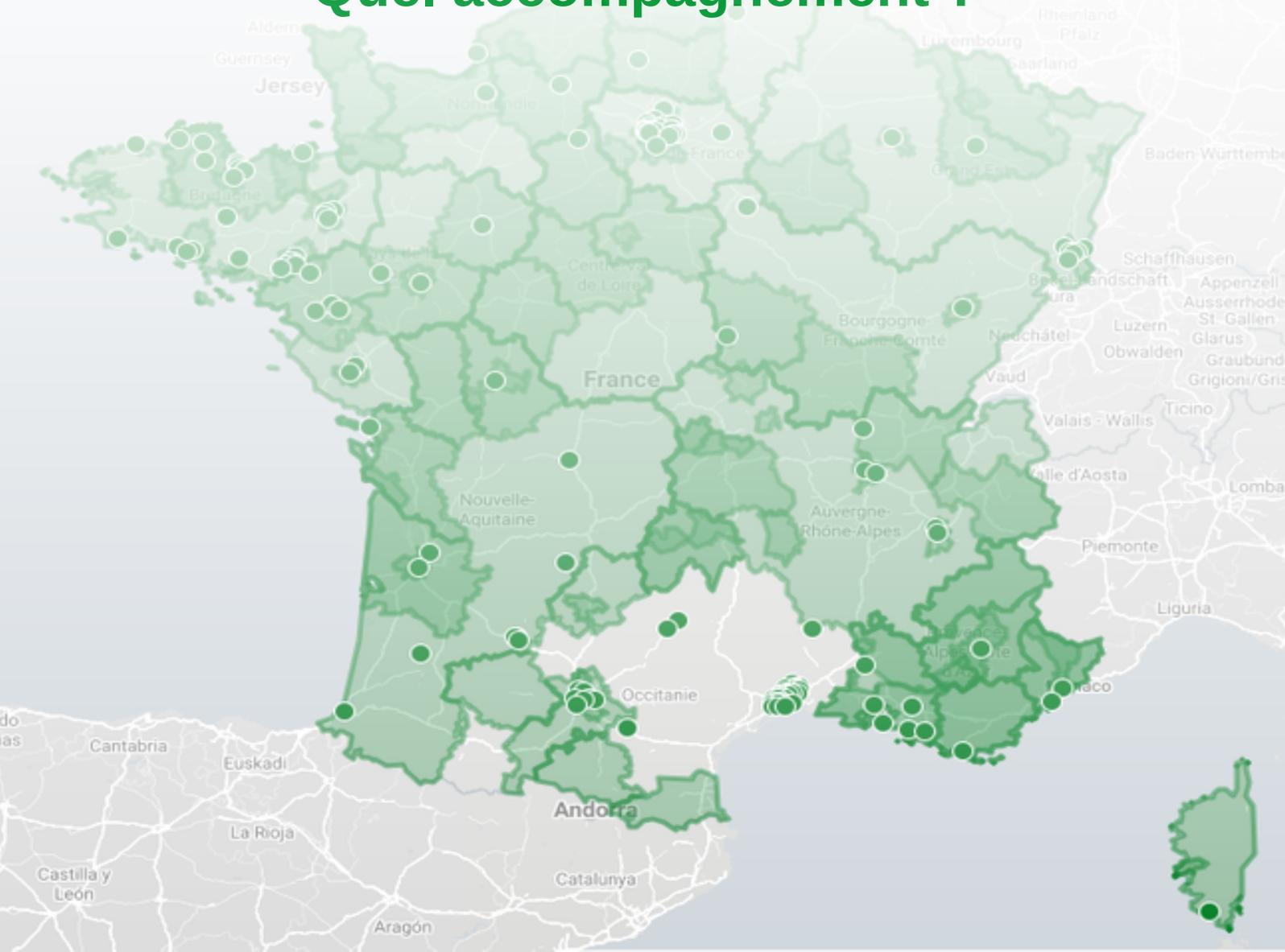


OPEN DATA

Quelle(s) réutilisation(s) ?
Quel accompagnement ?



Sommaire

Brèves

p 4

- Zoom sur l'open data transport
- L'internet des objets dans les collectivités : quelles données pour quels usages
- Eformation : l'Open Data et vous
- Supervision des données : feuille de route 2018-2019
- Qui a ouvert Quoi ? Les résultats du recensement de Dataactivist

Dossier

Introduction - Stratégie d'ouverture locale

p 16

- Entretien avec Laurence Comparat, Présidente d'OpenDataFrance
- « L'open data territoriale : où en est-on ? » article de Jean-Marie Bourgogne, Délégué Général d'OpenDataFrance
- L'exemple du Grand Lyon

1 L'accompagnement

p 25

- DatARA - Urb@Data, une IDG ouverte à l'ensemble des acteurs publics de la région Auvergne-Rhône-Alpes
- Retour sur la mise en place de GéoGuyane
- DataSud : rencontre avec Mathieu Kazmierski et Thomas Bekkers

2 Les réutilisations

p 36

- Entretien avec Mathieu Garnier
- Le Géocodage de la base SIRENE
- Entretien avec Jean-Marc Lazard, Fondateur d'OpenDataSoft
- Exemple entrepreneurial avec LKSpatialist

Petit glossaire réglementaire autour de l'open data

p 46

Lectures

Calendrier

L'ouverture des données n'a jamais autant été au centre des préoccupations de très nombreux acteurs. Ce numéro, comme un point d'étape pour profiter du paysage, montre tout le chemin parcouru et permet aussi d'avoir une idée de celui qui reste devant nous.

Au-delà de la transparence de l'action publique, il s'est agi très tôt de permettre la réutilisation de la donnée pour augmenter sa valeur d'usage tout en partageant ainsi son coût global pour la collectivité qui la produit. L'espoir d'une économie émergeant autour de ces données ouvertes a également été porté par cette vague puissante. Même si cela laisse parfois dubitatif les plus petites structures, la dimension économique conditionne d'une certaine manière le succès de certains projets. Copernicus, le programme européen d'observation de la Terre, avait fait le choix de la gratuité et de l'ouverture des données dès le départ, pour garantir une réutilisation massive et facilitée. Galileo, à l'inverse, disposait d'un modèle économique permettant un retour sur investissement (partiel) par des services payants adressés aux professionnels du positionnement de précision (inférieur au mètre) tout en offrant la gratuité pour le service de base (de l'ordre de 5m de précision). Le mouvement de libération des données pousse aujourd'hui à remettre en question ce modèle économique et à ouvrir les signaux de base de la constellation Galileo qui permettent la création de ces services de précision à valeur ajoutée. La commission européenne y travaille et l'offre concurrente l'y pousse également. Nous y reviendrons dans un prochain numéro.

L'ouverture des données rebat les cartes des modèles économiques, c'est une évidence. Mais avec suffisamment de recul, c'est certainement une belle opportunité, d'autant plus pour la sphère publique. A condition d'y voir clair et d'avoir une stratégie. C'est le cas par exemple du ministère de l'Ecologie qui se dote d'une feuille de route pour assurer la supervision générale de ses données.

Enfin, à titre individuel, nous avons la possibilité d'influer sur nos propres données à caractère personnel depuis le 25 mai dernier. Le côté répressif du RGPD devrait jouer son rôle pour faire évoluer les choses. Mais il y a aussi des démarches nouvelles, comme celle portée par la FING par exemple, qui nous proposent de radicalement changer notre point de vue sur les données que jusqu'ici nous consentions, résignés, à offrir sur un plateau, pire, sur des plate-formes. Ces démarches, #MyData, #SelfData ou autre #MesInfos, commencent à trouver un écho grandissant.

Ce numéro, spécial, a pris soin de faire le tour de nombreux aspects de l'ouverture des données et des mesures d'accompagnement qui lui sont nécessaires. Que tous ceux qui ont bien voulu lui consacrer de leur temps soient ici remerciés. En particulier Antoine Lemot, qui reprend le flambeau de l'assemblage final de votre lecture quadrimestrielle préférée, mais aussi Stéphane Lévêque qui veille à son contenu.

Bernard ALLOUCHE

Cerema Territoires et ville

Brèves

Zoom sur l'open data transport

Stéphane Lévêque

Directeur de projets
Ouverture et usages
des données au Cerema

En 2015, la loi pour la croissance, l'activité et l'égalité des chances économiques (dite loi Macron) rendait déjà obligatoire l'ouverture des données sur les transports publics et les services à la mobilité. Sur la demande du Ministère des Transports, le Cerema accompagne en pratique cette ouverture : nous nous en étions fait l'écho dans le numéro 61 de Sign@ture.

Un observatoire, mis en ligne en septembre 2016 sur la plateforme web Passim, a pour but de regrouper les données concernant les transports collectifs, ainsi que les vélos et véhicules en libre-service, ayant été publiées par les autorités organisatrices de la mobilité. Et afin d'intégrer toutes les données déjà ouvertes sur des plateformes locales, le Cerema a référencé sur data.gouv.fr celles qui n'y figuraient pas encore.

En matière d'open data dans le domaine du transport et de la mobilité, le Cerema a également d'autres actions :

- l'ouvrage "L'Open Data en collectivité à la lumière des données de mobilité" paru en 2015 ;
- le projet Edredon, plateforme open data des données de trafic routier.

En savoir plus sur ces projets : <https://www.cerema.fr/fr/actualites/deux-observatoires-open-data-mobilite-geres-cerema>

The image shows two overlapping screenshots of web portals. The top one is PASSIM, an 'Annuaire des offres de transport et services d'info pour se déplacer en France'. It features a search bar and several filters: 'Territoire', 'Type de transport', and 'Types de données'. Below the filters, it lists '186 fiche(s) Open Data'. The bottom screenshot is data.gouv.fr, the 'Plateforme ouverte des données publiques françaises'. It shows a search for 'Centre d'études et d'expertise sur les risques, l'environnement, la mobilité et l'aménagement (Pôle "Mobilité")'. The search results show a list of data files with icons and titles like 'Tracé des lignes du réseau de transport TGV', 'Horaires en temps réel du réseau de transport TCL Lyon', etc. On the right side of the data.gouv.fr screenshot, there is a sidebar with statistics: '97 jeux de données', '2 membres', '2 abonnés', and a 'Tableau de bord'.

Brèves

L'Internet des objets dans les collectivités

Florent Boithias

Directeur de projets
Villes et territoires
intelligents au Cerema

Internet des objets dans les collectivités : quelles données pour quels usages ?

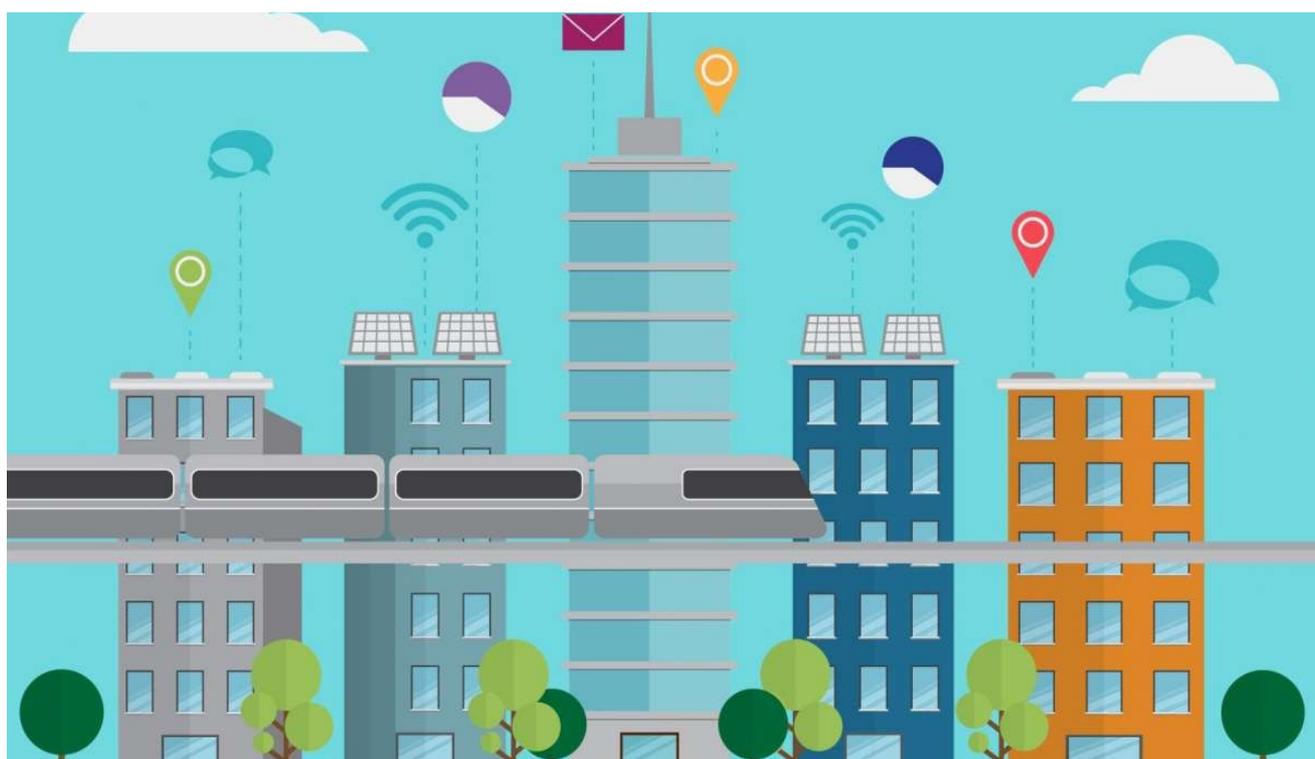
Optimisation de la gestion des réseaux urbains, information en temps réel pour mieux gérer les déplacements... Les promesses de l'Internet des objets suscitent un vif intérêt auprès des collectivités. Mais le marché des technologies n'est pas encore stabilisé et le passage à l'action est donc risqué. Le Cerema a étudié les premiers retours d'expérience, en partenariat avec l'École des Ponts ParisTech et l'association Les Interconnectés, et propose un premier état des lieux sur le sujet.

L'Internet des objets, ou « IoT » pour « Internet of things » en anglais, désigne les infrastructures basées sur des réseaux de capteurs communicants. La remontée d'information en provenance du terrain ainsi générée, la data, permet l'amélioration

de la qualité de nombreux services : aide à la maintenance et à l'exploitation des réseaux, information en temps réel aux habitants, évaluation de politiques publiques, etc.

L'IoT séduit aisément les collectivités, car il a l'avantage de proposer des retours sur investissement directs et mesurables, dès lors que des économies de fonctionnement sont générées par une gestion optimisée des réseaux d'eau, de gaz, de transport, de collecte des déchets, etc. Des modèles économiques plus complexes émergent aussi à travers la valorisation des données recueillies.

Les premiers retours d'expérience commencent à être disponibles, qui ont servi de base à l'étude conduite par le Cerema, en partenariat avec l'École des Ponts ParisTech et l'association Les Interconnectés.



Optimiser la gestion de la ville grâce à l'acquisition de données d'usage

L'usage le plus évident de l'IoT concerne l'optimisation de la gestion des réseaux urbains grâce à la télé-relève, qu'il s'agisse des réseaux d'électricité, de chauffage ou encore d'eau. Ces systèmes permettent notamment de repérer des phénomènes anormaux et coûteux à identifier autrement : fuites d'eau, défaillances d'isolation, mauvaise utilisation des équipements, gaspillage, etc. Au-delà des économies de fluides directement générées par une résolution accélérée des dysfonctionnements, l'information ainsi disponible permet d'optimiser l'organisation de la maintenance et de travailler à l'optimisation de l'usage des infrastructures. Par exemple, à Saint-Sulpice-la-Forêt, village de 1 400 habitants en Ile-et-Vilaine, la mesure des consommations énergétiques a permis une sensibilisation à la bonne gestion des bâtiments publics et une optimisation des agendas d'occupation pour minimiser les dépenses.

L'Internet des objets est également utilisé par les collectivités pour réguler le trafic, au service d'une meilleure qualité de vie et de l'attractivité des centres-villes. Par exemple, la mise en place de capteurs de présence de véhicules sur les places de stationnement constitue un levier pour améliorer la rotation sur les places dédiées au commerce de proximité et pour augmenter le taux de paiement. Ainsi, à Saint-Amand-Montrond, commune de 10 000 habitants dans le Cher, les services de la ville reçoivent une notification en cas de dépassement du temps de stationnement autorisé sur les places « minutes ». Ces capteurs permettent également une application optimisée du forfait post-stationnement, générant des revenus supplémentaires pour la collectivité.

Une connaissance fine des flux routiers en temps réel constitue en outre une aide à la décision pour de nouveaux aménagements de voirie et pour une meilleure régulation du trafic. La Ville de Rouen a quant à elle couplé des données de trafic et de qualité d'air devant une école pour optimiser le trafic en fonction des nuisances mesurées. Elle joue pour cela sur la régulation des

feux tricolores. Il faut cependant disposer d'un itinéraire de substitution afin de reporter le trafic pour que le dispositif fonctionne bien.

Sur la base des données recueillies en temps réel, des services aux usagers peuvent être développés, tels que des calculateurs d'itinéraires multimodaux et de prévision de trafic, comme le projet Optimod à Lyon. Plus simplement, l'information sur la position des transports en commun et donc sur les temps d'attente réels aux arrêts constitue en elle-même une aide à la décision précieuse pour les usagers, susceptible de favoriser leur utilisation des transports collectifs.

Angers, de son côté, a tenté de mesurer le taux de remplissage de ses silos à verre pour optimiser les tournées de ramassage. Les silos étant la plupart du temps pleins en centre-ville, il s'est avéré que le potentiel d'optimisation était faible. Mais en périphérie, le remplissage plus hétérogène et moins fréquent des silos a permis une réelle rationalisation des tournées.



Copie d'écran de l'application Optimod à Lyon

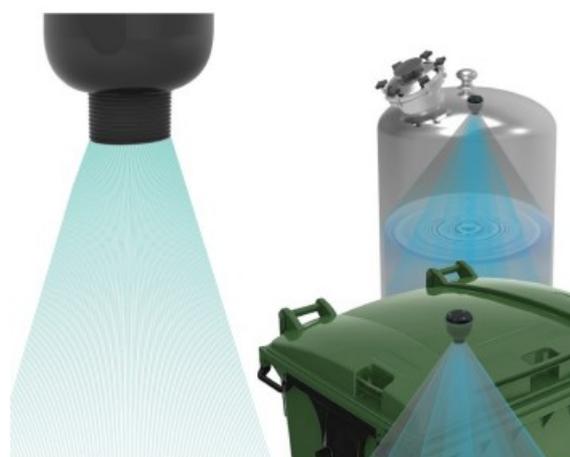


Schéma de fonctionnement d'un capteur de remplissage CleanCap - Site d'E-Cube Labs, <https://www.ecubelabs.com/fr/> (source : étude Cerema)

Pour faire des économies de matériaux et réduire l'impact sur l'environnement, des expérimentations concernant l'optimisation du salage des routes par l'introduction de capteurs dans la chaussée ont été menées par l'agglomération de Paris-Saclay et la Métropole de Lyon. Il s'agit de déterminer en temps réel s'il est plus efficace d'épandre du sable ou du sel, selon la température mesurée, et s'il est réellement nécessaire d'intervenir. Pour ce genre de dispositifs, le retour sur investissement dépend directement du climat local et de la fréquence des épisodes de gel.

D'autres cas d'usages sont également possibles : l'éclairage intelligent, qui permet d'économiser de l'énergie et de repérer les pannes, le contrôle des incivilités par télésurveillance ou encore la mesure de la qualité de l'air et du bruit ambiants, que ce soit de manière collaborative, grâce à des capteurs mobiles, ou par des capteurs fixes.

Obtenir des données de qualité

Pour que les données recueillies soient utiles, il faut avant tout qu'elles répondent à un besoin identifié en amont ou révélé a posteriori. Pour cela, s'appuyer sur les équipes de terrain, qui sont au contact des phénomènes à mesurer et qui sont susceptibles d'exploiter directement les données recueillies, constitue un facteur de succès. Mais après cette phase d'identification du besoin, il faut que les mesures réalisées soient fiables et représentatives des phénomènes physiques. Sur ce point, plusieurs paramètres sont à prendre en compte.

Tout d'abord, la précision intrinsèque du capteur joue sur la qualité de mesure, mais également sur son coût. Ainsi, des capteurs onéreux ne pourront pas être déployés en autant de points de mesure que des capteurs de moins bonne qualité, et un compromis sera alors nécessaire, en complément d'un algorithme d'extrapolation plus ou moins poussé pour couvrir les zones dépourvues de capteurs.

La fréquence d'enregistrement et la fréquence de transmission des données vers une centrale de télé-relève constituent également des para-

mètres importants. En effet, la fréquence d'enregistrement doit être cohérente avec la fluctuation naturelle du phénomène observé, tandis que la fréquence de transmission des données doit être fonction des performances du réseau de communication utilisé. Toutes deux, ainsi que le processus de mesure lui-même, consomment de l'énergie, et influent donc sur l'autonomie du capteur lorsque celui-ci fonctionne sur batterie.

Enfin, le positionnement des capteurs par rapport aux phénomènes observés est lui aussi un élément déterminant. Par exemple, sur la Place de la Nation, à Paris, les capteurs de bruit ont essentiellement mesuré un bruit de fond, ce qui n'a pas permis de cibler le bruit local de la voie de circulation à côté de laquelle ils étaient positionnés.



Photo d'un capteur de qualité de l'air (particules fines, O₃, NO₂) - Site du constructeur : Libelium <http://www.libelium.com> (source : étude Cerema)

En attendant des standards, miser sur l'interopérabilité des formats de données

Le marché de l'Internet des objets n'est pas encore stabilisé : les acteurs dominants ne sont pas encore connus, le potentiel d'innovation est toujours élevé et l'offre n'est pas encore totalement standardisée. De ce fait, il faut rester prudent dans le choix des technologies et éviter l'enfermement dans une unique solution.

Pour cela, il est recommandé de miser sur une interopérabilité maximale, que ce soit entre les réseaux de transmission ou entre les jeux de données recueillis, afin de pouvoir croiser ces derniers et, par exemple, comparer des données anciennes avec des données actuelles.

Concrètement, il faut chercher à conserver un maximum de données contextuelles relatives à la mesure, telles que l'heure de la mesure ou la position précise du capteur, par exemple, pour augmenter les chances de pouvoir s'adapter à un ou plusieurs futur(s) standard(s). Il faut également assurer la portabilité des données lorsque l'on contractualise, afin de les conserver en cas de changement d'opérateur.

Valoriser les données en les libérant, en créant de nouveaux services et en adaptant les tarifications à l'usage

Le développement de l'Internet des objets pose la question du modèle économique à développer pour pérenniser les infrastructures et les services mis en place. Au-delà du retour sur investissement direct lié aux économies de fluides et de maintenance, ainsi qu'à l'optimisation de l'application du forfait post-stationnement, décrits plus haut, des modèles économiques plus originaux apparaissent, s'appuyant sur la valorisation des données recueillies.

Tout d'abord, l'ouverture des données issues de l'Internet des objets (open data) est source de développement socio-économique pour la collectivité, grâce à la création de nouveaux services et à la meilleure diffusion de l'information qu'elle permet.

En parallèle, c'est une économie servicielle qui se développe, grâce notamment aux croisements de données. La collectivité est ainsi en mesure de proposer le service de réalisation d'un itinéraire le plus rapidement possible au lieu d'un simple moyen de transport. De même, pour la gestion des déchets, au lieu de fournir comme prestation le fait de débarrasser les déchets à dates fixes, il s'agit d'assurer un service de ville propre en ramassant les déchets quand cela est nécessaire. Au niveau de la contractualisation, cela se traduit par un développement des marchés portant sur des services et non sur des prestations.

Les modèles économiques évoluent aussi par la transformation de l'usager en producteur de données. En effet, nombre de modèles économiques dits de « plateforme » utilisent la donnée géné-

rée par les utilisateurs eux-mêmes pour assurer le service. La plateforme Waze, par exemple, optimise ses calculs d'itinéraires en fonction des informations en temps réel en provenance de la communauté de ses utilisateurs. De même, à Saint-Amand-Montrond, les habitants peuvent transmettre des mesures de bruit réalisées par leur smartphone directement à leur collectivité via une application dédiée.

Enfin, les objets connectés permettent d'avoir une connaissance plus fine et en temps réel des degrés de fréquentation ou d'utilisation d'un service, et il devient alors possible d'envisager une tarification du service au plus proche du coût social qu'il représente au moment où il est utilisé. Par exemple, dans les domaines de la distribution d'eau ou d'électricité, le modèle dit « d'effacement » se développe, visant à lisser les pics de consommation par une tarification incitative, plus élevée aux heures de pointe.

Tout ceci nécessite des prises de décisions transparentes, associant fortement les citoyens. L'enjeu est en effet la confiance à long terme des administrés dans les dispositifs mis en place, qu'il s'agisse de la protection des données personnelles ou de la transparence des algorithmes qui vont progressivement automatiser les villes.

Accéder à l'étude conduite par le Cerema, en partenariat avec l'École des Ponts ParisTech et l'association Les Interconnectés.



<https://www.cerema.fr/fr/actualites/internet-objets-smart-city>

Brèves

E-Formation

L'Open Data et vous

Cédric Spérandio
Directeur du service de la
modernisation et de la
coordination régionale,
Préfecture Auvergne-Rhône-
Alpes

Fabrice Thiébaux
Directeur de projets en
géomatique et formation à
distance au Cerema

L'open data et vous : répondez aux quiz et obtenez votre badge open data !

En ligne sur le site du Cerema (eformation.cerema.fr) depuis le 2 avril 2018, le nouveau module de E-formation « l'open data et vous » s'adresse à tous les agents publics qui souhaitent comprendre les principes de l'open data et son implication dans leur quotidien professionnel et privé.

Cette formation, proposée par la préfecture de région Auvergne-Rhône-Alpes, a pour objectif d'acculturer les agents publics à l'open data, enjeu devenu majeur notamment depuis la promulgation de la Loi pour une République numérique en octobre 2016, et l'obligation pour les administrations de publier leurs données publiques.

Or, si la réussite de l'administration numérique passe forcément par la formation des agents qui produisent quotidiennement de la donnée, force est de constater qu'aujourd'hui encore l'open data demeure une notion obscure pour une majorité et semble être cantonnée à un milieu d'experts et de spécialistes.

Objectif : donner un premier niveau de connaissance de l'open data aux agents publics

Ce programme de formation a pour objectif d'aider à comprendre les principes de l'open data, ses implications dans la fonction publique et le rôle des agents vis-à-vis de la donnée publique.

Pour répondre au plus près des attentes des utilisateurs, les objectifs pédagogiques de la formation "L'open data et vous" ont été co-construits avec des agents publics. Le module a ensuite fait l'objet d'un test utilisateur avant d'être mis en ligne.

PARCOURS DE FORMATION

**Moins de 3 heures pour comprendre l'open data
et son impact sur vos activités**

**L'OPEN
DATA
ET
VOUS**

e-formation

- ▶ POUR TOUS LES AGENTS DE LA FONCTION PUBLIQUE
- ▶ POUR COMPRENDRE L'OPEN DATA ET SES ENJEUX AU QUOTIDIEN
Pour vous, votre administration et pour la société.
- ▶ UN MODULE DE FORMATION EN LIGNE ADAPTÉ AU RYTHME DE CHACUN
Vidéo, quizz, infographies...

Pour suivre la formation :
eformation.cerema.fr

Logos: République Française, Auvergne-Rhône-Alpes, CEREMATA, Cerema

Le parcours de formation est découpé en cinq thèmes :

- La donnée au quotidien,
- L'open data : un enjeu de société,
- Open data et mission de service public,
- Vous et le cycle de vie de la donnée,
- Savoir trouver de la donnée.

Un module de formation simple, ludique et facilement accessible

D'une durée totale de trois heures, ce programme a été pensé pour permettre aux agents de le suivre à distance et en autonomie. Il peut aussi être proposé à un groupe d'agents en présentiel par un animateur afin de faciliter l'apprentissage, notamment pour les publics les plus éloignés de l'utilisation des outils numériques.

Atout majeur pour la réussite d'une formation en ligne, "L'open data et vous" mobilise plusieurs vecteurs pédagogiques. En plus des textes, des vidéos, des infographies et des quiz permettent en effet aux utilisateurs de tester leurs apprentissages tout au long du parcours.



À l'issue de la formation, les agents peuvent valoriser les compétences acquises grâce à l'obtention d'un badge numérique.

Ce module de formation en ligne a été conçu dans le cadre d'un partenariat entre

la Préfecture de Région et son laboratoire d'innovations territoriales @RCHIPEL, le Cerema, le CNFPT et l'association OpenData France, avec le concours financier du Fonds Innovation Ressources Humaines 2017.

La conception pédagogique, la production des ressources, la mise en forme et la publication de la formation ont été assurées par le groupe Digit@l du Cerema Territoires et Ville, sur la base des connaissances et des savoirs apportés par l'ensemble des partenaires.

Le volet communication du projet a été accompagné par un Junior Lab du master 2 Médiations urbaines, savoirs et expertises (MUSE) de l'Institut de la communication, Université Lyon 2.

Depuis l'ouverture de la formation...

500 inscriptions

407 quiz réalisés par les apprenants

73 badges attribués aux plus courageux qui ont répondu aux 5 questionnaires de validation

Les cadres de catégorie A et A+ sont les plus nombreux à suivre cette formation (63 % des inscrits), suivis par les agents de catégorie B (22 % des inscrits)

70 % des inscrits sont des agents de la fonction publique d'État et 16 % sont issus de la fonction publique territoriale

Le CNFPT héberge la formation sur ces propres outils de e-formation pour la valoriser directement auprès des collectivités

Le Ministère de l'Intérieur va faire de même

Pour accéder à la formation, inscrivez-vous sur eformation.cerema.fr

La superviseure générale des données du MTES et du MCT publie sa feuille de route 2018-2019

La supervision générale des données (SGD) du pôle ministériel MTES (Ministère de la Transition Ecologique et Solidaire) et MCT (Ministère de la Cohésion des Territoires) développe son action au travers de la publication de sa première feuille de route, 2 ans après la remise du rapport de préfiguration.

Pour mener à bien son action, Laurence Monnoyer-Smith, Commissaire générale au développement durable (CGDD) et Superviseure générale des données, s'appuie depuis 2 ans sur une task-force et un comité de correspondants « données » issus des directions générales, de directions régionales et des principaux opérateurs concernés.



Une journée dédiée à la donnée a été organisée le 17 novembre 2017, moment fort destiné à l'administration centrale ou déconcentrée, aux opérateurs, au réseau scientifique et technique et organismes sous tutelle, à tous les profils : du manager aux agents, profil technique, métier ou administratif. Chacun a pu se faire un programme de découverte des différentes facettes de la donnée au travers des deux tables rondes et des vingt ateliers proposés.

Au niveau de l'État, l'administrateur général des données marque dans son rapport au Premier Ministre sa volonté de renforcer le réseau des administrateurs ministériels des données. Le MTES et le MCT sont précurseurs dans ce mouvement.

C'est donc dans un contexte fertile que la feuille de route est mise au point. Son titre « La donnée : un nouveau bien essentiel pour la transition écologique et solidaire et la cohésion des territoires » fait écho au titre du rapport de l'Administrateur général des données (AGD) Henri Verdier parlant d'« infrastructure essentielle ».

La feuille de route « ambitieuse, réaliste et adaptée aux enjeux des services du pôle ministériel » présente 40 actions réparties dans 4 domaines que sont la gouvernance de la donnée, l'architecture de la donnée, l'infrastructure applicative et technique et enfin la sécurité et la conformité. La prise en compte du besoin de montée en compétence à tous les échelons, informatiques comme métiers, apparaît en fil conducteur dans ces actions.

Gouvernance de la donnée

Ce domaine traite des outils et moyens à déployer afin que la donnée, bien essentiel, rentre au cœur des métiers au travers d'une gouvernance à tous niveaux, en lien avec l'AGD. Au sein de la gouvernance à mettre en place, une task-force composée d'agents du SG et du CGDD constitue le bras agile de la SGD pour les

actions quotidiennes de la mise en œuvre de cette gouvernance. Un comité des correspondants « données » des directions générales, de directions régionales et de principaux opérateurs se positionne en relai dans les deux sens pour alimenter la réflexion et faciliter tous les échanges.

Un corpus de règles sera construit qui permettra de disposer d'un cadre de gestion et d'utilisation des données clair et partagé.

Architecture de la donnée

Une production et une utilisation optimales des données ne vont pas sans disposer d'un inventaire de celles-ci muni de critères de classification pour en faciliter la recherche. Pour ouvrir les usages de données métier, les valoriser et les enrichir au travers de croisements, il convient, au départ, aux producteurs d'en modéliser les concepts afin qu'ils soient appropriés par les autres métiers. Il convient aussi de travailler dans le cadre plus large de l'État, notamment en s'acculturant au « service public de la donnée » ou en insérant nos travaux au regard du principe « dites-le-nous une fois ».

Infrastructure applicative et technique

Pour faciliter la transmission, le partage, la valorisation des données par les métiers, il faut alléger au maximum le poids technologique de l'infrastructure tout en offrant des briques technologiques modernes et des exemples d'infrastructures qui donnent envie d'« y aller », de découvrir et de construire. De nouveaux outils sont à mettre en place pour faciliter l'usage des données par le « métier ».

Sécurité et conformité

L'ouverture des données ne va pas sans un minimum de précautions et de mise en conformité aux règles : tenir un registre des traitements dans le cadre de la mise en œuvre du RGPD, gérer les droits d'accès et tracer les accès aux données. La SGD accompagne les services métiers dans ces étapes qui ne doivent pas devenir des obstacles à l'ouverture des données.

Pour lire la feuille de route :

<https://www.ecologique-solidaire.gouv.fr/missions-du-superviseur-general-des-donnees#e0>

« Qui a ouvert quoi ? » le recensement des données des villes réalisé par Dataactivist

Entre avril et septembre 2017, la stratégie sur l'open data de 16 grandes villes françaises (Paris, Marseille, Grand Lyon, Toulouse, Nice, Rennes, Nantes, Montpellier, Bordeaux, Lille, Angers, Grenoble, Poitiers, Strasbourg, Martignes et Saint-Malo) a été étudiée et fait l'objet d'un recensement. Ce travail, réalisé sur plus de 400 jeux de données, a été mené par **Timothée Gidoin**, membre de Dataactivist.

Les objectifs de ce recensement étaient :

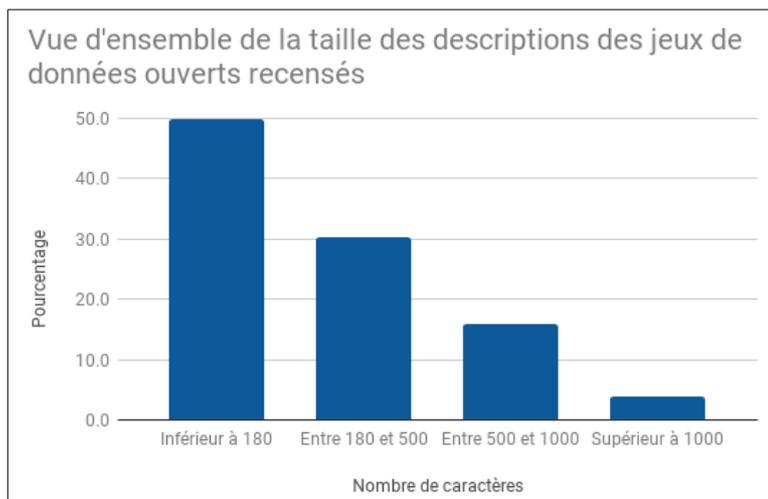
- d'obtenir un « premier panorama des données ouvrables par les villes »
- d'avoir « un outil d'évaluation de la "communicabilité" d'une donnée : le fait qu'une ville donnée a ouvert un jeu de données similaires prouve qu'il est possible légalement d'ouvrir ces données dans une autre collectivité »
- « de proposer un outil de pilotage d'une politique d'open data » qui permet de guider les chef-fe-s de projet d'open data vers de nouveaux jeux de données à ouvrir.

Sur la forme des stratégies d'open data, les résultats du recensement mettent en avant 3 sources de frictions principales et un manque de dialogue avec les réutilisateurs qui pourrait pour autant atténuer ces frictions. Sur leur fond, ils proposent de reconsidérer politiquement l'approche de l'open data.

Première source de friction, « **la grande hétérogénéité des données** » publiées montre la disparité des politiques d'open data entre municipalités et ce malgré le socle commun de données locales proposé par Open Data France (et qui « aide à la mise en œuvre d'un ensemble cohérent, interopérable et prioritaire de données produites au niveau local afin de constituer un gisement national de qualité »¹).

La deuxième source de friction révélée par le recensement est « **une documentation des données laconique** » qui a deux effets négatifs :

- une mauvaise indexation des données et donc des difficultés de "découvrabilité" des données ;
- des difficultés de réutilisation par le public par manque d'informations sur la donnée.



Graphique extrait du blog de dataactivist (<https://medium.com/dataactivist>)

1 Présentation du socle commun des données sur le site d'OpenDataLocal : <http://opendatalocale.net/presentation/>

Enfin la troisième source de friction est le nombre « **des données parfois inexploitable**s ». En cause, la qualité des données, leur actualité, leur complétude, leur structure (colonne en doublon, coordonnées géographiques erronées, etc) qui empêchent ou compliquent fortement leur réutilisation.

Tableau ne veut pas toujours dire données structurées - extrait du blog de dataactivist (<https://medium.com/dataactivist>)

	A	B	C	D	E	F	G
13							
14	Nombre de places assises.....			2 306 places			
15	(dont médiathèques centrales			1 665 places)			
16							
17	Entrées.....			982 793			
18							
19	Inscrits.....			56 821			
20							
21	Prêts.....			1 955 381			
22							
23	Ensemble des collections.....			plus d'1 000 000 de documents			
24	dont :						
25	disques et textes enregistrés			97 792			
26	vidéocassettes et DVD			42 839			
27	partitions			3 547			
28	documents adaptés aux personnes handicapées.....			10 075			
29	les collections patrimoniales :			321 398			
30	pages numérisés.....			226 300			
31							
32	Services informatiques et numérisations						
33	Accès à la vidéo à la demande,						
34	Accès au téléchargement de musique et de livres						

Malheureusement, pour palier à ces 3 sources de frictions, il n'existe que « **peu de mécanismes de remontée des problèmes par les usagers** ». Ainsi, des plateformes d'échanges, du type fil de commentaires, offrant la possibilité d'échanges et d'entraides entre usager n'ont été recensées que dans 6 villes.

Mais plus que « ces sources de frictions » le recensement soutient la conclusion que l'objectif principal aujourd'hui pour les collectivités doit être de « **passer de l'open data "de l'offre" à l'open data "de la demande"** ». En effet, à l'encontre de l'idéal des grands principes de l'open data, les administrations cadenassent trop l'ouverture (elles contrôlent l'ensemble du processus du choix de la donnée à ouvrir à sa publication) sans véritable consultation citoyenne.

Dataactivist propose donc à travers le projet DODOdata de « réveiller les données qui dorment ». Soit de laisser au citoyen la liberté de proposer quelle donnée il souhaite voir ouverte. Les subventions de la ville de Vitry-sur-Seine ou la mise à jour des autorisations de tournages de la ville de Paris en sont des exemples.

Enfin, il est à signaler que Dataactivist étend son champ d'investigation aux données ouvertes par les conseils départementaux et s'est lancé dans l'automatisation par le machine learning de ce travail de recensement.

Les résultats complets et l'analyse du recensement sont à retrouver sur :

<https://medium.com/dataactivist/qui-a-ouvert-quoi-le-recensement-des-donn%C3%A9es-des-villes-est-maintenant-ouvert-b7f697135c1f>

Pour découvrir Dataactivist, il suffit de suivre le lien <https://dataactivist.coop/>

Et le projet DODOdata est expliqué ici : <https://dododata.io/>

DOSSIER

L'ouverture des données publiques prend un essor considérable depuis la loi pour une République numérique, qui acte le principe de l'ouverture par défaut des données publiques et liste des jeux de données de référence via le service public de la donnée.

L'objectif de ce dossier est de faire **un point d'étape qualitatif** de cette dynamique, en mettant l'accent sur les réutilisations et l'accompagnement.

Ainsi nous commencerons par analyser les **aspects stratégiques de l'ouverture des données locales**, avec l'interview de **Laurence Comparat**, présidente d'OpenDataFrance, autour des valeurs et des enjeux actuels de l'open data. En regard de cet entretien, l'article de **Jean-Marie Bourgogne**, Délégué général d'OpenDataFrance, précise l'état d'avancement de l'ouverture des données et les actions en cours. L'illustration du **Grand Lyon** viendra compléter ce tour d'horizon.

Pour parler d'**accompagnement**, nous avons voulu mettre l'accent sur les plateformes régionales : les infrastructures mises en place en Auvergne-Rhône-Alpes – **DatARA** – et en Guyane – **GéoGuyane** – sont en effet en train d'évoluer vers le grand public et l'ouverture des données. En regard de ces articles plus techniques, nous avons voulu un retour qualitatif avec un entretien croisé autour de **DataSud**, questionnant notamment les réutilisations.

La **réutilisation** justement, enjeu fort à venir, est traitée avec deux interviews : celle de **Mathieu Garnier**, réutilisateur enthousiaste, et celle de **Jean-Marc Lazard**, fondateur d'OpenDataSoft. Des exemples de réutilisation – réalisés respectivement par **Christian Quest** et par **LKSpatialist** – illustreront encore plus les propos.

Introduction : stratégie d'ouverture locale	p. 16
1. L'accompagnement	p. 25
2. La réutilisation	p. 36

introduction

Stratégie d'ouverture locale

Rencontre avec
OpenDataFrance :
Laurence Comparat et
Jean-Marie Bourgogne

L'exemple du Grand Lyon

Entretien

Laurence Comparat

Présidente de l'association OpenDataFrance,
Adjointe Open data, logiciels libres et Administration générale
à la ville de Grenoble



Quelles sont les valeurs sous-tendues par l'open data ?

Historiquement, l'ouverture des données publiques s'inscrit dans la mouvance du libre et des communs. On y retrouve le principe d'avoir des biens communs qui vont aider à l'appropriation, la diffusion ou l'élaboration de la connaissance. Il s'agit de permettre aux acteurs, qu'ils soient citoyens, entrepreneurs, chercheurs, ou encore journalistes, de s'emparer de ces données et de faire sens au service de l'intérêt général.

C'est cela qui conduit à l'idée que l'on rend publiques des données, sous des formats qui sont ouverts, facilement réutilisables, et avec une liberté de réutilisation. Le service public peut se retrouver assez facilement dans ces valeurs-là, car il est garant de la transparence de l'utilisation de l'argent public, et de l'intérêt général.

Une fois posée cette philosophie de base, se pose la question de comment les institutions s'y intègrent concrètement. Il y a des enjeux de démocratie, de transparence de la vie publique, mais aussi, à l'échelle des collectivités territoriales, d'animation et de connaissance du territoire.

Cette animation est à destination des réutilisatrices et des réutilisateurs potentiels : citoyennes et citoyens, associations, chercheuses et chercheurs, journalistes, entreprises, etc. L'enjeu est de dialoguer avec ces acteurs-là pour que les données que l'on va rendre publiques puissent servir au territoire, et répondre à des besoins locaux.

Le fait de rendre les données disponibles gratuitement entraîne une perte de visibilité sur les réutilisations, et un dispositif d'animation permet d'avoir des retours : les données sont-elles réutilisées, par qui, dans quel but, conviennent-elles aux utilisateurs, etc.

En quoi ces valeurs s'inscrivent-elles dans le travail d'OpenDataFrance ?

OpenDataFrance a à la fois un rôle d'accompagnement des collectivités locales, en matière d'ouverture des données, un rôle de veille, et un rôle de contact avec différents acteurs. Nous dialoguons par exemple avec les éditeurs de logiciels, pour voir comment accéder aux données stockées dans les applications. Nous avons aussi été en discussion avec l'État, lors de l'élaboration de la Loi Lemaire notamment. Une fois obtenue l'obligation légale d'ouverture des données, la question s'est posée des moyens que l'on se donne pour aider les acteurs publics concernés à la mise en œuvre effective. Il y avait un véritable enjeu à ne pas laisser les acteurs publics seuls, et cela a été tout le sens de la démarche OpenData Locale en 2017.



Cette année, nous portons des projets autour de l'observatoire de l'open data et de la qualité des données (Validata), plus dans une démarche d'analyse globale et d'outillage des acteurs, aussi bien pour celles et ceux qui vont vouloir rendre de la donnée publique, que pour celles et ceux qui vont l'utiliser. Il s'agit d'avoir une vision consolidée de ce qu'est l'open data en France aujourd'hui : connaître ce qui est disponible, mais aussi sa qualité, au sens du format et de la facilité de réutilisation. L'information diffusée est de qualité, car elle sert quotidiennement au fonctionnement de l'institution, mais il s'agit de savoir si elle est réutilisable facilement.

OpenDataFrance a donc un accompagnement à la fois théorique – partager la philosophie de l'open data – et très pratique, avec des outils qui permettent d'ouvrir effectivement les données.

Il y a un véritable enjeu à donner envie d'ouvrir les données, car l'obligation réglementaire – sans moyen associé – n'est pas forcément une motivation suffisante. Il faut faire comprendre tout l'intérêt qu'il y a pour l'institution à entrer dans une dynamique, à s'emparer du sujet pour elle-même.

Avec un recul de 5 ans (date de la création d'OpenDataFrance), ces valeurs portées ont-elles avancées grâce à l'open data ?

Plus généralement où en est le mouvement ?

Il y a eu l'époque des pionniers, époque à laquelle Rennes, Toulouse, Lyon, et d'autres s'étaient lancés dans l'open data et avaient un petit peu essuyé les plâtres. C'était des gens qui connaissaient l'histoire et la philosophie de l'open data, et qui étaient très motivés pour ça.

Aujourd'hui, les choses ont évolué. De très nombreuses nouvelles collectivités sont entrées dans OpenData Locale l'année dernière : on a eu un pic, même si on est encore très loin d'avoir touché tout le monde.

Et on a un profil nouveau : ce sont des acteurs qui s'inscrivent dans le chemin qui a été ouvert, et qui ont peut-être moins de bagage historique et culturel sur la démarche, qui ont moins d'expertise technique sur ces questions-là, qui ont une approche moins « militante ». C'est intéressant parce qu'ils bénéficient de la dynamique qui a été créée par les pionniers, et qu'ils arrivent aussi avec une approche plus pragmatique, plus interrogative, sur le choix des données à rendre publiques par exemple. Ces regards croisés, entre pionniers à la vision stratégique et avec une expérience de plusieurs années, et nouveaux entrants aux questions très pratiques, sont intéressants et provoquent des échanges riches.

Il est difficile de dire si l'open data en France a servi une dynamique démocratique et de transparence : en tout cas elle ne l'a pas desservie.

Toute la question est de savoir ce qui est fait avec les données, et c'est quelque chose de difficile. Sur le principe, on n'a pas à demander de compte à l'utilisateur : les données sont en accès libre. Tout la difficulté réside dans le fait d'avoir des retours, y compris sur l'expertise citoyenne, sans brider la réutilisation. On a des contacts spontanés – à Grenoble par exemple, sur la qualité de l'air, il y a des débats sur les réseaux sociaux – mais c'est difficile à mesurer de manière systématique.

Quels sont les grands types de réutilisations ?

Il y a des réutilisations assez classiques sur les données de transport, qui sont réutilisées y compris par les collectivités elles-mêmes. A la Métropole de Grenoble par exemple, nous avons développé une application locale de mobilité grâce à nos données qui sont par ailleurs en open data. Ainsi, en parallèle à l'ouverture d'un lot de données, on offre aussi un service parce que c'est de notre compétence.

Les données cartographiques sont aussi beaucoup utilisées, y compris par des acteurs associatifs comme OpenStreetMap. Par exemple à Digne, il n'y avait pas de moyens suffisants pour un service SIG en interne, et un travail a été fait avec la communauté locale d'OpenStreetMap pour réaliser, avec les agents de la collectivité et les habitants, la cartographie du territoire. La collectivité utilise les données qui ont été collectées par OpenStreetMap pour gérer sa cartographie en interne.

On a aussi des usages plus commerciaux. Il y a par exemple l'application « Qui dit Miam » qui est partie des

cantines à Toulouse, et offre maintenant un service payant aux collectivités en proposant, sur la base des menus de la cantine, une information des parents sur les menus, les allergènes, etc.

Des utilisations peuvent être plus citoyennes ou militantes, avec les données sur la qualité de l'air, les données financières, etc.

La diversité de ces réutilisations est intéressante : l'open data n'est pas utile qu'à un seul type d'acteur.

L'open data peut-elle favoriser l'émergence de start ups, agir sur l'économie ? Est-ce une motivation des collectivités ?

Cela dépendra du type de collectivité : une commune, une métropole, un département ou une région n'auront pas les mêmes motivations.

Dans les territoires dans lesquels il y a une dynamique autour des start ups – par exemples les villes et régions labellisées French Tech – c'est un des axes de la stratégie open data que de s'adresser à ces acteurs-là.



Au début de l'open data, il y avait l'idée que l'on mettrait les données en ligne et que cela marcherait tout seul. On se rend compte que ce n'est pas vrai. Il faut entrer en dialogue avec les acteurs du territoire, car les données que l'on va mettre en ligne ne sont pas forcément celles que le territoire va attendre. Il faut dialoguer autour des besoins en données pour répondre à plusieurs questions : qui possèdent les données dont les acteurs ont besoin, peut-on les diffuser, faut-il les créer si elles n'existent pas, etc.



On peut alors rentrer dans un cercle vertueux : la mise à disposition des données entraîne l'expression de besoins par les acteurs du territoire, qui entraîne elle-même l'enrichissement des données diffusées par la collectivité.

Les conseils donnés par OpenDataFrance relèvent de la stratégie et de la définition de priorités. Il ne faut pas forcément vouloir diffuser toutes les données tout de suite, mais plutôt y aller progressivement et définir les premières données à ouvrir en fonction d'objectifs stratégiques, et de la cible. On ne rend pas publiques les mêmes données selon que l'on s'adresse au monde économique, aux citoyens, ou aux autres administrations.

Pour éclairer ces choix, OpenDataFrance a défini le socle commun des données locales, qui suggère une dizaine de lots de données que l'on peut ouvrir en premier. Ces jeux de données concernent des compétences répandues, et sont formatés pour rendre plus facile leur création. Une autre façon de démarrer : se baser sur les besoins émis à l'encontre de la collectivité ; de cette façon, on est sûr que les données vont être réutilisées.

Quels sont les dispositifs existants en matière d'animation des acteurs ?

Il n'y a rien de vraiment normé : cela dépend des ressources disponibles sur le territoire. Une commune peut par exemple se rapprocher de son intercommunalité si cette dernière a déjà une démarche open data. Les acteurs locaux peuvent aussi servir de relais, pour la réutilisation ou l'animation : acteurs FrenchTech, associations actives sur le territoire – par exemple Dataactivist ou OpenStreetMap – etc.

Il est important pour une collectivité qui se lance dans une démarche d'open data de ne pas rester seule. Il y a suffisamment d'acteurs de l'open data en France pour en trouver un à proximité, pour échanger des expériences. Ce sont ces réseaux locaux qui permettent un effet d'entraînement.

Il y a peu de personnes formées à l'open data dans les institutions. On a besoin d'une forte montée en compétence sur les enjeux de l'open data, qui posent beaucoup de questions. Se tourner vers des personnes qui ont déjà franchi le pas de l'open data, cela permet de ne pas rester seul avec ses questions. Et cela permet aussi de mutualiser.

C'est aussi l'intérêt d'une association comme OpenDataFrance : être une caisse de résonance et un lieu d'échanges d'expériences.

Y a-t-il des acteurs autres que les collectivités qui ouvrent leurs données ? Sont-ils accompagnés par OpenDataFrance ?

OpenDataFrance est exclusivement en appui des collectivités locales. On estime qu'il y en a 4000 qui tombent sous l'obligation de la Loi Le-maire.

Selon les cas, bien sûr, on rencontre d'autres acteurs. OpenDataFrance est notamment en relation avec Etalab, car les collectivités peuvent utiliser data.gouv.fr, le portail open data de l'Etat, pour publier leurs données.

On peut être en relation également avec les membres de l'enseignement supérieur et de la recherche, avec les services déconcentrés de L'État...

Par exemple, en Auvergne-Rhône-Alpes, dans le cadre d'OpenData Locale, l'un des animateurs de la donnée à été le laboratoire @RCHIPEL, de la Préfecture de Région. Même si ce laboratoire se concentrait sur l'accompagnement des services de L'État, les démarches d'open data sont vraiment des démarches de territoire, et des passerelles peuvent se faire entre les différents acteurs. Les données des uns peuvent servir aux autres.

Quels enjeux voyez vous pour la suite ?

Il y a clairement l'enjeu de la formation des agents et des élus, en lien avec le RGPD, et la transformation numérique des collectivités. Quand on arrive à rendre publique la donnée, c'est qu'il y a eu un travail en interne de structuration et de gestion de la donnée : l'open data est la partie immergée d'autres démarches numériques. Ce sont des questions qui montent fortement, mais il y a besoin d'une remise à niveau globale des agentes et des agents pour être capables de proposer les bons outils, les bonnes méthodes, etc. Ce type de formation à l'organisation numérique et à la donnée devrait faire partie d'une formation à l'arrivée en poste ou des préparations aux concours.

Pour le reste, OpenDataFrance continue la démarche OpenData locale, pour favoriser la montée en puissance des collectivités. De plus, on déploie des outils d'observation – quantitative et qualitative – et de mise en qualité des données diffusées, pour les rendre le plus propre possible afin qu'elles soit facilement réutilisables.

*Entretien réalisé le 11 juin 2018,
par Stéphane Lévêque*

Article

« L'open data territorial, où en est-on ? »

Jean-Marie Bourgogne

Délégué Général d'OpenDataFrance

Plus que quelques semaines et nous entrons dans le régime d'ouverture par défaut des données publiques tel que le prévoit la loi République Numérique, et dont la date d'application au plus tard est le 7 octobre 2018. Les collectivités sont-elles prêtes ? quels enseignements peut-on tirer des efforts menés ces derniers mois pour accompagner les acteurs publics ? quels sont les défis qui se présentent à elles ?

Les écosystèmes s'enrichissent, les offres d'accompagnement s'étoffent et pourtant les collectivités restent encore très démunies au regard des enjeux de la maîtrise des données.

Comme nous l'avons vu pour le RGPD, règlement déjà ancien puisqu'il remonte à 2015, les collectivités assument, ou plutôt affrontent, les échéances avec discipline. Elles essaient de comprendre les enjeux et de transformer les « obligations en opportunités ». Mais cette formule cache souvent un sentiment de contrariété vis-à-vis des priorités et des ressources disponibles de la collectivité. Les collectivités préfèrent-elles vraiment l'étreinte à la contrainte ? Ce n'est pas sûr et dans de très nombreux cas, ce qui était ressenti comme une obligation, une rupture, est perçu après sa mise en œuvre comme un progrès, un avantage, que l'on ne remettrait plus en question aujourd'hui.

Des collectivités pionnières se sont engagées dès 2010 dans l'ouverture des données. Plus de 300 collectivités font de l'open data aujourd'hui, c'est beaucoup et c'est trop peu. Beaucoup parce que les collectivités importantes ayant des compétences dans le champ des grands ser-

vices urbains et de l'économie sont prêtes. La quasi-totalité des régions, des métropoles et des grandes agglomérations ont ouvert leurs données depuis des années et poursuivent leurs efforts en étendant le périmètre de leur action vers la Ville intelligente (optimisation des services urbains par les nouvelles technologies) ou la maîtrise du patrimoine informationnel de la collectivité (nomination d'un Administrateur Général des Données, etc). Mais la situation est bien différente au niveau des petites et moyennes collectivités. Sur un total de 4400 collectivités concernées par l'open data par défaut, seulement 5 % sont engagées dans un projet open data, et ce taux tombe à 0,5 % pour les plus petites (voir les résultats de l'Observatoire open data des territoires). Hormis quelques notables exceptions, l'open data n'est pas le sujet des petites communes, ni des territoires ruraux.

Pourtant, sous l'effet d'entraînement des régions et de pôles métropolitains, l'offre de soutien de petites et moyennes collectivités s'est étoffée : les régions, les départements, les métropoles, les structures de mutualisation (syndicats mixtes et GIP) proposent souvent un accompagnement des petites communes à travers une offre

de formation, de coaching et d'hébergement des données (ex. DataSud, le département de Haute-Garonne, MontpellierMétropole, RedonAgglomération, etc). En 2017, sur les 130 nouvelles collectivités ayant ouvert leurs données, 110 l'ont fait grâce au soutien très actif d'une collectivité de rang supérieur. Cela témoigne de la maturité des collectivités les plus anciennes dans l'open data. Cela se traduit aussi par une capitalisation des bonnes pratiques, une rationalisation des plateformes de publication et une

plus grande homogénéité dans le type de données publiées. La systématisation des dispositifs d'accompagnement locaux permettra sans doute le passage à l'échelle des collectivités ouvrant leurs données dans les toutes prochaines années.

Mais cette factorisation de moyens locaux ne répond pas totalement à deux grands défis : l'interopérabilité des données et l'appropriation des données par les publics ciblés, c'est-à-dire la société civile (citoyens, entreprises, chercheurs et observateurs) et les acteurs publics eux-mêmes.

Des données structurées et compréhensibles

L'interopérabilité est vraiment un point essentiel pour permettre l'émergence de services nouveaux basés sur la donnée et pour faciliter leur appropriation. Pour répondre à ce besoin, il est nécessaire d'établir des normes, de fait ou réglementaires. C'est ainsi que l'État a encadré par décrets la publication des données essentielles des marchés publics et celles des subventions passées par les collectivités. Cela concourt à une publication homogène de données uniformisées au niveau national et cela offre ainsi la possibilité de créer un observatoire de la commande publique. C'est aussi le cas des Plans Locaux d'Urbanisme, des vitesses limites sur les voies, des bornes de recharges électriques, etc. Cela ressemblerait à une liste à la Prévert si nous ne reconnaissons pas que ces données sont indispensables pour nourrir des applications très utilisées (nos GPS), nos voitures électriques, etc

Seules des données structurées et normalisées permettent la constitution de tels référentiels.

La définition de ces standards nécessite une gouvernance pour encadrer la négociation entre les parties prenantes, pour définir les autorités de régulation et les règles de gestion. Cette gouvernance est en cours de mise en place à travers les travaux d'Etalab et d'OpenDataFrance, mais il reste encore beaucoup à faire. Notons que cette démarche révèle la maturité de la communauté « data » qui progresse vers le concept de bien commun informationnel.

La contrepartie de cette amélioration de la qualité « structurelle » des données est sans aucun doute l'augmentation de l'effort de production de telles données normalisées. D'une politique de publication « dans l'état », qui avait l'avantage de faciliter le travail de production, les collectivités doivent s'efforcer de traiter et convertir préalablement les données avant publication. Cette difficulté n'est pas négligeable et peu de collectivités se sont réellement attelées à cette exigence. Pour les aider dans leurs démarches, des travaux sont en cours pour intégrer des modules d'export open data normalisés dans les logiciels et les plateformes d'échange spécialisées pour les collectivités locales.

Des données, pour quoi faire ?

L'ouverture des données devient un élément de la transformation de l'action publique dans de nombreuses collectivités et ce mouvement va en croissant, c'est un fait établi par l'observatoire open data des territoires. En 2017, le nombre de collec-

tivités et de jeux de données ouvertes a cru de près de 40 % par rapport à 2016. La loi république numérique va soutenir et amplifier ce mouvement, on peut considérer qu'il est irréversible.

Pendant l'évaluation de l'impact de l'open data reste encore un sujet peu maîtrisé. Beaucoup d'acteurs publics s'intéressent aux changements concrets que cela engendre au sein des collectivités : Quels sont les usages les plus vertueux ? En quoi cela nourrit le monde économique, la ville intelligente ? Comment estimer le gain de productivité des services ?

Chaque acteur public peut témoigner de réutilisations exemplaires grâce aux données ouvertes : ici un service de tri de déchet collaboratif, là des applications de services de proximité, ailleurs des offres de mobilité multi-modales, informations des parents sur le menu des cantines, etc. Les administrations elles-mêmes exploitent leurs propres données : outils d'analyse des données budgétaire, agenda des événements municipaux, gestion des associations, etc.

De jeunes entreprises du numérique développent de nouveaux services, mais des entreprises classiques enrichissent leur offre grâce aux données en open data : immobilier, tourisme, agriculture. Les chercheurs, les observateurs, les associations utilisent tous les jours les données pour comprendre le potentiel des territoires et mesurer l'action publique.

Il existe de nombreuses études à ce sujet, mais il apparaît nécessaire de collecter et de documenter de façon plus méthodologique les différents usages qui naissent autour des données publiques. Il sera aussi utile, quand cela se justifie, de s'intéresser au retour sur investissement pour les collectivités. Ce chantier de valorisation des usages des données ouvertes est à peine entamé. En collaboration avec la Fondation internet nouvelle génération, OpenDataFrance proposera à la fin de l'année 2018 un chapitre entier sur ce sujet.

Quels sont les prochains défis de l'open data ?

De nombreux autres défis se cristallisent autour de la maîtrise des données : le partage et la gouvernance des données privées, la protection des données à caractère personnel, la donnée des villes et celles des « champs », les algorithmes de traitement, la souveraineté des acteurs publics dans un mode ouvert à tous les opérateurs mondiaux. Nous traiterons ces points lors de prochains articles dans cette revue. Le rendez-vous est donné...



Illustration

La stratégie du Grand Lyon

Encart réalisé par Christophe Badol, à partir des échanges ayant eu lieu lors de la COTITA Centre-Est du jeudi 1er février 2018, consacrée à « l'ouverture des données au service de l'aménagement et du développement durables des territoires ».

Nathalie Vernus-Prost, Administrateur général des données à la métropole de Lyon, anime la plateforme de données du Grand Lyon : www.data.grandlyon.com. La collectivité doit mettre en place un cadre de confiance autour de l'open data : le Grand Lyon est très attentif aux nouveaux usages qui sont faits de ces données, en animant des communautés de réutilisation des données. Le collectivité a mis en place deux licences particulières qui ont permis d'encourager les acteurs privés à diffuser leurs données au côté de la métropole notamment parce que ce dispositif permet d'identifier les réutilisations faites des données.

Pour Nathalie Vernus-Prost, il y a des enjeux politiques forts derrière une politique d'open data. Beaucoup d'acteurs se positionnent comme producteurs de données alternatifs aux institutions publiques. Celles-ci doivent regagner la maîtrise de leur territoire et de leurs politiques publiques, être un vrai contrepoids aux plateformes mondialisées pour offrir des services publics à la hauteur des attentes des citoyens, garantir la gouvernance de la donnée sur les territoires.

La donnée permet à une collectivité de renforcer son rôle de fédérateur des acteurs du territoire pour mener à bien les politiques publiques : ainsi Le Grand Lyon a renforcé l'exercice de sa compétence et sa capacité d'action autour de la maîtrise de la mobilité, à partir des données qu'il produit et rassemble, notamment les données que les délégataires de service public sont tenus de fournir au bénéficiaire ou encore les données du secteur privé. Il procède de même avec sa nouvelle compétence énergie.

Parce que l'open data est avant tout un choix politique, il n'y a pas une, mais différentes politiques d'open data, qui renvoient à différentes finalités de politiques publiques.

En savoir plus sur la COTITA Centre-Est du jeudi 1er février 2018, consacrée à « l'ouverture des données au service de l'aménagement et du développement durables des territoires ».

<https://www.cerema.fr/fr/actualites/ouverture-donnees-au-service-du-developpement-durable>

1 L'accompagnement

Les exemples de :

DatARA-Urb@Data
GéoGuyane
DataSud

Article

DatARA - Urb@Data, une IDG ouverte à l'ensemble des acteurs publics de la région Auvergne-Rhône-Alpes

Yorick Martin, chargé d'études « numérique », SGAR, Préfecture de la région Auvergne-Rhône-Alpes

Les infrastructures de données géographiques (IDG) PRODIGE Auvergne et GEORHONEALPES ont fusionné durant l'été 2017 pour former DatARA. Ce nouveau dispositif, bien que piloté par la Préfecture de région Auvergne-Rhône-Alpes est ouvert à l'ensemble des acteurs publics de la région. DatARA est conçu à l'usage des services de l'État, des établissements publics et des collectivités, mais aussi des organismes privés et des citoyens.

DatARA utilise les technologies PRODIGE (pilotage par la MIG du MTE²) et RESPIRE offrant notamment les fonctions propres aux IDG (catalogage de données géographiques ou géolocalisables, création de cartes, gestion de copies figées ou vivantes des ressources distantes, création de ressources dérivées, production de services web – visualisation et téléchargement, flux

ATOM pour s'abonner, gestion fine des droits, etc.).

Vers un véritable portail open data pour limiter les outils

Dans le contexte réglementaire actuel d'ouverture des données publiques, la Préfecture Auvergne-Rhône-Alpes souhaite ouvrir un portail open data pour ne pas démultiplier les outils mis à disposition du public. De ce fait, d'une infrastructure de donnée géographique, DatARA va devenir un véritable portail open data, un catalogue régional fédérateur des données publiques ouvertes en région Auvergne-Rhône-Alpes. Ces évolutions s'appuient sur le projet Urb@Data.



Données publiques ouvertes
en Auvergne-Rhône-Alpes

² Mission d'Information Géographique du Ministère de la Transition Écologique et Solidaire



Un laboratoire d'innovation pour relever le défi

Urb@Data est un projet incubé par le laboratoire d'innovation territoriale de la Préfecture de région Auvergne-Rhône-Alpes : @RCHIPΣL. Ce projet a pour but d'accompagner au plan technique, les acteurs d'un territoire contraint réglementairement, à libérer leurs données publiques (Directive Inspire, Loi pour une République numérique, etc). Il se veut complémentaire des démarches engagées par le gouvernement tel que le projet OpenData Locale piloté par OpenDataFrance et des initiatives locales.

Urb@Data, pour répondre à ces objectifs, ambitionne :

- d'étendre aux données non géographiques ou non géolocalisables les fonctionnalités des outils de catalogage des portails d'informations géographiques ;
- de proposer aux acteurs publics d'un territoire des solutions pour garantir la diffusion de leurs données en tenant compte des compétences, des moyens (humain et financier) et des différentes stratégies des acteurs ;
- de structurer et organiser (urbaniser) du niveau communal au niveau national, les relations entre les initiatives de ces différents acteurs d'un territoire.

Un chantier ouvert fin 2017 pour une mise en service prévue en octobre 2018

Pour atteindre cette organisation, la Préfecture de la région Auvergne-Rhône-Alpes s'appuie sur la plateforme DatARA. Plusieurs chantiers sont en cours et à venir pour atteindre les objectifs du projet :

- refonte de la charte graphique du site éditorial RESPIRE,
- migration vers la version 4.1 du progiciel PRODIGE,
- définition sous GeoNetwork d'un modèle de fiche de métadonnée ISO 19139 permettant de gérer des données ouvertes non géographiques,

- développement d'un module PRODIGE de moissonnage DCAT pour accéder :
 - dans un premier temps aux web-services de la plateforme data.gouv.fr,
 - et à l'API de PRODIGE permettant de recopier localement les données (gestion de copies vivantes ou figées autorisant des traitements non rélisables par des services web : jointure entre données, etc.).

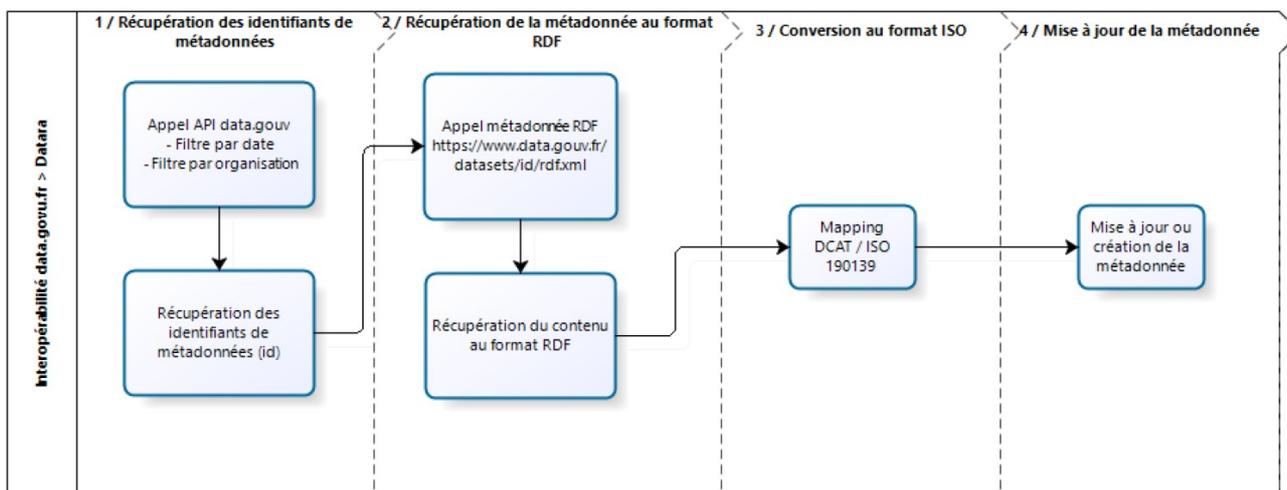
À l'issu de ces évolutions, il sera possible :

- de saisir des données non géographiques dans PRODIGE et de les consulter dans les catalogues des plateformes PRODIGE et data.gouv.fr,
- de saisir, moissonner des données non géographiques de data.gouv.fr et de les consulter dans le catalogue des plateformes PRODIGE,
- de générer des graphiques pour les données non géographiques ainsi intégrées, en plus de cartes pour les données géolocalisables.



Une interopérabilité optimisée au service des utilisateurs

Le module de moissonnage DCAT sera optionnel, au choix des instances PRODIGE. Son processus de moissonnage retenu permet le filtrage des données de data.gouv.fr par organisation et par date. Il permet également la conversion entre la normalisation DCAT utilisée pour les métadonnées dans le domaine de l'open data, et la norme ISO des métadonnées utilisée pour les données géographiques. Son principe peut être illustré de la manière suivante :

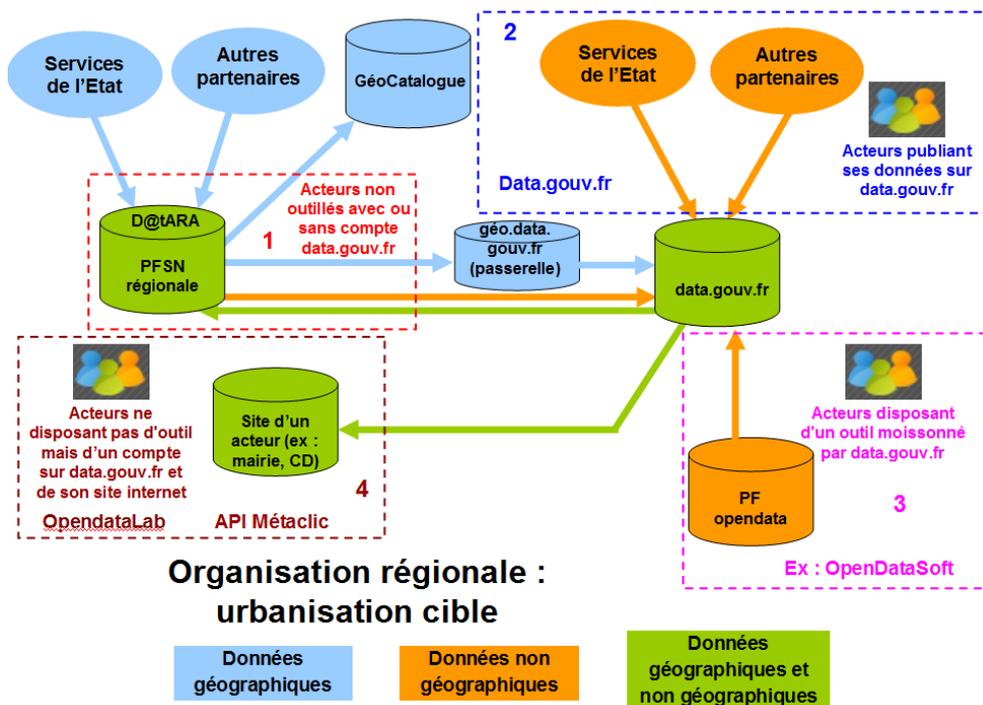


Un utilisateur aura ainsi une vision globale et unifiée de l'offre en données en Auvergne-Rhône-Alpes. Il pourra rechercher, découvrir à travers les métadonnées, télécharger ou représenter et analyser les données géographiques, non géolocalisables, Inspire, non Inspire et ce quelle que soit leur provenance (plateformes IDG territoriales ou thématiques ou data.gouv.fr).

Plus d'infos sur <http://www.datarara.gouv.fr>

Page suivante : schéma présentant l'organisation régionale de DatARA

Organisation régionale de DatARA



Organisation régionale : urbanisation cible

1 – Solution Plateforme de services numériques régionale de type PRODIGE (ex : DatARA)

- Services de l'État et partenaires : acteurs non outillés avec ou sans compte data.gouv.fr,
- Vecteur de publication, de diffusion des données (GeoNetwork : intégrant des nouveaux modèles de fiches de métadonnées ISO 19139 pour des données ouvertes non géographiques),
- Catalogue fédérateur régional (un nouveau module PRODIGE moissonne data.gouv.fr),
- Relai pour le GéoCatalogue et data.gouv.fr.

2 – Solution type data.gouv.fr

- Partenaires, acteurs publiant leurs données sur data.gouv.fr pour exposer ses données. Ils ne disposent ni d'outil ni de site internet.

3 – Solution type propriétaire ou open source

- Partenaires, acteurs publiant leurs données sur leurs propres plateformes moissonnées par data.gouv.fr.

4 – Solution type OpenDataLab (open source)

- Partenaires, acteurs ne disposant pas d'outil mais d'un compte sur data.gouv.fr et d'un site internet pour exposer ses données (API Métaclitic³ : moissonnage data.gouv.fr).

³ MetaClic est un catalogue de données ouvertes libre et gratuit. Il permet d'intégrer dans un site Internet un catalogue des données diffusées par la plateforme <http://www.data.gouv.fr>.

Ce logiciel a été développé dans le cadre du projet <http://www.opendatalab.fr>

Pour en savoir plus : <http://datacode.github.io/metaclic/>

Article

Retour sur la mise en place de GéoGuyane

Article rédigé par Antoine Lemot, suite à l'entretien du 21 juin 2018 avec Sébastien Linares, Responsable de l'unité information géographique, diffusion de la connaissance à la DEAL Guyane

GÉO GUYANE

En Guyane, le partage des données est une préoccupation ancienne partagée par les services de l'Etat et l'Agence d'Urbanisme et de Développement de la Guyane (AUDeG). La conjonction entre développement de SIG et réglementation européenne a abouti en 2013 au lancement de GéoGuyane, Portail géographique de la Guyane. Retour sur ce projet avec Sébastien Linares AJOUTER SES FONCTIONS

Cartographie, Inspire et Prodiges aux origines du projet

La cartographie assistée par ordinateur, parce elle est utilisatrice et productrice de bases de données, a été depuis le début des années 1990, très demandeuse de partage de données.

« Le partage des données est une préoccupation ancienne »

En Guyane, le développement des SIG, impulsé par la collaboration entre AUDeG et service de l'État (DAF et DDE) a ainsi posé les bases d'un projet de plateforme géographique.

Cette dynamique s'accélère en 2007 dans un contexte réglementaire incitant à la diffusion de l'information géographique au public (Directive INSPIRE et circulaire PRODIGE) ; et suite à un séminaire Inspire présentant la toute nouvelle directive européenne, un projet de plateforme est véritablement lancé. Ce projet se concrétise en 2009 avec une demande de financement au fonds européen de développement régional (FEDER)

Puis en 2011, la création de la DEAL s'accompagne du lancement de Prodiges Guyane auquel s'associe l'AUDeG. Si Prodiges est dans ces années encore peu mature, l'opportunité d'un tel outil accessible à moindre coût pousse le projet.

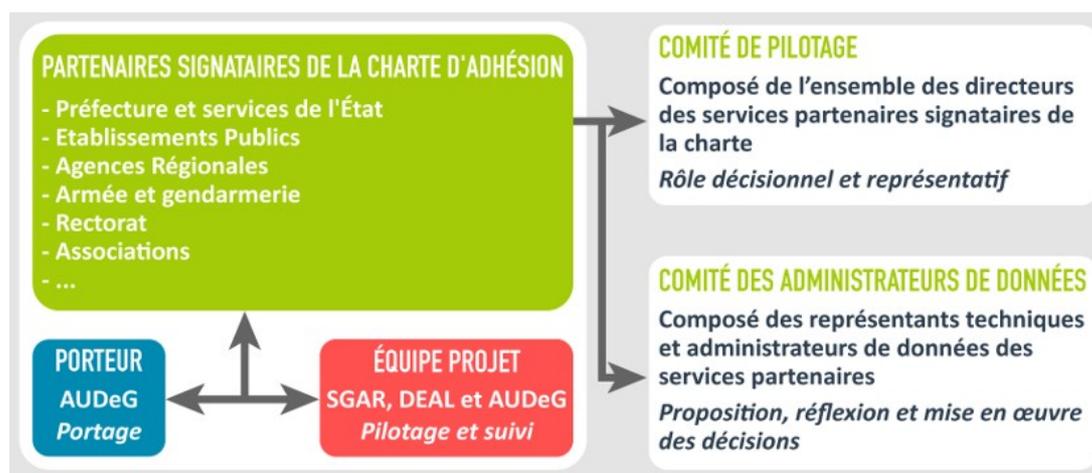
Enfin en 2013, Géoguyane est véritablement lancé avec la signature d'une charte par 20 partenaires.



Organisation

Les principes d'organisation et de fonctionnement du partenariat GéoGuyane sont définis dans sa charte. Ils se basent sur trois instances : un comité de pilotage, un comité des administrateurs de données, une équipe projet. De plus, des groupes de travail spécifiques et autonomes peuvent être créés pour définir, organiser et produire des données.

L' AudeG et la DEAL, acteurs historiques du projet constituent avec le SGAR, l'équipe projet, présidé par le Préfet de Région de la Guyane



De Report et Carmen à Prodiges ; des services de l'État au grand public

Géoguyane permet aujourd'hui de répondre, en un seul outil, à la fois à l'objectif de partage des données et à celui d'ouverture de ces données au grand public. Pour Sébastien Linares, cet aboutissement est venu en deux temps.

Chronologiquement, le partage des données entre services de l'État est le premier objectif, qui apparaît rapidement indispensable dans le domaine des Systèmes d'Information géographique. Pour y répondre, à la fin des années 1990, l'outil REPORTS, développé par le CERTU pour le compte du CNIG et de l'AFNOR, propose principalement la saisie, la consultation et l'import-export des fiches de métadonnées. Premier pas vers le partage des données, cet outil permet donc de répondre à la question : "existe-t-il des bases de données concernant un territoire particulier, dont l'échelle et le contenu répondent à des besoins précis ? Et si oui, où et comment se les procurer ?"⁴

Puis, l'objectif d'ouverture des données au public arrive plus tardivement, moins poussé par des exigences techniques et organisationnelle que par des contraintes réglementaires venues de l'Europe. En 2007 la directive INSPIRE appelle à la mise en place de véritables Infrastructures d'Information Géographique et la diffusion des données géographiques suivant des règles précises. L'application « CARTographie du Ministère de l'ENVironnement » (CARMEN) vient alors répondre à ces exigences en "favorisant l'échange et la réutilisation des données entre différentes sources d'informations (protocoles d'interopérabilité, catalogue, etc)"⁵.

Les outils REPORTS et CARMEN ont donc permis de répondre à ce double objectif de partage et de diffusion des données géographiques, mais séparément et chacun avec des lacunes. Le développement de PRODIGE est venu apporter un outil complet proposant à la fois l'hébergement des données géographiques, la création et le catalogue des métadonnées, la production de

⁴ Le catalogue et les métadonnées : http://archives.cnig.gouv.fr/Front/docs/cms/catalogue-et-metadonnees_124018157232197700.htm

⁵ Présentation de Podige : <http://www.geoinformations.developpement-durable.gouv.fr/presentation-de-prodige-a1568.html>

cartes, et la mise à disposition des données, métadonnées et cartes.

D'un outil aux fonctionnalités plutôt simples – REPORTS – et d'un public restreint – les services de l'État –, la Guyane a su passer judicieusement, en 13 ans, à un outil multifonctionnel et ouvert à tous.

Aujourd'hui sous l'enjeu de démocratisation de l'information, les objectifs du projet sont forts :

- porter à connaissance du public les données produites en Guyane,
- développer la mutualisation et le partage des informations existantes,
- accroître l'efficacité des services producteurs et utilisateurs d'information géographique,
- favoriser le développement de l'information géographique en Guyane.

L'ouverture des données : un paradigme fort à promouvoir

Si la mise à place de la plateforme a été confrontée à quelques problèmes techniques et organisationnels inhérents à tous projets de cette envergure, c'est peut-être la question de l'ouverture des données au grand public qui a apporté des freins à son développement.

Des périmètres métiers (carrières, ICPE), pour lesquels les données étaient considérées encore récemment comme sensibles et ne devant pas faire l'objet de diffusion au public, ont été réticents à participer. Pour Sébastien Linares, cela est donc passé par une « véritable révolution » et de gros efforts de sensibilisation de la part de la DEAL : organisation de journées INSPIRE présentant et vulgarisant la directive, accompagnements personnalisés aidant à la diffusion des données sur le portail.

L'interopérabilité plutôt que le monopole : un gage pour l'open data

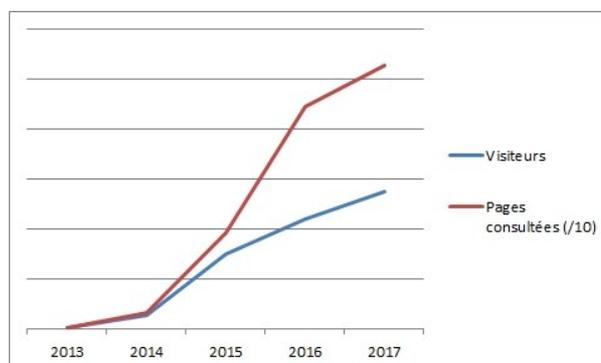
En Guyane, deux plateformes institutionnelles sont proposées : Géoguyane qui est celle d'État et Guyane-SIG celles des collectivités territoriales. Chacune d'elle a donc des modes de gouvernance et de financement différents. Mais, pour Sébastien Linares, loin d'être en concurrence ou de créer une redondance qui "perdrait" les citoyens, ces deux portails sont complémentaires dans un contexte d'interopérabilité dont la priorité est bien l'ouverture des données.

"Géoguyane se développe dans un contexte d'interopérabilité et non de non monopole avec comme priorité l'ouverture des données"

Ainsi depuis 2017 GeoGuyane permet un moissonnage vers Guyane-SIG, comme vers le Geocatalogue et geo.data.gouv.fr. En raison des exigences du format, le moissonnage vers INSPIRE et assuré par l'intermédiaire du Geocatalogue.

Une appropriation démocratique

Entre nombre de données ouvertes et nombre de consultations sur le portail, les résultats sont convaincants. Ainsi en 2018, Géoguyane affichent plus de 800 séries de données actuellement ouvertes dont 1/4 proviennent de la DEAL et 9 894 visiteurs.



une démarche à poursuivre

Les consultations du portail par le public sont un très bon indicateur de sa bonne santé et de son utilité démocratique. Grâce à Prodige, des outils de suivi fins permettent d'analyser l'utilisation de GéoGuyane. Les statistiques montrent que trois sujets sont très "écoutés" : le foncier, les risques majeurs et technologiques, les activités minières. Pour ces dernières, le portail enregistre un pic de téléchargement lors du versement des titres et autorisations d'activité.

Ce constat montre que le public a compris l'opportunité de l'open data, et que à la fois les professionnels du secteur, les associations de protection de l'environnement ou les citoyens ayant un projet immobilier savent se saisir de l'outil pour trouver les informations qu'ils souhaitent consulter et mettre à l'épreuve.

Aujourd'hui, pour Sébastien Linares, l'ambition de Geoguyane est forte dans sa simplicité : améliorer la qualité du service. Cela passe par une consolidation du patrimoine, une meilleure disponibilité et actualisation des données ainsi que par la prise en compte de tous les acteurs privés et publics, spécialistes ou profane et de leur participation.

Le projet se doit de communiquer sur l'open data tant vers les détenteurs de données pour les pousser à les diffuser, que vers le public pour l'inciter à s'en emparer. Ainsi la DEAL répond régulièrement aux sollicitations d'utilisateurs, un travail avec la DAF est en cours et une sollicitation de projet est venue du Rectorat de l'Académie de Guyane.

" Améliorer la qualité du service par "la consolidation du patrimoine, la disponibilité et l'actualisation des données "

Plus d'infos sur <https://www.geoguyane.fr/>



Rencontre avec

Mathieu Kazmierski et Thomas Bekkers

« Le point sur DataSud », propos recueillis et mis en forme par
Mathieu Rajerison

Mathieu Kazmierski est chargé de l'open data au Centre Régional d'Information GÉographique (CRIGE) PACA

et

Thomas Bekkers est en quelque sorte son homologue à la Région Sud Provence-Alpes-Côte d'Azur. Il est chef de projet Données ouvertes et intelligentes au sein du service Smart Région.

Nous les avons interrogés tous deux sur la création de la plateforme de données DataSud, une solution mutualisée de diffusion de données ouvertes et géographiques et intelligentes à l'échelle du territoire régional.

La fusion des plateformes de données géographiques et d'open data en Région Sud

Le 24 juin 2016 a été décidée la fusion de deux plateformes de diffusion de données PACA : celle du géoportail CRIGE PACA (identifié par la loi NOTRE), et celle de la Région appelée Open-PACA. Mathieu Kazmierski, chargé de mission open data au CRIGE PACA relate : « On a coutume de dire au CRIGE que 80 % des données ouvertes sont géographiques, et que 80 % des données géographiques sont ouvertes (enfin presque) ! »

« On a coutume de dire au CRIGE que 80% des données ouvertes sont géographiques, et que 80% des données géographiques sont ouvertes (enfin presque) ! »

PUBLIER
DES DONNÉES

DATA
SUD

TROUVER
DES DONNÉES



OPEN PACA ET CRIGE PACA EN QUELQUES CHIFFRES

+ 142

ORGANISATIONS

+ 766

JEUX DE DONNÉES

+ 700 000

TÉLÉCHARGEMENTS

+ 5 TO

DE DONNÉES

LES PORTAILS DATASUD



Les besoins de faire face aux demandes croissantes des réutilisateurs et de souscrire à des exigences réglementaires ont motivé cette convergence : « Celles de la directive INSPIRE en premier lieu, mais aussi celles qui concernent par exemple les données d'urbanisme qui alimentent les Schémas de Cohérence Territoriale ou les Plans Climat Air-Énergie Territorial, sans oublier les Plans Locaux d'Urbanisme ou les données cadastrales ».

En aparté, Mathieu Kazmierski précise d'ailleurs : « on trouvait déjà il y a 15 ans sur le géoportail du CRIGE les ingrédients d'un "open data avant l'heure" », à savoir : « la mise en place de la première "licence étendue" sur les données de l'IGN ou encore la diffusion gratuite des orthophotographies sur chaque département ».

Thomas Bekkers, chef de projet Open et Smart data au Conseil Régional PACA ajoute que « Les commanditaires partagent également des enjeux communs de première importance, comme les questions de souveraineté, de production de bien numérique commun ou encore de contribution OpenSource ». La question de simplifier et décloisonner l'accès aux données est au cœur de DataSud. M. Kazmierski parle ainsi « de faciliter la lisibilité des dispositifs de contribution en données à l'échelle régionale, en créant un guichet unique de dépôt mariant l'open et le geo ».

La prise en compte de réutilisateurs aux profils variés

« un citoyen ne pratique pas un catalogue de données de la même façon qu'une start-up ou un service de l'Etat. »

Unifier l'accès demande d'anticiper différents profils de réutilisateurs. À ce propos, Mathieu Kazmierski affirme qu'« un citoyen ne pratique pas un catalogue de données de la même façon qu'une start-up ou un

service de L'État. ». Pour ce faire, il ajoute que « (...), dans DATASUD, nous avons fait le choix d'avoir 2 niveaux de description : un niveau "simple" avec les champs basiques et un niveau "exhaustif" ».

En plus de niveaux d'accès différenciés, DataSud propose une palette très riche de services : « une API d'accès au catalogue, une API⁶ d'accès aux données statistiques, un datastore, la possibilité de diffusion des données sur accès restreint, de les synchroniser à distance, et surtout une approche "multi-utilisateurs". (...) Mais une API, même bien documentée comme c'est le cas dans DATASUD, n'est pas à proprement parler à la portée de tout le monde ! ». C'est pour cela que la plateforme propose de « créer des petites dataviz⁷ simples sur les données statistiques ». Toutes ces fonctionnalités ont été conçues avec un mot d'ordre : « simplifier ! ».

Encourager et documenter les réutilisations

Qui dit plate-forme d'échanges sous-entend de pouvoir capter les réutilisations faites des données. Au sujet de la traçabilité des réutilisateurs, Thomas Bekkers et Mathieu Kazmierski regrettent : « Une des forces et faiblesses de l'open data est de maintenir l'anonymat d'accès aux données ». Ainsi, ils ont connaissance des réutilisations principalement par « déclarations volontaires et de connaissance par animation du réseau des acteurs de la donnée en région ».

« Une des forces et faiblesses de l'Open Data est de maintenir l'anonymat d'accès aux données »

Tout sera en tout cas accompli afin de mettre en valeur les réutilisations. Selon eux, « déclarer une réutilisation doit être aussi logique et immédiat que renseigner la date de création d'une donnée ou son producteur », et « à la manière de data.gouv.fr, les réutilisations feront partie des métadonnées d'un jeu de données ». Des événements tels que les hackathons, ou les différents challenges OpenPaca mettent à l'honneur l'open data et les réutilisations, souvent grâce à des utilisateurs avertis. Mais l'open data ne doit pas être qu'affaire de spécialistes. Il est primordial d'assurer l'acculturation d'une frange non négligeable de la population à la dimension data.

⁶ Application Programming Interface (Interface de Programmation)
⁷ Datavisualisation (infographie)



Mathieu Kazmierski détaille à ce titre : « plusieurs projets financés par la Région incluent des actions de formations (des Espaces Publics Numériques) ou d'acculturation comme l'organisation de meetups, de carto-parties ou de projets de sensibilisation de différents publics ». « Nous intervenons à la demande dans des meet-ups, comme avec Nodya ou à l'ESDAC par exemple, ou dans le cadre d'initiatives nationales comme OpenDataLocale, pour parler de données ouvertes et géographiques. (...) Nous avons collaboré dans un Géocamp sur les équipements des gares, participé à l'organisation des State Of the Map 2017 en Avignon. Nous organisons régulièrement des ateliers autour des outils contributifs » pour enfin conclure au sujet de la data literacy : « Le pendant proche de la notion de " littératie numérique " dans notre métier, c'est l'animation de réseaux et d'ateliers pour l'acculturation des élus et de leurs techniciens ».

« Le pendant proche de la notion de " littératie numérique " dans notre métier, c'est l'animation de réseaux et d'ateliers pour l'acculturation des élus et de leurs techniciens »

qu'elle énonce à propos des utilisateurs : « Faire des demandes de données qui ne seraient pas encore diffusées, préciser les faiblesses et problèmes rencontrés en matière d'expérience utilisateur, tester et utiliser les APIs, ou encore nous signaler des erreurs dans les métadonnées ou les données » car « [l']objectif est de reproduire et d'innover dans l'animation autour de jeux de données, identifiées comme des données motrices par la communauté ».

Connaître la qualité des données

Il est essentiel d'assurer ou de préciser le niveau de fiabilité des données, car elles peuvent potentiellement être exploitées par tout type d'utilisateurs. Au-delà de la pure ouverture des données, « viennent ensuite la qualification des données puis l'accompagnement des produc-

teurs et des écosystèmes qui se construisent autour de leurs données. C'est le point d'entrée vers un cercle vertueux d'amélioration en continu des jeux de données ! ».⁸

Avenir de DataSud

Côté données, nous devrions voir ouverte : « (...) une partie des données touristiques émanant du système d'information touristique API-DAE ou [des données] dans les thématiques de la formation et de l'emploi, du patrimoine ou de la mobilité (...) une démarche transverse très active en matière de données Energie (...) ». Enfin, dans le domaine de la Smart City, DataSud espère récupérer « [des] données issues des réseaux et infrastructures intelligentes (capteurs) émanant des projets de Smart City et d'IoT⁹ (...) Smart City et Open Data convergent en termes d'outils et de besoins ». Côté outils : « Une seconde phase de développement prévue en 2018 doit apporter tout un lot de fonctionnalités géographiques et participatives avancées ».

Côté événements : « un hackathon est (...) prévu en 2019. Coorganisé par la Région et Orange, il portera sur la production de services à partir des données de l'Énergie, cela dans un contexte de portage d'un projet de Recherche et Développement "FlexGRID Data" ». « Nous sommes en cours de construction avec la Préfecture de région d'un projet ambitieux pour l'organisation d'un hackathon réunissant services déconcentrés de l'État et start-ups ».

Plus d'infos sur <https://www.datasud.fr/>

« C'est le point d'entrée vers un cercle vertueux d'amélioration en continu des jeux de données ! »

⁸ A ce sujet, voir la démarche QuaDoGeo, avec notamment le séminaire Cerema-CRIGE PACA : <https://www.cerema.fr/fr/actualites/seminaire-quadogeo-mise-pratique-qualification-donnees>

⁹ Internet of things (Internet des objets)

Rencontre avec Mathieu Garnier

Le Géocodage de la base SIRENE

Rencontre avec Jean-Marc Lazard

Exemple de LKSpatialist

Entretien avec

Mathieu Garnier « Matamix »

Mathieu Garnier est un scientifique de la donnée (data scientist) qui s'adonne à de multiples expériences sur la base de données ouvertes, souvent géographiques, et en partage le fruit sur son blog [<https://mtmx.github.io/>] pour le plus grand plaisir de ses lecteurs.

Comment t'est venue l'idée de bidouiller les données ?

La première plateforme d'open data qui a suscité mon intérêt est celle de Paris où j'habitais à l'époque. Elle était encore peu connue, mais on y trouvait quelques jeux de données à partir desquelles j'allais pouvoir effectuer des analyses, principalement des cartes, qui m'intriguaient d'un point de vue personnel. C'est un excellent prétexte pour acquérir de nouvelles compétences techniques, que ce soit dans le traitement des données ou leur visualisation. Grâce à ce portail, j'ai pu m'intéresser aux arbres à Paris, aux résultats des élections par bureau de vote ou encore aux noms de rue grâce à la base Fantoir. Auparavant ces infos-là n'étaient pas structurées et encore moins diffusées par les services municipaux qui ne devaient pas en percevoir l'intérêt potentiel pour le grand public. Mais certains croisements ont pourtant un intérêt, comme ceux de mon analyse sur les bibliothèques montrant dans quels quartiers les 'Tintin' ou les 'Titeufs' étaient les plus empruntés, et la corrélation de ces emprunts avec le niveau de vie. Puis il y a trois ans je me suis formé au logiciel open source R pour élargir mes compétences en statistique et en cartographie. L'ouverture de mon premier blog (qui est tombé dans les limbes d'internet entre-temps) a coïncidé avec cette période de manipulation de données, il me permet de garder une trace de mes travaux en même temps que de les partager.

Quelles réutilisations ont le plus de succès ?

La première réutilisation qui a été plus largement partagée concernait la géographie du sport réalisée à partir des données des licences sportives détenues par le Ministère des Sports. Cette carte (<http://www.slate.fr/story/79066/carte-france-sports-plus-pratiques-commune>) détaillait les



Blog de Matamix : <https://mtmx.github.io/>

sports les plus pratiqués par commune. Ce n'était pas aussi complet que l'atlas publié par le ministère des sports en parallèle de la diffusion des données, mais on apportait un autre regard sur le sujet : j'ai édité les graphiques et les cartes en collaboration avec le journaliste qui a écrit sur le sujet.

C'est le même genre de coopération qui a permis d'aboutir à une série d'articles sur les toponymes (<http://www.slate.fr/story/116421/rues-departements-heros-regionaux>) et les personnages qu'on retrouve le plus sur les panneaux de rues. La base des équidés a également été une belle découverte puisqu'elle a enfin répondu à une question qu'on était nombreux à se poser, à savoir quels sont les noms les plus courants des poneys.

Quelle forme de réutilisation suscite le plus d'intérêt, de curiosité ?

L'appétit pour les cartes a progressé aussi vite que leur réalisation a été facilitée par les outils informatiques. Elles ont généralement plus d'impact que les graphiques car tout le monde, à peu près, sait lire une carte. Lire une carte, l'interpréter est dans la culture d'à peu près chacun. Les gifs, ces images animées qui défilent en boucle, ont aussi du succès, car ils ont une dimension dynamique mais facilement lisible et partageable, j'essaie d'en produire de plus en plus.



LA CAMPAGNE AU POIL

📌 Blog Post

Application pour un coin de campagne idéal

Blog de Matamix : <https://mtmx.github.io/>

Les applications, enfin, représentent une forme très intéressante, mais elles demandent plus de temps en création.

J'ai par exemple créé une petite appli qui permet de distinguer différents types de campagnes à partir de données diverses (https://mtmx.github.io/blog/appli_campagnes/) : relief, élevage avec nombre de vaches et de moutons, points d'eau, données météorologiques Ceci afin de choisir à titre personnel des destinations de vacances. Il n'y a pas longtemps, je me suis rendu dans le Nord-Est dans un coin que m'avait suggéré mon appli à partir des critères que je lui avais fixés. Je m'attelle rarement au développement d'applications parce qu'il est plus chronophage : la manipulation ludique des données doit rester un plaisir et non une corvée...

Es-tu sollicité par des médias, des journaux, pour réaliser des articles basés sur de l'open data

Je ne réalise des articles sur mon blog que quand j'ai le temps, et ma connaissance des sujets est souvent superficielle. La partie bibliographique est souvent réalisée par quelqu'un d'autre. Dans le cas des médias, il s'agit de journalistes. Certains de mes articles sont apparus dans des journaux locaux. Par exemple, ma cartographie des rues de la soif, réalisée sous le titre de bistrographie (<https://mtmx.github.io/blog/bistrographie/>), a eu le droit à un encart dans le journal Ouest France car c'est la rue Saint-Michel à Rennes qui est championne de sa catégorie tandis que mon article sur les "trous perdus" (les endroits les plus éloignés de toute habitation) avait eu droit à une brève dans quelques titres de la presse quotidienne régionale dont un journal local ariégeois.

Comment qualifierais-tu tes lecteurs, si tant est que tu les connais ? Spécialistes de la donnée ou utilisateurs lambda

Je pense que mes lecteurs ne sont en majorité pas des spécialistes de la donnée, ni des techniciens. Les articles publiés sur mon blog ne rentrent pas dans les détails techniques qui pourraient être obscurs pour les non-initiés. Par contre, grâce à la plateforme github sur lequel est hébergé mon nouveau site, les scripts sont disponibles pour quiconque souhaite comprendre la méthode employée voire répliquer l'analyse et l'adapter. Ça m'oblige à documenter les scripts, à soigner le code, et cette rigueur est un bon réflexe à prendre.

Une anecdote assez marrante à propos de l'article sur les trous perdus : un lecteur racontait qu'il avait randonné dans son département pour atteindre ce fameux point et y avait planté un drapeau

Réutiliser de l'open data est-il selon toi à la portée de tout le monde ? Quel est le niveau requis ? Que conseillerais-tu comme méthodes, outils, à celles et ceux qui aimeraient se lancer ?

De plus en plus d'outils libres et accessibles à tous permettent de réaliser des graphiques et des cartes. L'outil magrit (<http://magrit.cnrs.fr/>) réalisé par le CNRS permet avec une grande facilité de réaliser des cartes, par exemple, tout comme kepler.gl (<https://uber.github.io/kepler.gl/#/>) lancé tout récemment. Aussi, sur certains portails open data, des visionneuses graphiques accompagnent dorénavant chaque jeu de donnée et rendent ainsi la donnée brute plus présentable et visuelle sans nécessité de manipulation préalable. L'utilisation de logiciels open-source, dont un des atouts est la gratuité, pour explorer les données ouvertes me paraît vraiment avantageuse aujourd'hui avec une amélioration continue de leur ergono-

mie et de nombreux supports de formation gratuits en ligne. L'enjeu est ce que les experts appellent la littératie des données : la culture de la donnée afin que tout un chacun dispose à minima des bases leur permettant d'accéder à l'univers open data, qu'ils ne soient pas démunis face à des chiffres et qu'ils puissent pourquoi pas déceler des interprétations erronées.

Est-ce que la disponibilité de données pivot, souveraines, a permis de faire de belles réutilisations ?

La base Sirene de l'Insee contient une multitude d'informations intéressantes, sa diffusion gratuite début 2017 était attendue avec impatience. De nombreux services ont été créés pour l'enrichir et la rendre plus facilement interrogeable mais les réutilisations ludiques ont été finalement rares. Je m'en suis servi pour quantifier la présence des bars dans le pays, mais aussi pour illustrer un papier d'un journaliste sur les nouveaux types de restauration et de commerces de bouche en région parisienne. Plus globalement, les données dites « de référence » (<https://www.data.gouv.fr/fr/reference>) ont pris une place considérable dans le paysage : les produits de l'IGN et de l'Insee font aussi partie de ce spectre et avec leur diffusion plus large qu'auparavant, de nouvelles formes d'utilisations ont émergé.



Quelles sont les données que tu utilises le plus ? Et celles dont tu penses qu'il serait intéressant de disposer ?

La base adresse nationale est un outil crucial que ce soit pour les administrations ou les entreprises du secteur privé, je l'utilise régulièrement. D'après les annonces faites à la suite de la Loi pour une République Numérique, on devrait voir arriver de nouvelles sources de données provenant notamment des Services Publics Industriels et Commerciaux. Disposer de données anonymisées d'opérateurs de téléphonie mobile ou de transports afin de pouvoir cartographier les flux de population est un enjeu majeur. Une autre source potentielle m'intéresse à titre personnel : la base des chiens et chats recensés en France (<https://www.i-cad.fr/>) dans laquelle sont renseignées un tas d'infos sur l'animal (nom, couleur, race...) et son propriétaire. Cette donnée est gérée par le Ministère de l'Agriculture mais administrée par un organisme privé. Il est évident qu'un accès à des données anonymisées issues de cette base serait prometteur en termes d'analyses de tous poils sur les liens entre les humains et les animaux.

De façon plus générale, je m'informe moins systématiquement qu'avant sur la publication de petits jeux de données locaux. À mon sens, l'enjeu porte dorénavant sur les gros socles de données structurants, sur leur qualité (documentation, actualisation...) ainsi que sur leur mode de diffusion via des API par exemple. OpenStreetMap, projet collaboratif de grande envergure, s'est rendu indispensable pour l'exploitation d'un tas d'informations géolocalisées. Le web scraping est aussi une alternative pour récupérer des données, je suis notamment un grand fan de wikipedia et de ces listes en tous genres.

Entretien réalisé par Mathieu Rajerison



Illustration

Le géocodage de la base SIRENE

par Christian Quest

Rédaction : Antoine Lemot

Depuis le 5 janvier 2017, la base SIRENE de l'INSEE est disponible en open data sur le site de data.gouv.fr, mais malheureusement pour les géomaticiens, cette livraison est au format CSV et non géolocalisable directement...

En proposant un géocodage minutieux et de qualité de cette source importante de données, **Christian Quest**, porte-parole d'OpenStreetMap (OSM) France et coordinateur du projet BAN chez Etalab, nous offre un exemple de transformation et de réutilisation de données ouvertes... ainsi qu'un cercle vertueux de la démarche open data.

Transformer pour améliorer, ...

Le travail de Christian Quest a, dans un premier temps, consisté à nettoyer le fichier CSV diffusé par l'INSEE : modification de l'encodage de l'ISO à l'UTF8, normalisation des libellés et des quotes. Cette première étape lui a permis de passer d'un CSV de 8,5Go à 4,2Go.

... réutiliser des données déjà ouvertes ...

Ensuite, étape clef, le géocodage des données est effectué à partir de trois bases elles-mêmes accessibles en open data : la BAN (Base Adresse Nationale), la BANO (Base d'Adresses Nationale Ouverte d'OSM) et la POI (Points d'Intérêt d'OSM).

... , au passage, en profiter pour comparer...

Pour arriver à un géocodage le plus exhaustif possible, Christian Quest a combiné les résultats obtenus par la BAN et la BANO. Ce double utilisation lui a permis de comparer et sélectionner l'adresse la plus proche de celle initialement proposée dans la base SIRENE.

...produire une nouvelle donnée,

Le résultat est donc un fichier CSV pour chaque département avec un champ longitude et un champ latitude permettant une géolocalisation ponctuelle directe.

et enfin ouvrir l'ensemble de la démarche.

Pour compléter cet important travail, Christian Quest propose :

- les résultats de son géocodage en Licence Ouverte sous http://data.cquest.org/geo_sirene/last/ (avant industrialisation),
- les scripts sous la plateforme d'hébergement GitHub <https://github.com/cquest/geocodage-spd/tree/master/insee-sirene>
- le retour des utilisateurs si des problèmes sont détectés <https://github.com/cquest/geocodage-spd/issues>

Des informations plus précises sur la démarche sont disponibles ici :

<https://medium.com/@cq94/g%C3%A9ocodage-de-la-base-sirene-2f0e14e87a8d>



Entretien

Jean-Marc Lazard

Fondateur et CEO de OpenDataSoft



Comment êtes-vous venu à l'open data ?

OpenDataSoft existe depuis 6 ans. L'entreprise a été créée avec l'objectif de développer et mettre sur le marché, via une plate-forme, une offre destinée au partage et à l'ouverture des données. Avec la volonté de s'inscrire dans un mouvement d'ouverture et de partage notamment sur la thématique des villes et des territoires, pour laquelle il y a un fort enjeu associé de transparence et d'innovation. Toutes les parties prenantes (publiques et privées) ont un besoin de partager, entre elles ou avec d'autres, les données qui constituent l'information et la base de la connaissance pour disposer des outils d'évaluation, pour prendre de meilleures décisions et développer de nouveaux services aux publics. Il y a 20 ans, pour diffuser et partager simplement des contenus en ligne, les CMS (systèmes de gestion de contenu) n'existaient pas. Nous avons fait le même constat pour les données il y a 6 ans et avons décidé de créer une plate-forme conçue spécifiquement pour le partage et l'échange de données par et pour le plus grand nombre.

L'objectif était de rendre le plus simple et rapide possible la réutilisation de données issues d'applications métiers diverses, afin de révéler leur valeur, en permettant de les restituer dans des formats standards, de les analyser ou de les agréger. Pour cela, car les données viennent d'horizons métier divers avec des différences de genèse, de format, de standard, il nous est apparu logique également d'apporter une solution de normalisation via les API. Cette normalisation permet d'aller au-delà des standards métiers existants, pour toucher le plus grand nombre avec des formats les plus répandus et reconnus sur le Web et le mobile. Notre vision de départ c'était de traiter cette chaîne logistique, de la collecte à la diffusion. Ensuite, au fil du temps, la typologie très variée des utilisateurs de notre solution a permis de faire évoluer continuellement cette dernière, tant sur les questions de collecte de données, de modèle économique que d'exploitation des données. Avec des clients dans 17 pays – surtout en Europe et moyen orient, un peu en Amérique du Nord – nous intervenons dans des secteurs variés, avec toutefois 40 % de collectivités territoriales, notamment des grandes villes telles Paris, Rennes, Toulouse ou Nantes. Nous accompagnons les clients sur les besoins et les bonnes pratiques de l'open data.

« L'objectif était de rendre le plus simple et rapide possible la réutilisation de données issues d'applications métiers diverses »

Pourquoi distinguer open data et SmartCity ?

« Lorsque l'on parle open data, les interlocuteurs ne se sentent pas tous concernés »

Lorsque l'on parle open data, les interlocuteurs ne se sentent pas tous concernés. Pour eux, le terme « données ouvertes » n'est pas forcément synonyme de qualité et de sécurité des données. On a donc créé l'offre SmartCity pour un type de clients. C'est la même plate-forme technique mais pas le même champ d'utilisations. De même pour le partage interne de données, les clients sont sensibles à notre capacité à gérer et valoriser leurs données en toute sécurité et des garanties de services élevées. Cela reste cependant la même plate-forme technique, disponible en "multi-clouds" pour répondre à des cas d'usages et architectures techniques variés.

Votre démarche est centrée sur la réutilisation. Pouvez-vous nous en dire plus ?

Les données de nos premiers clients étaient issues de capteurs. La question a rapidement été de savoir comment faciliter la réutilisation de ces données techniques par le grand public et tous ceux qui n'appartiennent pas au domaine métier initial.

Notre différence d'approche : nous ne sommes pas partis de la donnée mais nous nous sommes intéressés aux besoins des (ré)utilisateurs potentiels, du grand public, en essayant de savoir quel était leur niveau de compétence, de culture de la donnée ? On s'est rendu compte que l'approche par les API permettait de standardiser, d'apporter un accès performant, "à la demande", à tout un ensemble de données (pas uniquement géographiques), pour par exemple intégrer très simplement des réutilisations, des visualisations interactives dans un site web.

La mise en avant des réutilisations est allée de soi, en favorisant, comme tout contenu sur le web, la viralité de telle ou telle donnée à succès.

Et comme ces réutilisations sont "branchées" sur des API, on peut savoir à n'importe quel moment quelle donnée est la plus consultée à travers l'usage d'une visualisation cartographique par exemple. Cela permet au producteur de la donnée de connaître ce qui intéresse les gens, et de faire ainsi une sorte de marketing de ses données.

Également, en s'inspirant de ce que font les autres, une petite commune, avec ses propres données, pourra elle-même disposer des outils qu'une grande ville comme Paris. Cela contribue à rendre autonomes et agiles des petites villes pour créer leur propre base de connaissance.

En bref, notre credo est qu'il ne faut pas réserver la donnée aux spécialistes d'un domaine. Nous faisons évoluer notre plate-forme continuellement pour présenter les données le plus simplement possible, pour démocratiser leur accès.

« notre credo est qu'il ne faut pas réserver la donnée aux spécialistes d'un domaine »

Pouvez-vous nous indiquer des exemples phares de réutilisations qui en sont faites par les entreprises, les citoyens, les collectivités ?

Une jeune pousse Symaps.io a par exemple mis en place un service en ligne, pour les professionnels du commerce et les franchises, afin de les aider à identifier les zones idéales pour l'implantation de magasins. À partir d'un algorithme qui leur est propre, cette solution utilise, notamment à travers les API, des données open data : statistiques, socio-démographiques et celles propres à un territoire. Elle n'utilise pas que des données open data, mais sans données open data mises à disposition de manière industrielle, elle ne pourrait pas offrir ce type de services.



La société Handisco développe une canne blanche connectée pour les personnes non voyantes. La canne a un boîtier qui délivre des vibrations au niveau de la main et permet la localisation exacte ainsi que les principaux points d'intérêt à proximité comme les magasins, restaurants, lieux publics, arrêts de transports en commun et leurs horaires des prochains passages. Le prototypage de ce genre de solution et la création de l'entreprise n'est possible qu'avec des données ouvertes et si possible en temps réel comme les horaires des transports en commun.

« Il y a des réutilisations possibles par la collectivité elle-même, même si les "puristes" expliquent que ce n'est pas au producteur de la donnée de la rendre intelligible »

Il y a des réutilisations possibles par la collectivité elle-même, même si les "puristes" expliquent que ce n'est pas au producteur de la donnée de la rendre intelligible. La commune d'Issy-les-Moulineaux par exemple fait de la pédagogie sur son rapport d'activités qu'elle met à disposition depuis 4 ans. Elle

complète les données par des analyses, pas pour des professionnels mais pour ses citoyens dans une démarche participative. Il reste toujours la possibilité aux plus aguerris d'aller consulter les données sources en cas de besoin.

D'une manière générale les services indiquent que la démarche est très positive, car elle les expose et les oblige à avoir des relations différentes avec leurs administrés.

Votre livre blanc sur les 10 points essentiels pour ouvrir ses données propose « Pensez qualité, pas quantité ». N'est-ce pas à contre courant des obligations réglementaires d'ouverture de l'ensemble des données ?

À l'époque de la création du livre blanc, les gens ne cessaient de se comparer sur la quantité de jeux de données publiés, alors qu'il faut penser avant tout à la réutilisation qui va en être faite. La quantité n'est pas le critère de succès. Si la mise à disposition coûte que coûte peut être une première phase, dans un second temps, il faut passer à la montée en qualité des données. Un critère de succès c'est lorsque les collaborateurs d'une entreprise ou d'une collectivité utilisent leurs propres données ouvertes pour être plus efficace, pour innover. Là on peut alors dire qu'on a réussi, qu'on a franchi un pallier.

Quelles sont selon vous les évolutions futures de l'open data et celles de votre produit ?

On vient d'en parler, c'est quand on commence à fermer la boucle : on met à disposition des données de qualité moyenne, on est challengé pour les améliorer, on progresse en interne sur la qualité, et cette montée en qualité fait qu'on utilise soi-même les données mises à disposition. Nous travaillons par exemple à la mise en place de formulaires de saisie des données auprès des utilisateurs afin de pouvoir faire remonter les problèmes et les erreurs, et du côté des agents afin de faire un retour sur les remarques mais aussi de les corriger. On entre ainsi dans le cercle vertueux du collaboratif interne et externe. D'ailleurs

« On entre ainsi dans le cercle vertueux du collaboratif interne et externe »

la distinction entre interne et externe s'estompe. On pense surtout réutilisateurs : quand une société ou un partenaire utilise de manière professionnelle des données, il y a une sorte de contrat de service



implicite qui se met en place, à la manière du service publique de la donnée que met en place l'État français.

Nous avons également de plus en plus de clients pour lesquels le sujet est l'automatisation de la mise à disposition de données est prioritaire, dans cet esprit de "contrat de service" justement. Nous sommes au début d'un autre cercle vertueux grâce à la standardisation de la donnée, qui nous permet de travailler sur des API d'alertes à partir des flux de données, lors de changements ou de dépassement de seuils : j'ai comme exemple les personnes qui commencent à nous demander des API de "push" pour des alertes Météo France. La plate-forme ne met donc plus seulement à disposition des données, mais elle est capable d'envoyer des alertes à un système tiers. L'open data doit devenir proactif et nous prévient quand il y a changement significatif de la donnée ou d'un état de la donnée.

Entretien réalisé par Christophe Badol



Encart réalisé par Mathieu Rajerison, à partir des échanges ayant eu lieu lors de la COTITA Méditerranée du 27 mars 2018, consacrée à l' « Ouverture des données pour les territoires ».

La réutilisation des données open data peut s'illustrer avec l'exemple de LKSpatialist, une startup montpelliéraine de 34 salariés née en 2015 dans le giron de l'open data de l'État et de la métropole de Montpellier. Voici trois exemples de services qu'elle développe autour du foncier et qui montrent la diversité des développements, et des cibles potentielles de l'open data, dans un cadre entrepreneurial.

L'application appelée **LINA**®

qui accélère la prospection foncière en cartographiant le potentiel parcellaire bâti et non bâti sur un territoire, en donnant accès à des informations telles que les documents d'urbanisme, la forme de la parcelle, la réglementation, certaines contraintes liées à l'habitat ; cette application est plutôt à destination des professionnels de l'immobilier (promoteurs, aménageur, etc..)



NOAM®, une application gratuite et à destination des particuliers

Nouvelle Offre pour Acheter-Louer Malin, permet de rechercher un bien immobilier selon ses caractéristiques et sa proximité vis-à-vis de certains lieux (mobilité, commerces, loisirs, services et équipements)

L'application **FARA**® DPE+

qui permettra de connaître le gisement d'énergies renouvelables identifié sur une parcelle, qu'elle soit solaire ou géothermique



En savoir plus sur la COTITA Méditerranée « Ouverture des données pour les territoires ».

<https://www.cerema.fr/fr/actualites/retour-journee-ouverture-donnees-territoires-du-27-mars-2018>



Petit glossaire réglementaire autour de l'open data

La loi pour la croissance, l'activité et l'égalité des chances économiques (dite loi Macron) du 6 août 2015

Elle a rendu déjà obligatoire, dans son article 4, l'ouverture des données sur les transports publics et les services à la mobilité.

La loi pour une République numérique (dite loi Lemaire) du 7 octobre 2016

Elle acte les principes de l'ouverture par défaut des données publiques, et liste des jeux de données de référence via le service public de la donnée.

En savoir plus : <https://www.economie.gouv.fr/republique-numerique>

Le Règlement général de la protection des données (RGPD)

Il est un règlement européen entré en vigueur le 24 mai 2016. Il réglemente la protection des données personnelles, en renforçant les droits des personnes (droit à la portabilité des données personnelles), et responsabilisant les acteurs traitant des données.

En savoir plus : <https://www.cnil.fr/fr/comprendre-le-reglement-europeen>

L'Internet des objets dans la smart city



Le Cerema a commandé à l'Ecole des Ponts ParisTech un rapport sur les usages de l'Internet des objets dans les villes intelligentes françaises

L'étude propose un état des lieux des usages et des technologies, analyse les premiers retours d'expériences et formule des préconisations pour bien démarrer un projet.

A retrouver sur le site du Cerema :
<https://www.cerema.fr/fr/actualites/internet-objets-smart-city>

« Instaurer des données, instaurer des publics : une enquête sociologique dans les coulisses de l'open data »



Une thèse de Samuel Goeta sur « l'émergence et la mise en œuvre des politiques d'open data »

A lire au format PDF sur le site du CCSD (Centre pour la Communication Scientifique Directe) :
<https://pastel.archives-ouvertes.fr/tel-01458098>

« Ouvrir, partager, réutiliser - regards critiques sur les données numériques »



Sous la direction de Laurence Monnoyer-Smith, Clément Mabi et Jean-Christophe Plantin, cet ouvrage interroge « de manière critique les multiples processus sociotechniques au sein desquels les données s'insèrent, de leur production à leur usage »

Éditions de la Maison des sciences de l'homme :
<https://books.openedition.org/editionsmslh/9026?lang=es>

« Open data : où va-t-on ? »



16 min d'entretien avec Paula Forza, député LREM et Séverin Naudet, ancien conseiller numérique du Premier Ministre.

A écouter en podcast sur France Culture :
<https://www.franceculture.fr/emissions/le-numerique-est-politique/le-numerique-est-politique-du-dimanche-08-avril-2018>*

Salon de la data

le 12 septembre 2018 – Cité des congrès de Nantes

Salon de la Data

Cette année les thématiques seront :

- Maîtriser ses données : gouvernances des données, open data, sécurité des données, stockage des données...
- Communiquer et diffuser ses données : visualisation de données, cartographie, données géographiques, data journalism,
- Exploiter et valoriser ses données: Big Data, IoT, Data Science, Smart City,

Toutes les infos à retrouver sur : <https://salondata.fr/>

#OpenGov Forum Open d'Etat n°3 et 4

n°3 le 9 juillet et le n°4 le 17 septembre



« Conçus et impulsés par Etalab, Dataactivist et Vraiment Vraiment dans le cadre de la démarche de gouvernement ouvert en France, les forums « Open d'État » ont pour but de faire se rencontrer régulièrement des agents publics impliqués dans des projets d'innovation, mais aussi des acteurs de la société civile (acteurs privés, associations, ONG...) lors de forums ouverts. »

Pour en savoir plus :

<https://www.etalab.gouv.fr/les-rencontres-open-detat-le-forum-du-gouvernement-ouvert>

Sign@ture 66

Directeur de la publication
Pascal BERTEAUD

Directeur délégué de publication
Christian CURÉ

Rédacteur en chef
Bernard ALLOUCHE

Equipe de rédaction
Antoine LEMOT
Christophe BADOL
Mathieu RAJERISON
Stéphane LÉVÊQUE

www.cerema.fr

Photo couverture : Carte de France de l'ouverture des données – OpenDataFrance
<https://bit.ly/2u78XN7>