

Constitution d'une base de données nationale des arrêts de transport collectif

Laurent CHEVEREAU, chargé de projet « Information multimodale » - CERTU (CEREMA)

Patrick GENDRE, ingénieur d'études - CETE Méditerranée (CEREMA)

Yves BONIN, chef du service AGIL (Appui Géomatique et Infrastructures Linéaires), responsable du PCI « Géomatique appliquée aux infrastructures de transport » - CETE Méditerranée (CEREMA)

Au 1^{er} janvier 2014, les 8 CETE, le Certu, le Cetmef et le Sétra fusionnent pour donner naissance au Cerema

En France, la connaissance géographique de l'offre de transport collectif est aujourd'hui très limitée et disparate d'un réseau à un autre. Il en résulte des difficultés pour informer l'utilisateur à large échelle et pour faire des analyses statistiques territoriales.

S'inspirant des travaux menés en région Nord-Pas de Calais, le Certu a lancé en 2013 un travail de constitution progressive d'une base de données nationale des arrêts de Transport Collectif (ATC) urbains et départementaux. Cette base ATC vise à répondre aux besoins d'études des services du ministère du développement durable (Medde) : services centraux tels que le SOeS - Service de l'Observation et des Statistiques - mais aussi le Cerema et les services déconcentrés régionaux et départementaux.

L'historique du projet

En 2007, le Conseil National de l'Information Statistique (CNIS) a rédigé un rapport établissant un « *Inventaire des besoins d'information sur les transports* ». En 2012, un bilan « 5 ans après » a permis de mettre en lumière les recommandations qui n'ont pas encore été suivies d'effets. Parmi celles-ci, la recommandation n°27 proposait de « *Construire un système numérisé sur les réseaux de transport en commun urbains en y incluant des informations sur l'accessibilité physique et temporelle des points de desserte* ». Le Certu a été chargé de ce projet et s'est entouré d'un groupe de pilotage, regroupant la DGITM, l'AFIMB, le SOeS, le Cete Méditerranée, et le Cete Nord-Picardie. Le Gart s'est ensuite associé au projet, en co-signant – avec le SOeS et le Certu – un courrier adressé aux autorités organisatrices de transport (AOT) urbaines et départementales en septembre 2013, les invitant à participer au projet.

Les objectifs du projet

L'objectif principal est donc de constituer une base nationale d'arrêts de transport collectif (base ATC). Cette base visera à être la plus exhaustive possible sur le territoire français. Pour les réseaux de TC qui intégreront le projet, elle comportera la liste des arrêts géolocalisés, avec une description de chacun d'eux par un ensemble d'attributs dont le critère d'accessibilité physique aux arrêts si l'information existe.

Lorsque les données horaires seront disponibles, des indicateurs seront calculés afin d'éclairer l'accessibilité temporelle de l'arrêt : amplitude de desserte, nombre de passages par jour, nombre de passages à l'heure de pointe ...

Par ailleurs, au-delà de ces indicateurs demandés par le CNIS, d'autres seront calculés, sous réserve de disponibilité des données :

- la population desservie par chaque arrêt (à moins de 300m pour les bus urbains)
- la population desservie par chaque ligne (à moins de 300m d'un des arrêts de la ligne)
- la vitesse moyenne théorique par tronçon

Une fois ces indicateurs unitaires calculés, des indicateurs agrégés seront établis, par exemple par commune, par périmètre de transport urbain (PTU) ou par département.

Les étapes du projet

Les données de l'offre de transport sont du ressort des autorités organisatrices de transport (AOT). Celles-ci ont donc été sollicitées pour participer au projet en fournissant leurs données. De la collecte des données à la diffusion de la base, le projet se compose de 4 étapes :

1. **Collecter les données** : Cette étape s'entend nécessairement avec l'accord des AOT. Quand elles existent, les données Open Data sont privilégiées (avec un accord de fait, détaillé dans une licence) ; Dans le cas contraire, et si les données sont intégrées dans un SIM de taille départementale ou régionale, le responsable du SIM est sollicité pour effectuer un export des données depuis l'outil de gestion du SIM. Si le SIM ne peut être un canal de diffusion, il est proposé aux AOT d'envoyer directement leurs données au Certu.
2. **Homogénéiser les données** : Les données sont reçues sous des formes et formats divers. Les sources Open Data et SIM permettent généralement d'avoir un format standard (GTFS) ou normalisé (Neptune). Dans les autres cas, les formats sont divers (.shp, .tab, .kml, csv, .xls) avec des systèmes de projection potentiellement différents. La description des arrêts et lignes est également variable, à la fois dans le nombre de champs renseignés, dans le choix des critères renseignés (et les intitulés correspondants) et dans la qualité des attributs renseignés. Afin d'homogénéiser ces données, les arrêts sont ainsi mis en cohérence avec le modèle d'arrêt partagé issu des travaux engagés par l'AFIMB.
3. **Donner une valeur ajoutée aux données** : Une fois les données converties dans un format et modèle unique, il devient possible de calculer des indicateurs homogènes. Ces indicateurs constituent des premiers éléments d'analyse sur l'offre de transport elle-même, ou sur le croisement de l'offre et du territoire (population, services, emplois), à un niveau fin (un arrêt de bus) ou agrégé (un PTU par exemple).
4. **Diffuser données et indicateurs** : Une plate-forme de diffusion a ensuite été mise au point s'appuyant sur l'outil Cartélie, application de cartographie dynamique sur internet du Medde. Elle permet la visualisation et la consultation des données et indicateurs, avec un accès restreint pour certaines données. La base complète sera disponible pour les services du ministère du développement durable. Il est important de préciser qu'aucune donnée des AOT ne sera diffusée sans son accord explicite.

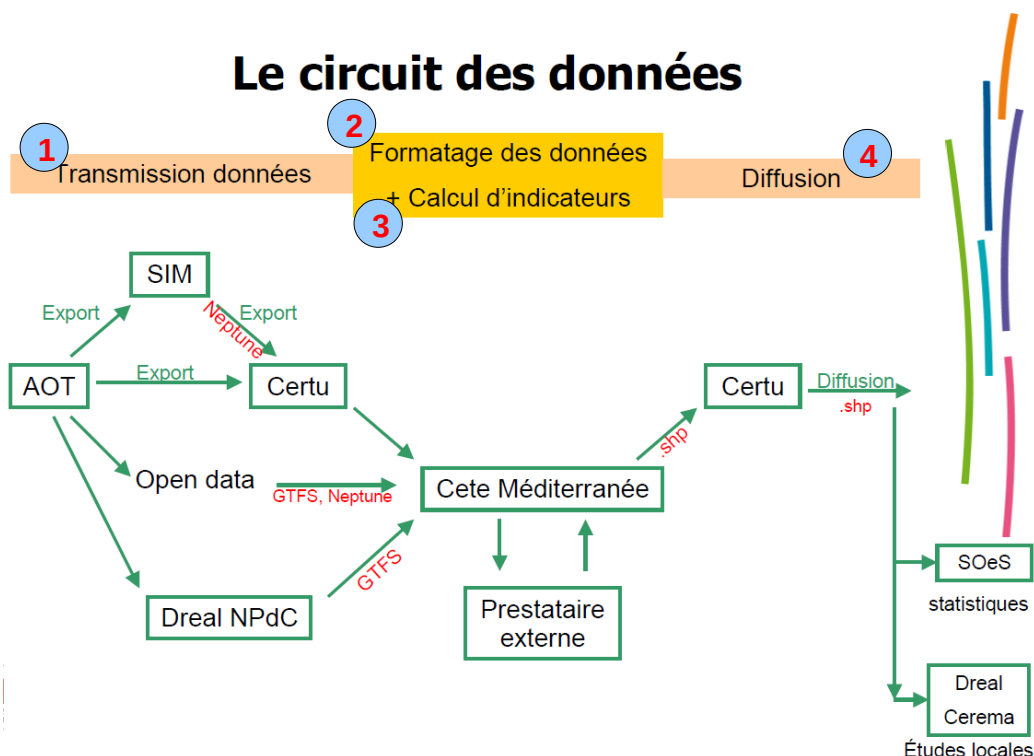


Figure n°1 : Le circuit des données d'offre de transport dans le projet de base ATC

Le calendrier du projet

Le projet vise à initier une base d'arrêts, puis à la pérenniser en la rendant la plus exhaustive possible.

Pour la version 2013 de la base ATC (son initialisation), la collecte des données (étape 1) s'est arrêtée fin novembre 2013 ; L'homogénéisation des données (étape 2) a débuté en septembre 2013 et se poursuivra jusqu'en début 2014. Le calcul d'indicateurs a commencé en octobre 2013 et se poursuivra au premier trimestre 2014. La mise au point de la plate-forme de diffusion a commencé en novembre 2013 ; Son ouverture avec diffusion des données et indicateurs est prévue au second trimestre 2014.

Pour la suite, le projet prévoit une actualisation a minima annuelle, période suffisante pour les analyses envisagées.

Les perspectives de disponibilité des données

Pour la version 2013 de la base ATC (son initialisation), nous n'avons pas proposé de convention spécifique aux AOT. Les données exploitées seront donc :

- les données Open Data (30 réseaux TC dont 26 avec lignes et 24 avec horaires)
- les 14 réseaux TC de la région Nord Pas-de-Calais (base déjà constituée)
- les 5 réseaux TC des Alpes Maritimes (convention existante)
- quelques AOT supplémentaires

La base 2013 comportera donc environ 50 réseaux de TC, avec plus de 40 réseaux avec information exploitable sur les horaires (format GTFS ou Neptune).

Pour la version 2014 de la base, des conventions sont proposées aux AOT. Suite au courrier envoyé en septembre 2013, plus de 60 AOT se sont manifestées pour participer au projet. Elles pourront ainsi intégrer le projet en 2014 avec mise en place de convention sur les usages et diffusions qui seront faits des données. On peut donc s'attendre à une centaine de réseaux TC dans la base 2014, dont une moitié avec des données horaires exploitables.

Les possibilités offertes par la base ATC

La base ATC vise à permettre des études statistiques et des analyses territoriales. Elle viendra ainsi compléter les différentes sources de données statistiques sur le transport en France. Elle pourrait à terme être utile pour rendre plus complète la Base Permanente des Équipements (BPE) de l'Insee. Du côté des analyses territoriales, ces données sur l'offre TC peuvent permettre aux Cete de réaliser des études locales (pour les Dreal par exemple) :

- mesurer l'impact d'un nouveau projet de transport, en intégrant l'offre TC dans le modèle ;
- mesurer la cohérence entre l'offre et la répartition des activités (population, emplois, activités) comme l'illustre la figure n°2 ;
- mesurer les disparités territoriales ;
- aider à analyser une offre intermodale et la cohérence entre les différents réseaux concernés ;
- aider à mettre en avant des zones plus congestionnées ou moins bien desservies ...

- Fréquence de passage, un dimanche, à Rennes (bus uniquement):

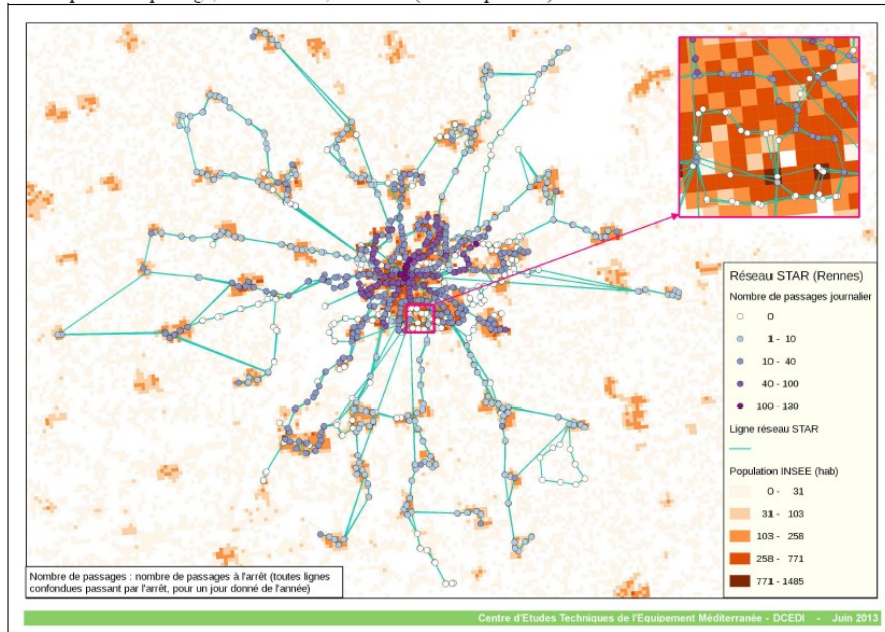


Figure n°2 : Exemple d'indicateur calculé sur la base des données Open Data de Rennes

Toutes ces études contribuent aux grandes politiques publiques, à savoir essentiellement :

- tendre vers une égalité des territoires
- répondre à l'impératif de transition énergétique

Chacune de ses politiques pouvant être traitée par la dimension sociale (vers un égal accès aux équipements et services), par la dimension économique (attractivité économique des territoires ; accessibilité à l'emploi ; vulnérabilité économique des ménages face au coût de l'énergie) ou par la dimension environnementale (développement de l'intermodalité, et des transports collectifs pour favoriser le report modal).

Conformément à l'élargissement des bassins de vie, ces problématiques doivent généralement être analysées sur des territoires larges, couvrant plusieurs réseaux TC. La base ATC viendra donc compléter l'offre routière et ferroviaire déjà bien connue, sur des territoires recouvrant plusieurs réseaux.

Conclusion

Le projet de constitution d'une base nationale d'arrêts de transport collectif (base ATC) est en phase d'initialisation. Ce projet est à la fois ambitieux et impliquant de nombreux acteurs. Il nécessite donc un travail technique important et un pilotage de projet qui tienne compte des attentes et des contraintes de chacun. Ces éléments impliquent un projet conséquent, mais les réelles difficultés du projet sont les questions de propriété et de diffusion des données.

En effet, si certaines AOT ont fait la démarche d'ouvrir leurs données, bon nombre d'entre elles préfèrent maîtriser entièrement l'information, et par conséquent s'opposent à toute diffusion de leurs données. Ceci vaut surtout pour les données sur les horaires et leur utilisation pour informer directement l'utilisateur. Il semble en effet que ces données, généralement constituées dans le cadre d'un service public industriel et commercial (SPIC), ne relèvent pas d'une obligation de mise à disposition. Les réflexions en cours sur l'Open Data vont néanmoins dans le sens d'une ouverture. Celle-ci participe à la mise en place de la directive européenne Inspire qui cite explicitement les arrêts TC comme devant faire l'objet d'une diffusion sans réserve auprès des acteurs publics comme privés.

Ce mouvement pourrait donc inciter les AOT à intégrer le projet par la suite. C'est tout l'enjeu de cette base ATC : fédérer les AOT, par soucis d'ouvrir leurs données, ou par intérêt pour elles-mêmes, grâce à l'uniformisation des données et au calcul d'indicateurs que propose le projet. A terme, cette base ATC, construite par la puissance publique, permettrait de repositionner celle-ci face aux acteurs privés déjà positionnés. Le fait de pouvoir disposer de ces données construites par la puissance publique semble effectivement nécessaire, mais cela n'empêche pas de réfléchir aux complémentarités à trouver avec les acteurs privés et les services qu'ils peuvent offrir.