentation. Elle a consisté à ajouter à l'activité ordinaire de production et de mise à jour des données une activité de vérification, qu'il a été convenu de confier à un prestataire. Même si le prestataire en question n'était pas à proprement parler « extérieur » à la communauté d'OSM, puisque c'est à nouveau à Carto'Cité qu'il a été fait appel, les bouleversements que cette nouvelle étape du projet provoquait étaient loin d'être anodins¹⁷. Au fil des discussions et des ajustements, c'est

¹⁶ Sensibles au risque pour la SNCF de s'aventurer seule dans la production d'une nouvelle infrastructure informationnelle, les responsables du projet au sein de l'opérateur de transport s'étaient déjà préoccupés de la question alors que le projet était en voie de consolidation. Cela s'est notamment joué autour du référentiel de description des gares qui, une fois stabilisé, a été partagé avec les organisations publiques et privées faisant partie de l'écosystème francilien.

¹⁷ Cette distribution autour du travail des données n'était pas complètement nouvelle. Le Stif l'avait expérimentée quelques temps auparavant autour des données de la cyclabilité des infrastructures routières de la région, elles aussi produites en partenariat avec OSM et dont la vérification a été confiée à une entreprise prestataire dont les fondateurs étaient, tout comme ceux de Carto'Cité, des contributeurs anciens et reconnus d'OSM. Les débats au sein de la communauté ont porté dans le même temps sur ces deux « premières ».

non seulement une certaine division du travail des données qui s'est instaurée, mais aussi, avec elle, une accommodation de la définition des tâches elles-mêmes. Du côté d'OSM, il était peu envisageable d'installer une procédure de surveillance presque hiérarchique des données. L'activité confiée à Carto'Cité a progressivement été respécifiée comme une série d'interventions ordinaires de maintenance qui obéissaient à un protocole ménageant les modalités de participation d'OSM :

« Ce n'est pas de la surveillance, qui fait penser qu'on surveillerait les contributeurs. On est amené à observer des contributions de personnes qui ne connaissent pas toujours bien les règles pour l'ajout de données sur ces objets, surtout en zone complexe. Il faut alors les contacter par mail, et faire comme on fait dans la communauté, leur proposer d'ajouter leurs données dans une relation, etc. [...]. Ce respect des règles permet la pérennisation des données » (cofondateur de Carto'Cité, entretien, mars 2017).

Ainsi est-ce en trouvant un point de rencontre à mi-chemin entre les exigences d'intégrité des données du côté des grandes entreprises publiques de transport et les formes de contributions habituelles d'OSM qu'a été organisée cette part du travail des données relatif à leur qualité et à leur pérennité.

Conclusion

En découvrant le fait que les infrastructures physiques des gares résistaient aux pratiques habituelles des cartographes bénévoles, les partenaires de ce projet se sont confrontés à une série d'épreuves qui les ont progressivement conduits à se réorganiser et à redéfinir les conditions de production des données jusqu'à la prise en compte des modalités de leur maintenance. Ces épreuves mettent en lumière l'épaisseur sociotechnique des données géographiques destinées à être « ouvertes », une épaisseur qui reste dans l'angle mort des travaux qui s'en tiennent à une analyse des cartes et des données au repos, sans s'intéresser aux conditions concrètes de leur production (Kitchin, Dodge, 2007 ; Kitchin, Gleeson, Dodge, 2012) et aux frictions que celle-ci engendre (Edwards, Mayernik, Batcheller *et al.*, 2011).

À l'issue de la description de ce processus au long cours, plusieurs dimensions semblent importantes à retenir. D'abord, ce cas invite à prendre au sérieux le caractère protéiforme de la relation qu'entretiennent les différents partenaires au fil du projet. Un grand nombre de travaux portant sur les transformations des territoires urbains associées à la multiplication des technologies numériques mobilisent la distinction entre secteurs public et privé comme principe explicatif posé *a priori* ou comme moyen d'inscrire des situations locales dans une tendance à long terme avec une portée qui reste souvent implicite : il serait possible d'attribuer à chaque acteur des objectifs propres, stables et clairement définis. La trajectoire du projet de base de données géographique initialement porté par OSM France et certaines équipes de Transilien invite clairement à ne pas céder à cette tentation simplificatrice. Si les deux organisations demeurent et ne se diluent évidemment pas l'une dans l'autre à l'issue du processus, elles se reconfigurent sensiblement au fil des épreuves qui ne s'en tiennent jamais à des problèmes « purement » techniques. C'est le cas du côté des équipes de Transilien, qui se trouvent dès le départ dans une démarche de transformation associée à la mise en place d'une politique

d'innovation ouverte. Mais c'est aussi le cas de la communauté d'OSM qui évolue avec les projets, intégrant de nouveaux objets à cartographier, de nouvelles méthodes, de nouveaux instruments, et acceptant après de nombreuses discussions la mise en place d'une procédure inédite de maintenance des données qui instaure des tâches parallèles à sa logique de contribution horizontale confiées à des acteurs en contrats avec leurs partenaires.

Qui plus est, les deux organisations ont tendance à s'hybrider. Nous l'avons vu, les liens qui se tissent entre Transilien et OSM sont vite très étroits ; c'est un objectif clairement affiché des deux côtés dès le démarrage du projet. Le travail en commun pour la constitution du premier dictionnaire de tags et les discussions permanentes au fil des deux projets se sont notamment traduits par plusieurs participations aux conférences *State of the Map* organisées par OSM France, faisant de Transilien un acteur important de la communauté OSM, et de ce projet un exemple phare des innovations dans le domaine. Au fil des opérations et des épreuves, nous avons plus généralement vu que la frontière entre les deux organisations devenait elle-même de plus en plus poreuse. Les uns forment les autres, un contributeur important d'OSM est recruté par Transilien et l'entreprise prestataire à laquelle il est fait appel à deux reprises est dirigée par un membre important de la communauté des cartographes volontaires.

Ce premier point invite à nuancer les travaux qui ont comparé les cartes et les bases de données d'OSM à ce que S. Leigh Star et ces collègues ont appelé des « objets frontières » (Lin, 2011 ; Soden, Palen, 2014). L'un des grands apports de l'article original de S. L. Star et James R. Griesemer (1989) réside dans le rééquilibrage de la perspective des premières analyses de Michel Callon et Bruno Latour qui décrivaient les innovations comme des processus d'intéressement au fil desquels les intérêts de différents acteurs étaient progressivement traduits au profit de ceux que défendent les innovateurs. Se pencher sur les objets frontières consistait à se détacher de cette posture pour prêter attention à toutes les formes de traduction et d'intéressement et décrire des situations dans lesquelles aucun « vainqueur » ne se dégage nettement. Ce premier geste correspond tout à fait au projet que nous venons de décrire, qui voit des traductions s'opérer de part et d'autre des protagonistes sans que l'un ne puisse être présenté comme parvenant à aligner les intérêts de tous les autres. Toutefois, comme l'a rappelé S. L. Star à de nombreuses reprises (Bowker, Star, 1999 ; Star, 2010), la notion d'objets frontières sert par ailleurs à décrire des instruments qui sont mobilisés à l'interface de mondes sociaux qui sont appelés à coopérer malgré des préoccupations très différentes, sans que leurs modes de fonctionnement ne s'en trouvent perturbés. Les objets frontières (ou les « infrastructures frontières », qui intéressent plus particulièrement G. C. Bowker et S. L. Star) sont des moyens qu'élaborent les partenaires pour faire tenir un espace de collaboration. Si la base de données géographique dont nous avons suivi les principales étapes de la constitution est bien vouée à jouer ce rôle (par exemple auprès des développeurs et des participants aux hackathons), cela n'est pas en tant que telle que nous l'avons appréhendée ici. Avant de pouvoir devenir l'instrument de coopérations possibles, l'infrastructure informationnelle est dans notre cas le résultat progressif qui émerge d'une collaboration préalable, au cours de laquelle les partenaires se transforment et s'hybrident.

Notre enquête, qui se place en amont des versions stabilisées de la carte et de la base de données, met aussi l'accent sur le travail que l'existence même de données géographiques ouvertes nécessite. L'importance de ce « travail des données » (Denis, 2018) se fait d'abord sentir à même des opérations très concrètes. Le projet, non seulement dans son ensemble

mais aussi dès ses prémices, montre que les données, si souvent décrites avec le vocabulaire de la ressource naturelle, et ce en particulier lorsqu'il s'agit de dépeindre l'avenir des villes et de la gestion des territoires par le « numérique », sont loin d'être un matériau de base qui serait « naturellement » disponible. Ni leur récolte, ni leur ouverture ne vont de soi (Denis, Goëta, 2017). Ici, nous l'avons vu, c'est une série d'opérations de génération qui a été mise en œuvre, série qui s'est progressivement appuyée sur des procédures communes, des savoir-faire spécifiques et des instruments qu'il a parfois fallu inventer. Qui plus est, une fois cette génération assurée, l'inscription des données dans l'écologie sociotechnique de l'information voyageur, de ses infrastructures et de ses standards a fait émerger le problème de la pérennité de la base. Aux enjeux propres à la génération des données, ce sont ajoutés ceux que soulèvent les conditions techniques et organisationnelles de leur maintenance. Au fil du projet, ces opérations se sont révélées complexes. Une grande partie des épreuves que nous avons décrites se sont cristallisées sur les difficultés qu'elles représentaient et ont conduit les deux organisations partenaires à reconfigurer la participation des uns et des autres, à intégrer de nouveaux acteurs dans le projet et à faire de la plupart de ces opérations des tâches rémunérées.

On pourrait voir dans cette émergence progressive de la question du travail des données et de sa rémunération un reniement des principes de la communauté d'OSM et de son modèle bénévole. Ce serait aller un peu vite. D'abord parce que, nous l'avons déjà évoqué, la situation n'est pas complètement nouvelle (Palen, Soden, Anderson, Barrenechea, 2015 ; Soden, Palen, 2014). Ensuite, parce que la division du travail qu'a installé ce projet et les effets de professionnalisation qu'elle a fait émerger au sein même des contributeurs d'OSM n'ont donné lieu à aucune crise existentielle. Dans le projet de produire toujours plus de données géographiques ouvertes de qualité, la possibilité de s'ouvrir à des formes de collaboration entre contributeurs bénévoles et professionnels semble même s'être installée comme une évidence, voire une nécessité. Ainsi la conférence *State of the Map* de 2017 a-t-elle hébergé une session intitulée « Professionnels d'OSM, unissons-nous ! » au cours de laquelle ont notamment été discutés les principes qui devaient permettre d'intégrer les deux formes de travail¹⁸. Cette ouverture aux enjeux techniques et organisationnels que soulèvent les difficultés et les coûts de la production et de la maintenance des données géographiques destinées à être ouvertes souligne les limites d'une problématisation de la *volunteered geography information*¹⁹ (Elwood, Goodchild, Sui, 2012) comme domaine à part et montre la nécessité de se départir de toute tentation de distinguer trop strictement données citoyennes et données professionnelles. Dans la même veine que les initiatives qui ont émergé du côté de la gestion des crises humanitaires, le projet initié par le partenariat entre OSM France et Transilien invite à prêter attention aux formes de coopération qui s'inventent dans la fabrique même des données. En s'attachant à faire des données un commun, sans nier le travail que leur génération et leur maintenance nécessitent, ces formes de coproduction font émerger des modèles alternatifs pour la gouvernance numérique des territoires, où la main mise des grands acteurs industriels du secteur n'a rien d'inexorable.

¹⁸ Trois principes ont été retenus : respect de la licence de données, celui des procédures de production des données et celui des manières d'interagir avec les contributeurs faisant des erreurs.

¹⁹ Terme stabilisé dans le monde académique anglophone, que l'on pourrait traduire par « - information géographique produite par des bénévoles ».

Références

- Baker K. S., Karasti H., 2018, *Data Care and Its Politics: Designing for Local Collective Data Management as a Neglected Thing*, in: Huybrechts L. Teli M, Light A. et al., eds, *Proceedings of the 15th Participatory Design Conference. Full Papers*, vol. 1, New York, ACM Digital Library. Accès : <https://doi.org/10.1145/3210586.3210587>.
- Bates J., Lin Y.-W., Goodale P., 2016, « Data Journeys: Capturing the Socio-Material Constitution of Data Objects and Flows », *Big Data & Society*, 3 (2). Accès : <https://doi.org/10.1177/2053951716654502>.
- Bowker G. C., Star S. L., 1999, *Sorting Things Out. Classification and Its Consequences*. Cambridge, MIT Press.
- Burns R., Meek D., 2015, « The Politics of Knowledge Production in the Geoweb », *ACME. An International Journal for Critical Geographies*, 14 (3), pp. 786-790. Accès: <https://acme-journal.org/index.php/acme/article/view/1233>.
- Crampton J., Krygier J., 2005, « An Introduction to Critical Cartography », *ACME. An International Journal for Critical Geographies*, 4 (1), pp. 11-33. Accès <https://acme-journal.org/index.php/acme/article/view/723>.
- Denis J., 2018, *Le Travail invisible des données. Éléments pour une sociologie des infrastructures scripturales*, Paris, Presses des Mines.
- Denis J., Goëta S., 2017, « Rawification and the Careful Generation of Open Government Data », *Social studies of science*, 47 (5), pp. 604-629.
- Denis J., Pontille D., 2013, « Une infrastructure évasive. Aménagements cyclables et troubles de la description dans OpenStreetMap », *Réseaux. Communication, technologie, société*, 178-179, pp. 92-125.
- Denis J., Pontille D., 2015, « Material Ordering and the Care of Things », *Science, Technology, & Human Values*, 40 (3), pp. 338-367.
- Dodge M., Kitchin R., 2013, « Crowdsourced Cartography: Mapping Experience and Knowledge », *Environment and Planning A: Economy and Space*, 45 (1), pp. 19-36.
- Edwards P., Mayernik M. S., Batcheller A., Bowker G., Borgman C., 2011, « Science Friction: Data, Metadata, and Collaboration », *Social Studies of Science*, 41 (5), pp. 667-690.
- Elwood S., Goodchild M. F., Sui D. Z., 2012, « Researching Volunteered Geographic Information: Spatial Data, Geographic Research, and New Social Practice », *Annals of the Association of American Geographers*, 102 (3), pp. 571-590.
- Goodchild M. F., 2007, « Citizens as Sensors: The World of Volunteered Geography », *GeoJournal. Spatially Integrated Social Sciences and Humanities*, 69 (4), pp. 211-221.
- Graham S., Thrift N., 2007, « Out of Order: Understanding Repair and Maintenance », *Theory, Culture & Society*, 24 (3), pp. 1-25.
- Haklay M., 2013, « Neogeography and the delusion of democratisation », *Environment and Planning A: Economy and Space*, 45 (1), pp. 55-69.
- Jackson S. J., 2014, « Rethinking Repair », pp. 221-240, in: Gillespie T., Boczkowski P. J., Foot K. A., eds, *Media Technologies. Essays on Communication, Materiality, and Society*, Cambridge, MIT Press.
- Kitchin R., Dodge M. , 2007, « Rethinking Maps », *Progress in Human Geography*, 31 (3), pp. 331-344.
- Kitchin R., Gleeson J., Dodge M., 2012, « Unfolding Mapping Practices: A New Epistemology for Cartography », *Transactions of the Institute of British Geographers*, 38 (3), pp. 480-496.
- Leszczynski A., 2012, « Situating the Geoweb in Political Economy », *Progress in Human Geography*, 36 (1), pp. 72-89.

- Lin Y.-W., 2011, « A Qualitative Enquiry into OpenStreetMap Making », *New Review of Hypermedia and Multimedia*, 17 (1), pp. 53-71.
- Marquet C., 2016, « Faire du *smartphone* un instrument de la relation de service ? », *Réseaux. Communication, technologie, société*, 200, pp. 145-177.
- Marquet C., à paraître, « Rendre accessible par les données : cartographie amateur et politiques d'accessibilité des transports », in : Popescu C., Rapegno N., eds, *Géographies du handicap*, La Plaine Saint-Denis, Éd. des Maisons des sciences de l'homme associées.
- Palen L., Soden R., Anderson T. J., Barrenechea M., 2015, « Success & Scale in a Data-Producing Organization: The Socio-Technical Evolution of OpenStreetMap in Response to Humanitarian Events », pp. 4113-4122, in: *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, New York, New York, ACM Digital Library. Accès : <https://doi.org/10.1145/2702123.2702294>.
- Picon A., 2018, *Smart Cities. Théorie et critique d'un idéal auto-réalisateur*, Paris, Éd. B2.
- Plantin J.-C., 2011, « The Map is the Debate: radiation Webmapping and public involvement during the Fukushima issue », in: *Proceedings from Symposium on the dynamic of Internet and the Society*, Oxford.
- Plantin J.-C., 2018, « Google Maps as Cartographic Infrastructure: From Participatory Mapmaking to Database Maintenance », *International Journal of Communication*, 12, pp. 489-506.
- Quest C., « Don't be Evil... Until », *Medium*, 6 mai. Accès : <https://medium.com/@cq94/dont-be-evil-until-95f2e8dfaaad>. Consulté le 13/06/2019.
- Ribes D., Finholt T. A., 2009, « The Long Now of Technology Infrastructure: Articulating Tensions in Development », *Journal of the Association for Information Systems*, 10 (5), pp. 375-398.
- Rigollier L., 2016, *Des données dans la ville : quelles intelligences pour la smart city ? Vers une « culture des données » au sein des collectivités ?*, mémoire de master, École d'urbanisme de Paris.
- Schuurman N., 2009, « The New Brave New World: Geography, GIS, and the Emergence of Ubiquitous Mapping and Data », *Environment and Planning D: Society and Space*, 27 (4), pp. 571-572.
- Soden R., Palen L., 2014, « From Crowdsourced Mapping to Community Mapping: The Post-earthquake Work of OpenStreetMap Haiti », pp. 311-444, in: Rossitto C., Cioffi L., Martin L., Conein, B., eds, *COOP 2014 - Proceedings of the 11th International 311 Conference on the Design of Cooperative Systems*, Lausanne, Springer.
- Smith D. E., 2005, *Institutional Ethnography. A Sociology For People*, New York, AltaMira Press.
- Star S. L., Griesemer J. R., 1989, « Institutional Ecology, "Translations" and Boundary Objects: Amateurs and Professionals in Berkeley's Museum of Vertebrate Zoology, 1907-39 », *Social Studies of Science*, 19 (3), pp. 387-420.
- Star S. L., 2010, « This is Not a Boundary Object: Reflections on the Origin of a Concept », *Science, Technology, & Human Values*, 35 (5), pp. 601-617.
- Townsend A. M., 2014, *Smart Cities. Big Data, Civic Hackers, and the Quest for a New Utopia*, New York, W. W. Norton.
- Turner A., 2006, *Introduction to Neogeography*, Sebastopol, O'Reilly.
- Zook M. A., Graham M., 2007, « Mapping DigiPlace: Geocoded Internet Data and the Representation of Place », *Environment and Planning B: Planning and Design*, 34 (2), pp. 466-483.