



Ministère de l'Écologie, du Développement durable et de l'Énergie

Agence française pour l'Information multimodale
et la Billettique (AFIMB)

***L'open data dans le domaine du transport :
analyse des premières initiatives
et recommandations***

Avril 2013

Sommaire

Introduction	3
1. Méthodologie de l'étude.....	4
1.1. Vue d'ensemble.....	4
1.2. Description de la méthodologie.....	5
1.2.1. Analyse de la présentation et du contenu des données	5
1.2.2. Analyse des aspects contractuels et des réutilisations.....	5
2. Analyse et recommandations	6
2.1. Objectifs affichés dans les initiatives open data	6
2.1.1. Analyse des objectifs.....	6
2.1.2. Recommandations	6
2.2. Structure des plates-formes open data.....	7
2.2.1. Synthèse de l'analyse.....	7
2.2.2. Recommandations	8
2.3. Contenu des données publiées	11
2.3.1. Synthèse de l'analyse.....	11
2.3.2. Recommandations	13
2.4. Aspects contractuels et licence de réutilisation	16
2.4.1. Synthèse de l'analyse.....	16
2.4.2. Recommandations	18
2.5. Réutilisations identifiées pour des services à l'utilisateur final.....	19
2.5.1. Synthèse de l'analyse.....	19
2.5.2. Recommandations	21
3. Conclusions.....	22
3.1.1. Synthèse de l'analyse.....	22
3.1.2. Synthèse des recommandations et perspectives	23
Annexe 1	27
Annexe 2	50

Introduction

En France, des collectivités territoriales, en majorité des autorités organisatrices des transports (AOT), se sont lancées dans différentes initiatives open data marquées par une grande diversité des données mises à disposition.

La présente étude porte sur le domaine des données liées au transport public, apportant une analyse des démarches open data existantes et proposant aux promoteurs de ces initiatives des pistes d'amélioration pour dynamiser et surtout pérenniser les réutilisations dans des services au public.

L'étude est réalisée par Kasia Bourée (KBIC) & Gildas Baudez (Carte Blanche Conseil).

Nos observations se fondent sur une analyse approfondie des plates-formes, des données et des réutilisations de cinq sites :

1. Rennes,
2. Bordeaux et TransGironde,
3. Toulouse,
4. Nantes,
5. Loire-Atlantique.

Nous avons procédé pour chacune d'elles à une analyse descriptive "de base" sous forme de « fiches de lecture », résumant des objectifs initiaux ou affichés et présentant les principales caractéristiques des données transport mises à disposition. Nous avons cherché à compléter ces fiches par des informations sur les réutilisations. La synthèse de cette analyse est présentée au début de chaque chapitre de ce rapport.

Nous présentons des recommandations à discuter avec les acteurs de terrain : en premier lieu, aux producteurs de données, les initiateurs et gestionnaires des plates-formes de mise à disposition. Nos recommandations pourraient également être discutées avec les principaux réutilisateurs de ces données. Ces propositions visent une publication améliorée et harmonisée de données TC, en précisant le rôle qui pourrait être dévolu à l'outil *open source* CHOUETTE.

En conclusion, nous proposons une synthèse des recommandations qui peuvent être considérées comme prioritaires et dont l'application présenterait une amélioration de l'état actuel, notamment en ce qui concerne la facilité de réutilisation des données publiées. D'autre part, nous indiquons les différents travaux et actions visant à améliorer les outils techniques qui pourraient être mis à disposition des fournisseurs des données, administrateurs des plates-formes et réutilisateurs.

Les analyses des plates-formes existantes, ainsi qu'un examen plus détaillé de quelques aspects complémentaires, ont donc permis l'élaboration de recommandations concernant :

- la structure du site (2.1 et 2.2), recommandations destinées principalement aux gestionnaires des plates-formes,
- les données : formats et contenus (2.3), à l'intention principalement des fournisseurs des données,
- les aspects contractuels (2.4)
- les réutilisations et l'adéquation avec les objectifs initiaux (2.5)

Ce rapport est suivi de 2 annexes :

Annexe 1 : Présentation des fiches open data

Annexe 2 : Résultats des tests sur les jeux de données

1. Méthodologie de l'étude

1.1. Vue d'ensemble

Les sites suivants ont été recensés au premier semestre 2012 :

Sites OPEN DATA/outil consultation	URL
Nantes SEMITAN	www.data.nantes.fr
Rennes STAR	www.data.rennes-metropole.fr http://data.keolis-rennes.com/fr/les-donnees/donnees-telechargeable.html
Toulouse TISSEO	http://data.grandtoulouse.fr
Bordeaux CUB	http://data.lacub.fr
Pays de la Loire	www.data.gouv.fr
Paris RATP	http://www.ratp.fr/opendata/
Arles ACCM	http://www.opendata.regionpaca.fr
Montpellier APITAM	http://modulaweb.fr/apitam/
CG44 Loire Atlantique	www.data.loire-atlantique.fr
CG33 TransGironde	http://www.datalocale.fr/
CG41 Loir et Cher	http://www.data.gouv.fr/DataSet/30383156?xtmc=transport+en+commun&xctcr=1
CG71 Saône et Loire	http://www.opendata71.fr/
FR SNCF Gares	http://data.sncf.com
TFL Transport for London	http://www.tfl.gov.uk/businessandpartners/syndication/16492.aspx

Les deux derniers ont été identifiés au cours de l'étude comme apportant un éclairage complémentaire sur les potentialités de la mise à disposition et de réutilisation de données transport public.

Un échantillon de 5 sites a été exploré plus en détail :

1. Bordeaux et TransGironde,
2. Toulouse,
3. Nantes,
4. Loire-Atlantique,
5. Rennes.

Cet échantillon a été choisi principalement en fonction du format des données disponibles au moment de l'import. En effet, il a été jugé utile de pouvoir étudier le contenu des imports en utilisant le format de référence NEPTUNE.

Le tableau ci-dessous présente la liste des plates-formes open data ainsi que les formats de données disponibles au début de cette étude (automne 2012).

Sites OPEN DATA / outil consultation		Chouette		SIG			SIG	SIG
	xml GFTS	xml Neptune	xml Trident	csv	kml-kmz	daisy	xls	shp
1 Nantes SEMITAN	X >	X		trad.WGS84				
2 Rennes STAR	X >			trad.WGS84				
3 Toulouse TISSEO			x					
4 Bordeaux TUB				x	x			x LAMBERT93
5 Paris RATP				x WGS84				
6 Arles ACCM								x LAMBERT93
7 Montpellier API TAM				trad.WGS84				x
8 CG44 Loire Atlantique	X >		x	trad.WGS84				
9 CG33 TransGironde	x >	x	x					x LAMBERT93
10 CG41 Loir et Cher								x LAMBERT93
11 CG71 Saône et Loire				x		x	x	

1.2. Description de la méthodologie

L'analyse des plates-formes open data a été menée sous différents angles en prenant en compte les aspects suivants :

- présentation des données,
- contenu des données,
- contrats et licences,
- réutilisations des données.

1.2.1. Analyse de la présentation et du contenu des données

Une grille d'analyse des plates-formes a été élaborée. Elle comporte principalement les rubriques suivantes :

A. Analyse des informations générales :

1. nom de l'AOT ou du réseau, adresse du site
2. objectifs initiaux (résumé des textes originaux)
3. type du descriptif de données (détaillé - comportant une liste de propriétés, définition précise ou général - nom de la donnée accompagné éventuellement d'un descriptif succinct)
4. type de données transports publiées
5. format d'échange
6. période de validité
7. fréquence de mise à jour
8. contributeurs : acteurs, services ayant fourni la donnée
9. réutilisation

B. Analyse du contenu des données :

Il s'agit de la présentation du résultat de l'import des données dans l'application CHOUETTE et du résumé des résultats des tests effectués sur les données.

Les données utilisant les formats TRIDENT et NEPTUNE ont été importées et consultées en priorité sous l'outil CHOUETTE. Les données au format GTFS (Rennes) pour les horaires ont pu être importées dans un deuxième temps lorsque l'import GTFS sous CHOUETTE a été rendu disponible.

Pour ce qui est des données SHP, le CSV ou le KML, la vérification s'est limitée à la vérification de la possibilité de l'import et de la visualisation (en utilisant, pour ce qui est des points d'arrêt, un SIG acceptant ces formats).

1.2.2. Analyse des aspects contractuels et des réutilisations

Quelques éléments de l'analyse des aspects contractuels ou des réutilisations, comme l'indication de la licence, figurent dans les différentes fiches présentant les plates-formes open data. L'analyse des différentes licences a été effectuée dans un deuxième temps. Pour ce qui est des réutilisations des données, elles ne sont généralement pas listées directement sur les plates-formes, mais peuvent être déduites à partir de l'analyse des forums de discussion, par exemple.

2. Analyse et recommandations

2.1. Objectifs affichés dans les initiatives open data

2.1.1. Analyse des objectifs

On peut classer les objectifs des dix plates-formes publiques de données transport en deux rubriques principales:

- d'une part des thèmes généraux, partagés avec d'autres domaines d'application (économie, état civil, construction, urbanisme, etc.)
- d'autre part, des thèmes spécifiques qui sont le produit du rôle spécifique des collectivités territoriales dans la gestion de la mobilité, car elles s'appuient très largement sur les moyens d'opérateurs à qui elles délèguent l'essentiel de la production des services et des données, voire des systèmes d'information. Plus particulièrement on exprime l'exigence de mise à disposition des informations complètes, fiables et mises à jour pour faciliter l'utilisation des infrastructures et des services mis à la disposition des usagers dans un territoire donné.

Les objectifs généraux exprimés sont essentiellement les suivants :

- renforcer la transparence :
 - donner un meilleur accès à l'information au public
 - libérer les données publiques
- favoriser l'innovation :
 - offrir de nouveaux services
 - s'inscrire dans la modernité
- promouvoir les savoir-faire locaux

Parmi les objectifs spécifiques au transport, on peut citer le souhait de :

- permettre à de nouveaux acteurs d'intégrer les données de tout service de transport public (au sens d'accessibles à tous) pour faciliter la comparaison des services sur des critères objectifs en offrant des services innovants, avec par exemple, un service d'information (vraiment) multimodal ;
- mettre à disposition du grand public ces informations sous des formes qui l'encouragent au transfert modal et à laisser sa voiture au garage ;
- abattre les obstacles empêchant des initiatives de toute nature et de toute envergure de voir le jour, de se développer techniquement et économiquement.

2.1.2. Recommandations

A la différence de nombreux secteurs pour lesquels sont produites des données publiques, le domaine des transports met en jeu une combinaison complexe d'acteurs qui mérite un traitement à part en terme d'objectifs comme de droits. Les recommandations qui suivent s'orientent vers tous ces acteurs, usagers compris.

F 01 Messages au public, messages aux réutilisateurs¹

Nous recommandons que les objectifs spécifiques au secteur transports/déplacements (où les collectivités publiques et locales sont mues non seulement par des motifs d'intérêt général mais aussi par des intérêts spécifiques de gestion des infrastructures et d'utilisation équilibrée des différents modes) soient précisés de façon distincte. Il convient en effet de faire partager les objectifs d'intérêt général non seulement par le public le plus large mais aussi par les acteurs de la mobilité. Les objectifs spécifiques s'adresseront prioritairement aux opérateurs privés et publics des différentes offres de transport mises à la disposition des usagers, mais aussi aux prestataires de service d'information voyageurs.

F 02 Propositions d'échanges pouvant aller jusqu'à un partenariat informel

Vis-à-vis des porteurs de service d'information aux voyageurs, il est souhaitable d'avoir une attitude ouverte afin de connaître puis aider à résoudre les difficultés rencontrées dans cette nouvelle approche des médias et de la communication avec l'utilisateur. Ceci impose un service d'assistance et donc des moyens humains pour traiter ces retours, identifier les acteurs et gérer le dialogue avec eux.

¹ Voir à ce sujet les travaux de l'Urban ITS expert group (<http://www.predim.org/spip.php?article4377>)

2.2. Structure des plates-formes open data

2.2.1. Synthèse de l'analyse

L'échantillon choisi des plates-formes permettant d'accéder à des données ouvertes permet un constat simple : il n'existe pas de structure type de site.

Certaines caractéristiques des sites sont communes, en particulier celles qui concernent les objectifs et la démarche suivie. Les sites y indiquent l'origine de la démarche et, en quelque sorte, une justification.

En préambule de cette analyse, à part le fait qu'il s'agit d'un échantillon restreint, il faut souligner que le temps imparti à l'analyse des sites a été (volontairement) limité. En effet, l'exercice consistait à vérifier si, en un temps limité, il était possible de se repérer facilement dans un site donné et d'en extraire l'information recherchée.

Les types d'information recherchés ont été rassemblés dans une grille d'analyse pour chaque cas pris en compte. Il faut se référer aux différentes grilles renseignées au fur et à mesure de l'étude. Cet état de l'art (réalisé au 4^e trimestre 2012) peut être résumé de la façon suivante :

- Type de données publiées

La grille d'analyse des sites, établie tout au début du processus d'analyse, exprime les caractéristiques que l'on pouvait rechercher en premier lieu :

- les domaines des données publiées : voirie, tourisme, transport, pistes cyclables, hôtels, centres équestres, etc.
- les différents types de données transport : lignes, horaires, etc.

Il s'avère que ces caractéristiques n'ont souvent pas été retrouvées de façon immédiate, car ces informations se trouvent au sein de descriptifs, ne sont pas mis en évidence, ni affichées systématiquement sous forme d'une liste ou d'un menu sur les premiers écrans.

- Détail des données transport

Le détail des données relatives au transport public se limite bien souvent à une énumération : par exemple « lignes ; horaires », sans en donner le descriptif, en particulier une définition ou des propriétés.

Dans chaque cas, le format utilisé est indiqué. Cependant, l'accès au descriptif de ce format n'est pas immédiat dans tous les cas (p. ex. on indique « TRIDENT », sans référence à un document ou d'URL où une documentation peut être retrouvée).

Dans le meilleur des cas des fiches sont produites, indiquant les quelques caractéristiques pour chaque donnée. Cependant, une définition claire des données publiées n'a pas été trouvée : s'agit-il des arrêts commerciaux ou d'arrêts physiques ? Des pôles d'échange ? Comment les correspondances sont-elles traitées ? Ces questions restent généralement sans réponse immédiate.

- Type de services de transport

On rencontre des informations sur les types de service disponibles sur un territoire (p. ex. modes de transport disponibles, existence de transport à la demande, etc.). Cependant, il n'est pas toujours explicite si les données publiées concernent l'ensemble de ces services.

- Emprise géographique

La question de l'emprise géographique ainsi que de l'exhaustivité des données publiées n'a pas de réponse précise, immédiatement compréhensible. On pourrait s'attendre à ce que chaque plate-forme donne une représentation cartographique du territoire desservi ou une énumération simple des communes, villes et lignes. Par ailleurs, un doute subsiste quant à la complétude des données sur un territoire donné.

- Validité et fréquence de mise à jour

La période de validité des données est souvent indiquée. Cependant la périodicité de mise-à-jour est souvent imprécise (« ponctuelle », « lorsque nécessaire », « variable »).

- Collecte/contributeurs

Seuls certains sites indiquent l'origine des données, ainsi que les contributeurs. Quelquefois, l'indication porte sur un accord entre deux acteurs sans indication précise sur les rôles ou responsabilités de chacun.

- Import des données : outils et formats

Généralement les formats d'import des données sont indiqués. Cependant, il semble manquer l'information sur les outils de visualisation et de publication des données importées. De tels outils ne

sont pas obligatoires mais hautement souhaitables, notamment pour les réutilisateurs qui ne disposent pas de grandes ressources de développement et souhaitent se concentrer sur leurs applications de services à destination des utilisateurs finaux.

2.2.2. Recommandations

Les recommandations qui suivent s'orientent principalement vers les gestionnaires des plates-formes « Open Data ». Elles sont d'ordre général (G) et expriment le fait qu'il serait utile de retrouver facilement dans la communication associée à la mise à disposition de données et de services les rubriques suivantes :

G 01 Démarche suivie

Fournir un descriptif succinct de la démarche en séparant les aspects techniques des démarches politiques et économiques, les objectifs généraux des objectifs spécifiques au domaine du transport.

Ce descriptif visera en particulier à donner une meilleure visibilité sur les données qui sont et seront fournies : pendant combien de temps, avec quelle fréquence, par quels acteurs (si cela est prévisible et connu), les éventuels services associés (webservices, API...).

G 02 Liste des données publiées

Donner une liste des données publiées avec leur définition courte, choisie parmi celles utilisées dans Neptune. Exemple : pour un arrêt, mettre zone d'arrêt ou arrêt physique, en disant si la position est celle du poteau ou de l'arrêt du véhicule, ou de l'accès à la zone d'arrêt.

G 03 Métadonnées

Prévoir une méthode de publication systématique des métadonnées (c'est-à-dire l'information sur les données publiées), par exemple, pour chaque donnée, une/des fiches concernant les métadonnées associées. Les recommandations G 06, G 08, G 09 explicitent les métadonnées à renseigner pour l'ensemble de données. Pour les données géographiques, on pourra s'appuyer sur les recommandations de la directive INSPIRE (<http://www.developpement-durable.gouv.fr/La-directive-europeenne-Inspire-de.html>) qui indiquent l'ensemble des éléments à renseigner (donnée, description, identifiant, thèmes ISO concernés, thème INSPIRE, extension géographique, rectangle de l'emprise en degrés, référence temporelle, etc). Un exemple de fiche à renseigner est donné ci-dessous.

	A	B	C	D
1	 géoportail le portail des territoires & des citoyens			
2	Saisie des métadonnées d'une donnée géographique			
3				
4	Description			
5				
6	*Titre de la donnée ^{INSPIRE} :		<input type="text"/>	
7				
8	*Description de la donnée ^{INSPIRE} :		<input type="text"/>	
9				
10				
11				
12	*Identifiant de la donnée ^{INSPIRE}		<input type="text"/>	
13				
14				
15	*Thèmes ISO concernés ^{INSPIRE} (1) :		<input type="text"/>	
16			(2) :	
17			(3) :	
18				
19				
20	*Thème INSPIRE ^{INSPIRE} (1) :		<input type="text"/>	
21				
22				
23	*Extension géographique ^{INSPIRE} : Rectangle de l'emprise des données en degrés décimaux (par défaut, France métropolitaine)			
24	Lat N / S	<input type="text" value="48.2623666939523"/>	<input type="text" value="47.9406449463391"/>	
25	Long O / E	<input type="text" value="-1.97550663912549"/>	<input type="text" value="-1.4362366672985"/>	
26				
27	*Référence temporelle ^{INSPIRE}			
28				
29	Date de la donnée (création):		<input type="text" value="j/mm/aaaa"/>	

Figure 1: Saisie métadonnées INSPIRE dans le fichier Excel du géoportail de l'IGN téléchargeable <http://admin.geocatologue.fr/geocatadmin/LogonTileForward.do?requestedURL=/geocatadmin/admin/;jsessionid=12DAEDA78B44A6D973F0BC40724CF692>

G 04 Type de services de transport

Enumérer les modes ou services pris en compte avec l'indication de leur exhaustivité (indication des modes ou services existants mais non pris en compte).

G 05 Emprise géographique

Fournir une représentation cartographique ou une énumération des communes, villes et lignes avec l'information sur la complétude des données sur le territoire.

G 06 Validité et fréquence de mise-à-jour,

Indiquer la période de validité et la fréquence de mise à jour de chaque type de donnée.

G 07 Calendriers

Fournir un descriptif de la gestion des calendriers du réseau (en vue d'une mutualisation des données sur une même plate-forme).

G 08 Collecte/contributeurs

Indiquer les rôles ou responsabilités des différents acteurs pour chaque type de donnée, en particulier pour ce qui est de la collecte, de la mise à jour et de la mise en ligne.

G 09 Qualité/précision

Fournir des indications sur la précision : indiquer, pour les données de localisation, la précision des relevés (p. ex. la méthode utilisée). Fournir des informations sur la fiabilité des données. Préciser, par exemple, si les données sont utilisées (ou non) en production, formuler un avertissement pour une utilisation à but commercial, indiquer un taux d'erreur (si connu).

G 10 Import des données : outils et formats

Fournir l'information sur les formats et outils de visualisation ou de publication des données publiées (indiquer l'existence de l'outil open source CHOUETTE et de la norme française NEPTUNE, documentés sur www.chouette.mobi, de la norme GTFS sur <https://developers.google.com/transit/gtfs/reference?hl=fr>).

G 11 Réutilisation/condition de réutilisation

Fournir l'information sur les réutilisations observées, connues ou proposées ainsi qu'une explicitation simple des conditions contractuelles de réutilisation (renvoi sur les textes et licences).

G 12 Forum de discussion

Etablir un forum de discussion (en complément de contacts courriel du ou des responsables de la plate-forme) permettant aux utilisateurs de confronter leurs expériences ou difficultés, d'exprimer le besoin d'enrichissement du site ou la proposition de publication de nouveaux types de données, etc. Cette recommandation facilite l'adhésion de nouveaux acteurs

G 13 Outils de publication de nouvelles données

Indiquer où se trouvent (ou bien mettre à disposition) des logiciels afin d'encourager le processus de publication de nouvelles données et l'adhésion de nouveaux acteurs. L'information sur des outils de saisie, d'archivage ou de publication et de visualisation de certaines données (par exemple la localisation de mise à disposition des vélos en libre service, emplacement et/ou caractéristiques des équipements aux arrêts, caractéristiques d'accessibilité des arrêts etc.) peuvent inciter les utilisateurs à participer à l'enrichissement des bases de données existantes. Par ailleurs il serait préférable que les outils de collecte de données facilitent une collecte harmonisée en suggérant un ensemble minimal et commun de propriétés des données à collecter. Les gestionnaires des plates-formes, par exemple, sont bien placés pour juger de la pertinence de tels outils en vue d'une intégration des nouvelles données dans une base communautaire existante en garantissant la complétude et la cohérence de l'ensemble. L'outil open source CHOUETTE peut être cité dans ce cadre : il facilite une saisie harmonisée et le stockage des données relatives à l'offre théorique (et la visualisation des données géolocalisées).

2.3. Contenu des données publiées

2.3.1. Synthèse de l'analyse

Les formats

Les formats majoritairement utilisés sur un échantillon élargi à une dizaine de cas sont les suivants :

- XML TRIDENT-NEPTUNE et GTFS
- CSV et XLS
- Kml SHP

Le tableau en 1.1 présente les formats utilisés dans les cas pris en compte.

Un compte rendu des imports est présenté dans l'annexe 2 (« Résultats des tests de fichiers téléchargés Open Data »).

Les volumes des données importées sont très variés, comme le montre le tableau ci- dessous :

	LILA	CG33	Tisséo complet
Réseaux	1	1	1
Transporteurs	0	1	2
Lignes	10	91	107
Courses	208	2127	30659
Calendriers	13	456	30503
Correspondances	722	0	2077
Arrêts	669	2413	5298
points d'embarquement	417	2186	16
quais	0	0	3505
arrêts commerciaux	252	227	1776
pôles d'échange	0	0	1
ITL	0	0	0
points d'accès	0	0	0

Les types de données

Les données de transport public publiées concernent généralement les données suivantes :

- arrêts localisés,
- lignes,
- correspondances,
- horaires théoriques (courses).

Les données sur les systèmes de vélos en libre service (VLS) concernent surtout :

- la localisation des stations,
- la disponibilité des vélos aux stations,
- la disponibilité des emplacements libres aux stations.

De façon générale, les flux de données actualisées (services supprimés ou arrêts déplacés) ou en temps réel (retards...) sont rares. Leur gestion est en effet plus complexe et nécessite en effet plus de ressources. En outre, la mise à disposition au travers de web-services se révèle, dans de nombreux cas, mieux adaptée.

Ces données donnent une vision de l'offre. Plus elles sont complètes, mieux elles informent les utilisateurs susceptibles d'accepter un transfert modal. Elles restent insuffisantes pour une demande occasionnelle qui a en plus besoin de nombreux détails sur les conditions pratiques et tarifaires des services proposés.

Le contenu des données

Les données importées ont été analysées, d'une part de façon aléatoire pour quelques types de données (les arrêts, les correspondances), d'autre part, dans certains cas, au moyen des fonctions de test de l'outil CHOUETTE. La vérification s'est concentrée sur quelques données représentatives dans les différents cas. Les grilles d'analyse indiquent les erreurs ou insuffisances constatées. On peut les résumer de la façon suivante :

- Positionnement

Les données concernant les « arrêts » s'avèrent quelquefois inexactes quant à leur position géographique. La visualisation sur la carte de l'outil CHOUETTE des arrêts choisis de façon aléatoire permet ce constat (p.ex. arrêt placé sur un cours d'eau). La vérification effectuée avec CHOUETTE ne permet pas la détection d'une telle erreur, tout au plus un « avertissement » quant aux coordonnées et lorsque l'erreur est très grossière.

- Identifiants : unicité et existence

Les imports des données peuvent être réussis mais incomplets. Une consultation détaillée du compte rendu de l'import doit être effectuée.

Par exemple, dans un cas, des erreurs du fichier de validation existent mais ne sont pas fatales. L'import est donc effectué, le rapport de l'import indique des avertissements. Dans ce cas, 50 arrêts n'ont pu être importés, car les identifiants de ces arrêts ne sont pas uniques et ne sont donc pas enregistrés dans la base.

Par ailleurs, dans un cas, l'import n'a pas eu lieu à cause de l'absence des identifiants.

- Sémantique des types de données publiées

La problématique concerne ici la sémantique non clarifiée des données publiées. Les données publiées ne sont pas systématiquement définies. Par définition on entend « un discours qui dit ce qu'est une chose ou ce que signifie un mot » (Wikipedia). Ainsi les utilisateurs peuvent avoir une définition propre d'une donnée publiée ou utiliser une définition normalisée (ex. définie au travers de la norme française NEPTUNE). C'est le cas, par exemple, pour les arrêts. Il s'agit dans la plupart des cas des arrêts commerciaux, des points d'embarquement ou quais. Cependant, aucune information sur ces données n'est présente. Par exemple, dans le cas des points d'embarquement, on n'indique pas s'il s'agit du positionnement des véhicules ou du poteau, abri. Les arrêts commerciaux semblent être constitués de deux points d'arrêts (sans doute des points ou zones d'embarquement). Cela n'est pas clairement précisé.

Dans un cas, on annonce la prise en compte des accès. Cependant, l'import qualifie les arrêts importés soit comme des points ou zones d'embarquement (quais) ou arrêts commerciaux.

Les correspondances importées semblent concerner uniquement les correspondances au sein d'un même arrêt commercial. Il n'est pas indiqué si cette liste est exhaustive. En effet on pourrait supposer que des correspondances entre les arrêts commerciaux puissent exister.

Les données concernant les pôles d'échange ou accès ne sont pas renseignées.

- Qualité et précision pour la géolocalisation ?

La précision de la position n'est pas indiquée.

- Doublons

Des doublons ont été détectés (tests aléatoires) : deux correspondances au sein d'un même arrêt commercial avec des temps de correspondance différents, pour un même type d'utilisateur. Cet exemple est d'ailleurs difficile à interpréter car il représente soit une contradiction avec la notion normalisée de correspondance, soit le résultat d'une définition spécifique de correspondance, soit un « doublon ».

- Incohérences

Certaines incohérences passent facilement inaperçues. C'est le cas, par exemple, des séquences d'arrêts « aller » qui sont supposées ne pas avoir de « retour ». Or, la séquence d'arrêts « retour » existe bien. Il s'agit ici d'une incohérence sémantique non détectable par des routines CHOUETTE (vérification de liaison obligatoire de sens aller-retour si les deux existent). Toute incohérence détectée peut mettre en péril tout un jeu de données jugé non fiable.

2.3.2. Recommandations

Les recommandations qui suivent s'orientent principalement vers les fournisseurs des données publiées. Elles ont un caractère technique (précédé de la lettre T).

Elles sont de deux types :

- les recommandations notées T Mx et représentent un ensemble d'exigences « minimales » : leur respect permettra aux utilisateurs extérieurs, c'est-à-dire non producteurs des données, une réutilisation plus aisée des données publiées ;
- les recommandations notées T Nx représentent des exigences nécessaires pour :
 - la mutualisation aisée des données multi-source,
 - la généralisation de la réutilisation des données,
 - la généralisation des échanges entre systèmes et applications.

Elles sont basées sur l'utilisation des normes, c'est-à-dire l'application des spécifications ouvertes et publiques favorisant ainsi une publication harmonisée des données.

Recommandations techniques minimales relatives aux données publiées

T M1 Descriptif systématique de la signification de chaque donnée

(en particulier en vue de la mutualisation-agrégation).

Fournir une description systématique et détaillée des données : définition & ensemble de propriétés.

T M2 Vérification systématique du contenu

Procéder à une vérification des données publiées suivant une méthode d'analyse de la qualité et indiquer son type (p. ex. manuelle, automatique). Joindre le rapport d'analyse de la qualité.

T M3 Propriétés obligatoires des données

Il s'agit de donner une définition précise (signification) de chaque propriété, son caractère (obligatoire ou optionnel). En particulier, les identifiants doivent constituer une propriété obligatoire.

T M4 Formats

Il s'agit d'indiquer les formats des données publiées, définis le plus souvent dans les spécifications des formats d'échange (p. ex. TRIDENT, NEPTUNE, GTFS, CSV, SHP etc.). Indiquer clairement l'accès à la documentation de référence ainsi que le lien vers les outils (open source s'ils existent) permettant de manipuler ces formats de données.

Recommandations basées sur l'application des normes

Lorsque la mutualisation des données multisource est souhaitée, il est indispensable pour que le processus de mutualisation ait un sens, que les données à mutualiser possèdent une sémantique partagée par l'ensemble des acteurs. Il faut, par exemple, que l'information concernant un itinéraire, produite par un fournisseur soit bien comprise comme étant distincte de l'information concernant un parcours, que la donnée arrêt commercial, soit univoque et son lien avec un arrêt physique ou un accès soit clarifié. De telles descriptions de données, sont fournies par les modèles de données de référence européens Transmodel (EN12896) et IFOPT (EN28701). Ces modèles ont donné lieu à des spécifications des messages d'échange, en particulier des formats précis pour chaque caractéristique de ces données (par exemple, pour la caractéristique « nom » ou « localisation » de la donnée « arrêt »...). La norme française NEPTUNE permet, par exemple, d'échanger les données de l'offre théorique de service de transport public. Le standard européen SIRI spécifie les formats d'échange des données temps réel relatives à l'offre de transport.

L'idée maîtresse dans le développement des normes relatives aux données pour le transport public est de garder la cohérence des différentes spécifications, en particulier celles relatives à la sémantique (modèles de données) et aux formats des données (formats d'échange). Les recommandations encourageant l'utilisation des normes sont basées sur une logique similaire. En particulier, une même interprétation des données est favorisée - et donc leur correcte mutualisation et réutilisation - lorsque

l'ensemble des acteurs prendra en compte les modèles de données de référence et les spécifications de format de référence directement associées.

En France, pour ce qui est des données concernant l'offre théorique par exemple, l'outil open source CHOUETTE est basé sur une approche normalisée. Il permet de saisir et d'entreposer des données dans une base de données fortement imprégnée de la sémantique des modèles de données de référence normalisés (Transmodel et IFOPT), utilise un format d'échange normalisé, directement associé, connu sous le nom de NEPTUNE. CHOUETTE permet, outre la publication et l'import des données de l'offre (horaires et description du réseau), la visualisation des données sur un fond cartographique ainsi qu'un contrôle de la qualité des données. Des logiciels de vérification de conformité facilitent, lors des échanges, la vérification des données échangées.

Pour ce qui est des données temps réel relatives à l'offre, le standard d'échange SIRI, basé également sur les structures de données normalisées de référence et intégré à l'outil CHOUETTE, complète la série cohérente des normes pour la « saisie, entrepôt, publication, vérification, échange ».

Il faut mentionner dans ce contexte l'existence des convertisseurs de format, permettant d'importer, sous l'outil CHOUETTE, des données sous d'autres formats (GTFS...). Une correspondance entre les concepts manipulés par GTFS et NEPTUNE est établie. Cependant, les structures de données définies à travers GTFS sont bien plus rudimentaires que celles des normes européennes. Par exemple, la richesse des différentes notions relatives aux arrêts et leur structuration claire, prévue dans la norme européenne IFOPT permet de différencier les lieux d'arrêt mono- ou multimodaux, les zones d'accès (attente, embarquement, arrêt de véhicules), ainsi que les accès aux lieux d'arrêt. Ce n'est pas le cas de la spécification GTFS, où la notion de « stop » rassemble un certain nombre de caractéristiques des arrêts sans fournir de structuration systématique. Les structures décrites par la norme NEPTUNE permettent l'ajout aisé d'autres ensembles de données, par exemple relatifs aux équipements ou à l'accessibilité et cela de façon très naturelle, étant donné que l'ensemble est conçu dès le départ de façon modulaire. Ceci constitue un net avantage des normes européennes par rapport à la spécification GTFS : l'extension d'une base de données à d'autres données et l'intégration des nouvelles données est un processus naturel et prévu lors de la conception de ces normes. C'est l'une des raisons pour laquelle un utilisateur ayant publié des données de base relatives aux arrêts sur un territoire et voulant dans un deuxième temps ajouter et intégrer des caractéristiques relatives à l'accessibilité des arrêts sera avantagé lors de l'utilisation du format NEPTUNE.

On peut dire que l'utilisation généralisée d'un même format basé sur Transmodel/IFOPT que représente NEPTUNE favorisera

- l'agrégation des données provenant de sources différentes,
- leur réutilisation par des systèmes/utilisateurs variés,
- le rapprochement avec les données de l'offre temps réel,
- les extensions à d'autres sous-domaines (p.ex. définition de l'accessibilité, gestion des équipements, gestion des tarifs, établissement des statistiques relatives à l'exploitation).

Les recommandations qui suivent sont donc rédigées dans cette logique :

T N1 Descriptif systématique de la signification de chaque donnée

Indiquer les définitions des données ainsi que l'ensemble des propriétés retenues en stipulant leur conformité (ou non conformité) aux normes Transmodel/IFOPT et NEPTUNE/SIRI. Une référence à ces normes est donc attendue et recommandée.

T N2 Vérification systématique du contenu

Dans le cas où les données sont fournies dans un format normalisé, procéder à des vérifications de cohérence des informations fournies, en particulier au moyen des routines de vérification proposés par l'outil CHOUETTE relatives aux données de l'offre théorique (par exemple, pour constater la présence de l'ensemble des arrêts d'un parcours, la cohérence des temps de passage des courses successives, etc).

T N3 Propriétés obligatoires des données

Dans le cas où les données sont fournies dans un format normalisé, vérifier, en consultant la documentation des normes de référence (Transmodel/IFOPT & NEPTUNE), la présence des propriétés obligatoires : cela s'applique en particulier aux identifiants qui doivent être présents et uniques (au moins au sein de chaque ensemble des données publiées).

T N4 Formats

Pour les données relatives à l'offre théorique, l'utilisation d'un même format normalisé NEPTUNE est recommandée, permettant ainsi une définition exacte et différenciée des données. En complément, le format GTFS pour certaines utilisations peut s'avérer utile. Pour des échanges définis en XML, préciser s'il s'agit du standard TRIDENT, de la norme NEPTUNE ou autre. Indiquer clairement l'accès à la documentation de référence.

T N5 Identification unique d'objets fixes

Une identification unique et pérenne d'objets fixes provenant de plusieurs territoires (arrêts, parkings, POI, etc.) facilite la mutualisation des données les concernant. Les données ainsi mutualisées facilitent à leur tour la réutilisation par divers acteurs (applications, organismes), et par conséquent l'interopérabilité. Pour ce qui est des arrêts du transport public, une démarche normative est en cours, visant à définir des identifiants uniques, en application des normes de référence IFOPT/Transmodel. Il est recommandé d'adhérer pleinement à cette démarche basée sur le modèle d'arrêt partagé (en concertation avec les acteurs impliqués dans ce processus).

2.4. Aspects contractuels et licence de réutilisation

2.4.1. Synthèse de l'analyse

Dans les sites d'Open Data étudiés, les licences de réutilisation varient toutes mais légèrement, car elles s'inspirent toutes de 3 modèles :

- OdbL (celles adoptées par la Ville de Paris et le Conseil Général de la Gironde ont été adaptées de la [licence ODbL \(Open Database Licence\)](#) et de l' [Open Knowledge Foundation](#)),
- APIE (Bordeaux comme Montpellier ont opté pour les CGR de l'APIE¹, Rennes en a fait une adaptation),
- Etalab (qui est notamment compatible avec les licences « Open Government Licence » (OGL) du Royaume-Uni, « Creative Commons Attribution 2.0 » (CC-BY 2.0) de Creative Commons et « Open Data Commons Attribution » (ODC-BY) de l'Open Knowledge Foundation).

On aura compris qu'il n'y a pas un modèle commun de licence française.

Elles sont plus ou moins ouvertes mais toutes ces licences reposent sur 3 principes essentiels :

1. Les données sont mises à disposition de tiers identifiés qui doivent respecter des règles minimales, dont l'effet pratique est quasi nul en dehors de procédures de recours judiciaires. Dans la plupart des cas, ces tiers sont libres de réutiliser « l'information ». Ils ne sont même pas identifiés de manière certaines. Ils sont autorisés à :
 - Reproduire, copier, publier et transmettre « l'Information » ;
 - Diffuser et redistribuer « l'Information » ;
 - Adapter, modifier, extraire et transformer à partir de « l'Information », notamment pour créer des « Informations dérivées » ;
 - Exploiter « l'Information » à titre commercial, par exemple en la combinant avec d'autres « Informations », ou en l'incluant dans son propre produit ou application.
 - sous la seule réserve de mentionner la paternité de « l'Information » : sa source (a minima le nom du « Producteur ») et la date de sa dernière mise à jour.
2. Cette mise à disposition peut être payante. Dans certains sites open data analysés, cette possibilité est évoquée dans le cas d'utilisations commerciales des données et de revenus imputables à ces utilisations, mais elle ne semble pas appliquée. Comment effectivement procéder autrement que sur un mode déclaratif ?
Comme pour les données publiques d'autres domaines, la mise à disposition des données est, dans les cas étudiés, de fait gratuite.
3. La propriété intellectuelle de ces données reste aux « producteurs », mais cela ne présage pas de la propriété du produit des traitements qu'on réalise avec elles. Une nuance importante est introduite par la notion de base de données et base de données dérivées (qu'utilise par exemple la licence open data ODbL). Cette licence considère les données fournies comme un tout distinct des données que chaque réutilisateur est susceptible de lui ajouter et un tout différent des données produites par traitement des données qui lui ont été fournies.
Ces deux bases de données peuvent être intégrées à une base de données collaborative.
L'effet de cette intégration (faite par le réutilisateur) est que « la » base de données comprend les données initiales et les données ajoutées et que les droits d'utilisation sont étendus à toute la communauté de ceux qui ont signé la licence et qui utilisent tout ou partie des données.

On voit ainsi que les producteurs mettant à disposition des données ont de multiples possibilités de restreindre ou d'élargir les droits qu'ils ont concédés en y englobant ou non le produit des traitements et des ajouts faits aux données transmises dans le cadre du contrat de licence. De cette façon, ils se donnent un droit réel de contrôler la bonne utilisation qui est faite de leurs données et même, s'ils le souhaitent, de bénéficier à titre de réciprocité des enrichissements qui pourraient les intéresser, même si c'est, bien sûr, de façon non exclusive.

Le premier principe peut assurer une grande liberté dans la conception de ces réutilisations. Mentionner la source, outre le respect de ces obligations, confère une certaine crédibilité au service, même s'il ne s'agit en aucun cas d'une garantie réelle, car la manière dont est réutilisée la donnée est l'apanage du réutilisateur. La date de dernière mise à jour peut être un facteur de confusion pour l'utilisateur final car il dépend beaucoup du type d'information et du rythme normal de ces révisions (ex. un arrêt peut être provisoirement déplacé en dehors des périodes de mise à jour des horaires).

La gratuité simplifie la construction du modèle économique de l'application, mais l'absence d'engagements réciproques (garantie du fournisseur sur la qualité et la pérennité de la source, accès aux données traitées ouvert par le réutilisateur permettant d'en juger la pertinence et la qualité) le complique.

Les différents modèles de droits et obligations liées à la réutilisation de données ou de base de données mettent les réutilisateurs dans des situations sensiblement différentes selon qu'ils partagent ses droits avec des entités de taille comparable ou pas. C'est ainsi qu'on observe que les plateformes publiques peuvent selon les formats mis à disposition et selon le contenu plus ou moins générique des données, favoriser des initiatives locales ou encourager la réutilisation par des acteurs globaux. Il est frappant de constater à ce sujet que la réutilisation dans Google Maps ne semble pas dissuader les prestataires locaux de développer eux-mêmes leur service avec les mêmes données et en utilisant l'API cartographique de Google. Ceci se fait sans doute au détriment du modèle économique de l'acteur local qui sait qu'il devra au-delà d'une certaine audience acquitter à Google une redevance assise sur le nombre de requêtes.

La démarche open data initiale des collectivités territoriales reposant sur le double pari de la libération des initiatives économiques et citoyennes et de la plus grande transparence vis à vis du public se déploie mais elle rencontre des difficultés à convaincre des initiatives locales pérennisables et des acteurs industriels s'appuyant sur un autre modèle économique que les portails à revenus publicitaires.

Sa transposition dans le champ de la mobilité qui répond aux incitations de la Directive INSPIRE et en matière de transport à celles de la LOTI qui a prévu l'obligation de l'information multimodale des voyageurs n'échappe pas à la règle. Les applications publiques de services d'information multimodale se heurtent effectivement à la difficulté de rester souvent marginales en audience. La mise à disposition de données (ou de mise à disposition de données ou de services en libre accès) a précisément pour objectif de justifier les investissements déjà faits en ouvrant aux services que ces systèmes rendent possibles des audiences élargies. Cette approche incite une collectivité territoriale, notamment celles ayant le statut d'AOT (communauté urbaine, Département, Région) à faire relayer ses efforts et ceux de son opérateur par des initiatives multiples profitant de la souplesse et de la disponibilité des nouveaux médias télématiques. Les effets majeurs de cette démarche restent à venir.

Les réactions des opérateurs de transport public qui craignent de perdre le contact direct avec leurs clients ne se font pas attendre. Elles cherchent à mettre en place des solutions leur redonnant une certaine maîtrise des évolutions à venir de la médiation avec leur clientèle au travers d'initiatives qui tendent à leur rendre le contrôle des flux mis à disposition des réutilisateurs, avant que les plateformes publiques se soient imposées comme les meilleurs moyens pour les tiers de faire des services complets et fiables dans un modèle économique marqué par la gratuité des données dont elles ne sont pas propriétaires.

Deux types de solutions alternatives sont apparus :

1. Les plateformes opendata des opérateurs (ex. Keolis, RATP et SNCF) qui souhaitent mieux maîtriser l'évolution des relations avec les réutilisateurs. Elles s'appuient parfois sur le modèle d'une licence (voir celle explicite de la SNCF) de base de données collaborative (à la manière de certaines licences de logiciel open source) où l'opérateur se replace en animateur d'une communauté de réutilisateurs avec qui il entretient des relations régulières et dont il profite des enrichissements par le jeu de la réciprocité des apports à la base initialement mise à disposition.
Ces initiatives sont pour l'instant partielles (TER/Transilien/TET pour la SNCF, quelques données stations et arrêts mais pas d'horaires pour la RATP, quelques-uns des réseaux qu'il gère pour Keolis), mais elles pourraient rapidement se développer;
2. La relance de la contractualisation directe dont un des exemples est celui des conventions directes passées par Mappy grand spécialiste de la navigation routière avec différents opérateurs de transport public. Dans ce cas, les deux signataires partagent leur expertise pour se doter des moyens d'adresser des services performants à leurs clients respectifs sur la base d'une forme d'exclusivité sur des services plus à base de données en temps réel. Cette pratique n'est pas de l'open data, ni sur le plan technique, ni sur le plan contractuel. Il s'agit au contraire d'une solution intégrée dans le cadre d'un contrat fermé et repose plus sur des API que sur des transferts exhaustifs de données.

Une gamme assez large de solutions apparaît qui va de la totale liberté d'utiliser et de traiter à des formes plus ou moins précises de contrôle de conformité à des objectifs d'intérêt public, à l'instar de ce qui s'est fait dans les formes de partenariat privé-public mises en place pour la diffusion d'information routière numérique aux équipements de bord de navigation.

La totale liberté laissée aux réutilisateurs a le même inconvénient que la gratuité. Si on commence par là, il est très difficile de revenir à des conditions plus restrictives sans perdre une grande partie des acteurs initiaux en cours de route et peut être parmi les plus intéressants. D'un autre côté, dire à ce stade quelles sont les restrictions qui s'imposent comme les meilleurs garants d'un avenir équilibré n'est pas plus facile. Les exemples analysés montrent que les trois exigences que les promoteurs de la

mise à disposition sont en droit d'attendre de leur réutilisateur doivent d'abord s'appliquer à eux-mêmes et aux producteurs qui les alimentent : qualité, garantie de disponibilité, actualité des données. La documentation annexée à la mise à disposition doit donner des définitions précises assorties de critères mesurables de ces trois dimensions.

Si ces exigences étaient scrupuleusement appliquées, les collectivités « producteurs » seraient en droit de l'exiger de leurs réutilisateurs. Le modèle économique y gagnerait en pérennité, la collectivité en crédibilité et en efficacité de sa politique de mobilité, l'utilisateur y gagnerait en qualité de service.

La définition de ces trois exigences pourrait utilement profiter de ce qui se fait actuellement dans les échanges bilatéraux entre acteurs homologues (entre opérateurs de transport, entre AOT responsables de territoires contigus, entre systèmes d'information voyageur). L'avenir d'un système d'information est en effet conditionné par l'homogénéité des données dont il dispose et des informations qu'il diffuse, quelles qu'en soient les sources. Ce cahier des charges serait systématiquement joint à la documentation des données fournies, charge à lui de prendre les dispositions qu'il souhaite pour le respecter. Il résoudrait la première des conditions : celle qui veut qu'au delà des expérimentations et des inévitables tâtonnements initiaux, tout service au public se doit de ne pas dégrader le niveau de qualité atteint par les dispositifs de collecte de données qui l'alimentent. Seuls la traduction pour rendre l'information intelligible et le délai nécessaire pour acheminer l'information à son destinataire sont des variables d'ajustement acceptables.

L'argument qui veut qu'un mauvais service sera toujours sanctionné un jour par l'utilisateur reste vrai, mais compte tenu des différences considérables qui marquent les modèles économiques en présence, compte tenu aussi du fait que nous sommes dans un domaine proche de la communication institutionnelle des autorités régulant la mobilité, qui engagent leur image comme celle de leurs opérateurs, il peut être légitime d'imposer des minima aux réutilisateurs de façon à éviter que des services peu performants décrédibilisent les sources qui les alimentent. Nous recommandons une forme plus ou moins suscitée de concurrence entre les réutilisateurs. Cela facilite la comparaison et la notation par les utilisateurs des services, mais cela justifie pleinement une démarche ouverte qui cherche à encourager les nouveaux entrants et les plus innovants pourvu qu'ils fassent au moins aussi bien que les « installés ».

2.4.2. Recommandations

Les recommandations qui suivent ont un caractère contractuel (C). Elles recherchent l'établissement de relations responsables et durables entre producteurs, diffuseurs et réutilisateurs.

C 01 Choix de la licence

Il doit être déterminé par le niveau de contrôle et d'interaction qu'on souhaite avoir avec les réutilisateurs. L'APIE (Agence du Patrimoine Immatériel de l'État) et Etalab ont proposé des licences-types. Par ailleurs, l'APIE exprime des réserves sur l'emploi de la licence ODBL pour encadrer la réutilisation de données publiques. Il convient d'attirer l'attention sur les conséquences du choix de la licence, en fonction des objectifs poursuivis et des relations concrètes souhaitées avec les réutilisateurs obligés de bâtir un modèle économique pour offrir durablement des services au public.

C 02 Respect du service de transport

L'intérêt de la collectivité est de voir se multiplier des services de qualité tout en veillant à ce que le service de transport ne soit ni dénaturé ni altéré. Il est recommandé d'annoncer ce principe aux réutilisateurs, les invitant à garantir une information qui ne porterait pas préjudice au service de transport réellement mis en place.

C 03 Alerte des fournisseurs de données

En cas de présence d'informations erronées, les utilisateurs d'un service d'information doivent avoir la possibilité d'alerter l'autorité ayant fourni les données. L'organisme ayant téléchargé les données doit être identifié, et être informé de la nécessité de rendre publiques les sources de données de façon à ce que l'utilisateur mécontent ait le moyen de faire connaître ses critiques au propriétaire des données sources.

C 04 Suivi des réutilisations

Le réutilisateur pourrait être invité à informer le fournisseur de données du contenu du service qu'il met en place. Le fournisseur de données peut trouver utile de créer un observatoire des services offerts à partir de ses données.

2.5. Réutilisations identifiées pour des services à l'utilisateur final

2.5.1. Synthèse de l'analyse

Les réutilisations des données transport diffusées par les plates-formes publiques étudiées sont assez diverses. Les services web et les applications mobiles sont nombreuses mais souvent décevantes, soit parce qu'elles se ressemblent beaucoup, soit parce qu'elles souffrent de la piètre qualité des données (données brutes fournies ou produits de traitement hasardeux) ou de leur manque de complétude.

En dehors des services délivrés par les collectivités territoriales non AOT, les AOT et les opérateurs eux-mêmes, qui sont de bonne qualité et couvrent parfois un champ plus large que celui de leurs responsabilités directes, les services des réutilisateurs restent pour l'instant trop timides, c'est à dire très semblables et ne bénéficiant manifestement pas des mêmes budgets ou très précisément ciblés, plus originaux, parfois amusants, mais souffrant des défauts liés à une connaissance partielle des données.

On peut caractériser les réutilisations que nous avons analysées (et listées dans le tableau en annexe) de la façon suivante :

- Les services TC sont mis en ligne à la fois par les opérateurs de transport (sites web et applications smartphone), par des collectivités territoriales et des "nouveaux entrants" dans ce marché : start-ups ou junior entreprises.
Les plus convaincants de ces derniers services sont ceux qui choisissent un ciblage précis de clientèles spécifiques de ces modes : PMR, scolaires, touristes, ou qui intègrent des données complémentaires aux transports dans la pratique et la découverte du territoire, POI et points d'arrêt pour desservir les lieux de délivrance des services urbains (musées, espaces verts, crèches, bureau de poste, etc.) via des itinéraires prolongés par des cheminements protégés piétons ou cyclistes , avec animal de compagnie, par exemple
- Calcul d'itinéraires : il est difficile de se rendre compte sans approfondir jusqu'à quel point les services proposés innovent dans ce domaine ou se contentent de réutiliser des résultats ou API de plates-formes producteurs. Ce rapide tour d'horizon met toutefois en évidence une fois de plus que cette forme d'aide à la mobilité est très largement partagée par l'offre et la demande comme l'est aussi la représentation cartographique des réseaux et des trajets.
- Pôles d'échange : n'apparaissent pas dans ces services comme un élément structurant visible mais plutôt comme un garant discret de la multimodalité (point de choix) et de l'intermodalité (point de correspondance). Les services vélos : ils ont une importance significative en nombre, là où les données sur les stations VLS (c'est rare), le tracé et la pente des pistes cyclables ou la localisation des parkings dédiés sont mises à disposition. Comme le mode lui-même, les applications, notamment mobiles s'adressent à une clientèle plutôt jeune et redonnent un sentiment de liberté aux modes alternatifs à la voiture. C'est sans doute pourquoi ils ont acquis une place dans le paysage de l'information sur la mobilité bien supérieure à leur part modale dans la mobilité locale.
- Les aides pratiques à la mobilité (temps d'attente en stations, occupation pendant cette attente ou à bord, trajets choisis en fonction de critères variés comme l'agrément de l'environnement visible sont clairement considérées comme la voie à suivre pour se démarquer des services « tout public » des opérateurs et des opérateurs globaux. Cette tendance montre ses limites quand on ne dispose pas de données suffisamment précises ou temps réel que certains producteurs répugnent à mettre à disposition. Les cibles jeunes, les touristes dans les villes ou sur le littoral bénéficient ainsi de petites applications qui n'ont ni la prétention à l'exhaustivité, ni celle de s'adresser à un public très large, mais explorent un concept (la visite en marchant, la découverte à bicyclette, etc.) ou les attentes d'une cible précise : les promeneurs, les amoureux des vieilles pierres, les sportifs, etc.).
- Le multimodal intégral Metro, Tram, Bus, VLS, TAD, etc. est l'autre grande tendance vers laquelle la communauté des réutilisations souhaiterait pouvoir contribuer. La non complétude des données, l'absence de référentiel géographique normalisé (réseaux routier et en site propre utilisés par chaque mode, points d'arrêt, parkings, POI, etc.) permettant de comparer et d'interconnecter toutes les informations sur tous les modes rendent les développements coûteux et on voit mal comment ces investissements seront faits en dehors de modèles économiques durables, dont la plupart des réutilisations actuelles, trop expérimentales, ne témoigne pas. De fait, les réutilisations multimodales sur un territoire donné sont actuellement le fait de collectivités territoriales, d'AOT ou de quelques opérateurs, notamment les groupes ferroviaires nationaux et étrangers.

- L'entrée en scène des opérateurs globaux internet (Google Transit, Apple Map, Voyages SNCF,...) se limitent pour l'instant à l'intégration des points d'arrêts ou de réseaux TC. Google Transit par exemple, a intégré Bayonne Chronoplus, BAYONNE RESEAU SUD, Bordeaux Info TBC, Brive STUB, Cholet Choletbus, Grand Dax RDTL, Grenoble Plan du réseau uniquement, Le Mans Plan du réseau uniquement. Il s'agit d'accords directs passés avec la collectivité ou l'opérateur TC : TBC à Bordeaux pour plan et calcul d'itinéraire à base d'horaire théorique, simple plan à Grenoble, ...
- Ces réutilisations restent en marge de notre sujet car elles ne se situent pas dans le cadre des licences ODbI ou équivalentes.
- La visite thématique ou touristique est une autre niche explorée par les initiatives actuelles de réutilisation. Cette demande d'information de clientèles occasionnelles pourrait se révéler intéressante pour des acteurs aux ambitions mesurées. Il leur faudra déployer beaucoup d'expertises et en faire la preuve dans leur service, car les usagers ont tendance à utiliser des services génériques qu'ils ont l'habitude d'utiliser et dont ils connaissent les secrets. Comme on le voit, il s'agit surtout de réutilisations locales modestes mais l'apparition d'applications, notamment mobiles tournant sur les données de plusieurs territoires, constitue peut-être les prémices d'un véritable développement de services originaux.
- Les services orientés PMR, bénéficiant d'une bonne image dans le public et susceptibles de recevoir des financements publics dans le cadre d'actions d'intérêt général se révèle un autre marché de niche pour les développeurs. L'intérêt est qu'ils sont également l'occasion d'assembler des technologies d'interface avec des contenus spécifiques.

Pour les besoins de l'analyse, en dehors de leur contenu, les sites internet et applications mobiles recevant les données des plates-formes sélectionnées peuvent être classés en types différents selon les organisations qui les portent :

- Les applications « internes » des AOT ou de leurs opérateurs,
- Les applications « juniors » des écoles d'ingénieurs, des développeurs individuels, des junior entreprises locales,
- Les applications locales des entreprises locales, voire des consortia d'entreprises spécialisées en cartographie, géolocalisation, récepteurs spécifiques,
- Les applications transverses TC, trafic, vélos,
- Les applications des acteurs globaux de la mobilité : Google Transit, TomTom, Nokia, Apple Plans, etc.

Les sites web sont moins foisonnants, plus « sérieux » que les applications smartphone s'adressant souvent à des cibles plus jeunes et plus restreintes. Les uns comme les autres s'appuient surtout sur des API cartographiques des grands éditeurs comme Google, Yahoo ou Microsoft.

De cette analyse, on peut tirer déjà un certain nombre d'enseignements :

- Les volumes connus (nombre de téléchargements d'application smartphone sur les sites Android Market et Applestore) montrent qu'il s'agit encore de retombées modestes pour les applications mobiles de réutilisateurs qui ne permettent pas d'en garantir la pérennité économique. Certains de ces services semblent démontrer que les objectifs qualitatifs peuvent être atteints, mais il paraît difficile de garantir qu'ils le seront si on fait appel à l'initiative privée sans visibilité sur des modèles économiques durables.
- L'offre de données est assez opportuniste (sous-produits de l'exploitation). Les informations mises à disposition sur les plates-formes publiques sont souvent des données très proches de celles qui ont été produites pour les besoins de l'exploitation et dans le meilleur des cas de l'information voyageurs. Les points d'arrêt et leur localisation, les horaires théoriques des services, par exemple méritent d'être normalisés dans leur format et complétés dans leurs attributs pour être totalement exploitables d'un territoire à un autre.
- Les innovations en services restent limitées : la représentation cartographique des arrêts et des réseaux, le calcul d'itinéraires, ciblés ou non sur une catégorie particulière de voyageurs restent les références. Le principal intérêt démontré est que le ciblage favorise l'utilisation des modes alternatifs à l'automobile reliée à d'autres activités : travail, tourisme, loisirs, sports, etc.
- A l'image de ce qui se fait dans le domaine des applications mobiles, les réutilisations semblent recueillir la sympathie du public qu'elles visent (voir à ce sujet les forums d'utilisateurs), mais leur audience semble limitée, en tous les cas à des niveaux ne garantissant pas leur survie économique.

- Les jeunes et actifs sont des clients très souvent visés par des applications très ciblées utilisant le langage et l'approche de la mobilité des lycéens, des étudiants, des jeunes salariés. Cela va bien dans le sens des objectifs des plates-formes publiques de mise à disposition, d'autant que ces applications sont le plus souvent d'équipes très jeunes et très enthousiastes, mais il faudrait une méthode d'accompagnement plus structurée que des concours qui consacrent une démarche initiale mais sont moins adaptés à résoudre les difficultés de développement d'un projet d'entreprise.
- Les API et web-services mis à disposition par ces plates-formes doivent être traités à part pour bien montrer en quoi il s'agit d'une démarche différente mais complémentaire : ces services constituent en effet une alternative pour tous les services réutilisateurs qui ne souhaitent pas devenir experts des données transport mais considèrent que l'information multimodale fait partie intégrante des requêtes que sont susceptibles d'adresser ses utilisateurs à leur service. Un bon exemple est donné par les API de temps d'attente aux arrêts alimentés par des données en temps réel. Ces services les plus intéressants sont ceux qui répondent aux préoccupations de leurs cibles : "que faire en attendant son bus ? Lire les nouvelles, connaître la météo du jour, savoir quand il arrive ...

2.5.2. Recommandations

Les recommandations qui suivent concernent le comportement à avoir à l'égard des réutilisateurs (R). Elles recherchent l'établissement de relations équilibrées et durables entre producteurs, diffuseurs et réutilisateurs.

R01 Intérêts des petits acteurs

Veiller particulièrement à ce que les petits acteurs multisites et les acteurs locaux, qui ne sont protégés que par leur expertise de services très ciblés ou spécialisés, ne soient pas pénalisés.

Le respect des recommandations G12 / G13 et la mise en place des manifestations comme l'organisation de concours périodiques est une manière de les attirer.

R02 Dispositif de suivi des réutilisations

Il est recommandé d'organiser la veille sur les réutilisations des données mobilité. Il serait souhaitable que chaque plate-forme, ou un regroupement de celles-ci, mette en place un dispositif de suivi des réutilisations et de veille sur les expériences comparables pour tirer les enseignements des expériences et réorienter l'action des différents acteurs dans la direction souhaitée. Pour éviter que le dispositif soit trop lourd pour les fournisseurs et les réutilisateurs, il serait souhaitable qu'il s'appuie sur des forums d'utilisateurs adressés par les (services) réutilisateurs de la source des données servant à alimenter le service.

3. Conclusions

3.1.1. Synthèse de l'analyse

L'importance de la mobilité dans l'univers des données publiques est avérée : les transports et la mobilité occupent une place de choix dans la mise à disposition de données, parce que ce sont des informations tout public, qu'elles peuvent alimenter des services pratiques et quotidiens et qu'elles concernent l'appropriation d'un territoire par ses résidents et ses visiteurs.

L'interdépendance des modes de transport se retrouve dans l'information multimodale, qui n'est finalement que l'adaptation de l'information voyageur à une combinaison de modes qui contribue à l'intérêt collectif et à la sauvegarde de l'environnement en se substituant sur la plus grande partie du trajet au véhicule individuel. Cette adaptation va nécessiter de plus en plus d'interopérabilité entre les SI, fondée sur des formats d'échange communs et normalisés.

La complétude des données n'est actuellement pas assurée, du fait du caractère non obligatoire et non systématique de l'ouverture des données de transport. L'émiettement des organisations agissant dans le domaine, le peu de coordination entre modes n'utilisant pas les mêmes infrastructures ou ne dépendant pas de la même autorité font que les services, notamment privés, restent incomplets et n'offrent pas de ce fait un niveau de service suffisant.

La qualité des données fournies et des données traitées par les réutilisateurs doit être la meilleure possible. Les services qui déçoivent leurs utilisateurs n'ont qu'une vie assez courte et peuvent dégrader durablement l'image des responsables des sources de données qui les alimentent. La diversité des formats actuellement utilisés rend quasi impossible l'interopérabilité de ces systèmes, ce qui constitue un obstacle majeur au déploiement de services d'information multimodale de qualité sur un territoire plus vaste que celui du producteur des données.

Beaucoup de sites et d'applications réutilisatrices ont recours aux API cartographiques en ligne (Google Maps, Yahoo Maps, Live Search Maps, Map24...) pour la représentation des réseaux, l'aide à la sélection des origines-destinations et la visualisation des résultats du calcul d'itinéraires. Ceci marque profondément le « *look and feel* » des applications réutilisatrices qui sont très proches des sites d'information voyageur des opérateurs ayant recours aux mêmes outils ou des sites « maps » des éditeurs de ces API. C'est aussi pour cette raison que les formats propriétaires comme GTFS sont utilisés, ce qui rend ces services dépendant de l'évolution des conditions consenties par leurs éditeurs vis à vis des utilisateurs.

L'existence de données en temps réel est rare et elles ne semblent, actuellement, mises à disposition qu'au travers d'API qui sortent du périmètre de notre analyse. Il ne s'agit pas à proprement d'open data, mais de services mis à disposition de réutilisateurs de façon assez libérale, sous le contrôle et la mesure du point d'accès au service de mise à disposition. Dans de tels cas, il est prévu des limites d'accès en nombre et en fréquence, qui lorsqu'elles sont dépassées, déclenchent la modification des conditions contractuelles. En termes clairs, l'accès est plafonné ou soumis à participation aux frais d'extension des accès.

Le recours à des contenus hétérogènes, d'une source à une autre, pour informer sur des services de transport identiques d'une ville ou d'une région à l'autre limite beaucoup la couverture territoriale des applications et le déploiement de services génériques.

Malgré toutes ces réserves, on voit émerger quelques applications multisites aussi bien sur internet que sous forme d'applications mobile pour Iphone, smartphones Android ou Windows mobile (voir tableau des réutilisations identifiées en annexe).

L'imprécision quant à l'identité du producteur, propriétaire, gestionnaire des données utilisées, la durée de validité des données, information sur la gestion des calendriers, nuit principalement à l'agrégation des données provenant des territoires distincts. L'imprécision de l'information sur la fréquence de mise à jour (« quand nécessaire ») peut être interprétée de multiples manières quand il s'agit de déplacements d'arrêt, de passage aux horaires d'été, de détours d'itinéraires en cas de travaux, etc.

Pour un réutilisateur :

- l'emprise géographique des données publiées n'est pas claire
- le descriptif du contenu des données est trop général
- une liste structurée des données publiées est manquante

- la précision sur le contenu des données publiées (ex. : quels attributs renseignés concernant les points d'arrêt), est absente.

3.1.2. Synthèse des recommandations et perspectives

Deux grandes approches de la mise à disposition des données sont possibles, illustrées par les exemples analysés :

1. La publication tel quel : faites ce que vous voulez, mettez à disposition ce dont vous disposez, mais dites très précisément ce que c'est, en veillant à ce que le réutilisateur comprenne parfaitement la définition, le périmètre et les limitations de ce qui lui est fourni. Dans ce cas, l'effort d'amélioration porte sur la documentation.
2. La mise aux normes : faites ce qui vous est conseillé, chaque fois que la communauté Open Data mobilité se met d'accord sur un référentiel commun. On entend par là : des définitions communes, des formats d'échange communs, une conformité vérifiée à des normes communes et à des profils partagés par des outils identifiés et disponibles à tous.

Dans la première approche, l'appairage des données fournies avec d'autres données de provenance diverse est sous la totale responsabilité et aux frais du réutilisateur. Il prend les données « en l'état » et est totalement responsable du résultat final. Cela ne dispense pas le fournisseur d'investir dans une documentation précise lui garantissant, au moins en partie, que les réutilisations ne seront pas d'une qualité amoindrie par la complexité et la difficulté de compréhension de ce qui est mis à disposition.

Il en résulte les recommandations suivantes :

- Systématiser la présentation du contenu des sites de mise à disposition (en s'inspirant d'exemples comme le CG33)
- Indiquer la liste des données publiées avec le(s) format(s) disponible(s) et les liens vers les documentations et outils de test en ligne de ces formats
- Indiquer clairement l'emprise géographique
- Indiquer, pour chaque donnée, la liste des informations effectivement renseignées (pas seulement le format d'échange)
- Généraliser ou encourager la présence de forums en vue d'un « réseau » d'utilisateurs
- Pour des réutilisations intersites/interrégionales :
 - gestion temporelle des données – calendriers – validités – fréquences de mise à jour,
 - vérification de la conformité des données mises à disposition avec les normes et profils évoqués dans la documentation,
 - inscription dans un annuaire des données disponibles, une fois cette vérification faite avec succès.

Dans la seconde approche, seule apte à favoriser l'interopérabilité technique des plates-formes comme le fait justement remarquer Simon Chignard², il faut identifier précisément tous les producteurs et réutilisateurs de la communauté française Open Data Transport, résoudre les problèmes de gouvernance, à partir des initiatives communes de l'AFIMB, du GART et de l'UTP, et mettre à disposition les outils partagés nécessaires.

Quelle que soit l'approche retenue, la mise en œuvre de plates-formes mutualisées sur le modèle de ce qu'ont commencé à faire la région Aquitaine, le département de la Gironde et la Communauté urbaine de Bordeaux permet de résoudre ensemble des problèmes qui se posent à chacun et de faire des économies d'échelle bienvenues.

Les réflexions suivantes peuvent être formulées sur l'opportunité de travaux techniques futurs, de développements logiciels afin de rendre plus aisée la publication et la réutilisation des données transport :

- faciliter la mutualisation et la gestion des arrêts
 - fournir un modèle d'arrêt partagé (définition/caractérisation harmonisée des arrêts),
 - fournir un outil de construction d'identifiants (construction des lieux d'arrêt et gestion des responsabilités sur les lieux d'arrêt),
- offrir les extensions de l'outil CHOUETTE
 - améliorer l'administration de la base de données en vue de la création d'une base de données mutualisée afin d'en faciliter l'alimentation par diverses sources open data,
 - en fonction des domaines nouveaux à couvrir (définition de l'accessibilité, gestion des équipements...), prévoir des modules de saisie des données concernées, de façon à

² Monétiser les données du transport public... chiche ? par Simon Chignard dans Villes et Transports du 11/12/2012

- pouvoir les mettre à disposition et en assurer la mise à jour (par exemple des données d'accessibilité),
 - prévoir un module de saisie, publication, échange de l'offre tarifaire lié aux données de l'offre théorique de service,
 - prévoir un module d'alimentation, archivage, export des données temps réel en vue de création d'indicateurs statistiques,
 - améliorer l'outil de vérification des données :
 - prévoir la vérification du contenu des données déposées dans la base de données (pour l'instant, seuls les « fichiers importés» subissent des vérifications)
 - améliorer la formulation des messages d'erreur
 - renforcer les algorithmes de vérification en ajoutant des critères plus stricts de vérification de cohérence
 - formaliser la partie « publication » des données open data, intégrant les données dans un ou plusieurs formats, les métadonnées, la documentation d'accompagnement, la licence, etc. (cf. les exigences de contenu énumérées en 2.2.2).
- généraliser la diffusion et la vulgarisation des documents normatifs pour promouvoir leur application, en particulier dans le cadre de la publication des données open data afin de favoriser une sémantique et un format partagés.

Cette étude constitue une étape dans la réflexion sur l'ouverture des données transport. Elle est destinée à alimenter les travaux des autorités organisatrices, des opérateurs et des candidats réutilisateurs. Il est prévu qu'elle serve de base à des échanges, pour tester les propositions et être à l'écoute des souhaits exprimés par les participants.

Annexe 1

Présentation des fiches open data

1. La CUB (Bordeaux)	27
1.1. Fiche analytique	27
1.2. Commentaires généraux	30
2. CG44 Réseau Lila	32
2.1. Fiche analytique	32
2.2. Commentaires généraux	33
3. Nantes - Semitan	34
3.1. Fiche analytique	34
3.2. Commentaires généraux	35
4. Rennes Métropole	36
4.1. Fiche analytique	36
4.2. Commentaires généraux	38
5. Toulouse TISSEO	39
5.1. Fiche analytique	39
5.2. Commentaires généraux	41
6. Transgironde	42
6.1. Fiche analytique	42
6.2. Commentaires généraux	45
7. Exemples d'initiatives Open Data à l'étranger	46
7.1. Exemple de Londres	46
7.2. Exemple de New York City	47
7.2.1. NYC Open Data	47
7.2.2. MTA Data	47
7.2.2.1. Flux de données en temps réel	47
MTA All-Agence	47
MTA NYC Transit	47
7.2.2.2. Flux de données horaires et statiques	48
MTA toutes-Agences	48
MTA NYC Transit	48
MTA ponts et tunnels	49

Les fiches 1 à 6 ci-dessous présentent un état de l'art du 2^e semestre 2012.

1. La CUB (Bordeaux)

1.1. Fiche analytique

CUB	
Nom réseau	TBC
Adresse site	http://data.lacub.fr/ et http://www.infotbc.com/
Objectifs initiaux (basés sur les textes extraits du site)	<p><u>1. Origine de la démarche</u></p> <p>« Si j'ai souhaité que la Communauté urbaine de Bordeaux (Cub) s'engage aussi volontairement dans la démarche OpenData, c'est qu'il s'agit pour moi d'une évidence : au-delà des incitations réglementaires, il est démocratiquement logique de mettre aussi largement que possible les données publiques détenues par les collectivités à disposition des citoyens-usagers.</p> <p>Entièrement anonymes et donc respectueuses de la vie privée, ce sont des informations d'usage général et d'intérêt local dont nous devons pouvoir nous saisir, partager, réutiliser.</p> <p>C'est finalement une façon assez simple de faciliter la vie des habitants de l'agglomération bordelaise, mais aussi et surtout de leur rendre compte de ce que fait pour eux la Communauté urbaine au quotidien, avec leur argent : travaux de voirie, lignes et horaires de passage des transports gérés par la Cub, disponibilités en direct des places de parking, zonage du PLU, emplacement des déchetteries ou des bornes d'apport volontaire... (rdv sur data.lacub.fr)</p> <p>Ce retour vers le contribuable constitue pour moi non seulement une évidence, mais aussi un impératif politique.</p> <p>C'est aussi un enjeu essentiel de transparence de l'action publique qui ne peut être que bénéfique, non seulement pour le citoyen, mais également pour les élus et la collectivité en charge de missions d'intérêt général. Savoir-faire et faire-savoir : si la CUB a depuis toujours pu se prévaloir du premier, il est heureux que l'OpenData puisse désormais participer du second, et ainsi contribuer à valoriser l'excellence du service public et l'attractivité de tout un territoire.</p> <p>Et puis, au-delà, on sait très bien que l'OpenData agit comme un formidable démultiplicateur d'initiatives, un facilitateur de croisements et de rencontres, un moteur d'innovations. Toutes ces déclinaisons permettent d'améliorer concrètement le quotidien métropolitain où l'essentiel ne consiste plus vraiment à rechercher l'information mais, parmi le « bruit » environnant, de savoir trouver la bonne info au bon moment et de pouvoir la traiter.</p> <p>L'OpenData de la Cub est la matière première que nous mettons à disposition de tous -des particuliers aux associations en passant par les créateurs d'applications pour smartphones- afin que chacun puisse s'en saisir comme bon lui semble et ainsi réinvestir la sphère publique.</p> <p>En bref, c'est un peu l'incarnation de l'idéal de liberté intégrale et de partage dont rêvaient les créateurs du Net. A une époque où celui-ci est battu en brèche, je me réjouis que les collectivités publiques puissent ainsi lui redonner un sens ».</p> <p>Vincent Feltesse Maire de Blanquefort, Président de la Communauté urbaine de Bordeaux</p> <p><u>2. Ville de Bordeaux, Cité Digitale</u></p> <p>Pour plus de transparence, de collaboration et d'innovation entre ses politiques publiques, ses services et ses citoyens, la Ville de Bordeaux a choisi d'ouvrir son portail opendata.bordeaux.fr.</p> <p>Bordeaux facilite l'accès à ses données municipales</p> <p>La démarche <i>open data</i> s'inscrit pleinement dans la dynamique du programme</p>

	<p>"Bordeaux Cité Digitale". Comme vous le savez, la Ville met à votre disposition depuis plusieurs années déjà des services numériques et des informations <i>via</i> son portail bordeaux.fr, ses sites internet culturels et événementiels ou encore ses applications mobiles. Dans cette continuité, la Ville de Bordeaux met à votre disposition un portail <i>open data</i> spécifiquement dédié à l'ouverture de ses données publiques. D'accès simple et pratique, ce portail vous permet de télécharger gratuitement et en toute liberté les lots de données de la Ville portant sur nombre de domaines : données budgétaires, cartographiques, urbaines, culturelles, de développement durable et bien d'autres encore. Dans l'esprit du mouvement international <i>open data</i>, la Ville garantit que ces données sont non nominatives, mises à jour régulièrement et proposées sous un format ouvert afin de faciliter leur réutilisation et leur interopérabilité avec d'autres données publiques ou privées. opendata.bordeaux.fr, un outil collaboratif pour une cité plus ouverte</p> <p>Grâce à cet outil destiné à la fois aux usagers, chercheurs, associations et entreprises innovantes, la Ville de Bordeaux souhaite participer à l'émergence de nouveaux usages. Elle s'engage ainsi dans une démarche de transparence démocratique et de valorisation de son patrimoine immatériel. Le portail opendata.bordeaux.fr présente ses données sous une forme normalisée à destination des réutilisateurs (analystes et développeurs). Un contact métier est associé à chaque flux de données présent sur le portail.</p> <p>Pour le grand public, elle apporte d'autres services, notamment des entrées thématiques, un moteur de recherche performant et des espaces dédiés aux échanges autour de ce projet citoyen:</p> <ul style="list-style-type: none"> • Les données : un espace de présentation et de téléchargement des bases de données accompagnées d'outils de visualisation ludiques permettant d'analyser en toute simplicité les informations municipales • Les applications : un espace de présentation des applications développées autour des données municipales au bénéfice de tous • Les supports : un espace dédié aux échanges entre les producteurs et les réutilisateurs, entre les services de la Ville et les citoyens, afin de vous accompagner dans l'utilisation du portail et créer, à terme, un véritable écosystème autour de la donnée publique à Bordeaux "
Données transport	<p>Réseau hiérarchisé de voirie : SIG Arrêt, ligne, couloir, dépôt, pts vente, station, parcours : SIG Parc relai : SIG Tronçon, couloir, ligne, parc relai : tables de relation Courbe circulation, états de trafic, temps de parcours Temps attente stations VCUB : temps réel</p>
Format d'échange	<p>XML : GTFS CSV, DXF, SHP</p>
Vérification de l'import	<p><i>Voir annexe 2</i></p>
Plus d'info	<p>http://data.lacub.fr et opendata.bordeaux.fr.</p>
Propriétaire de la donnée	<p>CUB Bordeaux pour les TC et le trafic routier Ville de Bordeaux pour le stationnement et les pistes cyclables</p>
Contenus	<p>SIG, tables de relation, temps réel</p>
Contributeurs	<p>TBC, CUB, Ville de Bordeaux</p>
Collecte	
Licence ou conditions de réutilisation	<p>ODbl</p>
Fréquence de mise à jour	<p>A chaque changement : « chaque fois que nécessaire » Temps réel pour trafic et VLS</p>
Réutilisations identifiées	<p>http://tourisme-accessible.bordeaux.fr/ voies en stationnement payant pistes cyclables communales</p> <p>http://arbres.bordeaux.fr/ soit 4 services édités par le mairie de Bordeaux Transports Bordeaux ed. YBonnell tel>10 000 android Info TBC ed. Keolis tel>10 000 Android Parking Direct ed. Matthieu Lorgeoux tel>500 Android</p>

Bordeaux onLive ed Louge/Baheu Windows Phone 7

Ibordeaux tram bus Sébastien Duguet (payant) iphone
Imove in bordeaux ed. Codega Studio (payant) iphone

Parking Direct

Trouvez immédiatement une place de parking libre dans l'un des parkings publics de Nantes et Bordeaux grâce à la mise à jour en temps réel. Téléphonnez au parking pour obtenir des informations détaillées ou lancez la navigation GPS pour y arriver sans encombre.

Détail des fonctionnalités :

- Affichage des places libres, parkings abonnés, parkings fermés
- Navigation vers le parking sélectionné via Google Navigation
- Recherche des parkings proches de l'adresse de destination

Auteur : Matthieu Lorgeoux

Date de publication : 07/03/2012



Bordeaux onLive

Bordeaux onLive est une application de référence sur WP7 pour tous les Bordelais se déplaçant.

V3

Itinéraires, cartes, favoris, correspondances ainsi que la disponibilité des cycles et emplacements.

Trafic

Le trafic intérieur ainsi que celui de la rocade en temps réel.

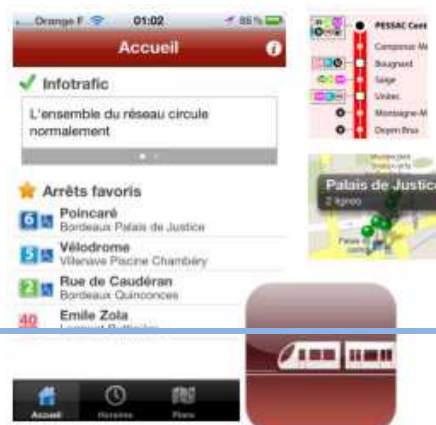
Parkings publics

Itinéraires, cartes, et disponibilités en temps réel.

Parcs relais

Itinéraires, cartes, et capacité détaillée (mobilité réduite, 2 roues, etc.).

TBC



iBordeaux Tram Bus

iBordeaux Tram Bus vous donne accès aux informations essentielles pour utiliser les trams et bus de la CUB :

- Prochains passages
- Plan du réseau
- Plan des lignes
- Carte des arrêts
- Affichage des perturbations

Les arrêts peuvent être ajoutés en favori et sont alors affichés en page d'accueil.

L'application sera progressivement enrichie de nouvelles fonctionnalités.

Auteur : Sébastien Duguet

Date de publication : 13/01/2012

iMove in Bordeaux

Application iPhone permettant de géolocaliser tous les moyens de transport et de stationnement, dans Bordeaux, au travers d'une interface originale et simple d'utilisation.

- Stations VCUB

Information temps réel sur la disponibilité des vélos et des bornes.

- Parcs Relais

Capacité d'accueil détaillée et interconnexion avec les lignes TBC.

- Parkings

Nombre de places libres en temps réel.

- Arrêts de Bus et de Tram

Affichage rapide des arrêts de Bus et de Tram.

- Lignes de Bus et de Tram

Affichage du tracé des lignes et respect des couleurs de TBC pour mieux vous y retrouver.

Auteur : CODEGA STUDIO

Date de publication : 28/11/2011



Transports Bordeaux

Application pour smartphones Android, fournissant un accès rapide aux données des transports de Bordeaux dans votre poche.

Cette application regroupe les fonctionnalités suivantes sur les transports de Bordeaux :

- Horaires de Bus et de Tram avec géo-localisation
- Arrêts de Bus à proximité
- Widgets pour permettre l'accès rapide aux horaires
- Vélos disponibles dans les stations VCUB
- Stations VCUB à proximité
- Gestion des arrêts de Bus favoris
- Gestion des stations de Velo favorites
- Perturbations et compte Twitter @tbc

Auteur : Yan Bonnel

Date de publication : 14/06/2011

1.2. Commentaires généraux

Points positifs

- Bonne vue d'ensemble des données disponibles

Points négatifs

- Pas d'indication de fréquence de mise-à-jour
- Pas de description de gestion des calendriers

Périmètre et évolutivité des données

La Cub étant en phase d'expérimentation, les données disponibles sur ce site ne représentent qu'une première partie des données que La Cub souhaite potentiellement ouvrir. Le portail devrait accueillir prochainement de nouvelles données. Il est possible de faire part de ses remarques et suggestions au travers d'un formulaire.

Formats

Dans le cadre de l'expérimentation, seront proposés « dans la mesure du possible » des données dans des formats courants et exploitables.

AFIMB

Évolutions fonctionnelles

A ce jour, le site met à disposition des données géographiques aux formats ESRI shapefile et KMZ, et permet leur visualisation au travers de WebServices (WMS et WFS). Il propose aussi des données alphanumériques en CSV ou dans les formats natifs dans lequel elles sont utilisées dans les services de la CUB.






Il est prévu de rendre prochainement les données géographiques accessibles au travers d'une API CUB.

Seront également intégrées, dès mise en place des moyens techniques le permettant, des données temps réel.

2. CG44 Réseau Lila

2.1. Fiche analytique

CG 44 (Loire-Atlantique)	
Nom réseau	LILA
Adresse site	http://data.loire-atlantique.fr
Objectifs initiaux (basé sur le texte du site)	<p>Le plan d'actions s'articule autour de 3 enjeux. Pour chacun, des objectifs des actions ont été définies et vont être mises en œuvre jusqu'en 2014.</p>
Type de descriptif des données	Très général
Type de données transport	<ul style="list-style-type: none"> Tracé des voies navigables gérées par le Département de Loire-Atlantique : canal de Nantes à Brest, rivière de la Sèvre navigable (Sèvre nantaise jusqu'à Monnières), rigole d'alimentation, rivière de l'Erdre Horaires et points d'arrêt des lignes Lila régulières. Éléments de la charte graphique du réseau de transport Lila (aires de covoiturage, lila à la demande, lila lignes régulières et lila scolaire) : logos, pictogrammes et nuancier des lignes Lila
Format d'échange	Les exports (horaires/arrêts) sont fournis aux formats GTFS et XML/Trident. Pas précisé pour les voies navigables
Période de validité	Horaires valables à compter du 1er septembre 2012 jusqu'au 5 juillet 2013 inclus pour le premier fichier GTFS, et du 1er juillet au 31 août 2012 pour le deuxième fichier GTFS
Mise à jour	« Lorsque nécessaire » : indication de la mise à jour trop imprécise pour une réutilisation (mutualisation) des données dans un contexte interrégional. Indication de la gestion des calendriers manquante.
Type de services	énumération des modes disponibles manquante
Résultats de l'import logiciel CHOUETTE	Effectué Cf. Annexe 2

	Eléments	Qté
	Réseaux	1
	Transporteurs	0
	Lignes	10
	Courses	208
	Calendriers	13
	Correspondances	722
	Arrêts	669
	 points d'embarquement	417
	 quais	0
	 arrêts commerciaux	252
	 pôles d'échange	0
	 ITL	0
	<u>Quelques incohérences ou erreurs</u>	
	Mode(s) non renseigné(s) ou « autre »	
	Lignes : erreurs de coordonnées (ex. ligne A : points au sud du Ghana)	
	Arrêts : informations manquantes (ex. adresse), coordonnées erronées (ex. arrêt Bois/Plage)	
	Arrêts commerciaux : différenciés mais superposés aux points d'embarquement (ex. arrêt Bole Eden : pas de différenciation de localisation).	
	Correspondances : uniquement entre paires d'arrêts identiques	
Emprise géographique	Voies navigables : Loire-Atlantique (ligne) Horaires et points d'arrêt : réseau Lila	
Liste des données	Lignes, arrêts (commerciaux, points d'embarquement, correspondances, calendriers, courses)	
Plus d'info	Forum permettant un dialogue (quelles donnée cherche-t-on ? quels services ?)	
Contributeurs	-	
Collecte	Pas d'information sur l'assurance qualité	
Réutilisation	Non indiquée	

2.2. Commentaires généraux

Points positifs

Bonne vue d'ensemble des données publiées

Points négatifs

Quelques manques d'information sur les formats utilisés.

Vérification insuffisante des données mises à disposition

3. Nantes - Semitan

3.1. Fiche analytique

NANTES	
Nom réseau	SEMITAN
Adresse site	data.nantes.fr
Objectifs initiaux-démarche (basé sur les textes extraits du site)	<p>Nantes Métropole et la Ville de Nantes ont pris l'engagement commun d'ouvrir leurs données en février 2011.</p> <p><u>Objectifs</u> : la volonté des deux collectivités est de mettre à disposition de tous les informations récoltées dans le cadre de leurs missions respectives. Il s'agit ici de mener une action</p> <ul style="list-style-type: none"> • de redistribution de ces données aux citoyens, sous la forme de nouveaux services, • de développer le tissu économique local, • d'améliorer les services rendus aux usagers et entreprises de la métropole nantaise. <p>La transmission, le partage et la réutilisation de ces données dématérialisées favoriseront, à terme, l'innovation et l'activité tout en améliorant la vie quotidienne des habitants.</p> <p><u>Collecte et mise à disposition</u> :</p> <p>Les informations disponibles sur ce site sont collectées par les institutions dans le cadre de leur activité. Elles sont mises à disposition gratuitement dans le respect des conditions de la licence et seront enrichies progressivement. Nantes Métropole, la Ville de Nantes et leurs partenaires veillent à protéger les données personnelles des usagers. Les informations proposées sur le site sont donc publiques, anonymes et ne relèvent en aucun cas de la vie privée ou de la sécurité des citoyens.</p> <p><u>Evolution souhaitée</u> :</p> <p>L'ouverture des données publiques permettra :</p> <ul style="list-style-type: none"> • le développement de nouvelles applications, • l'enrichissement d'études, • la création artistique • l'émergence de nouveaux acteurs • le développement de l'activité économique au profit d'une meilleure qualité de vie pour les habitants
Type de descriptif des données	Descriptif général – liste sans descriptif détaillé de chaque donnée
Type de données transport	Liste des arrêts, horaires et parcours de tous les bus et tramway de la SEMITAN circulant sur le territoire de Nantes Métropole.
Format d'échange	<p>Les jeux de données sont accessibles sur data.nantes.fr, soit en téléchargeant des fichiers, soit en utilisant une interface de programmation (API).</p> <p>Données Tabulaires</p> <ul style="list-style-type: none"> • CSV (Comma Separated Value) : format ouvert • XLS (Excel) : format propriétaire (Microsoft), standard de fait <p>Données complexes</p> <ul style="list-style-type: none"> • XML (<i>Extensible Markup Language</i>) : format ouvert, standard de fait, recommandé par le RGI • JSON (<i>JavaScript Object Notation</i>) : format ouvert, standard de fait, recommandé par le RGI dans le cas de services REST <p>Images</p> <ul style="list-style-type: none"> • JPEG (<i>Joint Photographic Experts Group</i>) : format propriétaire, standard de fait, recommandé par le RGI • PNG (<i>Portable Network Graphics</i>) : format ouvert, recommandé par le RGI • GIF (<i>Graphic Interchange Format</i>) : format ouvert, recommandé par le RGI

	<p>Données géographiques</p> <ul style="list-style-type: none"> • SHP (<i>Shape file</i>) : format propriétaire (ESRI), format ouvert, standard de fait • KML/KMZ (<i>Keyhole Markup Language</i>) : format ouvert, standard de fait <p>Les données géographiques sont proposées dans différents systèmes :</p> <ul style="list-style-type: none"> • CC47 (<i>Conique Conforme 47</i>) : projection associée au système géodésique RGF93, coordonnées planes > SHP • L93 (<i>Lambert 93</i>) : projection associée au système géodésique RGF93, coordonnées planes > SHP • WGS84 (<i>World Geodetic System 1984</i>) : système géodésique, coordonnées géographiques > KML/KMZ, CSV <p>Données de transport</p> <ul style="list-style-type: none"> • GTFS (General Transit Feed Specification) : format ouvert, standard de fait
Période de validité	Indiquée pour les horaires particuliers Descriptif des calendriers non présent
Mise à jour	Variable / Ponctuelle Données « TR » : MAJ toutes les 5 min Horaires 2011 (sans période de validité) Horaires été Horaires rentrée
Type de services	Réseaux de tramway (41 kilomètres de lignes exploitées depuis 1985), une ligne de bus en site propre (Busway), un réseau d'autobus performants (85 % des autobus fonctionnent au gaz naturel pour véhicules) et des services fluviaux (Navibus)
Résultats de l'import du fichier GTFS dans le logiciel CHOUETTE	cf Annexe 2
Emprise géographique	La Semitan a pour mission principale l'exploitation du réseau de transport en commun sur les 24 communes de l'agglomération nantaise.
Liste des données	Temps réel : arrêts, temps d'attente à une zone d'arrêt, horaires Théorique : arrêts, horaires, circuits : arrêts, horaires, parcours (bus, tramway) Disponibilité dans les parkings publics
Plus d'info	Forum http://data.nantes.fr/forum/
Contributeurs/ Partenaires	Pays de la Loire Semitan
Collecte	-
Réutilisation	Référence à la loi n° 78-753 du 17 juillet 1978 portant diverses mesures d'amélioration des relations entre l'administration et le public et diverses dispositions d'ordre administratif, social et fiscal

3.2. Commentaires généraux

Points positifs

Bonne vue d'ensemble des données publiées ainsi que des formats utilisés.

Points négatifs

Pas d'indication de la fréquence de mise à jour pour chaque donnée

Pas de gestion des calendriers

Pas d'information sur l'emprise géographique, carte

Absence de liste des données échangées avec leurs définitions et propriétés (sans entrer dans l'API)

4. Rennes Métropole

4.1. Fiche analytique

RENNES Métropole	
Nom réseau	STAR (Keolis)
Adresse site	www.data.rennes-metropole.fr et data.keolis-rennes.com
Objectifs initiaux (basé sur les textes extraits du site)	<p><u>Origine de la démarche</u></p> <p>Trois bonnes raisons d'ouvrir les données publiques :</p> <ol style="list-style-type: none"> ce patrimoine immatériel reste largement ignoré du grand public alors que cette somme d'informations est de nature à améliorer le quotidien des citoyens, à accroître la transparence, à créer de la valeur d'usage. Deux phénomènes vont accélérer ces développements : <ul style="list-style-type: none"> l'accès à Internet via les téléphones mobiles, qui connaît une très forte croissance, entraînera de nouveaux usages d'accès aux informations liées au territoire ; la demande participative des citoyens est l'expression d'un profond changement de société qui entraîne là aussi de nouveaux rapports avec le secteur public ; c'est aussi une formidable opportunité d'associer les habitants à une démarche de coélaboration et de participation ouverte, d'augmenter la capacité innovatrice, de libérer les forces créatives des acteurs rennais qu'ils soient associatifs, économiques, sociaux, culturels, professionnels, étudiants, usagers, habitants ou citoyens ; enfin, c'est participer au développement de l'attractivité de la métropole rennaise et poursuivre la construction d'une ville novatrice et audacieuse où s'inventent la société et les usages de demain.
Données transport	<p>Fichiers Horaires du réseau STAR (KEOLIS)</p> <p>Données géographiques :</p> <ul style="list-style-type: none"> données géographiques du réseau STAR <ul style="list-style-type: none"> CSV, DXF (WGS84), SHP (L93), SHP (WGS84) aires de stationnement <ul style="list-style-type: none"> DXF (WGS84), SHP (L93), SHP (WGS84) aires de stationnement payant <ul style="list-style-type: none"> DXF (WGS84), SHP (L93), SHP (WGS84) localisation des stations de prêt du système de vélo en libre service (VLS) "Le vélo STAR" de Rennes Métropole. <ul style="list-style-type: none"> DXF (WGS84), SHP (L93), SHP (WGS84), CSV <p><u>Autres données disponibles</u></p> <p>Données LE vélo STAR</p> <ul style="list-style-type: none"> Places disponibles aux stations - TR Vélos disponibles aux stations - TR Type d'équipement à la station (avec ou sans CB) - DS État de fonctionnement des stations - TR Actualités diffusées sur le site www.levelostar.fr - TR <p>Données du réseau STAR</p> <ul style="list-style-type: none"> Horaires de passage des bus à chaque arrêt - DS et TR Topologie du réseau (lignes, arrêts, sens, terminus) - DS Couleurs et types de lignes - DS Pictogrammes des lignes - DS Géolocalisation des arrêts - DS <p>La ligne a du Métro</p> <ul style="list-style-type: none"> Topologie de la ligne du métro : stations, direction, terminus - DS Géolocalisation des stations - DS État des stations de métro : ouverture-fermeture - TR Cartographie des équipements (ascenseurs, escalators...) - DS État de fonctionnement des ascenseurs, escalators aux stations - TR

	<p>Les Parcs Relais</p> <ul style="list-style-type: none"> • Places disponibles pour chaque parc - TR • Horaires de fonctionnement des parcs relais - TR • État des parcs relais : ouverture-fermeture - TR • Géolocalisation des parcs relais - DS <p>Les points de vente</p> <ul style="list-style-type: none"> • Géolocalisation des 3 agences commerciales STAR et du point KorriGo (gare) - DS • Horaires de fonctionnement des agences - DS • Coordonnées des agences - DS • Géolocalisation des points de ventes dépositaires - DS <p>L'information trafic</p> <ul style="list-style-type: none"> • Les déviations en cours - DS (avec mises à jour) • Les perturbations sur les lignes - DS (avec mises à jour) • Les reports d'arrêts - DS (avec mises à jour) <p><i>TR: temps réel</i> <i>DS : données statique</i></p>
Format d'échange	XML : GTFS CSV, DXF, SHP
Résultats de l'import du fichier GTFS dans le logiciel CHOUETTE	<i>CF Annexe 2</i>
API	<p><u>Fonctionnalités de l'API</u></p> <ul style="list-style-type: none"> • API Version 1 <ul style="list-style-type: none"> ◦ getstation ◦ getdistrict • API Version 2 <ul style="list-style-type: none"> ◦ getbikestations ◦ getbikedistricts ◦ getlinesalerts ◦ getlines ◦ getequipments ◦ getequipmentsstatus ◦ getmetrostations ◦ getmetrostationsstatus ◦ getrelayparks ◦ getpos ◦ getcities ◦ getcitydistricts • API Version 2.1 <ul style="list-style-type: none"> ◦ getbusnextdepartures
Plus d'info	http://www.data.rennes-metropole.fr
Propriétaire de la donnée	Rennes Métropole / Keolis
Contenus	Horaires théoriques et points d'arrêt des lignes du réseau STAR de Keolis Rennes.
Contributeurs	Keolis Rennes
Collecte	
Licence ou conditions de réutilisation	ODbl
Fréquence de mise à jour	A chaque changement : « chaque fois que nécessaire »
Réutilisations identifiées	ArceauVélo, EnVéloàRennes,, eo'City, Go2rennes, Handimap.org, l'arretPublic, itinerennes mobil'rennestout rennesboug transportsRennes RennesBus, VelocityRennes,

4.2. Commentaires généraux

Points positifs

- Bonne vue d'ensemble



















Points négatifs

- Pas d'indication de fréquence de mise à jour
- Pas de description de gestion des calendriers
- S'appuie exclusivement pour ce qui concerne les transports sur les fournitures Keolis : tram, métro, bus, vélo. C'est regrettable compte tenu de la richesse de données dont dispose la ville et qu'elle pourrait mixer avec Dor Breizh, les données du CIGT des voies rapides (autoroutes et périphérique) autour de Rennes.

5. Toulouse TISSEO

5.1. Fiche analytique

TOULOUSE	
Nom réseau	TISSEO
Adresse site	http://data.grandtoulouse.fr
Objectifs initiaux (basé sur les textes extraits du site)	<p><u>Objectifs :</u></p> <ul style="list-style-type: none"> • une plus grande transparence démocratique • un important gain de croissance car elle facilite, le développement d'applications utilisant ces données comme matériaux de base. <p><u>Justification :</u></p> <p>Partant du constat que l'ouverture des données créera entre les collectivités et les citoyens une relation nouvelle, qu'il s'agit d'une opportunité économique pour le tissu très riche en entreprises innovantes de Toulouse et de son agglomération et au-delà de l'obligation légale qui est la leur, les collectivités toulousaines (la Communauté Urbaine du Grand Toulouse et ses communes membres) ont décidé de libérer leurs données.</p> <p><u>Prise de décision :</u></p> <p>La mairie de Toulouse a délibéré en ce sens le 23 septembre 2011, la Communauté Urbaine du Grand Toulouse le 29 septembre. D'autres communes et partenaires vont s'inscrire dans cette dynamique prochainement. L'annonce publique de l'ouverture du portail Grand Toulouse Data a été faite au cours de la Novela le 22 octobre 2011.</p> <p><u>La démarche</u></p> <p>Le projet « Grand Toulouse Data » a démarré à la fin du premier trimestre 2011. Il a été souhaité de donner rapidement les outils nécessaires à la libération des données publiques toulousaines en commençant par le choix de la licence sous laquelle les données seront réutilisables. Le choix des jeux de données candidats à l'ouverture est fait de façon pragmatique en commençant par celles immédiatement disponibles, d'autres viendront enrichir le panel au fur et à mesure.</p> <p><u>Formats :</u></p> <p>Les données sont publiées principalement sous des formats libres ou à tout le moins très répandus, afin, là aussi, de faciliter leur réutilisation.</p> <p><u>Confidentialité :</u></p> <p>L'open data consiste donc pour les administrations publiques, ainsi que pour les entreprises privées exerçant une mission de service public, à mettre à la disposition de tous les publics l'ensemble de leurs données qui ne seraient pas protégées par ailleurs. Ces protections sont la protection des données à caractère personnel ou les documents protégés par un droit de propriété intellectuelle, ou encore mettant en jeu la sécurité publique.</p> <p>Il s'agit donc de tirer partie d'une obligation légale afin de favoriser la transparence démocratique imaginée par la Déclaration des Droits de l'Homme et du Citoyen de 1789, mais aussi de favoriser le développement économique de notre territoire.</p> <p><u>Accès aux décisions</u></p> <p>L'ensemble des délibérations de la Communauté Urbaine du Grand Toulouse depuis janvier 2000 est disponible sur le site www.grandtoulouse.fr ou directement en cliquant ici.</p> <p>L'ensemble des délibérations de la ville de Toulouse depuis mars 2002 est également disponible sur le site www.toulouse.fr.</p> <p>Toutes les délibérations sont accessibles via un moteur de recherche permettant de rechercher par date de conseil municipal ou communautaire et par mots clés. Afin de ne pas multiplier les espaces de stockage d'informations similaires elles ne sont pas dupliquées sur le portail Open Data. Comme il s'agit d'informations officielles ne devant pas être modifiées, le format utilisé est un format figé (pdf) contrairement à ce qui est mis en ligne sur ce portail.</p>

Type de descriptif des données	Général																														
Type de données transport	Données descriptives de l'offre de transports (lignes, arrêts, horaires) de Tisséo SMTC. Particularités des lignes : <ul style="list-style-type: none"> certains horaires lignes (A, B) sont exprimés sous forme de fréquences certaines lignes (TAD 105, 106, 118, 119, 120, 150, 201, 202, 204, 205) sous soumises à réservatio ; les arrêts ne sont desservis qu'en cas de réservation certaines lignes (106, 118, 119, 120) correspondent à des TAD "Zonaux". Les horaires au sein d'une zone sont identiques. 																														
Format d'échange	Les données sont publiées au format Trident. Présence d'un renvoi sur : http://www.chouette.mobi/spip.php?rubrique61																														
Résultats de l'import du fichier Trident dans le logiciel CHOUETTE	Cf. Annexe 2 Le zip ne peut pas être importé dans CHOUETTE sans faire une conversion de l'encodage des caractères et corriger l'en-tête des fichiers XML en ISO-8859-1. <u>Compte rendu de l'import</u> <table> <thead> <tr> <th>• Eléments</th> <th>Qté</th> </tr> </thead> <tbody> <tr> <td>Réseaux</td> <td>1</td> </tr> <tr> <td>Transporteurs</td> <td>1</td> </tr> <tr> <td>Groupes de lignes</td> <td>0</td> </tr> <tr> <td>Lignes</td> <td>3</td> </tr> <tr> <td>Courses</td> <td>136</td> </tr> <tr> <td>Calendriers</td> <td>136</td> </tr> <tr> <td>Correspondances</td> <td>65</td> </tr> <tr> <td>Arrêts</td> <td>142</td> </tr> <tr> <td> points d'embarquement</td> <td>0</td> </tr> <tr> <td> quais</td> <td>89</td> </tr> <tr> <td> arrêts commerciaux</td> <td>53</td> </tr> <tr> <td> pôles d'échange</td> <td>0</td> </tr> <tr> <td> ITL</td> <td>0</td> </tr> <tr> <td> points d'accès</td> <td>0</td> </tr> </tbody> </table> <p>Calendriers échus : 46</p> <p><u>Quelques problèmes du contenu :</u></p> <ul style="list-style-type: none"> Grande complexité de gestion de calendriers (1 calendrier par course ?) : mériterait une explication Parfois deux correspondances (« connection links ») apparaissent avec deux temps de correspondance (en fonction du type d'utilisateur ?) ; François Verdier – François Verdier, l'un des « connection links » semble erroné (composé d'un point seulement). 50 erreurs non fatales pour l'import et identiques apparaissent: certains arrêts ne sont pas enregistrés car une clé est considérée comme non unique. 	• Eléments	Qté	Réseaux	1	Transporteurs	1	Groupes de lignes	0	Lignes	3	Courses	136	Calendriers	136	Correspondances	65	Arrêts	142	 points d'embarquement	0	 quais	89	 arrêts commerciaux	53	 pôles d'échange	0	 ITL	0	 points d'accès	0
• Eléments	Qté																														
Réseaux	1																														
Transporteurs	1																														
Groupes de lignes	0																														
Lignes	3																														
Courses	136																														
Calendriers	136																														
Correspondances	65																														
Arrêts	142																														
 points d'embarquement	0																														
 quais	89																														
 arrêts commerciaux	53																														
 pôles d'échange	0																														
 ITL	0																														
 points d'accès	0																														
Emprise géographique	Pas explicitée																														
Liste des données	Sans descriptif détaillé																														

Plus d'info	Pour toute question, vous pouvez adresser vos mails à OpenData@tisseo.fr
Contributeurs	-
Collecte	-
Réutilisation	Licence libre ODbL qui sera applicable par défaut. L'acceptation de la licence est un préalable au téléchargement des données.

5.2. Commentaires généraux

Points positifs

- Bonne vue d'ensemble
- Données assez complètes

Points négatifs






- Pas de liste suffisamment précise
- Pas d'explicitation du territoire couvert ni de carte
- Des défauts de conformité aux formats proposés

6. TransGironde

6.1. Fiche analytique

TRANSGIRONDE	
Nom réseau	TransGironde
Adresse site	http://www.datalocale.fr/
Objectifs initiaux (basés sur les textes extraits du site)	<p><u>Objectif : interopérabilité informationnelle</u> Le Conseil général de la Gironde et le Conseil régional d'Aquitaine font donc le choix d'une exploitation durable et valorisante de leurs ressources informationnelles, par la mise en place d'une plate-forme de publication mutualisée, en cherchant à optimiser leur réutilisation et leur accessibilité.</p> <p>Les deux partenaires ont signé une convention les liant pour deux années d'expérimentation au terme desquelles une évaluation de la démarche sera conduite.</p> <p>La publication sur Internet des données publiques s'inscrit dans la continuité du mouvement de production puis de dématérialisation massive de données publiques par les administrations.</p> <p>Elle s'appuie sur l'adoption de standards ouverts et l'utilisation d'Internet comme vecteurs de communication.</p> <p>Elle amplifie le cadre légal qui s'impose aux collectivités publiques en matière de diffusion des données publiques. Toutefois, au-delà des simples obligations légales, on peut constater qu'un grand nombre de ces données n'est pas suffisamment exploité par le public alors que ce patrimoine pourrait améliorer son information, voire créer un environnement favorable à l'innovation et, plus largement, à l'économie.</p> <p><u>Enjeu : coconstruction durable de l'administration électronique</u> En terme d'usage, les collectivités bénéficieront de cette démarche par une amélioration de la connaissance, par exemple en matière de compréhension des dynamiques des territoires.</p> <p>Elles pourront également envisager une meilleure coconstruction de l'administration en encourageant le besoin d'enrichissement du débat citoyen grâce à une plus grande transparence de leur action publique. Cette démarche participera à la création de nouveaux services nés de la combinaison et d'une réutilisation originale de ces données. Enfin elle contribuera à l'appropriation par tout un chacun des outils dans un monde fortement numérisé.</p> <p><u>Programme d'action :</u></p> <ul style="list-style-type: none"> • organiser la diffusion d'innovations en impliquant tous • expérimenter sur l'administration électronique • favoriser l'ouverture et la découverte des richesses culturelles de la Gironde • favoriser l'accès pour tous aux savoirs et au débat public • renforcer la communication avec les Girondins via les supports existants • être exemplaire sur le mode de fonctionnement interne en termes de démarche de qualité <p><u>Expérimentation : apprendre et évaluer (...)</u> Les collectivités partenaires du projet ne prétendent pas avoir identifié ou mis en œuvre l'ensemble des critères permettant d'inscrire cette démarche dans la durée. Beaucoup de questions restent encore en suspens concernant la cible à atteindre: notamment sur les ressources humaines et financières dont elles disposent pour mener à bien cet objectif ambitieux, sur la volonté des ré-utilisateurs de s'inscrire dans une démarche de partage et de redistribution, sur la capacité des citoyens à saisir cette opportunité de dialogue constructif.</p> <p>Les partenaires sont conscients de l'état d'imperfection de certaines des données mises à disposition ainsi que des améliorations qui pourront être apportées à la plate-forme.</p> <p>Ils comptent sur la participation active des ré-utilisateurs pour parvenir à l'objectif de qualité affiché.</p>

	<p>A cet effet, des formulaires de contact sont à votre disposition pour interagir avec les services en charge de la gestion des données ainsi qu'un espace d'échange pour encourager la discussion.</p> <p>La feuille de route du projet est disponible sur cette page pour vous permettre de connaître les engagements des partenaires du projet ainsi que les liens avec d'autres initiatives locales ou internationales.</p> <p>Le Conseil général de la Gironde et le Conseil régional d'Aquitaine sont partenaires de la plate-forme d'information géographique PIGMA, et de la Banque numérique du Savoir d'Aquitaine BNSA.</p> <p>Le Conseil général de la Gironde est également partenaire de la bibliothèque numérique européenne EUROPEANA au sein de laquelle sont valorisées les ressources des Archives départementales de la Gironde (lien vers les collections moissonnées via le protocole OAI-PMH).</p>												
Type de descriptif des données	<p>Existence des métadonnées <i>présentées sous forme de fiches claires et complètes</i></p> <p>Liste des informations générales et métadonnées :</p> <ul style="list-style-type: none"> • Formulaire de contact • Contribution à la libération des jeux de données • Horaires des lignes régulières du réseau transgironde • Lignes régulières de transport public TransGironde • Localisation des points d'arrêts des lignes régulières du réseau TransGironde • Pré-visualisation du jeu de données • Liste des communes du département • Somme de contrôle jeu de données 												
Type de données transport	<p><i>horaires, points d'arrêt, lignes</i></p> <p>Les fiches/métadonnées donnent les définitions des différentes données :</p> <ul style="list-style-type: none"> • arrêts : il s'agit des points d'accès • horaires : il s'agit des lignes d'autocars • lignes : régulières et scolaires 												
Format d'échange	TRIDENT et NEPTUNE (lignes d'autocar)												
Période de validité	<i>Indiquée</i>												
Mise à jour	Fréquence indiquée si la donnée est mise à jour de façon régulière												
Type de services	<p><i>Il s'agit ici des « services » proposés par le site :</i></p> <p><u>Exemples de réutilisation des jeux de données</u> : une rubrique est prévue pour renseigner les réutilisations.</p> <ul style="list-style-type: none"> • Le site fournit des « outils » du type : <ul style="list-style-type: none"> ○ <u>Formulaire de contact</u> : le site fournit un canal en ligne de remontée d'information. ○ <u>Contribution à la libération des jeux de données</u> : il est possible de demander ou de proposer la mise en ligne de nouveaux jeux de données. 												
Résultats de l'import du fichier Trident dans le logiciel CHOUETTE	<p><i>Import globalement réussi. Cf. Annexe 2</i></p> <p><u>Résumé de l'import</u></p> <p>Code : cg33</p> <p>Préfixe des identifiants Neptune : CG33</p> <p>Système de référence spatiale optionnel (SRID) :</p> <p>Fuseau horaire : Paris</p> <p>Période de validité : du 30/01/2012 au 31/08/2012</p> <table> <tbody> <tr> <td>Éléments</td> <td>nté</td> </tr> <tr> <td>Réseaux</td> <td>1</td> </tr> <tr> <td>Transporteurs</td> <td>1</td> </tr> <tr> <td>Lignes</td> <td>91</td> </tr> <tr> <td>Courses</td> <td>2127</td> </tr> <tr> <td>Calendriers</td> <td>456</td> </tr> </tbody> </table>	Éléments	nté	Réseaux	1	Transporteurs	1	Lignes	91	Courses	2127	Calendriers	456
Éléments	nté												
Réseaux	1												
Transporteurs	1												
Lignes	91												
Courses	2127												
Calendriers	456												

	<p>Correspondances 0</p> <p>Arrêts 2413</p> <p> points d'embarquement 2186</p> <p> quais 0</p> <p> arrêts commerciaux 227</p> <p> pôles d'échange 0</p> <p> ITL 0</p> <p><i><u>Quelques incohérences du contenu détectées</u></i></p> <ul style="list-style-type: none"> • Exemple 1 : ligne BORDEAUX BUTTINIERE BLAYE : <ul style="list-style-type: none"> ○ aucune séquence d'arrêts en sens opposé ○ horaires : séquences d'arrêts présentes dans les tableaux horaires pour les deux sens (cependant la numérotation des courses sans doute erronée (p.ex. plusieurs courses 1 dans un même calendrier) • Exemple 2 : erreurs dans coordonnées (ex. arrêt AMBARES Bellevue : point d'embarquement sur cours d'eau)
Emprise géographique	379 communes desservies Liste des communes du département : indiquée et visualisée
Liste des données	<p>Horaires des lignes régulières du réseau TransGironde Les jeux de données mis à disposition permettent de disposer de la liste des arrêts d'autocar du réseau TransGironde avec pour chaque ligne du réseau, l'indication des points d'arrêts et des horaires théoriques de passage. Ces informations sont structurées au format Trident.</p> <p>Lignes régulières de transport public TransGironde Information géographique En sa qualité d'autorité organisatrice des transports publics interurbains, le Conseil général de la Gironde gère 64 lignes régulières et des centaines de services scolaires.</p> <p>Localisation des points d'arrêts des lignes régulières du réseau TransGironde Information géographique Ce jeu de données fournit la localisation des arrêts des lignes régulières du réseau de transport public TransGironde.</p>
Plus d'info	
Contributeurs	<p>Direction des transports terrestres - Conseil Général de la Gironde :</p> <p>Localisation des points d'arrêts des lignes régulières du réseau TransGironde</p> <p>Lignes régulières de transport public TransGironde</p> <p>Horaires des lignes régulières du réseau TransGironde</p> <p>(Les principaux contributeurs sont indiqués par service, type de données fourni).</p>
Collecte	Cf. contributeurs
Réutilisation / Qualité	<p>Les objectifs suivants sont mentionnés :</p> <ul style="list-style-type: none"> • démarche de coconstruction avec les utilisateurs • possibilité d'indiquer les réutilisations • liste structurée/référentiel des bonnes pratiques (13 rubriques, 3 priorités) pour faciliter la réutilisation

6.2. Commentaires généraux

Points positifs :

- Annonce ou liste des usages possibles et actuels
- warning sur l'imperfection des données
- outil pour permettre de mettre en ligne de nouveau jeux de données
- métadonnées présentées par des fiches descriptives des données (claires comportant une définition des données)
- liens vers les textes normatifs
- référentiel de bonnes pratiques avec 3 priorités.

Points négatifs :

- pas de description des calendriers
- quelques incohérences des contenus des données.

7. Exemples d'initiatives open data à l'étranger

L'open data dans les transports existe à l'étranger.

Quelques exemples illustrent bien tout l'intérêt mais aussi la diversité de cette démarche.

Londres montre l'intérêt économique pour l'AOT et l'exploitant, New York montre les conditions minimales d'utilisation et la complétude de l'information utile aux réutilisateurs.

Ces deux premiers exemples peuvent être considérés comme des succès et confirment une partie des analyses et des recommandations faites dans le rapport.

7.1. Exemple de Londres

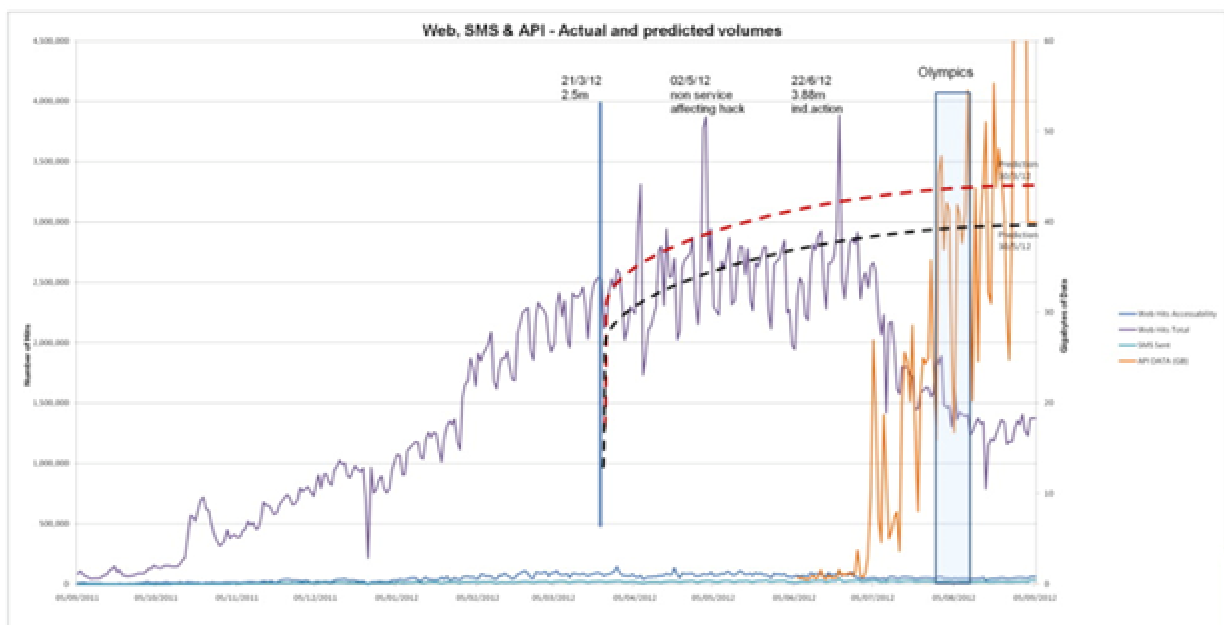
<http://www.tfl.gov.uk/businessandpartners/syndication/16492.aspx>

TfL (Transport for London) a commencé à ouvrir ses données en juin 2010 et a réussi à attirer de nombreuses réutilisations de ses API en temps réel par divers sites web et applications mobiles qui ont généré un nombre de requêtes tel que, pendant les jeux olympiques de 2012, le surcroît de demandes d'information a pu être traité à moindre coût par TfL, c'est à dire sans réinvestir massivement dans ses systèmes d'information. Trente applications ont été développées par des tierces parties qui ont généré une audience supplémentaire qui a continué à croître depuis la fin des Jeux. Le site TfL en revanche a retrouvé ces étages de février 2012 avant le début de la forte croissance d'information sur les bus en temps réel.

Le lien ci-dessus illustre bien les informations demandées aux réutilisateurs avant qu'ils puissent télécharger les jeux de données.

Le graphique qui suit montre l'évolution des connexions web et mobile à Londres depuis l'ouverture des données par TfL. Il montre selon Simon Reed de TfL qui l'a présenté au Congrès ITS de Vienne que le succès rencontré pour les JO d'été ne s'est pas démenti depuis, mais qu'il est largement imputable aux succès des nouvelles applications inaugurées lors de cet événement. C'est selon lui que les londoniens s'apercevant que mieux informés, ils peuvent traverser une période critique (+ 30 % d'affluence dans les transports publics) sans trop de nuisances.

Live Bus Information - Demand



7.2. Exemple de New York City

7.2.1. NYC Open Data

NYC Open Data met à disposition des données publiques riches générées par plusieurs agences de la ville de New York pour une utilisation publique.

N'importe qui peut utiliser ces ensembles de données pour participer et améliorer le gouvernement en menant des recherches et des analyses ou la création d'applications, ce qui permet une meilleure compréhension des services fournis par les agences, l'amélioration de la vie des citoyens et de la façon dont le gouvernement les sert.

Les jeux de données sont disponibles dans une variété de formats et sont mis à jour lorsque de nouvelles données deviennent disponibles. Les données sont présentées par catégorie, par agence, ou autre organisation de la ville. La description des données, la méthode de collecte, et d'autres documents contextuels appelés métadonnées, rendent ces jeux de données plus faciles à comprendre et à utiliser. L'utilisation de ces jeux de données et la façon dont ils sont générés peuvent être mieux comprise à la lecture des conditions d'utilisation.

Conditions d'utilisation

Les conditions d'utilisation suivantes s'appliquent aux visiteurs du portail NYC Open Data et les développeurs d'applications qui obtiennent des données de la ville à travers ce portail web unique.

En téléchargeant des ensembles de données accessibles, l'utilisateur accepte les conditions d'utilisation du NYC.gov ainsi que sa politique de confidentialité. L'utilisateur s'engage également à toutes autres conditions d'utilisation définies par les entités qui fournissent des données ou des services à travers le site. Les entités fournissant des données comprennent, sans s'y limiter, les agences, les bureaux, les bureaux, départements et autres entités discrètes de New York City. Les données publiques mises à disposition sur le portail NYC Open Data sont fournies à titre informatif.

La ville ne garantit pas l'exhaustivité, l'exactitude, le contenu, ou l'adéquation à un usage particulier des données publiques mises à disposition sur le portail NYC Open Data. Ce refus de garanties doit être étendu à l'ensemble des données publiques fournies par ce portail ou par une tierce partie.

Les agences de la ville constituent la source officielle des données disponibles sur Open Data New York. Ces entités sont responsables de la qualité des données et du contrôle de la version des ensembles de données et fournitures disponibles sur le site. Les données peuvent être mises à jour, corrigées, écrasées à tout moment. La fréquence de mise à jour attendue est indiquée pour chaque ensemble de données sur le site. Les anciennes versions de jeux de données ne seront pas conservées.

7.2.2. MTA Data

<https://nycopendata.socrata.com/d/mmu8-8w8b>

Le jeu de données transport mis à disposition est très complet.

[Description](#)

Information concernant MTA (Metropolitan Transportation Authority of the State of New York) métro, bus, trains RER, ponts et tunnels

Activité : visites 6 204, téléchargements 3 243.

7.2.2.1. Flux de données en temps réel

7.2.3. MTA All-Agence

- [Statut du service](#) - actualisés toutes les minutes

7.2.4. MTA NYC Transit

- [Métro](#) (beta)
- [MTA Bus Time](#)™

- [Ascenseur / Statut Escalator](#) - rafraîchie toutes les 5 minutes
 - [Tag / Définitions de champs pour XML Feed](#)

7.2.4.1. Flux de données horaires et statiques

au 18/03/13

7.2.5. MTA toutes-Agences

MTA métro, de bus et de trains de banlieue GTFS calendrier de données est mise à jour en fonction des évolutions de services, en moyenne, environ tous les 4 mois. Nous vous recommandons de télécharger des fichiers GTFS dans Internet Explorer v8 et plus, Firefox, Safari, Chrome ou Opera.

- [GTFS](#) annexe données
 - [New York City Transit Subway](#) - Mise à jour Février 11, 2013
 - New York City Transit Bus
 - [Bronx](#) - Mise à jour Janvier 17, 2013
 - [Brooklyn](#) - Mise à jour Janvier 17, 2013
 - [Manhattan](#) - Mise à jour Janvier 17, 2013
 - [Queens](#) - Mise à jour Janvier 17, 2013
 - [Staten Island](#) - Mise à jour Janvier 17, 2013
 - [Long Island Rail Road](#) - Mise à jour Mars 18, 2013
 - [GTFS données au format JSON](#)
 - [Programmer seule version - format JSON](#)
 - [GTFS données au format XML](#)
 - [Programmer seule version - format XML](#)
 - [Metro-North Railroad](#) - Mise à jour Février 12, 2013
 - [Bus Company](#) - Mise à jour Mars 12, 2013
- [Les données du Programme de capital tableau de bord](#)
- [MTA Arts transit pour les ressources](#)
- [Couleurs Ligne MTA](#) - Mise à jour Novembre 21, 2011
- [Les données de performance](#) - Mise à jour Juillet 11, 2011
 - [Liste des ensembles de données dans ce paquet](#) (117KB)

7.2.6. MTA NYC Transit

- [Les données tarifaires](#) - rafraîchie hebdomadaires
- Mouvements de trains historiques pour 1/2/3 4/5/6 et trains pour mai, 2011
 - [Documentation](#)
 - [Données historiques mai 2011](#) (23 Mo)
- Shapefiles
 - [Les fichiers de formes pour circuits d'autobus](#) - Mise à jour Juillet 26, 2012

AFIMB

- [Données SIG entrées et sortie](#) du métro - Mise à jour Juillet 19, 2011
- [Les données d'utilisation tourniquet](#) – mises à jour hebdomadaires

7.2.7. MTA ponts et tunnels

- [Données quotidiennes trafic Plaza](#)
- [Emplacements des détaillants E-ZPass](#)

Lien

<https://nycopendata.socrata.com/Transportation/MTA-Data/mmu8-8w8b>

Annexe 2

Résultats des tests sur les jeux de données

Contenu

1. Résultats des tests d'import de données géographiques.....	52
1.1. Import de fichiers au format SHP décrivant des points d'arrêt	52
1.1.1. CG 33 Transgironde	52
1.1.2. Loir et Cher : CG41.....	52
1.1.3. ACCM Arles.....	52
1.1.4. CUB Bordeaux.....	53
1.2. Import de fichier au format CSV décrivant des points d'arrêt.....	53
1.2.1. Paris-Ile de France RATP	53
1.2.2. Montpellier TAM.....	53
1.2.3. SNCF gares (France)	53
1.2.4. LILA (CG 44).....	54
1.2.5. Nantes TAN.....	54
1.2.6. Rennes STAR	54
1.3. Import de fichiers au format SHP décrivant un linéaire	54
1.3.1. CG33 pistes cyclables.....	54
2. Résultats des test d'import de données transports, déplacement dans un validateur 55	
2.1. Validation de données au format GTFS par le GTFS FeedValidator de Google.....	55
2.1.1. CG33 Transgironde	55
2.1.2. Toulouse TISSEO:.....	55
2.1.3. LILA - CG 44	55
2.1.4. NANTES TAN	55
2.1.5. RENNES STAR	56
2.1.6. LA CUB.....	56
2.1.7. Conclusion	56
2.2. Validation des données au format Neptune par la validateur CHOUETTE	57
2.2.1. CG33 Transgironde	57
2.2.2. Toulouse TISSÉO	57
2.2.3. CG44 LILA	57
2.2.4. NANTES TAN	57
2.2.5. RENNES STAR	58
2.2.6. La CUB :.....	58
2.2.7. Conclusion.....	58
2.3. Conversion de données au format GTFS en données au format NEPTUNE par l'import-export de CHOUETTE	59
2.3.1. LA CUB :	59

A l'appui de l'étude des initiatives « Open Data » dans le domaine de la mobilité, nous avons réalisé quelques tests sur les jeux de données mis à disposition.

En conclusion il s'avère que les jeux de données mis à disposition sont rarement exploitables, car ils ne passent pas les tests de validation.

Plusieurs plates-formes renvoient les réutilisateurs vers les sites de Google Transit et chouette.mobi pour se familiariser avec les formats de données, afin peut être de procéder aux modifications eux-mêmes.

1. Résultats des tests d'import de données géographiques

Nous n'avons pas traité de fichiers XML pour ces premiers exemples, car il aurait fallu passer par des convertisseurs pour exploiter ces formats dans un SIG.

Pour le KML spécifiquement, qui est un format XML, il est possible de le traiter avec des outils libres tels que QGIS, et de réexporter la couche en Shapefile.

Les CSV, textes tabulés, ou DBF peuvent être importés sans problèmes avec les outils standards de MacMap® et de la plupart des SIG. Pour un format XLS, il faut convertir au préalable dans un des trois formats précités.

1.1. Import de fichiers au format SHP décrivant des points d'arrêt

1.1.1. CG 33 TransGironde

Catégorie	Résultat
Nature des données	Localisation des points d'arrêts des lignes régulières du réseau TransGironde
URL	http://www.datalocale.fr/dataset/ig_transgironde_pa
Format	ShapeFile en Lambert93
Info donnée sur la qualité	Bonne
Résultat de l'import	L'import est réalisé sans problème. Les données sont cohérentes.

1.1.2. Loir et Cher : CG41

Catégorie	Résultat
Nature des données	Points d'arrêt du réseau de transport départemental "Route41"
URL	http://www.data.gouv.fr/donnees/view/Points-d%27arr%C3%AAt-du-r%C3%A9seau-de-transport-d%C3%A9partemental-%22Route41%22-30383156
Format	ShapeFile en Lambert93
Info donnée sur la qualité	Aucune
Résultat de l'import	L'import est réalisé sans problème. Les données ne sont pas complètes.

1.1.3. ACCM Arles

Catégorie	Résultat
Nature des données	Points d'arrêts du réseau de transport de l'ACCM
URL	http://opendata.regionpaca.fr/donnees/detail/points-darrets-du-reseau-de-transport-de-laccm.html
Format	ShapeFile en Lambert93
Info donnée sur la qualité	Aucune
Résultat de l'import	L'import est réalisé sans problème. Les données sont cohérentes.

1.1.4. CUB Bordeaux

Catégorie	Résultat
Nature des données	Arrêt physique sur le réseau
URL	http://data.lacub.fr/data.php?themes=10
Format	ShapeFile en Lambert93
Info donnée sur la qualité	Aucune, mais visualisation directe possible
Résultat de l'import	L'import est réalisé sans problème. Les données sont cohérentes.

1.2. Import de fichier au format CSV décrivant des points d'arrêt

1.2.1. Paris-Ile de France RATP

Catégorie	Résultat
Nature des données	Positions géographiques des stations du réseau RATP
URL	http://www.ratp.fr/opendata/
Format	CSV en longitude/latitude (WGS84)
Info donnée sur la qualité	
Résultat de l'import	Le contenu des données est exploitable. Pour plus de précisions, voir : http://openstreetmap.fr/blogs/cquest/stations-ratp

1.2.2. Montpellier TAM

Catégorie	Résultat
Nature des données	Localisation des points d'arrêts du réseau bus et tramways
URL	http://modulaweb.fr/apitam/
Format	Format texte spécifique, où les coordonnées sont fournies en longitude et latitude (WGS84).
Info donnée sur la qualité	
Résultat de l'import	Le format étant relativement bien structuré, on peut le modifier à la main et récupérer les données.

1.2.3. SNCF gares (France)

Catégorie	Résultat
Nature des données	Liste des gares de voyageurs du RFN avec coordonnées et adresses postales
URL	http://test.data-sncf.com/index.php/gares-connexions.html
Format	Format CSV, enregistré en texte tabulé pour import dans un SIG comme MacMap©, coordonnées en Longitude/Latitude (WGS84)
Info donnée sur la qualité	Aucune
Résultat de l'import	L'import est possible. Le fichier n'est pas complet.

1.2.4. LILA (CG 44)

Catégorie	Résultat
Nature des données	points d'arrêts du réseau de transport Lila lignes régulières
URL	http://data.loire-atlantique.fr/donnees/detail/horaires-et-points-darrets-du-reseau-de-transport-lila-lignes-regulieres/
Format	GTFS, import de la table stops (csv en WGS 84)
Info donnée sur la qualité	Aucune
Résultat de l'import	Import possible, pas d'erreur bloquante constatée.

1.2.5. Nantes TAN

Catégorie	Résultat
Nature des données	Liste des arrêts de tous les bus et tramways de la SEMITAN circulant sur le territoire de Nantes Métropole.
URL	http://data.loire-atlantique.fr/donnees/detail/arrets-horaires-et-circuits-tan/
Format	GTFS, import de la table stops (csv en WGS 84)
Info donnée sur la qualité	Aucune
Résultat de l'import	Import difficile, plusieurs arrêts sont incohérents.

1.2.6. Rennes STAR

Catégorie	Résultat
Nature des données	Données théoriques du réseau STAR
URL	http://data.keolis-rennes.com/fr/les-donnees/donnees-telechargeables.html
Format	GTFS. Import de la table stops.txt, qui contient le descriptif des points d'arrêts. Les coordonnées sont fournies en longitude/latitude (WGS84).
Info donnée sur la qualité	Aucune
Résultat de l'import	Import possible, une erreur constatée.

1.3. Import de fichiers au format SHP décrivant un linéaire

1.3.1. CG33 pistes cyclables

Catégorie	Résultat
Nature des données	Réseau des pistes cyclables départementales
URL	http://www.datalocale.fr/dataset/ig-pc-CG33
Format	ShapeFile en Lambert93
Info donnée sur la qualité	Moyenne
Résultat de l'import	L'import est réalisé sans problème. Les données ne sont pas complètes (coupures du linéaire). Notamment on voit que les coordonnées ne sont pas toutes complètes, renvoyant les points très loin quand on veut les positionner dans une application avec visualisation cartographique.

2. Résultats des test d'import de données transports, déplacement dans un validateur

2.1. Validation de données au format GTFS par le GTFS FeedValidator de Google

2.1.1. CG33 TransGironde

Catégorie	Résultat
Nature des données	Horaires des lignes régulières du réseau TransGironde
URL	http://www.datalocale.fr/dataset/liste-lignereguliere-transgironde
Format	GTFS
Info donnée sur la qualité	Exhaustive
Résultat de la validation	D'après le validateur, 105 erreurs et 1464 alertes. Les erreurs ont été remontées par les réutilisateurs via la plate-forme. Réponse de l'administrateur de la plate-forme : « Nous essayons de corriger ces problèmes rapidement mais nous n'avons pas identifié jusqu'à présent de script permettant de transformer en GTFS les fichiers produits en XML TRIDENT. L'outil CHOUETTE que nous utilisons actuellement ne semble pas compatible avec l'export généré par notre éditeur. Merci pour votre patience. »

2.1.2. Toulouse TISSEO:

Pas de fichier GTFS mis à disposition.

2.1.3. LILA - CG 44

Catégorie	Résultat
Nature des données	Horaires du réseau de transport Lila lignes régulières
URL	http://data.loire-atlantique.fr/donnees/detail/horaires-et-points-darrets-du-reseau-de-transport-lila-lignes-regulieres/
Format	GTFS
Info donnée sur la qualité	Aucune
Résultat de la validation	D'après le validateur, 6 erreurs et 63 alertes.

2.1.4. NANTES TAN

Catégorie	Résultat
Nature des données	Horaires et circuits de tous les bus et tramways de la SEMITAN circulant sur le territoire de Nantes Métropole.
URL	http://data.loire-atlantique.fr/donnees/detail/arrets-horaires-et-circuits-tan/
Format	GTFS
Info donnée sur la qualité	Aucune
Résultat de l'import	Le fichier nécessitait de connaître les prérequis de l'utilisation du validateur de GTFS pour être passé dans le validateur. D'après le validateur, 7 erreurs et 722 alertes.

2.1.5. RENNES STAR

Catégorie	Résultat
Nature des données	Données théoriques du réseau STAR
URL	http://data.keolis-rennes.com/fr/les-donnees/donnees-telechargeables.html
Format	GTFS.
Info donnée sur la qualité	Aucune
Résultat de l'import	D'après le validateur, 1 erreurs et 20 alertes. Par ailleurs des erreurs d'encodage de caractères perturbent le validateur (plusieurs crashes successifs).

2.1.6. LA CUB

Catégorie	Résultat
Nature des données	Offre de services bus et tramways
URL	http://data.lacub.fr/data.php?themes=10
Format	GTFS
Info donnée sur la qualité	Aucune, mais visualisation directe possible
Résultat de l'import	D'après le validateur, 0 erreurs et 1325 alertes. La réutilisation est possible.

2.1.7. Conclusion

On dénote un grand nombre de jeux de données au format GTFS comportant plusieurs erreurs et de nombreuses incohérences ou alertes (valeur invalide, arrêt trop rapproché, arrêt trop éloigné, trajet trop rapide, arrêt inutilisé) qui mettent en doute la qualité des données.

Seul le jeu de données mis à disposition par la CUB ne comporte pas d'erreur bloquante.

2.2. Validation des données au format Neptune par la validateur CHOUETTE

2.2.1. CG33 TransGironde

Catégorie	Résultat
Nature des données	Horaires des lignes régulières du réseau TransGironde
URL	http://www.datalocale.fr/dataset/liste-lignereguliere-transgironde
Format	NEPTUNE
Info donnée sur la qualité	Exhaustive
Résultat de la validation	Le fichier du CG33 est validé sans erreur par CHOUETTE.

2.2.2. Toulouse TISSÉO

Catégorie	Résultat
Nature des données	Offre de transports (lignes, arrêts, horaires) de Tisséo SMTC
URL	http://data.grandtoulouse.fr/les-donnees/-/opendata/card/14114-reseau-tisseo-metro-bus-tram-
Format	Trident
Info donnée sur la qualité	Exhaustive
Résultat de l'import/ validation	Après vérifications, il s'avère que le fichier zip ne peut pas être importé dans CHOUETTE sans faire une conversion de l'encodage des caractères et corriger l'en-tête des fichiers XML de UTF-8 en ISO-8859-1. <ul style="list-style-type: none"> 50 erreurs non fatales pour l'import et identiques apparaissent: certains arrêts ne sont pas enregistrés car une clé est considérée comme non unique.

2.2.3. CG44 LILA

Catégorie	Résultat
Nature des données	Horaires du réseau de transport Lila lignes régulières
URL	http://data.loire-atlantique.fr/donnees/detail/horaires-et-points-darrets-du-reseau-de-transport-lila-lignes-regulieres/
Format	GTFS
Info donnée sur la qualité	Aucune
Résultat de la validation	D'après le validateur, 20 tests inapplicables, 14 tests réussis, 7 alertes, 8 erreurs, aucune erreur fatale.

2.2.4. NANTES TAN

Catégorie	Résultat
Nature des données	Horaires et circuits de tous les bus et tramways de la SEMITAN circulant sur le territoire de Nantes Métropole.
URL	http://data.loire-atlantique.fr/donnees/detail/arrets-horaires-et-circuits-tan/
Format	NEPTUNE
Info donnée sur la qualité	Aucune
Résultat de l'import	Le fichier ne passe pas le premier test consistant à vérifier la conformité du schéma XML. Le fichier n'est pas exploitable.

2.2.5. RENNES STAR

Pas de fichier NEPTUNE mis à disposition.

2.2.6. La CUB

Pas de fichier NEPTUNE mis à disposition.

2.2.7. Conclusion

On dénote un grand nombre de jeux de données au format NEPTUNE comportant plusieurs erreurs et de nombreuses incohérences ou alertes qui mettent en doute la qualité des données.

Seul le jeu de donnée mis à disposition par le CG33 TransGironde a été validé par le validateur de CHOUETTE.

2.3. Conversion de données au format GTFS en données au format NEPTUNE par l'import-export de CHOUETTE

A travers les fonctions d'imports en GTFS puis d'exports en Neptune, on peut vérifier le contenu et la qualité des fichiers mis à disposition.

Seuls les données GTFS de la CUB sont importables, les autres présentant trop d'erreur.

2.3.1. LA CUB

Voici le résultat de l'import-export :

Catégorie	Résultat
Erreur	Erreur d'import
Problème	java.lang.NullPointerException (erreur fatale)
Information	Exhaustive
Résultat :	Sur le seul jeu de données validé par le GTFS FeedValidator, l'import dans CHOUETTE n'a pas fonctionné. L'export n'a pas pu être réalisé ensuite. Les raisons sont inconnues entre la validité du jeu de données, ou le bon fonctionnement de l'import GTFS par le logiciel CHOUETTE.