

Atelier Temps Réel (20 septembre 2018) : améliorer l'information voyageur sur les transports collectifs partout en France

Compte-rendu, version du 12 octobre 2018

Résumé des discussions

L'ouverture des données temps réel est un véritable défi technique à relever pour simplifier au maximum la diffusion d'une information fiable et actualisée sur les déplacements de tous les usagers. Avec des centaines d'opérateurs de transport aux capacités financières et techniques différenciées, l'ouverture ne se fera pas rapidement sans un accompagnement national. Les producteurs de données craignent le coût de mise aux normes des données et le coût humain et technique de mise à disposition (charge sur les serveurs locaux) malgré les avantages de l'ouverture des données (opportunité pour diffuser l'offre de transport public et attirer de nouveaux usagers dans les véhicules). La possibilité de demander une compensation aux réutilisateurs est parfois considérée par les producteurs comme une usine à gaz étant donnée la complexité de la mise en place d'un système de paiement fonctionnel.

Pour lever ces freins, une solution proposée est que le Point d'Accès National assure :

- la conversion des données dans le format normé (SIRI) selon le profil français en cours de définition, ce qui n'empêche pas une diffusion GTFS RT ;
- la gestion de la charge en termes de requêtes pour les producteurs qui le souhaitent, afin de réduire la complexité des relations entre producteurs et réutilisateurs.

A très court terme, l'équipe du Point d'Accès National va dès lors expérimenter durant les prochains mois le référencement de données sur le PAN en format GTFS RT ou en format propriétaire afin d'en exposer dans un premier temps une API SIRI Lite donnant les prochains passages à un arrêt. Les producteurs de données exposant déjà du format SIRI (ou SIRI Lite) pourront être référencés sur la plateforme dès lors qu'un travail d'harmonisation juridique aura été effectué.

Rappel sur la plateforme transport.data.gouv.fr

transport.data.gouv.fr est le Point d'Accès National français, interface numérique exigée pour chaque État membre par le [règlement européen 2017/1926](#). La plateforme vise à référencer l'ensemble des données (statiques et dynamiques) publiées en open data et relatives à l'information voyageur, pour tous les modes de transport (transport collectif régulier, transport à la demande, vélos, automobiles, etc.) Son objectif est de faciliter la réutilisation de ces données pour favoriser le déploiement de services de mobilité innovants au bénéfice des usagers (planification de trajets, disponibilités des moyens de transport à proximité, etc.)

Afin de construire cette plateforme, le Ministère chargé des transports s'est associé à beta.gouv.fr pour appliquer la [méthode Startup d'État](#) : transport.data.gouv.fr est développée de manière incrémentale par une équipe autonome, qui travaille au plus près des besoins des utilisateurs et qui est guidée par le seul objectif **de rendre les déplacements des usagers plus faciles grâce à une information voyageur fiable et disponible sur tous les supports**.

Calendrier de déploiement

Un premier chantier a été lancée en septembre 2017 : les données statiques décrivant les réseaux urbains, interurbains et nationaux de transports réguliers de personnes (lignes de bus, de métro, de tram, de train, etc.) La plateforme, opérationnelle dès janvier 2018, se repose sur data.gouv.fr et a d'ores et déjà permis à 49 autorités organisatrices de mobilité et 3 régions de publier les données statiques décrivant les réseaux de leur ressort territorial, au format GTFS et en licence ODbL. En parallèle, 8 réutilisateurs (Here Technologies, Mappy, Handisco, Blablacar, MobiGIS, Transit, Urban Pulse, Kisio Digital) consomment les données disponibles sur transport.data.gouv.fr pour améliorer le service rendu aux usagers ou déployer leur solution dans de nouveaux territoires plus facilement. Ce chantier se poursuivra jusqu'à atteindre l'exhaustivité du référencement.

Un deuxième chantier s'ouvre à l'occasion de cet atelier du 20 septembre : la mise à disposition des données dynamiques (temps réel) via la plateforme transport.data.gouv.fr. Ce compte-rendu reprend les échanges ayant eu lieu à cette occasion et détaille les prochaines étapes des travaux des équipes transport.data.gouv.fr.

L'atelier ayant eu lieu le 20 septembre a permis d'échanger sur les manières de lever les freins à la mise à disposition des données, pour permettre une ouverture utile et peu coûteuse. C'est un véritable défi technique, juridique et opérationnel qui s'ouvre pour diffuser au maximum une information voyageur temps réel fiable et améliorer les déplacements des usagers partout en France.

Nombre de participants : 48

Producteurs représentés : Agglomération de Chartres ; Auvergne-Rhône-Alpes ; Bordeaux Métropole ; Grand Lyon ; IDF Mobilités ; Isère ; Keolis Besançon ; Nantes Métropole ; RATP ; SNCF ; Tisséo ; Transdev

Réutilisateurs représentés : Apple Maps ; Here Technologies ; Kisio Digital ; Mappy ; Trainline ; Nextérité

Autres participants : AF83 ; CEREMA ; CRI Paris ; Etalab ; FNTV ; Vertone

Format des données

Aujourd'hui, selon le CEREMA (base [Passim](#)), une vingtaine de producteurs de données ont mis à disposition des API temps réel. Le plus souvent, ce sont des développements propriétaires, non standardisés.

Deux formats de diffusion de données temps réel pourraient être généralisés aujourd'hui, bien qu'aucun des deux ne s'impose totalement en raison de difficultés :

- GTFS RT :
 - o standard de fait (90% des données temps réel à l'international selon Here Technologies, lisible par la plupart des gros réutilisateurs)
 - o expose l'état du réseau dans son ensemble
 - o difficulté majeure : harmonisation des identifiants entre GTFS statique et GTFS RT (Tisséo et Besançon témoignent de leur volonté de faire du GTFS RT mais connaissent des problèmes de référentiels : les producteurs voudraient faire du GTFS RT mais l'info qu'ils ont ne le permet pas (info identifiants qui n'existent pas : bus qui arrivent en renfort et qui ne sont pas dans le système d'information : toute la chaîne d'informations est fausse).
- SIRI, norme européenne :
 - o norme européenne : obligation selon le règlement (UE) 2017/1926 (repris dans le projet de loi de la LOM)
 - o expose différents types de services :
 - donne des informations sur un arrêt précis (*stop-monitoring*)
 - transmet des alertes générales sur un réseau (*general-message*)
 - donne des informations sur l'ensemble d'une ligne (*estimated-timetable*)
 - donne la localisation des véhicules en temps réel (*vehicle-monitoring*)
 - etc.
 - o difficulté : même dans ses déclinaisons locales beaucoup de champs optionnels, ce qui n'est pas pratique et rend l'interopérabilité difficile. Un profil national plus léger (fondé sur **SIRI Lite, une variante de SIRI avec requêtes REST**), inspiré par l'Île-de-France, est en cours de définition mais présente de nombreux modules optionnels.

→ Des outils libres ici : <http://www.chouette.mobi/irys/developpeurs/>

- IRYS = application client / serveur
- Client de validation SIRI
- Adaptateur SIRI-Lite

SIRI et GTFS RT sont idéaux pour des communications machine machine, et donc pour alimenter les calculateurs d'itinéraires. SIRI Lite convient davantage à des éditeurs d'application qui n'ont pas leur propre infrastructure de gestion de temps réel et permet d'exhiber les prochains passages à un arrêt spécifique facilement.

Plus d'informations sur les formats : [Guide édité par l'AFIMB](#) et plus largement : <http://www.normes-donnees-tc.org/category/donnees-temps-reel/>

Certains réutilisateurs nous ont fait part de leurs préférences en termes de format :

Here Technologies : GTFS RT

Google Maps : GTFS RT

Citymapper : GTFS RT

Transit : GTFS RT

Handisco : non supporté pour l'instant

Conclusion : le groupe n'est pas parvenu à trancher. Ni l'un, ni l'autre des formats ne semble être une solution facile – même si les deux présentent chacun des avantages :

- **GTFS RT** : permet la meilleure diffusion de l'information voyageur sur les applications grand public mais nécessite un travail de fond par chaque AOM.
- **SIRI** : machine-machine en attendant mieux mais problèmes d'interopérabilité et complexité.
- **SIRI Lite** (et profil France en cours de définition) : bon compromis même si ne permet pas une diffusion massive de l'info voyageur.

Craintes et espoirs des participants

L'objectif commun de tous les participants est d'améliorer l'information voyageur au bénéfice de tous les usagers.

Pendant, en termes de mise à disposition des données temps réel, plusieurs craintes subsistent :

- Le Grand Lyon souhaite maîtriser les données pour que les réutilisateurs ne délivrent pas des services contraires à l'orientation des politiques publiques.
- D'autres producteurs (Transdev) insistent sur la nécessité d'identifier les réutilisateurs afin de contrôler les flux et de pouvoir limiter l'accès aux données en cas d'utilisation abusive
- Inquiet du spam (requetage massif) il faudra couper l'accès

Sur la question de la licence, le groupe s'interroge sur la pertinence d'une licence ODbL : a-t-elle du sens, sachant que le partage à l'identique de données périmées n'a pas grand intérêt pour les usagers des transports collectifs ?

Il est proposé de lever les freins à une diffusion maximale des données, tout en gardant en tête que le Point d'accès national se fait de manière itérative : tout ne sera pas traité d'un coup.

Conclusion : il serait préférable de proposer une licence ouverte sur les données temps réel ouvertes. Des conditions d'utilisation harmonisées pourront être proposées aux autorités organisatrices pilotes dans les prochaines semaines.

Accompagnement technique par le PAN

Avant de parvenir à l'étape ultime d'un flux unique vers l'ensemble des données temps réel de la France (beaucoup trop complexe), il est proposé par le groupe de travail de se contenter dans un premier temps **des prochains passages à un arrêt**.

Concrètement, une API en Siri Lite pourrait être constituée avec les autorités organisatrices qui jouent le jeu. Cette API serait coupée verticalement par réseau (par exemple : appel à Toulouse puis appel à Rennes, etc). transport.data.gouv.fr est le point final de l'API qui supporterait la charge pour ne pas la faire reposer sur les AO.

Quelques propositions de travaux pour l'équipe de transport.data.gouv.fr :

- Lexique commun : mettre un vocabulaire à plat (exemples : transport commandé, transport avancé, à partir de quelle unité de temps parle-t-on de temps réel, qu'est-ce qu'on entend par temps réel) ;
- Travail sur la question des référentiels d'arrêts, cruciale pour avancer sur nos sujets.

Modèle économique

Il est évident que la diffusion des données relatives à l'information voyageur représentera un coût, à la fois pour les autorités organisatrices et les opérateurs de transport. Cependant, ce coût est également un investissement (conquête de nouveaux clients) au même titre que les coûts engagés sur les guides d'horaires imprimés en version papier.

2 types de coûts pouvant freiner l'ouverture sont identifiés :

1. Investissement initial pour la création de la donnée (une crainte majeure est le surcoût que pourraient demander les éditeurs de SAEIV) ;
2. Coût de mise à disposition et maintenance (il ne s'agit pas uniquement des frais de serveur car ces aspects techniques peuvent se résoudre. C'est aussi du temps homme pour créer, déposer, maintenir.)

Il est prévu dans le cadre du projet de LOM de permettre aux producteurs de données de demander une compensation financière aux réutilisateurs à hauteur du coût de mise à disposition, mais les systèmes de paiement sont souvent difficiles et coûteux à mettre en place, ce qui rend cette solution compliquée dans les faits.

Le groupe a travaillé sur des pistes de solutions pour pallier ces deux freins à l'ouverture des données temps réel.

Transport.data.gouv.fr pourrait accompagner les autorités organisatrices ne pouvant pas supporter le surcoût associé à la génération de fichiers aux normes en récupérant le plus simplement possible des fichiers décrivant l'état de l'ensemble du réseau correspondant ("venez comme vous êtes"), à une fréquence qui reste à définir. Cette solution permet de baisser significativement les frais de serveur (vs requêtes à un arrêt). La plateforme nationale agirait ensuite comme un convertisseur vers des formats faciles à exploiter (GTFS RT) ou normés (SIRI).

Exemple (à expérimenter) : le Grand Besançon met à disposition un fichier décrivant l'ensemble de l'état de son réseau mis à jour toutes les minutes. L'équipe transport.data.gouv.fr le consomme pour exposer une interface en SIRI Lite (avec une première étape concernant simplement le service "stop monitoring").

→ la plupart des AO n'auront même pas à se poser la question du coût d'exploitation. Mais pendant les renouvellements de marché, il ne faudra pas omettre la mention de GTFS RT et/ou SIRI.

Si jamais il se pose la question du coût d'exploitation au dessus d'un niveau significatif, le seuil devra probablement permettre de récupérer l'état du réseau au moins toutes les minutes de manière gratuite, en une requête uniquement (max de 1440 requêtes par jour pour la plupart des réseaux).

Plan d'action :

3 acteurs se sont manifestés à la fin de la réunion pour proposer de faire partie des producteurs pilotes sur le sujet :

- le département de l'Isère (dont 4 autorités organisatrices) et la métropole de Grenoble ;
- le Grand Besançon ;
- le groupement Réunionir.

Il sera mené trois types d'actions par l'équipe transport.data.gouv.fr pour chaque catégorie :

(1) **les producteurs qui produisent déjà du temps réel ouvert** : travail d'accompagnement et d'outillage pour communiquer dans un langage commun avec la plateforme transport.data.gouv.fr. Les autorités organisatrices diffusant déjà du temps réel seront contactés prochainement dans une liste de diffusion spécifique pour commencer le travail d'harmonisation juridique et technique.

- a) API SIRI Lite : sera référencée sur transport.data.gouv.fr
- b) GTFS RT : sera référencé + exhibé en SIRI Lite sur le PAN

c) Format propriétaire : solution ad hoc pour exhiber du SIRI Lite sur le PAN

(2) **les producteurs qui n'ont pas d'API de temps réel ouvert mais qui ont un SAE** (ou outils tels Pysae et Joule (Zenbus)) : nous travaillerons à très court terme avec certaines autorités pilotes comme le Grand Besançon pour comprendre de quoi ils ont besoin pour ouvrir, leurs difficultés, le bon format, etc.

(3) **les producteurs qui réfléchissent à un nouveau SAE** : ils pourront, lorsque l'équipe transport.data.gouv.fr aura acquis l'expérience nécessaire au fil de l'eau, être accompagnés pour les aider à préciser les choses dans leur cahier des charges et à leur faciliter la mise à disposition.

Nous faisons un appel aux réutilisateurs pour expérimenter avec certains pilotes la mise à disposition de l'API SIRI Lite. Les réutilisateurs intéressés par les prochains passages à un arrêt peuvent donc prendre contact avec l'équipe du PAN dès à présent.

Nous invitons ceux qui le souhaitent à [rejoindre le Slack de transport.data.gouv.fr](#) pour échanger plus facilement sur ces sujets.

Prochain RDV le 15 janvier 2019 à 10h à Paris, pour faire le point.

Contribuez à la construction du PAN :

L'équipe transport.data.gouv.fr cherche à comprendre les besoins de tous les experts du sujet, ceux qui ont travaillé sur la question sur le terrain depuis de nombreuses années : si vous êtes dans ce cas et avez du temps à nous consacrer pour nous aider à construire la démarche la plus pertinente, n'hésitez pas à prendre contact.