



projet WWW.chouette.MOBI

Maintenance, accompagnement et recette de logiciels pour les échanges de données multimodales

GTFS et Neptune

Comparaison des formats d'échange de données d'offre TC

1 Introduction

Cette note présente les points clés pour comparer les 2 formats ouverts les plus utilisés pour échanger des données d'offre théorique de TC en France : GTFS et Neptune.

On propose, pour chaque format, de décrire sommairement : le contexte et les objectifs du format, les outils associés, puis de comparer d'une part les données qui peuvent être échangées avec GTFS et avec Neptune, et enfin d'indiquer la manière dont les 2 formats sont gérés dans le logiciel Chouette.

2 GTFS

(voir <http://www.normes-donnees-tc.org/spip.php?article64>)

GTFS a été conçu pour permettre à des développeurs pas forcément experts du domaine métier TC d'alimenter des calculs d'itinéraires TC et multimodaux dans un contexte open data, ou de réaliser simplement tout type d'application utilisant des données TC.

Les données décrivant l'offre d'un réseau TC sont fournies sous la forme d'un dossier compressé (.ZIP) contenant plusieurs fichiers texte séparé par des virgules (très facile à traiter), dont 6 sont obligatoires et 7 sont optionnels.

Bien que soutenu par Google qui l'utilise pour son service Google Transit, GTFS est un standard ouvert qui évolue selon une procédure de décision rendue dans un groupe de discussion :

<https://developers.google.com/transit/gtfs/changes>

Le standard est disponible ici :

<https://developers.google.com/transit/gtfs/reference>

GTFS est bien documenté en anglais, mais il existe aussi des présentations en français, par exemple :

http://data.ratp.fr/?eID=ics_od_datastoredownload&file=88



3 Neptune

Neptune est une norme française (AFNOR NFP 99506) publiée en 2009. C'est l'évolution d'un profil mis en place pour assurer les échanges entre le STIF et la RATP, issu des résultats du projet européen TRIDENT au début des années 2000, donc bien antérieur à GTFS.

Neptune s'appuie sur XML pour les échanges de données : par rapport à un simple format texte tabulé, l'avantage est que l'on peut décrire plus finement et surtout de manière moins ambiguë les relations entre les éléments de données et les contraintes qui s'y appliquent, donc en principe mieux valider la conformité d'un jeu de données. L'inconvénient est que les données sont plus complexes et plus difficiles à traiter.

Neptune a été développé dans le cadre du groupe de travail GT7 de la commission de normalisation CN03, et donc a priori par et pour des professionnels du métier des TC.

Compte tenu des procédures de normalisation, en pratique Neptune n'a pas évolué depuis 2009. Toutefois il mériterait d'être complété sur plusieurs points de détail, ce qu'il est prévu de faire en 2014.

Le devenir « naturel » de Neptune est d'évoluer vers un profil de la spécification technique NETEX, appelée à devenir la future norme européenne en matière de données de TC, ce qui permettra de décrire une offre TC avec des éléments de données complètement cohérents avec un ensemble plus large couvrant d'autres applications du TC : gestion des lieux d'arrêt (IFOPT), information temps réel (SIRI), gestion d'équipements, tarification... dans une logique de système d'information nécessitant des échanges entre applications, et entre organisations.

4 Les données

Plusieurs difficultés apparaissent lorsque l'on veut comparer 2 formats d'échange :

- d'une part **les éléments de données échangeables** (colonnes des tableaux GTFS, data elements et attributs des schémas XML de Neptune) **doivent être mis en correspondance** : ils ne correspondent pas forcément à la même notion. Grâce à l'existence de la norme européenne Transmodel, modèle de données conceptuel de référence pour les transports publics, il est possible en principe de comparer dans le détail les données échangées avec l'un ou l'autre des formats d'échange. Pour cela, la solution consisterait à « mapper » d'abord chaque format d'échange avec Transmodel, puis à faire le rapprochement entre les deux ; cet exercice a été effectué en 2008¹ pour comparer GTFS et TransXchange, l'équivalent britannique de Neptune. Il n'a pas été effectué dans le détail pour Neptune, même si des tableaux de correspondance ont été ébauchés.

- d'autre part, les données échangeables en principe ne sont pas forcément en pratique faciles à obtenir, à créer, à gérer ou à traiter. S'il n'existe pas d'outil libre ou gratuit associé à ces données, cela en réserve l'utilisation à des acteurs du domaine ou ayant des moyens importants. Exemple

1 <http://www.rtig.org.uk/web/Portals/0/090206%20-%20PTIC%20-%20Transmodel%20GTFS%20-%20v4%20%28Issued%29.pdf>



typique, l'information sur l'accessibilité des TC aux handicapés (PMR) peut être en principe présente dans un profil d'échange, encore faut-il pouvoir obtenir, renseigner et gérer cette information.

Voici les principaux points de comparaison entre termes de données décrites :

dans GTFS et pas dans Neptune	tracés de lignes
dans Neptune et pas dans GTFS	groupes de lignes, équipements, accès, versions des objets, itinéraires complexes (fourches, boucles...)

Par ailleurs, GTFS et Neptune ne décrivent pas tous les types de services de transport : par exemple, ils ne permettent de décrire le transport à la demande zonal.

5 Gestion dans Chouette

Le logiciel libre Chouette a été développé à l'initiative du ministère français chargé des transports afin de promouvoir la norme Neptune, et plus largement l'interopérabilité des systèmes d'information TC et multimodale. Il présente le gros intérêt de savoir convertir des données GTFS en Neptune et réciproquement, et de s'inscrire dans le passage vers NETEX.

Le détail de la manière dont GTFS est géré est décrit ici

http://www.normes-donnees-tc.org/IMG/pdf/NT-A130130-Mapping_Chouette_GTFS-20130422.pdf

Il faut savoir néanmoins que Chouette ne gère pas 100 % des données Neptune, ni 100 % des données GTFS. Les limitations actuelles sont précisées dans le tableau ci-dessous ; elles sont susceptibles d'être résolues progressivement dans le cadre du marché de maintenance et d'accompagnement de ce logiciel, si des demandes en ce sens sont exprimées.

limites de Chouette vis-à-vis de	(cf. http://forum.chouette.mobi/index.php?board=3.0)
GTFS	ne traite pas l'encodage UTF-8 « BOM » Byte-Order-Mark (doit être édité en amont dans par exemple notepad++); ne gère pas plusieurs données optionnelles de GTFS : shapes (tracés) ; sait importer en revanche les fréquences (les traduit en courses individuelles)
Neptune	aucun traitement des Facilities (équipements) ; aucun traitement des horaires en fréquences ; les données d'accessibilité peuvent être importées, exportées, validées, mais pas gérées (pas d'écran pour voir et éditer ces données)



6 En résumé

Le périmètre de données est assez similaire entre les deux formats d'échange, en principe un peu plus large pour Neptune mais en pratique assez peu, compte tenu des données existantes.

GTFS est simple d'utilisation, s'appuie sur une communauté d'utilisateurs ouverte et très dynamique, et sur de nombreux logiciels libres (à condition de parler l'anglais).

Neptune est plus complexe, mais s'inscrit dans la perspective d'une norme européenne couvrant assez bien tout le champ des applications métier TC. Il est plutôt destiné à des échanges entre systèmes d'information et applications métier, implémenté par des acteurs du métier TC, mais peut être utilisé pour des applications destinées à des utilisateurs finaux, par exemple dans un cadre open data, en s'appuyant sur le logiciel libre Chouette pour valider des données, les modifier à la marge ou saisir quelques lignes, et les convertir d'un format à l'autre, ou pour bénéficier du simulateur temps réel IRYS diffusé avec Chouette.