



**ATEC
ITS FRANCE**

FEUILLE DE ROUTE

mars 2020

Big Data & Mobilité

Exemples d'usages et pistes d'amélioration

ADER Thomas, Orange
AUTISSIER Jean-Baptiste, DGITM
BARTH Dominique, Laboratoire David (SI, SHS)
BERTRAND Guillaume, Colas
BRACQUEMOND Annie, Vedecom
BRIAND Yann, IRT Systemx
CHEBASSIER Charlotte, Kisio
COCOS Eric, gfi Informatique
COLDEFY Jean, ATEC-ITS
COURET Antoine, Geo4cast
DUNCKEL Teha, gfi Informatique
FAUCHEUX Pascal, gfi Informatique
GEBHART-LEVY Cathy, Eurométropole Strasbourg
GENDRE Patrick, ATEC-ITS
GOHON Thierry, Sopra Steria
HUYNEN Jean-Louis, gfi Informatique
JANSEM Bruno, Eurométropole Strasbourg
LEROUGE David, Colas
NUYTTENS Nicolas, CeremaTV
ORTIZ Manuel, De Gaulle Fleurance et Associés, cabinet
avocats
PLANTIER Axel, gfi Informatique
PENCIOLELLI Jean-François, gfi Informatique
QUATRE Grégoire, Ingerop
ROUSSEAU Christian, Renault
THEARD-JALLU Cécile, De Gaulle Fleurance et Associés, cabinet
avocats
YAHIAOUI Gérard, Nexya

Un document



Labellisé



Table des matières

1. Précisions sur l'étude	3
1.1. Le Big Data, c'est quoi ?	3
1.2. Le Big Data appliqué à la mobilité.....	4
1.3. Contexte et enjeux de l'étude	6
1.4. Objectifs et contenu de l'étude	7
1.5. Constitution du groupe de travail GT	7
2. Prise en compte des réflexions antérieures	8
2.1 Rapport Villani (2018).....	8
2.2 Rapport académie des technologies.....	8
2.3 Rapport ministère de l'industrie	9
2.4 Rapport FNCCR.....	9
2.5 Le point de vue des chercheurs	10
2.6 Article IFSTTAR 2019 « Les Big Data et la mobilité des personnes : les enjeux et défis d'un cadre émergent »	11
2.7 Rapport ITF OCDE	12
2.8 Rapport AIPCR.....	13
2.9 Synthèse des différents rapports	14
3. Contexte législatif et réglementaire en évolution	15
3.1 La loi d'orientation des mobilités du 24 décembre 2019 (dite « LOM »)	15
3.1.1 Articles 8 et 15 de la LOM sur les modalités d'organisation de la mobilité entre collectivités.....	16
3.1.2 Article 25 de la LOM sur l'ouverture des données nécessaires au développement de services numériques de mobilité	17
3.1.3 Article 27 de la LOM sur l'ouverture des données relatives aux déplacements des personnes handicapées ou en mobilité réduite	18
3.1.4 Article 28 sur les systèmes d'information voyageur et de billettique multimodaux (MaaS)	18
3.1.5 Article 32 de la LOM sur les données véhicules connectés et véhicules autonomes	18
3.2 Contexte européen.....	19
3.2.1 RGPD applicable depuis le 25 mai 2018	19
3.2.2 Proposition de Règlement e-privacy	20
3.2.3 Le Règlement « free flow of Data »	20
3.2.4 La nouvelle Directive « Open Data » sur la réutilisation des informations du secteur public	21
3.2.5 Lignes directrices de l'UE sur l'éthique de l'intelligence artificielle	21
3.2.6 Règlement délégué « MMTIS » du 31 mai 2017	22
3.2.7 Le paquet cybersécurité	23
4. Valeur ajoutée du Big Data selon les cas d'usages.....	24
4.1 Typologie des cas d'usages étudiés	24

4.2	Thématiques de la mobilité supports des cas d'usages étudiés	24
4.3	Descriptif des cas d'usages étudiés	25
4.3.1	Connaissance de la mobilité/évaluation de politiques publiques	25
4.3.2	Multimodalité/information et calculateurs d'itinéraires/billettique	27
4.3.3	Exploitation de réseaux/suivi et gestion événements/optimisation capacités et résilience/connexion et coordination multimodale	28
4.3.4	Gestion flottes de véhicules	31
4.3.5	Sécurité routière et sûreté/aides à la conduite	33
4.3.6	Autres fiches solutions France Mobilités illustratives	34
4.3.7	Synthèse des illustrations.....	35
5.	Difficultés encore à lever, recommandations et questions ouvertes	36
5.1	Volet technique.....	36
5.1.1	Difficultés	36
5.1.2	Recommandations	37
5.1.3	Questions	37
5.2	Volet organisationnel	37
5.2.1	Difficultés	37
5.2.2	Recommandations	38
5.2.3	Questions	39
5.3	Volet réglementaire	39
5.3.1	Difficultés	39
5.3.2	Recommandations.....	40
5.3.2	Questions	40
5.4	Volet humain et social	40
5.4.1	Difficultés	40
5.4.2	Recommandations.....	40
5.4.3	Questions	41
5.5	Volet économique	41
5.5.1	Difficultés	41
5.5.2	Recommandations.....	41
5.5.2	Questions	42

1. Précisions sur l'étude

1.1. Le Big Data, c'est quoi ?

Le Big Data désigne des ensembles de données (mégadonnées) devenus si volumineux qu'ils dépassent l'intuition et les capacités humaines d'analyse et même celles des outils informatiques classiques de gestion de bases de données relationnelles ou de l'information (définition wikipedia).

IBM donne une définition intéressante du Big Data au travers « the five Vs of Big Data » :

- Volume : un volume de données en perpétuelle augmentation
- Velocity : une vitesse de développement et une augmentation croissante de la donnée avec une tendance de plus en plus forte à de la donnée temps réel
- Variety : une diversité énorme de types de données
- Variability : des formats et des définitions différents dans le temps et dans l'espace qui posent des problèmes d'interopérabilité majeurs
- Value : une valeur croissante de la donnée de par son coût de production et de traitement avec la possibilité de nombreux enrichissements

La grande majorité des projets de modernisation et de transformation numérique passe par une valorisation accrue de la donnée.

La conclusion de toutes les études actuelles montre que la donnée est au cœur de la réflexion des entreprises, que celles-ci décident de partir de la donnée pour élaborer de nouveaux services (économie servicielle), ou bien qu'elles imaginent ce qu'elles peuvent devenir en tant qu'entreprise (pivotage) en s'appuyant sur les données disponibles.

Pour autant, la donnée existait avant le Big Data, mais les flux de données massifs sont désormais disponibles plus facilement. Désormais, à peu près tout est informatisé (numérique) : le ticket, le trafic et les déplacements, les infrastructures, les formulaires et bases de données administratives, les véhicules, et les données produites peuvent être stockées et mises à disposition dans le monde entier via l'internet, et donc potentiellement croisées entre elles et analysées pour des usages multiples.

"Traditionnellement", le système d'information était construit de manière descendante, à partir des objectifs de chaque organisation, pour lesquels on définit des indicateurs à suivre, des outils d'analyse pour les calculer et les étudier, on spécifie et recueille les données nécessaires pour produire ces indicateurs, et on organise les processus qui permettent ce recueil de données et d'information.

De plus en plus, la chaîne est en partie renversée, et c'est la disponibilité des données qui implique de mettre en place certains outils d'analyse et de concevoir les indicateurs qui en tirent parti, et peuvent même induire les objectifs d'une organisation (cas des startups notamment).

Les données sont de différentes natures :

- Des données externes privées (de plus souvent par API),
- Des données publiques et de plus en plus souvent en open Data,
- Des données internes, soit en provenance de capteurs sur le terrain (IOT), soit en provenance d'applications avec des utilisateurs internes ou externes.

Parmi ces données, il est essentiel de distinguer les données portant des informations à caractère personnel.

Un cloud est essentiel car quand les données sont vraiment massives elles ne sont en général pas stockées localement mais via un cloud.

Si le Big Data nécessite de mettre en œuvre des logiciels et des systèmes, des algorithmes, des métiers et des compétences, il demande aussi des organisations et des processus adaptés qui favorisent les échanges de données et leur ouverture.

1.2. Le Big Data appliqué à la mobilité

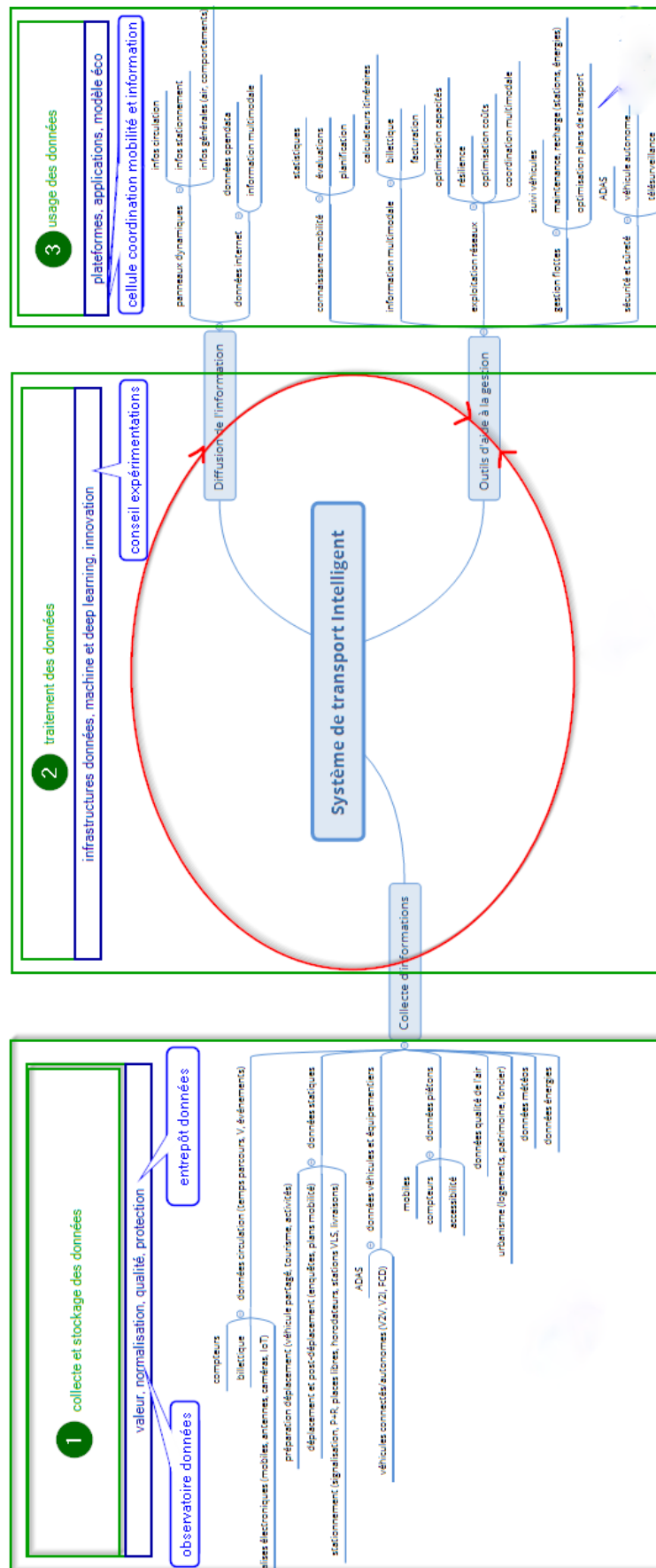
Faisant partie des secteurs "traditionnels" comme le bâtiment ou l'agriculture, les transports et les services de mobilité ont été informatisés relativement tard par rapport à d'autres secteurs comme la banque, les télécoms ou la pharmacie. Ce secteur rattrape le retard, surtout depuis une dizaine d'années, et a fait sa transition numérique. Les systèmes d'information et applications sont toutefois d'une grande diversité selon les domaines d'application, les opérateurs, et les territoires. L'objet de ce document est d'analyser quels sont les apports du Big Data, mais aussi les écueils. Il formalise quelles recommandations proposer et les questions ouvertes. Pour cela, il nous semble essentiel de tenir compte de cette grande diversité d'applications et de situations.

De nombreux usages du Big Data sont décrits dans les 3 étapes suivantes :

- Collecte des données de mobilité, ou associées à la mobilité (génération, collecte, stockage, sécurisation, traitement, agrégation avec d'autres données)
- Outils d'aide à la décision de la mobilité
- Diffusion de l'information de mobilité

La 1ère concerne la donnée en elle-même et son traitement.

Les deux dernières concernent davantage son utilisation et sa diffusion.



1.3. Contexte et enjeux de l'étude

Les aires urbaines françaises font face à des enjeux majeurs en matière d'accessibilité et d'environnement avec des saturations récurrentes sur les réseaux routiers et de transports en commun structurants. Les outils du numérique, le développement des données et leur disponibilité semblent offrir un potentiel d'optimisation de nos systèmes de mobilité tout en répondant de manière plus ciblée aux besoins. C'est une problématique forte que les membres du comité stratégique de Mobilité 3.0, le Ministère des transports et ATEC ITS France ont choisi d'adresser, compte tenu des enjeux économiques (nos entreprises doivent être à la pointe des technologies du traitement du Big Data), environnementaux (enjeux rappelés par la COP21), et d'accessibilité de nos territoires. Cependant, si la presse se fait tous les jours l'écho d'annonces sur les thématiques autour du Big Data (prévisions à 1h, intelligence artificielle, etc.), les solutions concrètes ne sont pas encore pleinement opérationnelles et il n'est pas certain que les entreprises françaises soient bien positionnées, ni que les gains escomptés soient au rendez-vous.

C'est l'enjeu de la démarche Mobilité 3.0 d'analyser et proposer des actions permettant le développement de tout le potentiel offert par les mobilités numériques, en particulier via l'accès à des données toujours plus nombreuses et des traitements toujours plus puissants et rapides.

En parallèle, le Ministère des Transports a mis en place un vaste plan d'action pour développer les innovations en matière de mobilité : la démarche France Mobilités.

Elle comporte différents volets :

- Un guichet unique d'entrée (DGITM)
- Une plateforme nationale = centre de ressources faisant le lien entre les ambitions des collectivités et les innovations des start'ups comportant des fiches produits d'innovations classées suivant des domaines fonctionnels (DGITM/ATEC/Cerema)
- La formation (IFSTTAR)
- Les dispositifs de financement
- Le soutien à l'ingénierie pour les zones peu denses (CEREMA, avec appui de la Caisse des Dépôts & Consignations et de l'ADEME) : des cellules régionales d'appui, de l'animation, de l'aide aux projets et des centres de ressources

Le CEREMA soutient le Ministère des Transports pour la mise en œuvre de la plateforme collaborative France Mobilités qui doit favoriser les retours d'expériences sur les innovations dans le domaine des mobilités et la mise en relation des acteurs (entreprises et associations proposant des solutions et collectivités cherchant à implanter des projets de mobilité). Les exemples d'usages identifiés et approfondis dans le cadre de ce livrable ont ainsi permis une alimentation de la plateforme France Mobilités avec plusieurs exemples issus de l'application de la technologie Big Data.

Pour répondre à ces attentes, et dans un contexte d'émergence d'une nouvelle loi sur l'orientation des mobilités (LOM), un premier groupe de travail GT a été constitué en 2017-2018 par l'ATEC-ITS dans le cadre de mobilité 3.0 et des assises de la mobilité sous pilotage du CEREMA et de gfi afin de définir la feuille de route à 10 ans du Big Data appliqué à la mobilité.

Un premier livrable a été élaboré par le groupe de travail et certaines de ses propositions intégrées dans la liste faite dans le cadre des assises de la mobilité et dans le projet de loi LOM.

1.4. Objectifs et contenu de l'étude

Le présent 2nd livrable constitue un approfondissement en illustrant sur des cas d'usages mobilité concrets (usagers, opérateurs et autorités organisatrices de mobilité) les gains obtenus ou attendus par le traitement de données de masse (Big Data) et en identifiant au travers de leur analyse les manquements/adaptations réglementaires, techniques, organisationnels, juridiques, économique à approfondir pour faciliter demain le déploiement des services identifiés.

1.5. Constitution du groupe de travail GT

Les personnes suivantes ont participé à tout ou partie du GT mis en place (4 réunions en tout). Parmi elles, la grande majorité avaient déjà participé à l'élaboration du premier livrable.

Services	Prénoms – Noms	Adresses mails
CeremaTV	Nicolas Nuyttens	nicolas.nuyttens@cerema.fr
gfi Informatique	Jean-François Penciolelli	jean-francois.penciolelli@gfi.fr
	Eric Cocos	eric.cocos@gfi.fr
	Jean-Louis Huynen	jean-louis.huynen@gfi.fr
	Pascal Fauchaux	pascal.fauchaux@gfi.fr
	Axel Plantier	axel.plantier@gfi.fr
	Teha Dunckel	teha.dunckel-barbier@gfi.fr
DGITM	Jean-Baptiste Autissier	jb.autissier@developpement-durable.gouv.fr
Sopra Steria	Thierry Gohon	thierry.gohon@soprasteria.com
Orange	Thomas Ader	thomas.ader@orange.com
Colas	David Lerouge	david.lerouge@colas.com
	Guillaume Bertrand	guillaume.bertrand@colas.com
Ingerop	Grégoire Quatre	gregoire.quatre@gmail.com
Geo4cast	Antoine Couret	antoine.couret@geo4cast.ai
Dgfla cabinet avocats	Cécile Théard-Jallu	ctheardjallu@dgfla.com
	Manuel Ortiz	mortiz@dgfla.com
Kisio	Charlotte Chebassier	Charlotte.chebassier@kisio.com
Eurométropole Strasbourg	Bruno Jansem	bruno.jansem@strasbourg.eu
	Cathy Gebhart-Levy	cathy.gebhart-levy@strasbourg.eu

Renault	Christian Rousseau	christian.rousseau@renault.com
Nexyad	Gérard Yahiaoui	
Vedecom	Annie Bracquemond	annie.bracquemond@vedecom.fr
Laboratoire David (SI, SHS)	Dominique Barth	dominique.barth@uvsq.fr
ATEC-ITS	Patrick Gendre	patgendre94@gmail.com
	Jean Coldefy	c3i@sfr.fr
IRT Systemx	Yann Briand	yann.briand@irt-systemx.fr

2. Prise en compte des réflexions antérieures

2.1 Rapport Villani (2018)

La transformation numérique de la société est un sujet majeur qui a fait l'objet de nombreux rapports publics et autres livres blancs produits par des think tanks ou des entreprises.

Les rapports suivants nous ont semblé pertinents :

Rapport de mission parlementaire coordonné par Cédric Villani, Donner un sens à l'intelligence artificielle, Pour une stratégie nationale et européenne, mars 2018, https://www.aiforhumanity.fr/pdfs/9782111457089_Rapport_Villani_accessible.pdf

Ce rapport, très médiatisé, met l'accent sur plusieurs points :

- Le rôle central de la donnée
- Les enjeux économiques et la nécessité de développer un écosystème associant la recherche publique, les entreprises, les start-ups et les administrations (l'Etat devant être "exemplaire")
- La nécessité de prendre en compte dès la conception les questions éthiques et sociétales, et la transformation des métiers

La mobilité et les transports font partie des 4 secteurs stratégiques identifiés.

Une recommandation au moins semble s'être concrétisée : celle de créer 4 instituts interdisciplinaires en IA (à Grenoble, Nice, Toulouse et Paris, ces 2 derniers comprenant une composante transport affirmée) dans le cadre du Plan IA financé par le PIA. Les budgets annoncés (1 milliard d'Euros) font que la communauté ITS doit s'y intéresser et peut en attendre des résultats concrets.

2.2 Rapport de l'Académie des Technologies

Rapport de l'Académie des Technologies, un changement de paradigme peut en cacher un autre. Rapport de l'Académie des technologies, coordonné par Yves Caseau, EDP Sciences, 2015.

http://academie-technologies-prod.s3.amazonaws.com/2015/12/08/16/18/21/693/BigData_internet.pdf

Ce rapport synthétique très pédagogique analyse bien pourquoi le Big Data est important et ce qu'il a changé : l'explosion des données (de toutes natures) liée à l'effondrement des coûts de production et de stockage, l'arrivée de technologies matérielles et logicielles permettant de gérer et traiter ces masses de données (les "Big Data" proprement dites), l'arrivée de nouveaux métiers autour de la donnée, qui correspondent à une nouvelle façon de faire de la statistique, et même une nouvelle façon de programmer, avec des cycles très courts permettant une grande agilité. En termes d'enjeux, outre la question de la protection des données personnelles, le rapport met l'accent sur le risque de "disruption" par les nouveaux entrants, et la nécessité de développer un écosystème et de s'adapter aux nouveaux outils et aux nouvelles façons de travailler. Les recommandations sont centrées sur le besoin de se former et d'accompagner le changement dans les entreprises et les pouvoirs publics, de créer des centres de ressources (infrastructures numériques) et des jeux de données utilisables par la communauté.

2.3 Rapport du Ministère de l'Industrie

Intelligence artificielle – Etat de l'art et perspectives pour la France, étude PIPAME / Direction Générale des Entreprises, Atawao consulting, 2019.

https://www.entreprises.gouv.fr/files/files/directions_services/etudes-et-statistiques/prospective/Intelligence_artificielle/2019-02-intelligence-artificielle-etat-de-l-art-et-perspectives.pdf

Ce rapport inclut quelques éléments sur le secteur des transports et de la mobilité. Parmi les recommandations, constituer et donner accès à de très grands jeux de données de qualité, l'accès à des données massives, corrélées, complètes, qualifiées, historisées, est une clé technologique majeure de mise au point de technologies d'intelligence artificielle aujourd'hui. Ce rapport insiste donc sur le lien très fort voire constitutif entre Big Data et IA. Il insiste aussi sur l'importance de territoires d'expérimentations assez grands pour que la masse de données soit suffisante, et sur l'existence d'infrastructures numériques solides pour ces projets. Le transport est placé en haut de la liste des secteurs prioritaires, après la santé, aux côtés de l'automobile, de l'environnement et des services d'utilité publique.

2.4 Rapport FNCCR

Collecte et gestion des données numériques pour le pilotage des politiques publiques, Vers un Big Data territorial, FNCCR, Cabinet Tactis et Parme Avocats, 2017

<http://www.fnccr.asso.fr/article/Big-Data-territorial-publication-de-letude-de-la-fnccr/>

C'est sans doute le rapport non sectoriel le plus pertinent pour notre groupe, parce qu'il concerne explicitement le rôle des collectivités, qui est essentiel dans Mobilité 3.0.

Ce rapport n'est pas centré sur l'IA et l'high tech, mais sur le point de vue des collectivités, et inclut l'Open Data, dont les jeux de données sont petits, mais très variés et peuvent être inclus dans le Big Data au sens large. Il distingue d'ailleurs dès l'introduction "Big Data", "Open Data" et "Shared Data".

La présentation distingue clairement les grands cas d'utilisation : prévoir et prédire (niveau études et conception), tracer et cibler (au niveau individuel de l'utilisateur), détecter et optimiser (au niveau de l'exploitation temps réel).

Le partage des données est vu comme apportant les valeurs ajoutées suivantes : réduction des coûts, innovation, transparence, changement de comportement, collaboration. Le rapport fait un bon état des lieux 2017 sur les projets et services autour des données dans les administrations, et sur les questions juridiques.

Pour la FNCCR, une clarification du cadre de gouvernance local, de la problématique des échanges inter-administration et des relations Etat / Collectivités devra progressivement s'imposer afin d'organiser aux différentes échelles un service de la donnée reposant sur la collecte, la diffusion et l'exploitation mutualisée des données.

Parmi les recommandations de la FNCCR, la plupart semblent très pertinentes pour notre domaine :

- Poursuivre la dynamique de la loi pour une République numérique (données de référence, mission de service public de la donnée...) ;
- Affirmer, dans les contrats, la maîtrise des collectivités sur leurs données (les collectivités ont en charge de nombreuses délégations de service public) ;
- Préparer la création de missions locales de service public de la donnée
- Créer un maillage national de structures de mutualisation informatique ;
- Mettre les Data sciences au service des politiques publiques locales
- Renforcer la coopération au niveau européen (normalisation, expérimentation...).

Le rapport inclut en annexe un glossaire (donnée de référence, méta-donnée, etc.) et des fiches pratiques pour les collectivités.

2.5 Le point de vue des chercheurs

Il est important de pouvoir raisonner à moyen terme sur les retombées possibles du Big Data, et il est ainsi pertinent d'associer les chercheurs à la réflexion de l'ATEC. L'existence du projet fédérateur de l'IFSTTAR "mobilité et transition numérique", qui a tenu un Séminaire sur traces et Big Data le 4 avril 2019 dans le cadre groupe "Ville et Numérique" du LabeX Futurs Urbains est à souligner.

Les présentations font ressortir :

- Le besoin reconnu de partager/mutualiser/ouvrir les données (des collectivités et de leurs délégataires) sur chaque territoire
- Les difficultés dues aux frontières institutionnelles entre collectivités (commune, intercommunalité, département, région)
- Les difficultés dues aux frontières dans l'organisation de chaque système d'information (mission numérique, open Data, DSI, services voirie, transport, etc.)
- Les difficultés juridiques (interdiction de croiser des données individuelles, secret statistique) mais exceptions possibles pour les travaux de recherche via le Centre d'Accès Sécurisé aux Données
- Les services mobiles privés (et notamment Waze qui a pris une part de marché dominante en France depuis 2017) remettent en question le monopole de fait qu'ont longtemps détenu les administrations en matière de statistique (et d'information

routière) ; cela pose aussi la question de pourquoi il n'existe pas de Waze français (ou européen), ou dit autrement, de comment faire pour qu'émergent des concurrents à Waze.

- Les smartphones apportent une solution complémentaire aux enquêtes pour mesurer la mobilité individuelle via les traces numériques de déplacements (connaissance des itinéraires), utile à 3 niveaux : comportement de mobilité, fonctionnement du territoire, performance du système de transport.
- Le Big Data pose la question de la transparence des algorithmes et du rôle des experts métiers humains dans les décisions et dans la définition/validation des algorithmes
- Le Big Data peut améliorer beaucoup d'applications existantes : maintenance de flottes et d'infrastructures, gestion du trafic et exploitation des réseaux
- Le Big Data change en profondeur nos outils d'observation et d'analyse (Dataviz, modèles) des mobilités
- Le Big Data rend possible de nouvelles applications s'appuyant sur des données individuelles et fournissant un service personnalisé (info temps réel, paiement...), déployées à un niveau local (péage urbain...) national (écotaxe, compte mobilité...) voire global (MaaS, Waze, assurance pay as you drive...)
- Le Big Data est indispensable au développement des véhicules autonomes
- Références :
 - Big Data et politiques publiques dans les transports, PALMA (de) André, DANTAN Sophie, éditions Economica, 2017. <https://www.economica.fr/livre-Big-Data-et-politiques-publiques-dans-les-transport,fr,4,9782717869439.cfm>
 - « Urban mobility and the smartphone : Transportation, travel behavior and public policy », Anne Aguilera, Virginie Boutueil, Elsevier, 222 pages, 2018.
 - Journées d'Etudes Mobilité Urbaine Cerema 2019, Ces journées annuelles sont consacrées au domaine de la connaissance de la mobilité et notamment des enquêtes, un domaine où le Big Data a un impact important en complémentarité de ces enquêtes.
<https://www.cerema.fr/fr/actualites/journees-echanges-mobilite-urbaine-13-14-mai-paris>

2.6 Article IFSTTAR 2019 « Les Big Data et la mobilité des personnes : les enjeux et défis d'un cadre émergent »

Cet article a été écrit par Thomas Ermans, Céline Brandeleer et Michel Hubert.

Quelques recommandations particulières :

- Pour l'étude de la mobilité, les Big Data sont complémentaires aux données traditionnelles : elles ont une meilleure exhaustivité et une granulosité spatio-temporelle plus fine (elles permettent de mesurer la variabilité temporelle de l'occupation des territoires par des individus, gérer des réseaux en des temps très courts) mais elles accusent un déficit d'attributs (elles se prêtent moins à l'analyse des causalités et des caractéristiques des personnes). Leur usage se heurte à différents obstacles : représentativité, respect de la vie privée, compétences techniques, propriété d'acteurs privés, qui imposent une validation régulière et faussent la

comparabilité dans le temps. La production de volumes de flux fiables et représentatifs nécessite des procédures de redressement et de validation des résultats qui s'appuient nécessairement sur des données externes

- Le facteur humain est bien présent à chacune des étapes de production de savoir et le Big Data ne peut se prévaloir d'une objectivité supérieure
- Le volume n'est pas forcément synonyme de fiabilité (des omissions, aléas,...) : exemple d'utilisateurs de téléphonie mobile localisés dans une cellule de communication différente de celle où ils se trouvent pour des soucis de charge de l'antenne, météo, topographie,... et les traitements effectués par l'opérateur sont souvent opaques ; les technologies évoluent fortement et empêchent des comparaisons dans le temps ou exemple des Floating Car Data (FCD) qui représentent encore à ce jour de faibles échantillons (utiles toutefois pour approcher la congestion ou les temps de parcours)
- La collecte des données implique une perte de certains facteurs explicatifs, notamment d'informations sur les individus (âge, genre, catégories socio-professionnelles) ou sur les déplacements (modes, motifs, vécus...)
- Les Big Data permettent de trouver des solutions techniques rapides et efficaces, mais sans forcément chercher la cause des problèmes, ni agir sur celle-ci, minorant ainsi certaines problématiques sociales (inégalités socio-économiques, problématiques environnementales...)
- Les déficits de contexte ne sont pas sans poser problème pour les exercices à visée opérationnelle et peuvent mener à des classifications et prédictions faussées
- Pour remédier à ces manquements, l'inférence de données manquantes le croisement avec d'autres sources de données est toujours possible, mais pose la question de la préservation de l'anonymat et la protection de la vie privée
- L'émergence des Big Data représente un enjeu de gouvernance important pour les pouvoirs publics ; ils nécessitent d'investir dans des compétences de Data scientist permettant de maîtriser une palette large d'outils techniques (programmation, mathématiques, statistiques, Data mining, machine learning,...) ; le processus d'extraction d'informations des Big Data est souvent réalisé par le privé, donc avec absence de droit de regard de la puissance publique ; l'opérateur public n'est pas propriétaire des données et n'a pas toujours un droit de regard sur les données, une validation préalable est nécessaire avec des données d'enquêtes et de comptages ; elle aurait intérêt à faire entendre ses besoins et mettre en œuvre conjointement le processus de production des données
- La conformité à des standards internationaux peut représenter une forme d'harmonisation facilitant les comparaisons.

2.7 Rapport ITF OCDE

Big Data and Transport, Understanding and assessing options, International Transport Forum, OCDE, 2015 <https://www.itf-oecd.org/Big-Data-and-transport>

Comme l'indique son sous-titre, ce rapport a plus un rôle d'explications que de propositions. Pour autant, il est vraiment un état de l'art et est illustré par de nombreux exemples.

Il présente les différentes étapes du cycle de vie des données, du recueil à l'exploitation (analyse et Data mining, modélisation, visualisation, diffusion), en passant par les traitements intermédiaires (qualification, nettoyage, fusion, stockage...), puis développe les questions de localisation et de protection de la vie privée.

Les éléments ayant retenus l'attention du groupe de travail sont :

- Les données localisées sont difficiles à anonymiser, il faut adapter le droit aux nouvelles pratiques
- La sécurité routière sera améliorée par la standardisation de quelques éléments de données clé concernant les véhicules
- Les AOM devront auditer les données utilisées (les Big Data ont aussi des problèmes de qualité) et se doter des compétences permettant d'interpréter les résultats des analyses utilisant des Big Data, pour s'en servir correctement comme aide à la décision.
- Inventer de nouveaux modèles de partage de données, notamment public-privé ; beaucoup de données utiles sont détenues par le privé et la collectivité en aurait besoin pour remplir ses missions de service public (cf. article 25 de la LOM sur le partage de données FCD)
- La visualisation de données sera amenée à jouer un rôle croissant dans l'élaboration des politiques publiques.

2.8 Rapport AIPCR

Big Data pour l'exploitation du réseau routier, Comité Technique B1, AIPCR, 2019

<https://www.piarc.org/en/Technical-Committees-World-Road-Association/Strategic-Themes-Access-Mobility/Technical-Committee-Road-Network-Operations-Intelligent-Transportation-Systems/>

La commission technique qui a produit ce rapport était présidée par Jacques Ehrlich, Sylvain Belloche (CEREMA) ayant assuré le secrétariat francophone.

Le rapport présente de manière claire la vision Data centrique qui se dessine et comment le Big Data s'articule avec d'autres tendances du moment (IOT, cloud), avec une vision du périmètre très proche de celle de notre GT.

Il est illustré de plusieurs exemples internationaux assez détaillés en conception et exploitation de réseaux routiers et de transports publics, ainsi que pour les services de mobilité.

Le rapport identifie des leviers ("éléments facilitateurs") pour réussir le Big Data : accompagner et se former, les normes, une vision d'ensemble partagée de l'architecture et des rôles, le processus et les indicateurs de qualité, des licences claires, la protection de la vie privée, la sécurité (avec des visions nationales très différentes d'un pays à l'autre au niveau de l'AIPCR).

Les recommandations essentielles sont :

- Les données massives ne nécessitent pas nécessairement d'importants investissements dans l'infrastructure informatique de l'entreprise. Les données massives peuvent être collectées et traitées par des tiers

- Il est recommandé aux exploitants de réseaux routiers de s'approprier leurs propres données, même lorsqu'elles sont collectées par des tiers, et de ne pas racheter les données à des entreprises privées.
- Les données en général et les données massives en particulier ont le potentiel d'améliorer les tâches de conception et d'exploitation des exploitants de réseaux routiers.
- Il est conseillé à tous les exploitants de réseaux d'infrastructures d'évaluer comment et où les données massives peuvent soutenir leurs tâches et de l'inclure dans leur stratégie
- Il est recommandé de commencer par ce qui est facilement disponible et de se concentrer sur les solutions à portée de main.
- Il faut souvent du temps pour faire mûrir et affiner les applications. Si l'analyse des données massives n'est pas correctement intégrée dans la stratégie de l'organisation, la déception peut conduire à l'annulation prématurée des budgets nécessaires à son intégration.
- L'avantage le plus évident des données massives est l'étalonnage des modèles de trafic.
- Des compétences en matière de données massives doivent être intégrées dans l'organisation de l'exploitant du réseau routier, avec des ressources internes ou externes.
- La coopération dans le domaine des données massives est essentielle.
- Respecter le principe de la " protection de la vie privée dès la conception ".

2.9 Synthèse des différents rapports

Voici une proposition de synthèse des recommandations formulées dans ces différents rapports :

- Il faut inclure dans notre réflexion non seulement les données massives, mais aussi les données de référence et l'Open Data
- Une large partie des données Big Data mobilité provient des collectivités et opérateurs de mobilité, les promesses de l'IA dépendent de ces données ; il faut donc que les collectivités et opérateurs puissent bénéficier d'une partie des ressources du plan IA, en coopération avec les labos et entreprises qui en bénéficieront pour développer de nouveaux algorithmes et services associés ; plus largement, il faut investir dans les infrastructures numériques de la mobilité
- Dans le même ordre d'idées, il faudra du temps pour vraiment tirer parti des Big Data, en attendant, il faut continuer d'investir dans la production et la qualification des données "de base" ; par exemple, on est encore loin d'avoir une information temps réel sur les horaires de passage des TC sur tout le territoire
- Points clés pour la réussite du Big Data : accompagnement et formation, processus et qualité des données, des licences claires, prise en compte des questions difficiles dès la conception (protection de la vie privée, éthique, transformation des métiers, sécurité), coopération concrète des acteurs de l'écosystème (collectivités, opérateurs, entreprises et startups, chercheurs) dans des projets pilotes mais aussi pour

mutualiser des connaissances et partager une vision commune de l'architecture et des rôles

- Les collectivités AOM et les exploitants de réseaux et de transport doivent avoir une stratégie sur les données massives, intégrant la question des compétences nécessaires, le domaine n'est pas réservé aux grands industriels et aux start-ups
- Les données massives ne nécessitent pas nécessairement d'importants investissements dans l'infrastructure informatique de l'entreprise. Les données massives peuvent être collectées et traitées par des tiers. Il faut commencer par ce qui est facilement disponible et se concentrer sur les solutions à portée de main
- La mise en application de l'article 25 de la LOM sur la mise à disposition des données FCD peut être un cas d'usage concret sur lequel porter l'effort.
- Les AOM ou opérateurs de mobilité devront auditer les données utilisées (les Big Data ont aussi des problèmes de qualité) et se doter des compétences permettant d'interpréter les résultats des analyses utilisant des Big Data, pour s'en servir correctement comme aide à la décision.
- La visualisation de données sera amenée à jouer un rôle croissant dans l'élaboration des politiques publiques, et dans les mobilités en particulier

3. Contexte législatif et réglementaire en évolution

3.1 La Loi d'Orientation des Mobilités du 24 décembre 2019 (dite « LOM »)

La loi n°2019-1428 d'orientation des mobilités, dite « LOM »¹ a été promulguée le 24 décembre 2019. Elle ambitionne de favoriser l'émergence de nouvelles solutions de mobilité et la création de plates-formes multimodales.

« Cette loi transforme en profondeur la politique des mobilités, avec un objectif simple : des transports du quotidien à la fois plus faciles, moins coûteux et plus propres »

Élisabeth Borne, Ministre de la transition Ecologique et Solidaire

La mesure phare de la LOM est l'ouverture des données des offres de mobilité, accompagnée de façon opérationnelle sur tout le territoire. Cette ouverture vise en particulier :

- Les données statiques (arrêts, horaires, tarifs...) et en temps réel (perturbations, disponibilités...),
- Les données des transports en commun ou à la demande,
- Les données des réseaux routiers et de stationnements, pour permettre un accès facile et en temps réel aux informations sur les solutions de transport disponibles.

Le texte met en place les autorités organisatrices de mobilité (« AOM ») qui reposent sur un couple intercommunalités – région, avec pour mission, notamment de proposer des solutions alternatives à la voiture individuelle. Les AOM devront faire en sorte qu'une application

¹ 1 LOI n° 2019-1428 du 24 décembre 2019 d'orientation des mobilités, JORF n°0299 du 26 décembre 2019.

multimodale existe sur leur territoire, permettant d'y regrouper l'ensemble des moyens de transport disponibles et les informations en temps réel sur leur situation.

Les entreprises devront également partager certaines données (par ex. plateformes de covoiturage, ou systèmes de billettique des services de transport) dans le cadre d'une information multimodale et d'une interconnexion entre plateformes.

Entre autres innovations, la circulation des navettes autonomes est permise à partir de 2020, le covoiturage est développé comme solution de transport au quotidien et les offres de mobilité en libre-service (trotinettes, vélos ou scooters sans station d'attache) se voient fournir un nouveau cadre de régulation.

La LOM prévoit ainsi des dispositions majeures pour rendre accessibles de nombreuses données et permettre d'accélérer le déploiement de services de type MaaS (Mobility as a Service) en particulier :

- Une obligation d'ouvrir la quasi-totalité des données d'offre de mobilité dans un calendrier accéléré, avec une démarche animée par les régions et les métropoles, une autorité de contrôle : l'Autorité de régulation des transports (anciennement Arafer) (Article 25 de la LOM) ;
- Une obligation de créer de nouvelles données (accessibilité personnes handicapées ou personnes à mobilité réduite) (Article 27 de la LOM) ;
- Une obligation d'ouvrir les canaux de vente et de partager les informations en mode multimodal pour le MaaS (Article 28 de la LOM) ;
- Une possibilité d'utiliser des ordonnances pour mettre en place une obligation de transmission aux acteurs désignés (gestionnaires, AOM et acteurs de la sphère « sécurité routière »), des données des véhicules connectés, des assistants de conduite et des systèmes numériques d'aide au déplacement à des fins de sécurité routière (prévention, intervention ou détermination des responsabilités), de connaissance de l'infrastructure, de connaissance du trafic routier (Article 32 de la LOM).

3.1.1 Articles 8 et 15 de la LOM sur les modalités d'organisation de la mobilité entre collectivités

Les articles 8 et 15 de la LOM définissent les modalités d'organisation de la mobilité entre collectivités.

Les textes organisent la suppression des zones qui ne sont pas couvertes par une autorité organisatrice de la mobilité (« AOM ») (à savoir notamment les communautés d'agglomération, les communautés urbaines, les métropoles, les communautés de communes voire les régions), en accordant de nouvelles compétences aux collectivités territoriales pour organiser notamment des services tels que l'autopartage, le covoiturage et le transport à la demande.

Les communautés de communes ont jusqu'au 1er juillet 2021 pour devenir AOM et exercer parmi 5 compétences :

- Transport public ;
- Transport à la demande (TAD) ;
- Transports scolaires ;

- Mobilités actives ;
- Mobilités partagées.

Passé ce délai, c'est la Région qui devient AOM pour les compétences non exercées.

Des plans mobilités qui couvrent l'ensemble des formes de mobilité vont remplacer les actuels « plans de déplacement urbain ». Le rôle des régions comme chefs de file de la mobilité est renforcé. Les autorités organisatrices de la mobilité peuvent plus facilement agir en matière de mobilité solidaire (par exemple accompagnement individualisé pour les demandeurs d'emploi, les apprentis et les alternants).

En tant que chefs de file, les régions concluent des contrats opérationnels de mobilité à l'échelle des bassins de vie avec : les AOM, les syndicats mixtes de transports, les départements et les gestionnaires de gares de voyageurs ou de PEM concernés.

Ces dispositions doivent répondre au déficit actuel de gouvernance à l'échelle des aires urbaines.

3.1.2 Article 25 de la LOM sur l'ouverture des données nécessaires au développement de services numériques de mobilité

L'article 25 de la LOM consacre l'ouverture des données nécessaires au développement de services numériques de mobilité : les données statiques et dynamiques sur les déplacements et la circulation sont rendues accessibles et réutilisables. Les règles d'ouverture des données seront celles du Règlement européen 2017/1926².

La LOM intègre les données en temps réel, notamment les horaires en temps réel, les perturbations des services réguliers de transport et la disponibilité des vélos en libre-service et de l'autopartage, la localisation des engins de déplacement personnel partagés disponibles (trottinettes, vélos ou scooters en free-floating), les données relatives aux points de recharge publics pour les véhicules électriques et hybrides.

Les régions et les métropoles ont un rôle d'animation de la démarche d'ouverture des données et de transmission vers l'interface numérique unique (point d'accès national).

Les autorités organisatrices sont responsables de la fourniture des données des transports qu'elles organisent, elles peuvent confier cette mission à leur délégataire. Les données devront être fournies aux normes prévues par le Règlement européen. L'ARAFER, devenue l'Autorité de Régulation des Transports en octobre 2019, étend son champ d'intervention, et assure notamment le contrôle des dispositions des articles 25 et 28.

² Règlement délégué (UE) 2017/1926 de la Commission du 31 mai 2017 complétant la directive 2010/40/UE du Parlement européen et du Conseil en ce qui concerne la mise à disposition, dans l'ensemble de l'Union, de services d'informations sur les déplacements multimodaux

3.1.3 Article 27 de la LOM sur l'ouverture des données relatives aux déplacements des personnes handicapées ou en mobilité réduite

Les déplacements des personnes handicapées ou à mobilité réduite sont favorisés avec notamment une meilleure information sur les solutions accessibles et une possible réutilisation des données.

L'article 27 de la LOM concerne l'obligation de diffusion de données relatives à l'accessibilité et l'obligation de les créer lorsqu'elles manquent. Ces données d'accessibilité de l'ensemble des services réguliers de transports publics, routiers comme ferroviaires, permettront le développement de services d'aide aux déplacements des personnes handicapées. Sont également concernées les données de localisation des dispositifs et balises diffusant des informations, y compris dans les bâtiments ouverts au public, ainsi que les données des itinéraires pédestres dans un rayon de 200 m autour des points d'arrêts prioritaires.

3.1.4 Article 28 sur les systèmes d'information voyageur et de billettique multimodaux (MaaS)

L'article 28 de la LOM porte sur les systèmes d'information voyageur et de billettique multimodaux (MaaS). Les autorités organisatrices doivent veiller à l'existence d'un service d'information à l'intention des usagers portant sur l'ensemble des modes de déplacements dans leur ressort territorial.

L'ensemble des services de mobilités sur un territoire donné devront ouvrir leurs canaux de vente pour tous les titres.

Toute personne, publique comme privée, peut constituer un service multimodal de réservation, vente et revente. Les aspects techniques et financiers de ces services sont précisés dans un contrat entre le fournisseur du service numérique multimodal et le gestionnaire de chacun des services, qui doit traduire des conditions financières raisonnables, équitables, transparentes et proportionnées.

3.1.5 Article 32 de la LOM sur les données véhicules connectés et véhicules autonomes

L'article 32 de la LOM sur les données de véhicules connectés et véhicules autonomes prévoit une possible obligation de fourniture de données de mobilité à certains acteurs.

D'ici le 26 décembre 2020, le Gouvernement peut prendre par voie d'ordonnance, toute mesure relevant du domaine de la loi pour rendre accessible les données pertinentes des systèmes intégrés aux véhicules terrestres à moteur, équipés de dispositifs permettant d'échanger des données avec l'extérieur du véhicule, dès lors qu'elles sont nécessaires aux :

- Gestionnaires d'infrastructures routières ;
- Forces de l'ordre ;
- Services d'incendie et de secours ;
- AOM.

Les finalités sont diverses : détection des accidents et incidents ou conditions de circulation génératrices d'accidents, prévention des accidents ou amélioration de l'intervention en cas d'accident ; connaissance de l'infrastructure routière, de son état et de son équipement ; connaissance du trafic routier ; détermination des responsabilités en cas d'accidents ou aux fins d'indemnisation.

Ces mesures pourront aussi viser à :

- Rendre accessibles, en cas d'accident de la route, les données des dispositifs d'enregistrement de données d'accident et les données d'état de délégation de conduite enregistrées dans la période qui a précédé l'accident, aux officiers et agents de police judiciaire aux fins de détermination des responsabilités ainsi qu'aux organismes chargés de l'enquête technique et de l'enquête de sécurité prévues à l'article L. 1621-2 du code des transports ;
- Rendre accessibles, en cas d'accident de la route, les données d'état de délégation de conduite enregistrées pendant la période précédant l'accident :
 - Aux entreprises d'assurance qui garantissent les véhicules impliqués dans l'accident, aux fins de déterminer les indemnisations, exclusivement lorsque le traitement de ces données est nécessaire à l'exécution du contrat d'assurance concerné ;
 - Aux fonds de garantie des assurances obligatoires de dommages, lorsqu'aucune entreprise d'assurance n'est en mesure de procéder aux indemnisations dans le cadre de l'exécution d'un contrat d'assurance. Sont rendues accessibles les données strictement nécessaires pour déterminer l'activation ou non de la délégation de conduite du véhicule aux fins d'indemniser les victimes en application de la loi n° 85-677 du 5 juillet 1985 tendant à l'amélioration de la situation des victimes d'accidents de la circulation et à l'accélération des procédures d'indemnisation ;

3.2 Contexte européen

3.2.1 RGPD applicable depuis le 25 mai 2018

Le Règlement n°2016/679, du 27 avril 2016 dit règlement général sur la protection des données (« RGPD ») encadre la circulation et le traitement des données à caractère personnel (sécurité et confidentialité).

Le RGPD s'inscrit dans la continuité de la Loi française Informatique et Libertés n°78-17 du 6 janvier 1978 et renforce le contrôle par les citoyens de l'utilisation qui peut être faite des données les concernant. Les obligations déclaratives sont toutes supprimées, sauf exceptions prévues par le droit interne, le RGPD s'inscrivant dans une logique de responsabilisation des acteurs, soumis à des obligations de compliance par diverses mesures d'ordre technique et organisationnel.

Il harmonise les règles sur le territoire de l'Union Européenne (UE) en offrant un cadre juridique unique aux professionnels. Il permet de développer leurs activités numériques au

sein de l'UE en se fondant sur la confiance des utilisateurs. Le RGPD s'applique à toute organisation, publique et privée, qui traite des données personnelles pour son compte ou non, dès lors qu'elle est établie sur le territoire de l'UE ou que son activité cible directement des résidents européens (monitoring ou offre de biens ou services, même à titre gratuit).

3.2.2 Proposition de Règlement e-privacy

La proposition de Règlement³ « vie privée et communications électroniques » du 10 janvier 2017 dite « e-privacy » abrogerait la Directive 2002/58/CE et complèterait le RGPD pour les données de communications électroniques pouvant être considérées comme des données à caractère personnel. La proposition de Règlement concerne également les données des personnes morales et la protection du secret des affaires.

La révolution RGPD ne sera complète qu'après l'adoption dudit règlement destiné à remplacer la directive de 2002. Ce règlement est supposé encadrer notamment toutes les activités de marketing dans la société de l'information.

Pour l'heure, les Etats membres de l'UE ne parviennent pas à trouver de compromis quant au contenu du futur règlement e-privacy, dont l'adoption est toujours en suspens⁴.

3.2.3 Le Règlement « free flow of Data »

Le Règlement 2018/1807 du 14 novembre 2018⁵, entré en application le 28 mai 2019, consacre un principe de libre accessibilité des données non personnelles générées ou traitées au sein de l'UE. Hors sécurité publique dans le respect d'un principe de proportionnalité, les exigences de localisation des données sont interdites. Une exigence de localisation de données signifie toute obligation, interdiction, condition, limite ou autre exigence prévue par le droit d'un Etat membre ou résultant des pratiques administratives générales et cohérentes dans un Etat membre et les organismes de droit public, notamment dans le domaine des marchés publics.

Ce texte prévoit également une coopération intracommunautaire entre les autorités des différents Etats membres et crée à cet effet un mécanisme de guichet unique, correspondant à un point de contact, que chaque Etat membre devra désigner et à qui toute autorité d'un autre Etat membre pourra s'adresser pour accéder à des données.

Le 29 mai 2019, la Commission Européenne a également publié des lignes directrices⁶ en grande partie consacrées à l'interaction entre le règlement Free Flow of Data et le RGPD, en

3 Proposition de règlement du parlement européen et du conseil concernant le respect de la vie privée et la protection des données à caractère personnel dans les communications électroniques et abrogeant la directive 2002/58/CE (règlement « vie privée et communications électroniques »)

4 <https://www.droit-technologie.org/actualites/le-reglement-e-privacy-a-vraiment-du-plomb-dans-laile/>

5 Règlement (UE) 2018/1807 du parlement européen et du conseil du 14 novembre 2018 établissant un cadre applicable au libre flux des données à caractère non personnel dans l'Union européenne

6 <https://eur-lex.europa.eu/legal-content/FR/TXT/PDF/?uri=CELEX:52019DC0250&from=EN>

particulier en ce qui concerne les ensembles de données « mixtes », c'est-à-dire composés de données personnelles et non personnelles.

3.2.4 La nouvelle Directive « Open Data » sur la réutilisation des informations du secteur public

Afin de favoriser l'open Data, le Parlement européen a adopté le 20 juin 2019, une nouvelle Directive⁷ concernant les données ouvertes et la réutilisation des informations du secteur public. Elle remplace la directive 2003/98 / CE de 2003 modifiée sur les informations sur le secteur public, également connue sous le nom de « directive PSI ».

Une fois entièrement transposées au niveau national, les nouvelles règles permettront de :

- Stimuler la publication de données dynamiques et l'adoption d'interfaces de programmes d'application ;
- Limiter les exceptions qui permettent actuellement aux organismes publics de facturer la réutilisation de leurs données au-delà des coûts marginaux de diffusion ;
- Élargir le champ d'application de la directive aux :
 - Données détenues par des entreprises publiques, sous réserve de règles spécifiques. En principe, la directive ne s'appliquera qu'aux données que les entreprises mettent à disposition en vue de leur réutilisation. Les redevances pour la réutilisation de ces données peuvent être supérieures aux coûts marginaux de diffusion ;
 - Les données de recherche résultant d'un financement public - les États membres seront invités à élaborer des politiques de libre accès aux données de recherche financées par des fonds publics. De nouvelles règles faciliteront également la réutilisation des données de recherche qui sont déjà contenues dans des dépôts ouverts.
- Renforcer les exigences de transparence pour les accords public-privé impliquant des informations du secteur public, en évitant les accords d'exclusivité.

Les États membres doivent transposer cette directive (UE) 2019/1024 avant le 16 juillet 2021.

3.2.5 Lignes directrices de l'UE sur l'éthique de l'intelligence artificielle

La Commission européenne propose sept principes⁸ à suivre dans ses lignes directrices publiées le 8 avril 2019⁹ :

- Facteur humain et contrôle humain : les systèmes d'IA devraient être les vecteurs de sociétés équitables en se mettant au service de l'humain et des droits fondamentaux, sans restreindre ou dévoyer l'autonomie humaine.
- Robustesse et sécurité : une IA digne de confiance nécessite des algorithmes suffisamment sûrs, fiables et robustes pour gérer les erreurs ou les incohérences dans toutes les phases du cycle de vie des systèmes d'IA.

7 Directive (UE) 2019/1024 du parlement européen et du conseil du 20 juin 2019 concernant les données ouvertes et la réutilisation des informations du secteur public

8 https://ec.europa.eu/commission/presscorner/detail/fr/IP_19_1893

9 <https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai>

[https://www.europarl.europa.eu/RegData/etudes/BRIE/2019/640163/EPRS_BRI\(2019\)640163_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/BRIE/2019/640163/EPRS_BRI(2019)640163_EN.pdf)

- Respect de la vie privée et gouvernance des données : il faut que les citoyens aient la maîtrise totale de leurs données personnelles et que les données les concernant ne soient pas utilisées contre eux à des fins préjudiciables ou discriminatoires.
- Transparence : la traçabilité des systèmes d'IA doit être assurée.
- Diversité, non-discrimination et équité : les systèmes d'IA devraient prendre en compte tout l'éventail des capacités, aptitudes et besoins humains, et leur accessibilité devrait être garantie.
- Bien-être sociétal et environnemental : les systèmes d'IA devraient être utilisés pour soutenir des évolutions sociales positives et renforcer la durabilité et la responsabilité écologique.
- Responsabilisation : il convient de mettre en place des mécanismes pour garantir la responsabilité à l'égard des systèmes d'IA et de leurs résultats, et de les soumettre à une obligation de rendre des comptes.

Le 26 février 2020¹⁰, la Commission européenne est venue rappeler sa position, en proposant une approche européenne de l'IA fondée sur trois piliers :

- Être en avance sur les développements technologiques et encourager son adoption par les secteurs public et privé,
- Se préparer aux changements socio-économiques induits par l'IA,
- Assurer un cadre éthique et juridique approprié.

La Commission souhaite une IA « made in Europe » et centrée sur l'humain.

Deux nouveaux livrables dignes d'intérêt ont été publiés le 19 février 2020 : un livre blanc¹¹, ouvert à consultation, visant à favoriser un écosystème européen d'excellence et de confiance dans l'IA et un rapport sur les aspects de sécurité et de responsabilité de l'IA¹². Le livre blanc propose :

- Des mesures qui rationaliseront la recherche, favoriseront la collaboration entre les États membres et augmenteront les investissements dans le développement et le déploiement de l'IA ;
- Des options politiques pour un futur cadre réglementaire de l'UE qui déterminera les types d'exigences juridiques s'appliquant aux acteurs concernés, avec un accent particulier sur les applications à haut risque.

3.2.6 Règlement délégué « MMTIS » du 31 mai 2017

Ce règlement n°2017/1926 du 31 mai 2017¹³ encadre l'ouverture des données d'offre de transports pour mettre en place l'information multimodale sur les déplacements.

10 <https://ec.europa.eu/digital-single-market/en/artificial-intelligence>

11 https://ec.europa.eu/info/files/white-paper-artificial-intelligence-european-approach-excellence-and-trust_en

12 https://ec.europa.eu/info/sites/info/files/report-safety-liability-artificial-intelligence-feb2020_en_1.pdf

13 Règlement délégué (UE) 2017/1926 de la Commission du 31 mai 2017 complétant la directive 2010/40/UE du Parlement européen et du Conseil en ce qui concerne la mise à disposition, dans l'ensemble de l'Union, de services d'informations sur les déplacements multimodaux.

Il est d'ores et déjà applicable et a vocation à être décliné dans les droits des Etats Membres, comme le montrent les références faites à l'article 25 de la LOM.

Ce Règlement met en œuvre concrètement la Directive 2010/40 du 7 juillet 2010 (dite « Directive ITS ») pour la mise à disposition dans l'ensemble de l'UE, des données nécessaires à la création des services d'informations sur les déplacements multimodaux. L'ouverture des données théoriques est obligatoire tandis que le choix de mettre à disposition les données dynamiques est laissé à la discrétion des Etats membres. Une série de principes communs sont reconnus, comme l'importance des licences, la prise en compte des coûts de mise à disposition, la nécessaire neutralité et non-discrimination dans la réutilisation des données sont reconnues.

Tous les modes de transport sont visés (publics ou privés, collectifs ou individuels) et une mise au format normalisée est exigée pour favoriser l'interopérabilité.

3.2.7 Le paquet cybersécurité

La Directive (UE) 2016/1148 du Parlement et du Conseil du 6 juillet 2016 dite « Network and Information Security » (« NIS »), a lancé une nouvelle phase dans la politique de l'UE pour renforcer la cybersécurité des réseaux et des systèmes d'information. Elle a été transposée en droit français par la loi du 15 février 2018 promulguée le 26 février 2018.

Au-delà de cette Directive, la Commission européenne a annoncé mi-septembre 2017 sa décision de doter l'UE de moyens supplémentaires pour répondre aux cyberattaques (« paquet cybersécurité »)¹⁴. Dans ce cadre, le Règlement UE sur la cybersécurité du 17 avril 2019¹⁵, a été adopté, prévoyant :

- Un ensemble de systèmes de certification à l'échelle de l'UE ;
- Une agence de l'UE pour la cybersécurité, qui succédera à l'actuelle Agence de l'Union européenne chargée de la sécurité des réseaux et de l'information (ENISA) et en étendra les compétences.

L'UE pourra également imposer des sanctions aux cyber hackers, ceux qui les soutiennent ou ceux sont impliqués à leurs côtés. Un centre industriel, technologique et de recherche en matière de cybersécurité, sera mis en place, soutenu par un réseau de centres nationaux de coordination, le tout pour contribuer à sécuriser le marché unique numérique et à accroître l'autonomie de l'UE dans le domaine de la cybersécurité.

14 <https://ec.europa.eu/digital-single-market/en/cyber-security>.

15 Règlement (UE) 2019/881 du parlement européen et du conseil du 17 avril 2019 relatif à l'ENISA (Agence de l'Union européenne pour la cybersécurité) et à la certification de cybersécurité des technologies de l'information et des communications, et abrogeant le règlement (UE) n o 526/2013.

4. Valeur ajoutée du Big Data selon les cas d'usages

4.1 Typologie des cas d'usages étudiés

Les illustrations proposées proposent des services pour les 3 acteurs de la mobilité suivants :

- Usagers
- Opérateurs privés ou publics
- Autorités organisatrices de mobilité
- Elles concernent les différents domaines fonctionnels :
 - Transports en commun
 - Partage de la voiture
 - Modes actifs (piétons et cyclistes)
- Logistique
- Intermodalité
- Mobilité connectée
- Gestion espace public
- Mobilité inclusive
- Management de la mobilité
- Gouvernance et planification
- Connaissance de la mobilité
- Modélisation

Les exemples proposés sont tirés en bonne partie du recensement de solutions ITS par France Mobilités, auquel ont largement contribué les participants à notre groupe de travail ; ces solutions ont des niveaux de maturité variables, depuis la solution opérationnelle, jusqu'à l'idée encore au stade de la preuve de concept et qui pourra à terme devenir une solution de mobilité.

4.2 Thématiques de la mobilité supports des cas d'usages étudiés

Les thématiques de la mobilité supports des cas d'usages étudiés sont les suivants :

- Connaissance de la mobilité / évaluation de politiques publiques
Collecte et mutualisation de données de mobilité à partir de plusieurs sources, dont des applications mobiles intégrant une technologie de collecte de géolocalisation et des données de téléphonie mobile, pour comprendre les déplacements des personnes, quantifier les flux de déplacements, simuler l'impact sur les réseaux de transports, ...
- Systèmes d'information voyageurs multimodaux/calculateurs d'itinéraires/systèmes de billettique
Collecte et analyse de données de différents services de mobilité en vue d'informer l'utilisateur sur les différentes possibilités de transports et de combinaisons de transports qui s'offrent à lui, leurs caractéristiques en particulier temps de parcours, voire de payer les services utilisés, la panoplie des usages constituant le MaaS ou mobilité servicielle.

- Exploitation de réseaux/suivi et gestion événements/optimisation capacités et résilience/connexion et coordination multimodale : Usages du Big Data pour la gestion des réseaux et de leurs perturbations.
- Gestion flottes de véhicules : Usages du Big Data pour optimiser l'offre (véhicules comme énergies) à la demande.
- Sécurité routière et sûreté/aides à la conduite/autonomie des véhicules : Usages du Big Data pour la prévention des risques.

4.3 Descriptif des cas d'usages étudiés

4.3.1 Connaissance de la mobilité/évaluation de politiques publiques

Le recueil de données des téléphones (données cellulaires ou GPS), des données de véhicules ou de la billettique apporte des compléments aux enquêtes de mobilité qui restent utiles pour l'analyse de la mobilité et la compréhension des processus de déplacements (de par la connaissance fine des caractéristiques socio-démographiques des interviewés et leurs liens avec les trajets effectués, les modes et les motifs).

Les 2 services en exemple incluent des services clé en main et ne se limitent pas au recueil de données, donc comprennent la partie "analyse". Ils touchent la mobilité présente et les grands flux de déplacements.

Nom cas d'usage	Geo4Cast	Flux Vision
Descriptif succinct	Collecte et mutualisation de données de mobilité à partir de plusieurs sources, dont des applications mobiles intégrant une technologie de collecte de géolocalisation, pour comprendre les déplacements des personnes, modéliser les origines destinations, simuler l'impact sur les réseaux de transports,	Quantifier et qualifier les flux de déplacements à partir des données du réseau mobile
Porteur	IT4PME	Orange
Fonctionnement	Permet une collecte massive, l'intégration de différentes sources, des associations / corrélations, l'optimisation d'un réseau de transport (donc des investissements ou coûts d'exploitation associés) L'information voyageur temps réel et personnalisée grâce à des données brutes (API open source, SDK mobile, FCD), des algorithmes de détection de	Transforme les données de réseau mobile en indicateurs et informations pertinentes pour appréhender la mobilité sur un territoire, représentatif de la demande terrain (Data science dans le redressement des données), permet : - de créer un observatoire des déplacements sur un territoire,

	mode et de redressement et des outputs définis avec les métiers exploitants	avec des données disponibles rapidement - d'identifier l'origine géographique des visiteurs et segmenter les usagers par profil (profils sociodémographiques, récurrence des voyages sur les 15 derniers jours) - d'estimer la temporalité des flux (heure, jour, mois) - d'adapter les infrastructures et les horaires de transport, de prioriser les investissements, de mesurer l'impact d'événements sur les déplacements - de segmenter les voyageurs selon les principaux modes de déplacement (échelle macro) : axes routier, ferré LGV / non GV ou aérien.
Usagers concernés	Toute ville souhaitant optimiser son réseau de transport, sa congestion, pollution. Tout opérateur souhaitant optimiser ses investissements, ses coûts d'exploitation.	Collectivités / Entreprises
Territoires concernés	Territoires urbains / Périurbain / Territoires peu denses	Territoires urbains
Année création	2016	2013 – 2014
Niveau maturité	Phase de déploiement / commercialisation - certaines fonctionnalités sur mesure	5 ans de Commercialisation
Temps déploiement	2 à 3 mois pour des premiers rendus (1 semaine d'intégration pour le SDK Mobile. 1 mois de définition des paramètres en lien avec la problématique à résoudre).	Très rapide dès validation des zones d'étude > couples O/D (quelques semaines)
Coûts	Coût de collecte et d'accès aux données brutes. Abonnement à la solution en fonction des cas d'usages.	Sur devis en fonction du cahier des charges

Références	Optimisation de réseaux de transport (Reims, Rennes) Optimisation des déplacements pour le monde sportif (plus de 2,5M de rencontres annuelles) Part de marché multimodale pour des opérateurs (ferroviaires, aériens) Etude de report modal (Bordeaux)	+ de 200 clients dont les principaux acteurs de la mobilité (RATP...) + présence à l'international : Espagne, Belgique, Roumanie, Slovaquie
-------------------	--	--

4.3.2 Multimodalité/information et calculateurs d'itinéraires/billettique

Un calculateur d'itinéraires relève du Big Data si on considère qu'il intègre une grande diversité de modes et de services de mobilité, ou que le service est sur le cloud vu le nombre de données, ou qu'il utilise les données d'usage du service (requêtes) pour faire de l'analyse.

La recherche d'itinéraires temps réel peut inclure des prévisions de temps de parcours qui impliquent de construire des modèles avec des Big Data de données historisées.

Le MaaS relève davantage du Big Data dans la mesure où il y a la billettique et où l'appli MAAS ou trafic peut inclure une fonction tracking utile pour le guidage temps réel mais aussi pour remonter des données.

Nom cas d'usage	Superhub	Urban Cyclers	Moov'Hub
Descriptif succinct	Calculateur d'itinéraires personnalisé multimodal	Calculateur d'itinéraires personnalisé vélo et incitations par gamification	1ère brique MaaS
Porteur	Gfi	Gfi	Mobility by Colas
Fonctionnement	Permet de combiner tous les modes de transports, replanifier l'itinéraire en temps réel, encourager les trajets intelligents et écologiques et personnaliser le profil utilisateur via l'algorithme de routage multicritères (Prix, temps, distance, CO2, etc) et l'infrastructure	Application mobile basée sur la collecte et l'analyse des données sur le cycliste, permet, par la suite, une aide à la décision pour les collectivités et pour les usagers cette offre représente un service facile, pratique, qui permet un accès à l'information de manière instantanée	Solution digitale intégrée de services de mobilité et de stationnement, permet à l'utilisateur d'avoir accès à tous les services de mobilité via 1 application unique, site web et application à destination des usagers qui permettent de s'informer, d'être guidés, de réserver et de payer simplement, un

	européenne de données sur la mobilité ouverte		observatoire pour la collectivité, permettant d'analyser et de piloter dynamiquement les incitations au changement de comportement
Usagers concernés	Collectivités, usagers	Habitants de la ville par le biais des collectivités, Collectivités locales	Autorités organisatrices, aménageurs publics et privés
Territoires concernés	Territoires urbains / Périurbain / Territoires peu denses	Territoires urbains / Périurbain	Territoires urbains / Périurbain / Territoires peu denses
Année création	2011 - 2013	2016 - 2017	2018
Niveau maturité	POC	Phase de déploiement	En expérimentation
Temps déploiement	+ d'1 an	- 3 mois	De 3 à 6 mois
Coûts	Sur devis	10 centimes / habitant	Fonction du contexte territorial, de la mise en place des services de mobilité
Références	Helsinki, Milan, Barcelone	Prague, Londres et en test à Bogota (en 2020)	Paris Saclay

4.3.3 Exploitation de réseaux/suivi et gestion événements/optimisation capacités et résilience/connexion et coordination multimodale

Coordination multimodale :

Ce type d'applis inclut des données gérées par les opérateurs et les collectivités, et archivées pour construire des modèles.

Nom cas d'usage	City Streaming Toulouse	Nice Smart City	Smart Transport
Descriptif succinct	Modèle prédictif temps réel permettant de prévenir les congestions urbaines par la mise en œuvre de mesures concertées entre les différents acteurs (ceux qui sont à l'origine des congestions comme ceux qui peuvent apporter des	Expérimentation sur Nice métropole pour anticiper, mieux comprendre les congestions sur le territoire et anticiper les mesures d'exploitation associées des PC circulation	Plateforme permettant l'optimisation des performances des transports en commun dans une ville donnée, via l'intermodalité des transports

	solutions : employeurs, AOM, logistique urbaine, ...) du type décalage des horaires, télétravail, augmentations de capacités, synchronisation des modes,		
Porteur	Sopra Steria	Ingerop	Gfi
Fonctionnement	Une plateforme de "Data Streaming" entre les différents acteurs (ville, employeurs, AOM, données environnementales, ...) pouvant influencer sur les congestions et proposer des mesures incitatives qui correspondent à la situation du "temps présent". Exemple de mesures => Décalage du temps de travail, adaptation horaire dans les transports, Incitation à rejoindre des espaces de coworking en zones péri-urbaines ...	Plateforme BigData métropolitaine qui centralise les données des différents services (IoT, environnementales, météo et d'autres modes (tram, bus) pour trouver des explications sur les variations de trafic), accès à ces données grâce à des API, mise en place d'un hyperviseur au niveau du PC d'exploitation permettant d'agréger en TR les données. Mise en œuvre de plans d'actions croisant ces différentes données.	Elle se base sur une plateforme Big Data qui collecte, normalise, analyse, visualise et administre les données.
Usagers concernés	Agglomérations, AOM, employeurs	Collectivités, habitants	Usagers, collectivités
Territoires concernés	Territoires urbains / Périurbain	Territoires urbains / Périurbain	Territoires urbains / Périurbain
Année création	2017	2018	2007
Niveau maturité	Phase d'implémentation	Phase de déploiement	POC
Temps déploiement		1,5 ans	6 mois et +
Coûts	Sur devis	Coût important de l'infrastructure et des outils mis en œuvre	Sur devis
Références	Toulouse	Prague, Nice	Belfort

Gestion d'itinéraires :

Nom cas d'usage	Smart Parking ++	Rechargement borne électrique
Descriptif succinct	Proposition trajet porte à porte avec stationnement temps réel garanti et dernier kilomètre avec des transports verts personnalisé	ADAS professionnelle pour les véhicules électriques permettant de choisir des emplacements pour l'implantation des dispositifs de recharge rapide et d'orienter vers les dispositifs disponibles au plus proche d'une position
Porteur	Sopra Steria	Sopra Steria
Fonctionnement	Application mobile qui propose des places de parking libres et la poursuite du trajet par des moyens de locomotions "verts" : transports en commun et/ou free floating, l'application permet de proposer des parcours selon certains critères (coûts / durées ...). Les données de l'open Data "Transport" sont exploitées pour proposer des parcours multi-modaux adaptés en fonction des critères sélectionnés. L'application cible les trajets de zones péri-urbaines vers les zones urbaines	Choix d'emplacements en fonction des flux de transports les plus utilisés
Usagers concernés	Le grand public, les gestionnaires de parking, les AOM	
Territoires concernés	Territoires urbains / Périurbain	
Année création	2018	2018
Niveau maturité	POC	
Temps déploiement		
Coûts		
Références	Paris	

Gestion des travaux :

Nom cas d'usage	Info Travaux	Qievo
Descriptif succinct	Application mobile fournissant une information pertinente aux citoyens sur les zones de travaux de la ville, la nature des travaux, leur durée, et les désagréments/perturbations engendrés, ...	Plateforme en ligne permettant de mieux maîtriser les flux logistiques des grands chantiers urbains en liant 3 métiers grâce au digital : l'infrastructure (aires de régulation), la signalisation (dynamique pour les itinéraires), et la logistique (gestion des flux/créneaux)
Porteur	Sopra Steria	Mobility by Colas
Fonctionnement	Elle exploite les données de la ville pour en fournir une information précise sur une application mobile	Qievo est un service qui organise les flux de logistique générés par les chantiers urbains en : - instaurant des circuits dédiés (aires de régulation en périphérie urbaine, itinéraires spécifiques modulables amenant aux chantiers) - offrant aux responsables logistiques des chantiers une plateforme unique permettant de planifier leurs créneaux de livraison Qievo donne ainsi la vision et la maîtrise globale de ces flux à l'autorité organisatrice
Usagers concernés	Collectivités	Aménageurs publics/privés en charge de cadrer plusieurs chantiers sur une zone donnée, intervenants de chantier (entreprise de construction, transporteur, logisticien), citoyen
Territoires concernés	Territoires urbains / Périurbain	Milieu urbain contraint
Année création	2018	2018
Niveau maturité	POC	En expérimentation
Temps déploiement		De 3 à 6 mois
Références	Ville de Paris	

4.3.4 Gestion flottes de véhicules

La gestion d'une flotte est du Big Data dans la mesure où le service est vendu par un opérateur global qui gère des milliers de véhicules, ou si on regarde la partie analyse faite sur ces données.

Nom cas d'usage			Electrific
Descriptif succinct	Mesures de régulation sur le réseau de bus en prenant en compte les conditions de circulation et la demande des usagers	Plateforme d'interopérabilité de services de covoiturage permettant d'interconnecter les offres de covoiturage des fournisseurs de services	Suite applicative pour le déploiement de la mobilité électrique permettant d'optimiser la production, la fourniture, le stockage et la consommation d'énergie pour les véhicules électriques
Porteur	Ingerop	IRT System X	Gfi
Fonctionnement	Collecte et historisation des données du trafic routier, profilage par type de jours, collecte de la fréquentation des bus (validation, montée, descentes), application d'une offre de transport associé au profil, Comparaison en TR de profil de l'offre avec les circonstances TR de déplacement, et proposition de mesures de régulation	Combinaison des offres, mise en relation, appariement et paiement pour des usagers inscrits sur des plateformes différentes. Se base sur la technologie blockchain : cette plateforme est invisible pour l'utilisateur ; ce dernier utilise son service habituel au sein duquel sont présentées de façon complémentaire les offres issues des opérateurs partenaires	Cette plateforme technologique collecte et analyse les données liées à la conduite du chauffeur lui permettant de mieux conduire la voiture en optimisant sa conduite et l'énergie, permet une meilleure coordination avec les stations de recharge, optimise la vie des batteries et les trajets, réduit les coûts de charge
Usagers concernés	Opérateurs / AOM / entreprises	Conducteurs et passagers de véhicules légers dont les parcours sont compatibles avec des trajets covoiturés, collectivités locales, entreprises (startup, TPE, PME), opérateurs / groupements	Usagers / Collectivités / Entreprises
Territoires concernés	Territoires urbains / Périurbain	Territoires urbains (quartiers politique de la	Territoires urbains

Année création	2018	ville, agglomérations, métropoles...) Périurbain	
Niveau maturité	Idée	2018	
Temps déploiement	2 ans	POC, expérimentation (2020)	Phase de déploiement (projet H2020)
Coûts		3 à 6 mois	1 an, voire +
Références		Projet Lyon Covoiturage Expérimentation [LCE] Date de mise en œuvre : 2018-2021 axe A6-A7 incluant 4 voies dynamiques réservées aux véhicules à occupation multiple	Sur devis E-wald (forêt bavaroise), Rép Tchèque, e-bus à Barcelone

4.3.5 Sécurité routière et sûreté/aides à la conduite

Le Big Data relève de la partie analyse des données historiques.

Nom cas d'usage		Anais
Descriptif succinct	Observatoire des risques routiers et de la mobilité : plateforme permettant le suivi de conducteurs afin de comprendre et de prévenir les risques d'accidents routiers dans leurs diversités	Service d'aide à la décision pour la gestion du patrimoine routier améliorant la sécurité des usagers de la route sur les territoires en identifiant des zones d'alerte sur les routes par traitement des données issues de capteurs embarqués dans des véhicules
Porteur	Gfi	Mobility byColas
Fonctionnement	Elle s'appuie sur les techniques d'Analytics et Big Data pour analyser les données déclaratives d'un panel de 1500 personnes	Des données de conduite anonymisées, issues d'une communauté de conducteurs volontaires, permettent de détecter des zones d'alerte pour formuler des recommandations d'interventions/aménagements
Usagers concernés	Les collectivités, les assurances, les usagers, les entreprises, les associations	Gestionnaires de patrimoine routier

Territoires concernés	Territoires urbains / Périurbain / Territoires peu denses / Autres	Territoires urbains / Périurbain / Territoires peu denses / Autres
Année création		2018
Niveau maturité	POC/Exp	En phase de déploiement et commercialisation
Temps déploiement	- de 3 mois	De 3 à 6 mois
Coûts	200 k€	Investissements : Setup – paramétrage du logiciel Fonctionnements / Exploitations : Coûts de licence (licence annuelle) - Software as a Service
Références		Lauréat du Prix de l'Innovation de la Sécurité Routière 2019 dans la catégorie "Intervention d'urgence et aménagements de l'espace routier"

4.3.6 Autres fiches solutions France Mobilités illustratives

- Connaissance de la mobilité/évaluation de politiques publiques

<https://www.francemobilites.fr/solutions/roofstreet>

Analyse les déplacements et trajets de 27 millions de porteurs de smartphone en France

<https://www.francemobilites.fr/solutions/quetzal> SDK modélisation transport

La modélisation des déplacements est typiquement un domaine "Data intensif"

- Systèmes d'information voyageurs multimodaux/calculateurs d'itinéraires/systèmes de billettique

<https://www.francemobilites.fr/solutions/roote>

API données services mobilité temps réel en France, avec bientôt la billettique associée

<https://www.francemobilites.fr/solutions/move>

MaaS

<https://www.francemobilites.fr/solutions/compte-mobilite>

MaaS

<https://www.francemobilites.fr/solutions/plateforme-dautopartage-en-marque-blanche>

Service MaaS autopartage

- Exploitation de réseaux/suivi et gestion événements/optimisation capacités et résilience/connexion et coordination multimodale

<https://www.francemobilites.fr/solutions/cognimooove>

Analyse de données usagers TC (billettique SAE)

<https://www.francemobilites.fr/solutions/octocity>

Analyse données fraude TC et parking

<https://www.francemobilites.fr/solutions/chronotruck>

Optimisation fret, comme pour la recherche d'itinéraires

- Gestion flottes de véhicules

<https://www.francemobilites.fr/solutions/bikepredict-redistribution>

Analyse de données VLS pour adapter au mieux l'offre à la demande

- Sécurité routière et sûreté/aides à la conduite/autonomie des véhicules

<https://www.francemobilites.fr/solutions/ellis>

Analyse de trajets automobiles pour réduire les accidents

<https://www.francemobilites.fr/solutions/toucango>

Détecteur d'endormissement par analyses faciales

4.3.7 Synthèse des illustrations

Le Big Data engendre des besoins de technologies de gestion et d'analyse de données qui vont au-delà de la gestion et de l'analyse de données "traditionnelles" (SGBD, tableur...). Ces technologies impliquent des compétences particulières qui sont rares.

Ce qui fait le "Big Data" relève majoritairement soit du volume, soit de la variété (données hétérogènes pour décrire les différents services de mobilité), plus rarement de la vitesse (qui concerne surtout la conduite automatisée - le "temps réel" des ITS est de l'ordre de la minute voire 10 secondes pour la régulation du trafic, guère moins).

Un critère clé est la nature de la solution :

- Fourniture de données plus ou moins brutes / agrégées (de plus en plus via une API)
- Logiciel standard ("progiciel"), de plus en plus en SAAS ou appli mobile : exemple, modèle de simulation, MaaS, outil de Datavis/analytics
- Logiciel sur mesure (exemple : gestion de trafic, observatoire)
- Etude (utilisant des outils de machine learning ou autres modèles)

Un autre critère clé pour classer les solutions Big Data est de savoir quelles données récupère l'utilisateur de la solution (opérateur de mobilité, collectivité).

La solution peut être "Big Data" pour le fournisseur de la solution qui gère les données à un niveau global ou en tout cas mutualisé (cellulaire, FCD, paiements...), mais seule une fraction des données (sur un territoire) est fournie à l'utilisateur, sur un territoire typiquement. A la limite, pour une étude, le client peut n'avoir qu'un document pour seul livrable, et pas de données. A l'inverse, pour un MAAS ou pour un système de gestion de trafic, ou un observatoire, l'opérateur ou la collectivité va devoir gérer les données produites pour son réseau/territoire, et même s'il y en a moins au niveau d'un territoire ça peut relever du Big Data, notamment à cause des données archivées.

Dans toutes les thématiques, celle qui semble le plus directement concernée est la connaissance de la mobilité (observatoire, modèles...).

Pour la plupart des thématiques, le principe est le même : des données peuvent être produites en temps réel (traces de validation, de paiement, de déplacements, mesures diverses), exploitées pour des besoins de régulation, mais aussi archivées et analysées pour des besoins de conception et d'évaluation.

En termes de territoire, la question du Big Data se pose avant tout pour les collectivités et opérateurs des plus grandes métropoles, ou pour des acteurs régionaux ou nationaux.

Pour des territoires moins peuplés, dans la mesure où certains services locaux sont des déclinaisons de services déployés sur France entière voire globaux (via le cloud), indirectement ils peuvent être utilisateurs de services relevant du Big Data, mais a priori les AOM ou opérateurs locaux n'ont pas vraiment de réseaux ni de ressources pour gérer en propre des problématiques Big Data.

5. Difficultés encore à lever, recommandations et questions ouvertes

5.1 Volet technique

5.1.1 Difficultés

Des difficultés résident dans la qualité des données et des systèmes d'informations. Il faut comprendre et maîtriser les données qu'on analyse.

La qualité, la complétude, l'actualité et la pertinence des données (exemple de la qualification des données de localisation d'un arrêt de bus, des horaires, ...) sont importantes pour la fiabilité des usages et des applications.

Le Big Data s'inscrit dans le SI des organisations (AOM, opérateurs, fournisseurs de solutions), et cela pose des problèmes de maîtrise technique, de maintenance, etc.

5.1.2 Recommandations

Il faut investir dans des infrastructures.

Comme tout système d'information, les systèmes d'information liés à la mobilité ouverts au grand public et aux professionnels nécessitent des infrastructures réseau, des serveurs de stockages et de traitement et des systèmes d'échange de données robustes et sécurisés pouvant supporter des modes dégradés. Une blockchain constitue une base de données qui contient l'historique de tous les échanges effectués entre ses utilisateurs depuis sa création. Cette base de données est sécurisée et distribuée : elle est partagée par ses différents utilisateurs, sans intermédiaire, ce qui permet à chacun de vérifier la validité de la chaîne.

Il faut affirmer, dans les contrats, la maîtrise des collectivités sur leurs données (les collectivités ont en charge de nombreuses délégations de service public).

5.1.3 Questions

Que peut-on capitaliser et mutualiser ?

Avec le Big Data vient l'idée que les données de mobilité sont sur le cloud et peuvent être hébergées par des opérateurs globaux, au premier rang desquels les géants de l'internet, ce qui met potentiellement en danger les industriels et fournisseurs de services français et européens, mais également pourrait limiter la maîtrise des AOM et collectivités dans leurs choix, voire leur souveraineté.

Une des pistes est de mutualiser certains efforts dans la communauté ITS pour faire face à ce risque et être plus réactifs : open source, architectures types, standardisation. La transparence des algorithmes constitue un challenge nouveau car nombreuses applications y ont recours sans pour autant en contrôler le bon fonctionnement ou l'impartialité (ex des calculateurs d'itinéraires ou de la sélection des modalités). La loi permet l'audit des applications conduisant à une décision individuelle. Les capacités d'audit des collectivités ou des usagers restent limitées. Les AOM devront auditer les données utilisées (les Big Data ont aussi des problèmes de qualité).

Que font les autres pays européens ? Qu'a-t-on à apprendre, avec qui coopérer plus ?

5.2 Volet organisationnel

5.2.1 Difficultés

L'ouverture à de nouvelles pratiques technologiques avec de nouvelles données (cf Art 13 de la LOM sur l'ouverture des données FCD, données des aides à la conduite, données d'autres véhicules...) apportera de nouvelles réponses techniques pour différents usages de mobilité (comme par exemple la sécurité routière ou la gestion des grands flux...). Cette définition de nouvelles méthodes et de nouveaux outils nécessite un accompagnement au changement d'organismes en charge de ces usages comme le CEREMA ou d'autres, des besoins de R&D

pour les validations scientifiques (complémentarité des données Big Data et des données traditionnelles également pour les validations, domaines de pertinence des données Big Data davantage pertinentes pour mesurer la variabilité temporelle de l'occupation des territoires par des individus, gérer des réseaux en des temps très courts mais se prêtant moins à l'analyse des causalités et des caractéristiques des personnes, représentativité, compétences techniques pour des données souvent opaques, validation régulière face à des technologies très évolutives...

5.2.2 Recommandations

Il faut mettre en place des processus de validation scientifique des données et de leurs usages.

Il faut mettre les moyens pour pérenniser des partenariats pour faire fonctionner les observatoires et autres plates-formes publiques, qu'il faut considérer comme de véritables infrastructures numériques (et pas comme des projets).

Il faut être vigilant aussi sur la représentativité des données (ex : données issues des véhicules connectés détenus par des CSP+, données issues d'acteurs privés ou publics), la neutralité des pratiques... ce qui sous-entend globalement la nécessité de mettre autour de la table les différents métiers (informatiques, transports,...) pour mettre en place une gouvernance de la donnée à l'échelle territoriale. Puissance publique et fournisseurs de données privées ont intérêt à mettre en œuvre conjointement le processus de production et d'usage des données (droit de regard de la puissance publique mais aussi des données utiles car en réponse à ses besoins).

L'association de chercheurs notamment économistes des transports aux projets d'observatoire permettra de définir ensemble des indicateurs pertinents pour les observatoires de mobilité.

La mise en place d'observatoires de données, comme service public de la donnée, permettant le recensement des données, la collecte des données de référence, l'élargissement à des données hors mobilité, et de façon générale l'open Data public mais également privé (openData privée pour des données d'intérêt général par exemple) facilitera le développement de systèmes d'analyse et d'aide à la décision mais aussi de nombreuses applications pour les opérateurs et les usagers ou d'autres bénéficiaires. Le point d'accès national aux données de transport a initié la démarche d'ouverture des données en recensant l'ensemble de l'information voyageur disponible en France <https://transport.Data.gouv.fr/> et l'ARAFER (autorité régulation des activités ferroviaires et routières) <https://www.arafer.fr/> développe un certificat de conformité des données de mobilité. La complétude de la donnée est importante. Il en est de même de la disponibilité, la continuité et la pérennité des données, les évolutions technologiques étant souvent très rapides.

La mise en place d'entrepôts pour l'appariement de données implique sans doute : une gestion par la collectivité garante de la qualité de la collecte, du partage, du traitement et de l'usage des données, plateforme urbaine de services applicatifs, voire régionale ?

Il faut une coopération concrète des acteurs de l'écosystème (collectivités, opérateurs, entreprises et startups, chercheurs) dans des projets pilotes mais aussi pour mutualiser des connaissances et partager une vision commune de l'architecture et des rôles.

La mise en place d'une gestion de la donnée transversale aux différents services de la collectivité est une difficulté organisationnelle à surmonter (décloisonnement des services). La visualisation de données sera amenée à jouer un rôle croissant dans l'élaboration des politiques publiques.

Le domaine a besoin d'expérimentations bout-à-bout (achat/collecte, traitement, usage, revente), de living labs, de plateformes d'expérimentations et d'échanges, d'hackatons suivis et évalués (appariement de données mobilité et hors mobilité, temps réel/temps prédictif, multimode...) fréquents, guichet de financement unique (AAP, PIA...).

Les résultats (documents, données, logiciels) de ces expérimentations doivent être plus ouverts afin de profiter autant que possible à la communauté et évidemment en restant compatibles avec l'intérêt commercial des participants ; on a connu trop de projets pilotes soutenus sur fonds publics qui n'ont eu aucune suite et dont les résultats n'ont finalement pas été rendus publics.

Il faut favoriser le partage de bonnes pratiques et d'outillage.

5.2.3 Questions

Comment évaluer ?

Une évaluation sur l'efficacité, la pertinence et les complémentarités éventuelles des solutions de la plateforme France Mobilités est à mettre en place : un comité scientifique pourrait être envisagé.

La mutualisation est un point clé (missions locales de service public de la donnée, maillage national de structures de mutualisation informatique). Quel est le bon niveau territorial pour mutualiser les compétences et les ressources en matière de données de mobilité ? (Métropoles et Régions, France, Europe ?). Renforcer la coopération au niveau européen : pour quoi faire ? Des projets pilotes, du partage d'expériences, mutualiser des infrastructures, standardiser, des benchmarks, du lobbying ?

5.3 Volet réglementaire

5.3.1 Difficultés

Faire émerger des standards équilibrés (impliquant industriels et utilisateurs) dans des délais raisonnables pour accélérer les services multimodaux et ne pas trop se les faire imposer par des acteurs globaux dominants n'est pas chose simple.

L'absence d'interopérabilité entre les données des différents acteurs transports publics, transports routiers, voiries, véhicules, PC feux, plateformes autopartage, covoiturage, autres... rend certains usages ou applications MaaS difficiles ou peu fiables, voire impossibles. Il faut donc des formats, des standards précis, ce qui passe par la réunion d'experts et d'utilisateurs, la participation aux instances de normalisation française, européenne et internationale, de la normalisation actée par la loi, un décret ou une ordonnance. La conformité à des standards y compris internationaux représente une harmonisation qui facilite les comparaisons.

5.3.2 Recommandations

Il faut suivre l'application du RGPD dans les ITS.

S'agissant de données personnelles (identité, géolocalisation, etc.), le consentement ou autre base légale appropriée, reste de rigueur. S'il s'applique, le recueil du consentement reste une difficulté importante aussi bien d'un point de vue organisationnel que technique.

La pseudonymisation ou quand cela est possible, l'anonymisation des données doivent être recherchées. A noter que l'anonymisation fait sortir du RGPD, elle est donc strictement encadrée. Quoiqu'il en soit, les données localisées sont toutefois difficiles à anonymiser. Le droit ne doit pas empêcher l'innovation notamment dans l'IA, face aux acteurs chinois ou américains qui font course en tête.

5.3.2 Questions

Quelle acceptabilité des Big Data pour les usagers ?

Les utilisateurs finaux sont de plus en plus informés de l'usage qui est fait de leurs données. La situation peut encore largement évoluer en matière d'acceptabilité et d'usage des services.

5.4 Volet humain et social

5.4.1 Difficultés

Gérer les risques liés aux données reste difficile.

Les aspects liés à la sécurité des dispositifs et à la cybersécurité sont pour autant essentiels (protection des architectures, pare-feux, etc.).

La propriété intellectuelle des algorithmes et des données brutes ou retraitées, la responsabilité des traitements des données, les assurances couvrant les risques associés à l'usage des données de mobilité sont autant de problématiques qui nécessitent de nouvelles Compétences et une gestion sociale adaptée.

5.4.2 Recommandations

Il faut former et faire évoluer les compétences, recruter.

C'est une question qui se pose globalement pour toute la filière ITS (pas que pour les industriels ou les AOM par exemple).

Cette difficulté doit être adressée par le développement de nouvelles compétences, que ce soit pour l'individu lui-même, l'entreprise, ou la collectivité : compétences informatiques (machine learning, deep learning, etc.), accompagnement des nouveaux métiers, formation de base pour éviter l'exclusion numérique, réduction fracture numérique dans les territoires peu denses, etc. Les collectivités devront se doter de compétences permettant d'interpréter les résultats des analyses utilisant des Big Data, pour s'en servir correctement comme aide à la décision.

5.4.3 Questions

Comment inscrire les ITS dans la réflexion actuelle sur l'éthique de l'IA ?

Au-delà des compétences, une réflexion plus globale sur la place de l'humain face aux risques de déqualification et de perte d'emploi. Le Big Data et l'IA font peur. Il faut donc améliorer la perception et l'acceptabilité de l'ensemble des systèmes mis en place aussi bien au niveau des usagers que des opérateurs publics ou privés, et ne pas négliger que chaque individu peut aussi contribuer à l'émergence d'innovations (crowdsourcing, etc.).

Il faut remettre l'humain au centre des réflexions et veiller à ce que l'IA reste contrôlé par l'humain.

Faut-il investir dans des formations pratiques sur Big Data mobilité ? Similaires par exemple à ce que fait beta Gouv : <https://beta.gouv.fr/alpha/> ?

5.5 Volet économique

5.5.1 Difficultés

L'achat public reste la principale difficulté.

Des travaux sur des dispositifs de marchés davantage adaptés à l'innovation sont à poursuivre (partenariats d'innovation, etc.). Des achats groupés au travers de marchés de type UGAP pourraient être envisagés.

5.5.2 Recommandations

L'état et les AOM doivent financer les infrastructures.

La promesse d'ITS améliorant vraiment les performances de notre système de mobilité repose sur des services numériques fiables mis en place aux niveaux local, régional, national voire européen, selon le type de service.

Les infrastructures nécessaires au déploiement de la collecte, du stockage, du traitement et de la diffusion des données de mobilité peuvent représenter un coût non négligeable pour

une collectivité et bénéficieraient d'une mutualisation à l'échelon régional avec d'autres systèmes d'information (eau, déchets, énergie, éclairage public, informatique de gestion).

Une infrastructure numérique est désormais à intégrer en sus de l'infrastructure « physique » (ports, gares, rails, routes, etc.). Les coûts sont d'un à deux ordres de grandeur inférieurs mais les budgets doivent être sanctuarisés (typiquement la promesse des ITS est d'améliorer la productivité des infrastructures de 10% pour 1% de leur coût).

Il faut également investir sur la performance énergétique globale de systèmes incluant les serveurs et les calculateurs nécessaires pour les traitements en temps réel des données.

Il y a souvent nécessité d'investir dans du traitement de données, du logiciel, des études, même si les données sont gratuites, afin d'en tirer une information utile. De plus, passer du POC à l'expérimentation est loin d'être accessible à tous.

Il faut faciliter et aider les tests de modèles économiques de MaaS, le partage de risques financiers privé/public.

5.5.2 Questions

Comment évaluer l'apport du Big Data ? Quels indicateurs ?

La chaîne de valeur de la mobilité se trouve bouleversée par la disponibilité et le partage des données (voir Uber, Blablacar et bien sûr les GAFA). De nouveaux modèles économiques sont à envisager. Des travaux avec des économistes sont à mener, autour du partage de la valeur, d'une juste rémunération, de la valeur réelle de la donnée, en particulier lorsqu'elle est valorisée, sans oublier les réflexions sur une fiscalité locale, la fiscalité verte, une redevance sur la donnée, ... Il faudra inventer de nouveaux modèles de partage de données, notamment public-privé ; beaucoup de données utiles sont détenues par le privé et la collectivité en aurait besoin pour remplir ses missions de service public (cf. article 13 LOM sur les données FCD)

Les engagements de performance actuellement intégrés dans les contrats de délégation de service public devront être réexaminés à l'aune de l'efficacité amenée par le Big Data et l'Intelligence artificielle avec le temps réel/prédictif, la possibilité d'une adaptation permanente de l'offre et de la demande.... Il faut accélérer l'émergence des Contrats à Performance Globale.

Enfin, la rentabilité de l'ensemble des systèmes d'information de mobilité reste à démontrer (coût des usages avec les nouvelles méthodes et outils vs les usages développés avec les anciennes pratiques). Il faut construire une argumentation autour des leviers économiques comme la baisse des coûts d'exploitation d'un réseau de transport public ...