

**EXAMEN PROFESSIONNEL DE PROMOTION INTERNE  
D'INGÉNIEUR TERRITORIAL**

**SESSION 2020**

**ÉPREUVE DE PROJET OU ÉTUDE**

ÉPREUVE D'ADMISSIBILITÉ :

**L'établissement d'un projet ou étude portant sur l'une des options choisie par le candidat, au moment de son inscription.**

Durée : 4 heures  
Coefficient : 5

**SPÉCIALITÉ : INFORMATIQUE ET SYSTÈMES D'INFORMATION**

**OPTION : SYSTÈMES D'INFORMATION GÉOGRAPHIQUE, TOPOGRAPHIE**

**À LIRE ATTENTIVEMENT AVANT DE TRAITER LE SUJET :**

- ♦ Vous ne devez faire apparaître aucun signe distinctif dans votre copie, ni votre nom ou un nom fictif, ni initiales, ni votre numéro de convocation, ni le nom de votre collectivité employeur, de la commune où vous résidez ou du lieu de la salle d'examen où vous composez, ni nom de collectivité fictif non indiqué dans le sujet, ni signature ou paraphe.
- ♦ Sauf consignes particulières figurant dans le sujet, vous devez impérativement utiliser une seule et même couleur non effaçable pour écrire et/ou souligner. Seule l'encre noire ou l'encre bleue est autorisée. L'utilisation de plus d'une couleur, d'une couleur non autorisée, d'un surligneur pourra être considérée comme un signe distinctif.
- ♦ Le non-respect des règles ci-dessus peut entraîner l'annulation de la copie par le jury.
- ♦ Les feuilles de brouillon ne seront en aucun cas prises en compte.

**Ce sujet comprend 59 pages dont 2 annexes.**

**Il appartient au candidat de vérifier que le document comprend  
le nombre de pages indiqué.**

*S'il est incomplet, en avertir le surveillant.*

- ♦ Vous répondrez aux questions suivantes dans l'ordre qui vous convient, en indiquant impérativement leur numéro.
- ♦ Vous répondrez aux questions à l'aide des documents et de vos connaissances.
- ♦ Des réponses rédigées sont attendues et peuvent être accompagnées si besoin de tableaux, graphiques, schémas ...

Vous exercez en tant qu'ingénieur territorial responsable du service Information Géographique composé de 10 agents, dans une communauté d'agglomération (INGEAGGLO) de 45 communes, totalisant 305 000 habitants sur 1 200 km<sup>2</sup>. La commune centre regroupe 120 000 habitants.

Vous êtes rattaché à la Direction Générale des services. Le nouveau Directeur Général des Services (DGS) s'interroge sur la gestion de la donnée dans la collectivité. Suite à une première vague d'ouverture de données géographiques par la collectivité, il souhaite apporter une nouvelle impulsion en mettant en place une véritable stratégie de la donnée.

À l'aide des documents et annexes, le DGS vous demande de répondre aux questions suivantes :

**Question 1 (2 points)**

Quels sont les enjeux de la mise en place d'une gouvernance de l'information géographique pour les collectivités ?

**Question 2 (6 points)**

Dans une note, vous présenterez l'organisation que vous mettriez en place pour la gouvernance de l'information géographique à INGEAGGLO, en détaillant les opportunités et les freins ainsi que les partenariats potentiels.

**Question 3 (3 points)**

Quelles sont les forces et faiblesses de l'organisation actuelle de votre SIG et de la première mise en œuvre de l'Open data ?

**Question 4 (4 points)**

- a) Vous développerez les objectifs, enjeux et points d'attention pour la mise en place de métadonnées dans la collectivité. (2 points)
- b) Vous rédigerez les clauses techniques d'un cahier des charges pour l'acquisition d'une solution de catalogage adaptée à la stratégie de la donnée d'INGEAGGLO. (2 points)

### Question 5 (5 points)

- a) Donnez les deux scénarios organisationnels les plus pertinents qu'il serait possible de mettre en œuvre pour coordonner la stratégie de la donnée à INGEAGGLO en détaillant les avantages et inconvénients de chacun. (2 points)
- b) Expliquez les forces et faiblesses à prendre en considération pour confier cette mission à votre service. Vous détaillerez les prérequis nécessaires : moyens financiers, financiers, organisationnels. (3 points)

#### Liste des documents :

- Document 1 :** « La donnée comme infrastructure essentielle » (extraits) – *Rapport au premier ministre sur la donnée dans les administrations 2016-2017* – Henri Verdier – 2017 – 18 pages
- Document 2 :** « Open Data : Le mouvement prend de l'ampleur » – Michel Bernard – *SIGMAG* – juin 2019 – 5 pages
- Document 3 :** « Pour une vision stratégique de la géoinformation à Genève (extraits) – *SITG* – janvier 2018 – 14 pages
- Document 4 :** « L'Information Géographique et l'Open Data » (extraits) – Groupe de travail Open Data – *AFIGEO* – janvier 2017 – 5 pages
- Document 5 :** « Les données géographiques souveraines » (extraits) – Valéria FAURE-MUNTIAN – *ecologique-solidaires.gouv.fr* – juillet 2018 – 4 pages
- Document 6 :** « Portail de données géographiques - Lorient agglomération » – *geocatalogue.lorient-agglo.bzh* – février 2020 – 3 pages
- Document 7 :** « Portail Open Data – Orléans Métropole » – *outil Open Data Soft* – février 2020 – 4 pages

#### Liste des annexes :

- Annexe 1 :** « L'organisation de la collectivité INGEAGGLO » – *INGEAGGLO* – 2020 – 1 page
- Annexe 2 :** « L'information géographique et l'open data d'INGEAGGLO » – *INGEAGGLO* – 2020 – 2 pages

#### Documents reproduits avec l'autorisation du C.F.C.

*Certains documents peuvent comporter des renvois à des notes ou à des documents non fournis car non indispensables à la compréhension du sujet.*

# DOCUMENT 1

« La donnée comme infrastructure essentielle » (extraits) –  
*Rapport au premier ministre sur la donnée dans les administrations 2016-2017*  
– Henri Verdier – 2017

Neuf données de référence ont été identifiées à ce stade<sup>5</sup> : cinq d'entre elles constituent un corpus cohérent de données géographiques, et quatre autres permettent d'identifier des entreprises, des associations, des administrations, des métiers et emplois. Toutes ces bases de données sont désormais accessibles via la plateforme data.gouv.fr sur un espace dédié<sup>6</sup>.

Le **Répertoire des entreprises et de leurs établissements** est produit par l'Institut national de la statistique et des études économiques (INSEE). Le répertoire Sirene (Système informatique pour le répertoire des entreprises et des établissements) enregistre l'état civil de toutes les entreprises et leurs établissements, quels que soient leur forme juridique et leur secteur d'activité (industriel, commerçants, artisans, professions libérales, agriculteurs, collectivités territoriales, banques, assurances, associations...). Il comprend à ce jour plus de 10 millions d'entreprises et d'établissements. Les services enregistrent quotidiennement près de 10 000 modifications.

La base Sirene est disponible librement et gratuitement<sup>7</sup>, en open data, depuis le 4 janvier 2017, en application des dispositions de la loi pour une République numérique promulguée en octobre 2017. Une mise à jour quotidienne est diffusée.

Le **Répertoire national des associations** (base RNA) est produit par le ministère de l'Intérieur. Il permet d'identifier de manière unique chaque association déclarée en préfecture (numéro RNA, anciennement numéro Waldec) et précise pour chaque association le titre, l'objet, le siège social de l'association et adresse de ses établissements, la durée, la nature juridique de l'association, le code d'objet social. Cette base comprend aujourd'hui 1,5 million d'associations.

La **Base de l'organisation administrative** de l'État et des services publics est produite par la Direction de l'information légale et administrative (DILA, services du Premier ministre). Ce référentiel constitue la base de référence pour identifier et contacter les structures des administrations centrales.

Le **Code officiel géographique** (COG) est la nomenclature officielle des cantons, des arrondissements, des communes, des départements et des régions, il en contient les libellés et les codes. L'INSEE est officiellement en charge de sa production depuis 2003. Le COG est mis à jour annuellement, notamment pour intégrer les fusions de collectivités. Il peut être considéré comme le fichier primaire de nombreuses sources de données géographiques.

Le **Plan cadastral** informatisé est produit par la Direction générale des finances publiques (DGFIP) du ministère de l'Économie et des Finances. Les parcelles sont identifiées de manière unique par le code INSEE de la

<sup>5</sup> Article R. 321-5 du Code des relations du public et de l'administration : [https://www.legifrance.gouv.fr/affichCodeArticle.do;jsessionid=99AAC7F5C97B5DE451DC16F85E6E2A10.tplgfr34s\\_3?idArticle=LEGIARTI000034196073&cidTexte=LEGITEXT000031366350&dateTexte=20171127&categorieLien=id&oldAction=&nbResultRech=](https://www.legifrance.gouv.fr/affichCodeArticle.do;jsessionid=99AAC7F5C97B5DE451DC16F85E6E2A10.tplgfr34s_3?idArticle=LEGIARTI000034196073&cidTexte=LEGITEXT000031366350&dateTexte=20171127&categorieLien=id&oldAction=&nbResultRech=)

<sup>6</sup> <https://www.data.gouv.fr/fr/referance>

<sup>7</sup> Voir article L. 324-6 du Code des relations du public et de l'administration.

commune, le numéro de section et le numéro de parcelle. Le plan cadastral français disponible en ligne est composé d'environ 600 000 feuilles de plan aux formats image ou vecteur.

Le **Registre parcellaire graphique** (RPG), produit par l'Agence de services et de paiement (ASP), est la base de référence concernant l'occupation des terres agricoles. Il contient 7 millions d'objets graphiques et îlots.

Le **Référentiel à grande échelle** (RGE) est produit par l'Institut national de l'Information géographique et forestière (IGN). Il comprend cinq composantes (orthophotographique – BD ortho, topographique – BD topo, altimétrique – BD Alti, parcellaire – BD parcellaire et adresse – BD adresse).

La **Base adresse nationale** (BAN) est coproduite par l'Institut national de l'information géographique et forestière, le groupe La Poste, l'association Openstreetmap France et la mission Etalab (DINSIC). Cette base a pour but de référencer l'intégralité des adresses du territoire français. Elle contient la position géographique de plus de 25 millions d'adresses.

Pôle Emploi produit le **Répertoire opérationnel des métiers et des emplois** (ROME). Cette base est l'outil de référence pour les questions d'emploi et d'évolution des compétences. À titre d'exemple, le service Bob Emploi utilise ce référentiel pour calculer des proximités entre les métiers et proposer aux demandeurs d'emploi des trajectoires d'évolution.

La liste des données de référence a vocation à s'enrichir au cours des prochains mois, en lien avec la nomination des administrateurs ministériels des données.

### ***S'ouvrir à de nouvelles formes de collaboration et de production des données***

Les modes de production des données connaissent eux aussi des évolutions que l'État ne peut ignorer. Les données d'**acteurs privés** peuvent être utiles pour la puissance publique, à l'image de l'utilisation, par l'INSEE, des données des tickets de caisse de la grande distribution, pour mesurer l'inflation.

La loi pour une République numérique a ainsi introduit la notion de **donnée d'intérêt général**, c'est-à-dire des données pertinentes pour le public, au-delà parfois d'un lien direct avec une mission de service public, dont l'ouverture présente un bénéfice social général. Outre le champ des subventions publiques et délégations de service public, cette notion a pu être mobilisée pour ouvrir, secteur par secteur, des données devant permettre au consommateur d'être mieux informé et acteur de ses décisions ou à l'utilisateur de bénéficier d'un service plus fluide (données sur les transports, l'énergie, la couverture mobile notamment), mais il est certain qu'une telle notion pourrait s'appliquer à terme de manière plus large. Cette question figure d'ailleurs parmi les thèmes de la révision de la directive sur les informations du secteur public que la Commission européenne vient d'engager.

Les **communs numériques** sont aussi porteurs d'un autre modèle de production et de gouvernance des données, qui met l'accent sur la collaboration et le partage.

La Base adresse nationale en est un exemple. Initiée dès 2015 par l'IGN, La Poste, OpenStreetMap France et Etalab avec l'appui de l'administrateur général des données, cette base est originale et unique non seulement par son contenu (qui en fait la base la plus exhaustive à ce jour concernant les adresses en France) mais aussi par sa gouvernance qui associe des administrations, des entreprises publiques et des contributeurs d'une association.

Ce modèle, fréquemment cité en exemple à l'étranger, peut être reproduit dans d'autres secteurs : la santé, l'aménagement du territoire, etc. Il peut être envisagé dès que la disponibilité d'une donnée est essentielle pour plusieurs acteurs (État, collectivités, acteurs privés ou associatifs) mais qu'aucun d'entre eux ne peut prétendre en assurer seul la production, la maintenance plus généralement la gouvernance.

### Définir de nouveaux standards de données

L'importance des *standards de fait* ne signifie pas pour autant que l'État doit renoncer à proposer de nouveaux standards, *a fortiori* quand ceux-ci viennent concrétiser des priorités de politique publique. Mais les conditions de réussite, et d'appropriation, reposent sur la capacité de la puissance publique à faire émerger et à associer un écosystème solide en lien avec ce standard<sup>8</sup>.

L'État contribue ainsi à définir non seulement des principes généraux (« la transparence de la commande publique ») mais aussi des **standards de données** pour les mettre en œuvre concrètement. Certains travaux visent à définir des normes et obligations de format, afin de rendre comparables et agrégables des données qui sont produites par une grande variété d'acteurs. C'est le cas, par exemple, des données des infrastructures de recharge de véhicules électriques, **des données essentielles de la commande publique** et des conventions de subvention.

L'obligation de mettre à disposition les données essentielles de la commande publique (marchés et concessions) est issue de deux directives européennes transposées par deux ordonnances en 2016. Ces textes renvoient à la formule consacrée du « *format ouvert et librement réutilisable* »<sup>9</sup>. Ainsi, à partir du 1<sup>er</sup> octobre 2018, les acheteurs publics devront publier les données relatives à leurs marchés publics d'un montant supérieur à 25 000 euros, ainsi qu'aux concessions, en open data.

---

<sup>8</sup> Ce principe est aussi valable pour les standards issus du secteur privé, comme l'illustre le cas des données de transport. La participation active de Google à la création du standard d'échange *General Transit Feed Specification* (GTFS) a permis non seulement de faciliter la réutilisation des données de transports, mais a aussi assuré à Google une position privilégiée pour intégrer des données de transport public dans ses services (et notamment Google Maps).

<sup>9</sup> Voir les articles 107 du décret n° 2016-360, 34 du décret n° 2016-86 et 94 du décret n° 2016-361.

Afin d'encourager l'**écosystème naissant autour des données** de la commande publique, il est apparu nécessaire de standardiser celles-ci. Une réflexion importante a été menée sur les formats : l'État a conduit une expérimentation avec quelques administrations pilotes afin de les élaborer, ce qui a abouti à un référentiel standard fixé par un arrêté du 14 avril 2017<sup>10</sup>. Le même travail de standardisation a été effectué pour les **données relatives aux subventions**<sup>11</sup>.

Parallèlement, dans le cadre du Partenariat pour un gouvernement ouvert (ou *Open Government Partnership*), la France s'est engagée avec cinq autres pays au sein de l'**Open Contracting Partnership** pour développer et promouvoir un format standard de données ouvertes liées aux marchés publics<sup>12</sup> (*Open Contracting Data Standard*), afin de travailler conjointement à l'ouverture et la mise à disposition des données relatives à la commande publique et faire émerger du même coup un standard international. **La France a pris la présidence de cette organisation** le 28 novembre 2017 afin de le promouvoir, tout en développant des usages concrets autour de la réutilisation des données relatives à la commande publique.

### **Une meilleure coordination de la production des données serait source d'économie et d'efficience**

Il y a encore des **gisements d'économie et d'efficience** dans la production des données de référence. Dans la situation actuelle, la production des données de référence repose sur quelques opérateurs ou directions ministérielles. Leur savoir-faire et leur expertise dans cette production doivent être soulignés, mais des progrès restent possibles dans **la gouvernance de la production** de ces données.

En effet, la plupart de ces producteurs agissent de manière assez autonome, notamment vis-à-vis de leurs tutelles ministérielles. Les principaux référentiels, bien que liés<sup>13</sup>, sont ainsi produits sans nécessairement partager **un cadre stratégique commun**. Une meilleure gouvernance de la production serait ainsi un facteur d'économie et d'efficience.

Par exemple, la mise en place d'un numéro unique d'identification des associations est évoquée depuis longtemps comme un facteur de simplification en faveur du secteur associatif<sup>14</sup>. Aujourd'hui les associations qui demandent des subventions, paient des impôts ou emploient des salariés

<sup>10</sup> Voir <http://www.data.gouv.fr/fr/datasets/referentiel-de-donnees-marches-publics/> et <https://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000034492587&categorieLien=id>

<sup>11</sup> Décret n° 2017-779 du 5 mai 2017 relatif à l'accès sous forme électronique aux données essentielles des conventions de subvention : <https://www.legifrance.gouv.fr/eli/decret/2017/5/5/PRMJ1636989D/jo/texte> et arrêté du 17 novembre 2017 relatif aux conditions de mises à disposition des données essentielles des conventions de subvention : <https://www.legifrance.gouv.fr/eli/arrete/2017/11/17/PRMJ1713918A/jo/texte>

<sup>12</sup> <https://www.open-contracting.org/2016/12/07/open-contracting-version-francaise/>

<sup>13</sup> Un exemple de lien entre les bases : la composante adresse, qui est utilisée par de multiples bases de données de référence (notamment la base Sirene ou le Répertoire national des associations).

<sup>14</sup> Voir le rapport du député Yves Blein au Premier ministre : *50 mesures de simplification pour les associations*, octobre 2014.

doivent souvent mentionner deux identifiants, l'un dans la base Sirene et l'autre de la Répertoire national des associations.

## 2. Améliorer la circulation de la donnée

Il faut encourager la circulation des données, dans le respect des secrets légaux et de la vie privée des individus. **La circulation doit devenir la règle**, et la non-circulation l'exception justifiée.

Les actions entreprises depuis la création de la fonction d'administrateur général des données répondent à deux objectifs complémentaires :

- adapter et **faire évoluer le cadre juridique** pour limiter les freins techniques, économiques et juridiques à la circulation des données ;
- concevoir et **opérer les outils et les dispositifs** (plateformes, API...) permettant de faciliter la circulation des données, en cohérence avec l'approche de l'État plateforme.



### Ouverture des données : des progrès significatifs

Depuis la publication du premier rapport de l'administrateur général des données, de nombreuses bases de données essentielles et très détaillées ont été publiées dans différents domaines. Si l'objectif d'open data par défaut de la loi pour une République numérique n'est pas encore atteint, il y a eu depuis deux ans un changement d'échelle dans de nombreux secteurs.

Dans le domaine des données de santé, la Caisse nationale d'Assurance Maladie a poursuivi l'effort de publication des données de santé engagé avec la publication de la base des dépenses d'assurance maladie interrégime (DAMIR) en 2015. La CNAM a ainsi publié la base sur les prescriptions hospitalières de médicaments délivrées en ville (juin 2017) et la base sur les dépenses de biologie médicale interrégimes (mars 2017).

Dans le domaine économique, l'ouverture du répertoire Sirene en janvier 2017 constitue une avancée majeure qui a donné lieu à un très grand nombre de réutilisations.

Dans le domaine des données géographiques, la publication du plan cadastral informatisé (septembre 2017) est un bon exemple de publication d'un jeu de données essentiel.

Dans le domaine du logement, la publication du Répertoire des logements locatifs des bailleurs sociaux (décembre 2017) contenant des données détaillées sur 4,9 millions de logements sociaux constitue un autre exemple de la publication de données granulaires à haute valeur ajoutée.

## 1. La donnée doit être considérée comme une infrastructure

**Uber, Airbnb, Booking** : chacun dans leur domaine (transports, tourisme, hôtellerie), ces acteurs illustrent le **basculement des modèles d'affaires**. La détention d'un actif matériel a longtemps constitué un élément clé pour la structuration de ces marchés. Pour prétendre être un acteur de l'hôtellerie, par exemple, il fallait posséder une capacité d'investissement dans des actifs immobiliers. La place déterminante qu'occupe aujourd'hui dans ce secteur un acteur tel que Booking.com illustre bien que ce paradigme de la détention d'actifs matériels est largement remis en cause.

Dans des environnements économiques dominés par la notion de plateformes, **la possession et l'exploitation d'actifs immatériels** font la différence. Au premier rang de ces actifs, bien avant la marque, se situe les données et la capacité à les traiter. Si certains de ces acteurs possèdent aujourd'hui un pouvoir de marché important, ils le doivent en grande partie à la masse de données qu'ils ont été capables de se constituer dans leur courte existence.

On constate ainsi un **décalage croissant** entre l'immatérialité des ressources mobilisées pour rendre un service et la matérialité des effets de ce même service.

L'impact de ces activités, en ligne, sur le réel est majeur. La plupart des grandes métropoles sont aujourd'hui confrontées à ces impacts. New York, Barcelone ou encore Paris mettent en place des régulations pour limiter l'assèchement de l'offre locative dans les zones touristiques sous l'effet des plateformes de location courte durée. À San Francisco, Uber transporte autant de passagers que le réseau de transport public de la ville, et dans certaines villes, l'entreprise de véhicules de tourisme avec chauffeur se présente comme une offre de transport complémentaire de l'offre publique. Une partie importante du succès de ces initiatives privées – qui interpellent et parfois bousculent l'initiative publique – tient à l'utilisation des données.

Deux traits permettent de caractériser **le rôle que jouent les données** dans les stratégies des plateformes. Le premier tient à la collecte : *tout ce qui peut être mis en donnée l'est*. **La collecte est massive, et continue**. Aucune interaction avec l'utilisateur, aucune requête, aucun clic n'échappe à cette mise en données. L'historique de consommation, la localisation de l'utilisateur, mais aussi les notations alimentent ainsi en permanence les algorithmes. Le second trait distinctif tient à l'exploitation des données, elle aussi continue et massive. Dans une entreprise **data-driven**, les données sont utilisées à chaque phase de la conception et de la fourniture du service. Mieux, grâce aux données, il devient difficile de séparer les phases de conception et de production, à l'image des tests A/B qui optimisent en temps réel les pages (ou les titres de presse en ligne) d'un site Web.

L'État est **concerné** à plus d'un titre par les débats autour des plateformes et de l'usage des données. Dans son rôle de régulateur tout d'abord : début

2018 le Parlement étudiera l'adaptation de la loi « Informatique et Libertés » aux dispositions introduites par le Règlement européen sur la protection des données (RGPD). De même, la France joue un rôle moteur en Europe pour impulser la discussion sur la taxation des entreprises du numérique.

L'État peut aussi apporter une réponse, **offensive et non pas défensive**, aux défis et opportunités soulevés par les plateformes en constituant une infrastructure publique de données.

### **L'infrastructure publique du *xxi*<sup>e</sup> siècle**

« *Data is a new class of public infrastructure for the 21st century. It is all around us and easy to miss. We need to view it as an infrastructure that is as fundamental to modern society as power and transport, and which requires investment, curation and protection*<sup>1</sup>. » (Nigel Shadbolt, vice-président de l'Open Data Institute).

Le développement d'un pays est étroitement lié à la présence d'une **infrastructure performante et de qualité**, qu'il s'agisse de routes, de lignes ferroviaires, de réseaux d'énergie ou de télécommunications. L'État a d'ailleurs longtemps consacré une partie importante de ses investissements à construire et maintenir ces infrastructures. À titre d'illustration, 1 km d'autoroute représente un investissement de 6 millions d'euros, et 1 km de ligne ferroviaire à grande vitesse 16 millions d'euros. En moyenne, les pays membres de l'Union européenne y consacrent plus d'**un tiers de leurs investissements publics**<sup>2</sup>. Ces infrastructures contribuent à aménager le territoire et facilitent les échanges et les déplacements des biens et des personnes. Les externalités générées sont très largement positives. A *contrario* la défaillance des infrastructures est l'un des facteurs explicatifs du retard de développement de certains pays.

Il faut aujourd'hui considérer les données comme l'une de ces **infrastructures essentielles et critiques**. Essentielles car, dans une économie de l'information, l'accès à la donnée de référence fiable et à jour est la **condition du développement** des services numériques. Critiques car il faudra s'assurer que la fourniture de ces données ne puisse être interrompue, qu'il s'agisse de défaillances involontaires ou d'actes malveillants. En ce sens, il serait possible de considérer que ces infrastructures font partie des activités d'importance vitale<sup>3</sup>. Le développement des villes intelligentes ou des *smart grids* repose en grande partie sur la sécurisation de la mise à disposition et de l'accès aux données.

---

**1** « Les données constituent un nouveau type d'infrastructure publique pour le *xxi*<sup>e</sup> siècle. Elles sont tout autour de nous et pourtant facile à manquer. Nous devons considérer l'infrastructure de données comme une infrastructure aussi fondamentale pour la société moderne que l'énergie et les transports, et qui nécessite des investissements, des services de conservation et de protection. »

**2** Infrastructure in the EU : Developments and Impact on Growth, Commission européenne, 2014.

**3** Selon la définition du Secrétariat général pour la Défense et la sécurité nationale (SGDSN) : « *Parce qu'elles concourent à la production et à la distribution de biens ou de services indispensables à l'exercice de l'autorité de l'État, au fonctionnement de l'économie, au maintien du potentiel de défense ou à la sécurité de la Nation, certaines activités sont considérées comme "d'importance vitale".* »

Plusieurs pays européens ont saisi la nécessité de considérer la donnée comme une infrastructure publique essentielle, au même rang que les infrastructures physiques.

Depuis 2013, le gouvernement fédéral allemand intègre un ministère en charge des transports et des infrastructures numériques. Les investissements en matière d'infrastructures physiques et informationnelles sont gérés de manière intégrée, par exemple sur le sujet de la voiture connectée<sup>4</sup>. Outre-manche, la Commission nationale des infrastructures (*national infrastructure commission*), qui a pour objet de formaliser la stratégie nationale à long terme et d'orienter les investissements dans le domaine des infrastructures critiques pour le pays, a intégré la problématique des données à son périmètre d'action. De même, l'OCDE a souligné en 2015 (« Data-driven innovation report ») l'importance de telles infrastructures pour le développement économique et social.

## 2. Les objectifs d'une infrastructure de données

Quelles que soient les différences d'approche au niveau européen (nous détaillerons ci-après les démarches initiées au Royaume-Uni, au Danemark et en Estonie pour les comparer aux initiatives françaises), il est frappant de constater que **les constats, les objectifs** sont partagés. Le manque de circulation et d'exploitation des données, leur indisponibilité représentent une perte nette pour l'ensemble de la société.

Des données qui ne circulent pas, ou qui sont sous-utilisées sont des données dont on ne tire pas la totalité de **la valeur d'usage**. Elles peuvent même tendre à perdre de leur valeur au fil du temps, faute d'être confrontées à leurs utilisateurs et de bénéficier d'un retour qui permet leur amélioration. Leur qualité se **détérior**e d'autant plus rapidement.

Les conséquences financières de l'absence d'une infrastructure de données de qualité sont bien réelles. Le non-acheminement de courriers postaux dus à une erreur d'adressage représente un surcoût de l'ordre de 300 millions d'euros par an.

Le rapport Fouilleron a identifié plusieurs cas où des administrations dupliquent ou recopient des bases de données existantes, voire même **créent leurs propres bases** faute de pouvoir accéder librement, gratuitement et de manière sécurisée aux données produites par d'autres administrations.

Les coûts de la mauvaise circulation et de la non-qualité des données peuvent se détailler de la manière suivante :

- les pertes directes et indirectes liées à l'utilisation de **données inex**actes ;
- le maintien de **bases de données redondantes** et le coût de double-saisie quand la base de référence n'est pas diffusée à tous ceux, acteurs publics et privés, qui ont en besoin, c'est par exemple le cas des collectivités

---

<sup>4</sup> « Policies for mobility and modernity », Federal Ministry of Transport and Digital Infrastructure (BMVI), 2014.

- qui n'avaient pas accès, jusqu'à une date récente, à la base officielle des associations alors qu'elles sont le premier financeur du secteur associatif;
- le fonctionnement en **mode dégradé** faute d'accès à la donnée la plus récente et/ou la plus précise;
  - les **coûts de transaction** liés à la recherche, l'acquisition et le traitement de données publiques, qui peuvent être très élevés pour des industries ayant un besoin critique de données.

### **Une infrastructure pour permettre la meilleure exploitation des données**

Une route sert à faciliter les échanges et le transport des biens et des personnes. Un réseau de télécommunications vise à faire circuler l'information et à faciliter la coordination des individus. Une infrastructure de données a elle aussi une finalité : permettre la **meilleure exploitation possible des données**.

Elle s'adresse bien sûr en premier lieu à l'administration elle-même, mais concerne *in fine* l'ensemble des acteurs (entreprises, associations, société civile) qui utilisent des données publiques. Voilà **l'objectif qui est partagé au niveau de l'ensemble des pays européens** qui se sont engagés dans des démarches de construction d'infrastructure de données.

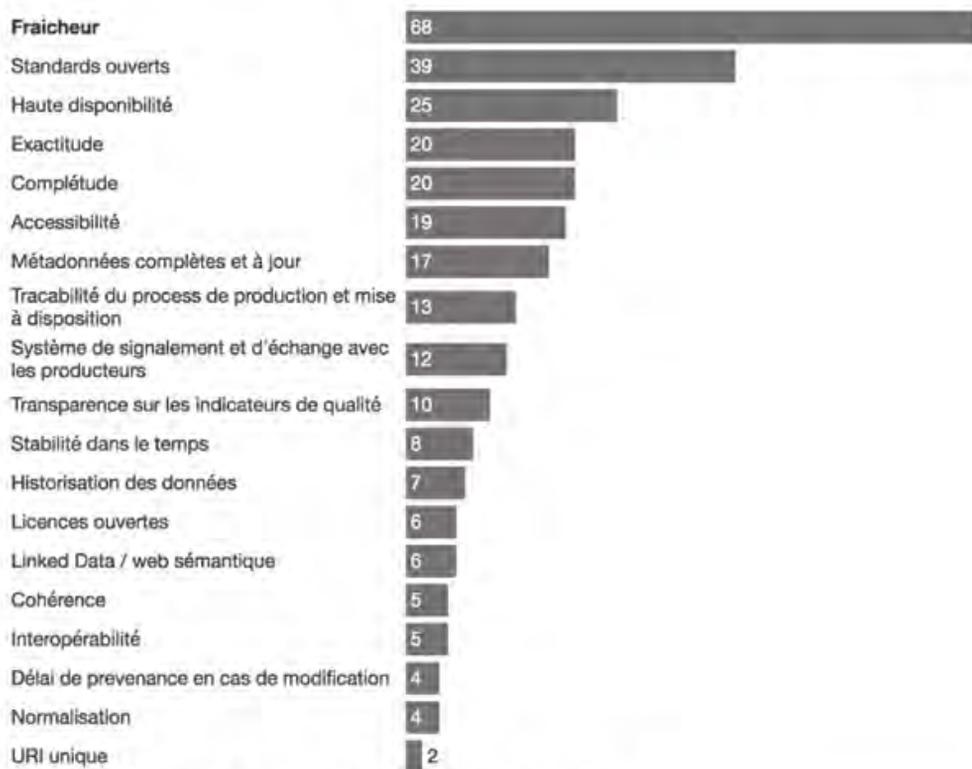
Comment faire en sorte que l'ensemble des acteurs s'empare des données mises à disposition ? Quels sont **les engagements** que la puissance publique doit prendre ? Pour le comprendre, Etalab a mené une consultation publique<sup>5</sup> auprès des utilisateurs potentiels des données de référence. Cette consultation, à laquelle 160 répondants (acteurs publics et privés, associations) ont contribué, a notamment permis d'identifier précisément les attentes, en particulier sur les critères de qualité des données de référence.

---

**5** Consultation menée du 29 septembre au 20 octobre 2016, synthèse des résultats disponible en ligne : [www.etalab.gouv.fr/consultation-spd](http://www.etalab.gouv.fr/consultation-spd)

## Les critères de qualité attendus pour les données de référence

Quels sont selon vous les critères de qualité que devraient respecter les données de référence ? (en nombre de citations)



Base: 160 répondants, les répondants peuvent avoir cités plusieurs critères

Source: Etalab - consultation service public de la donnée Récupérez les données

Créé avec: Datawrapper

### Des données à jour, disponibles et facilement réutilisables

La **fraîcheur** apparaît ainsi très nettement comme la principale dimension attendue (mise à jour des données, délai entre la survenance d'un fait, par exemple l'enregistrement d'une association, et son apparition dans la base diffusée). La **haute disponibilité** de l'infrastructure (de l'ordre de 99,5 % mensuels) est un autre élément attendu. L'utilisation de **standards ouverts** (second critère le plus fréquemment cité) a été introduite dans le Code des relations entre le public et les administrations à l'occasion de la loi pour une République numérique.

La complétude, l'exactitude et la fraîcheur constituent des dimensions classiques de **la qualité des données**. Mais les utilisateurs attendent aussi de la traçabilité sur le processus de production et de mise à disposition, la possibilité d'interagir avec le producteur (pour signaler une erreur ou proposer une amélioration d'une donnée) ou encore la transparence sur les indicateurs de qualité des données et de leur mise à disposition.

## «Des données sur lesquelles on peut compter»

L'une des caractéristiques d'une infrastructure c'est d'**être invisible**.

N'importe quel individu qui tourne un robinet s'attend à ce que l'eau potable y coule instantanément. Seule l'interruption temporaire du service d'eau (ou d'électricité/télécommunications/transports) nous permet de prendre conscience de l'ensemble des efforts et des moyens qui ont dû être mis en œuvre pour fournir le service de manière instantanée.

La fourniture de données publiques n'a pas pour le moment atteint ce niveau de qualité. Certaines données ne sont pas suffisamment mises à jour. D'autres sont peu ou mal documentées. Le schéma des données peut être modifié par leur producteur pour ses besoins internes et sans que les réutilisateurs aient été consultés ni même informés des changements à venir.

Cela est d'autant plus dommageable que l'infrastructure de données partage pourtant les mêmes caractéristiques que les autres infrastructures. Ses utilisateurs expriment le même **besoin de confiance dans l'infrastructure**<sup>6</sup> : je dois connaître *a priori* la qualité de la mise à disposition des données pour être en mesure d'appuyer mon service ou mon analyse sur ces éléments.

Si le service fourni est interrompu ou dégradé, si la qualité se dégrade alors la **confiance** des utilisateurs est rompue et l'infrastructure n'aura pas atteint ses objectifs.

---

<sup>6</sup> Un utilisateur compare notamment la mise à disposition des données de référence avec la standardisation existant dans le domaine hôtelier (« quand je choisis un hôtel de marque Novotel, je m'assure que mon expérience client sera identique dans tous les pays. Il n'y a pas de surprises et je sais précisément ce que je suis en droit d'attendre »).



## Plan de route de l'infrastructure de données

Une infrastructure de données est composée :

- de données de qualité, et en particulier de celles qui répondent aux critères des données de référence ;
- d'infrastructures de mise à disposition, en API et en téléchargement ;
- de mécanismes d'identification, de sécurisation et de contrôle ;
- de mécanismes de participation des utilisateurs à la montée en qualité des données.

Certaines **briques de cette infrastructure** sont déjà en place : le portail data.gouv.fr offre ainsi un service de téléchargement des données, ainsi que la possibilité d'interagir avec les producteurs des données (via le signalement et un forum de discussion pour chaque jeu de données). De même, certaines données de référence sont aujourd'hui exposées par des API, en particulièrement les données géographiques ou d'entreprises (API entreprise). Ces API sont référencées dans un catalogue : **api.gouv.fr** où les consommateurs peuvent les découvrir puis prendre contact avec les producteurs.

Un financement dédié dans le cadre du **Programme d'investissements d'avenir** (PIA) permet de renforcer les dispositifs existants (notamment en termes de performance et de sécurisation) et de mettre en place progressivement les éléments manquants. L'infrastructure de données doit permettre d'assurer la plus large circulation possible des données. Cela signifie concrètement que lorsque rien ne s'oppose à la diffusion à tous de ces données, elles doivent être proposées selon les principes des données ouvertes.

Cependant, quand certaines bases de données – ou certaines parties d'entre elles, par exemple les éléments d'identification d'un responsable associatif – contiennent des données à caractère personnel ou sont couvertes par d'autres secrets légaux, l'infrastructure de données doit permettre de s'assurer que seuls ceux qui ont le droit d'en connaître y accèdent. Pour ce faire, des **mécanismes d'identification et de contrôle d'accès** seront progressivement développés, en s'appuyant notamment sur les développements de FranceConnect Identité et de FranceConnect Agents.

Les utilisateurs de cette infrastructure participent de manière active à sa réussite et à sa gouvernance. Ils participent notamment à la montée en qualité des données en signalant des anomalies et en proposant des mises à jour. Cette brique de l'infrastructure de données est essentielle. A contrario, une plus large circulation des données sans possibilité d'interagir avec les producteurs se traduirait par une propagation des défauts des données et par une multiplication des coûts de contrôle et de correction des erreurs.

[...]

Depuis sa création en 2014, l'administrateur général des données a accompagné la mise en place d'une véritable **politique de la donnée** dans ses multiples dimensions : techniques, économiques et juridiques. Les lois ont évolué. Des outils, des plateformes, des API ont été développés et sont maintenant utilisés par de très nombreux utilisateurs. Des données essentielles sont mises à disposition en open data. Des réalisations très concrètes ont permis de valider le **bénéfice des datasciences** au service de l'action publique. Les mentalités aussi ont évolué, et plusieurs ministères commencent aujourd'hui à intégrer la donnée dans leur stratégie et leurs actions.

**Il faut maintenant transformer l'essai.** L'heure n'est plus aujourd'hui à démontrer l'intérêt d'une meilleure exploitation des données. Il faut à présent **mettre en ordre de marche** les administrations, en commençant par les ministères et les grands opérateurs. De par son rattachement aux services du Premier ministre et à la DINSIC, l'AGD a pour rôle d'accompagner l'appropriation de la révolution de la donnée par les administrations.

La feuille de route 2018 de l'administrateur général des données comporte cinq volets :

- mettre à disposition les **données et les infrastructures** mutualisées, les faire changer d'échelle ;
- développer une **doctrine de la circulation** des données au sein de la sphère publique ;
- renforcer le **réseau des administrateurs ministériels** des données, en faire un levier pour la politique de la donnée ;
- développer une expertise en matière d'**intelligence artificielle** au service de l'action publique pour faire de l'État l'un des premiers utilisateurs de ces outils ;
- **soutenir l'écosystème** des utilisateurs des données produites par l'administration, mesurer l'impact en termes sociaux, économiques et de transformation de l'action publique.

## 1. Mettre à disposition les données, les ressources et les infrastructures

### *Les données à fort impact économique ou social*

Les deux dernières années ont été marquées par la mise à disposition de jeux de données à fort impact : base des entreprises, cadastre, base adresse nationale... Parallèlement, des approches sectorielles ont été développées pour faciliter la réutilisation des données géographiques ou des données des entreprises.

En 2018, et en lien avec les administrateurs ministériels des données (cf. *infra*), nous identifierons, dans chaque ministère et chaque domaine de l'action publique, les données qui peuvent être qualifiées de **données de**

[...]

**référence**, au sens du Code des relations entre le public et l'administration. Certaines d'entre elles sont déjà disponibles en ligne sans toutefois atteindre le niveau de qualité de mise à disposition attendu du service public de la donnée. D'autres ne sont pas encore « ouvertes ». Dans les deux cas, il s'agira de travailler, en lien avec les producteurs, pour les rendre encore plus facilement *découvrables* et réutilisables. Cela pourra notamment inclure le développement d'API ou de toute autre ressource nécessaire.

Une attention particulière sera accordée à **la qualité des données** mises en ligne, en lien avec d'autres initiatives comme le projet Qualidata, lauréat du Programme d'investissements d'avenir. Les utilisateurs des données seront mobilisés pour participer à la montée en qualité des données, notamment par le développement d'un outil de signalement d'erreurs et de proposition d'amélioration.

### **Les standards de données et les infrastructures**

En 2018, le premier chantier consistera à **mettre en œuvre les standards** définis sur la commande publique et les conventions de subventions. Cela passera tout d'abord par la mobilisation des écosystèmes concernés : les administrations (État et collectivités) mais aussi les éditeurs de solutions logicielles et l'ensemble des réutilisateurs de ces données. Cet effort sera réalisé au niveau international dans le cadre de l'**Open Contracting Partnership** dont la France a pris la présidence fin 2017 (*cf. supra*). La lutte contre la corruption via l'analyse des données de la commande publique figure parmi les premiers cas d'usage étudié dans le cadre de ce partenariat.

Par ailleurs, nous poursuivrons notre effort pour **définir de nouveaux standards de données**. Traduire en standard de données des obligations légales et réglementaires cela permet non seulement de rendre l'application des règles plus simples, mais aussi de faciliter l'émergence d'écosystèmes de réutilisation de ces données.

Concernant les **infrastructures**, la DINSIC va poursuivre la mise à disposition d'outils mutualisés qui facilitent la circulation de la donnée : plateforme ouverte des données publiques, verticales thématiques (géographie, entreprise, transports), API, dispositif France Connect Identité. Ces outils facilitent la circulation et l'exploitation des données. Ils contribuent à la mise en place d'une véritable infrastructure de données telle que nous l'avons exposé dans la seconde partie de ce rapport.

## **2. Développer la doctrine de la circulation des données au sein de la sphère publique**

Le principe d'ouverture des données publiques par défaut figure maintenant dans la loi. Les outils – plateformes, API – existent pour permettre la circulation la plus large des données qui ne sont aujourd'hui couvertes par aucun secret et peuvent donc librement être partagées avec le plus

grand nombre. Cela n'épuise pas totalement cependant la question de la circulation de la donnée dans son ensemble.

### **Fournir la bonne donnée à la bonne personne, gérer le droit d'en connaître**

En 2018, l'administrateur général des données contribuera à faire **évoluer la doctrine sur la circulation des données**, y compris des données protégées par des secrets et en premier lieu la vie privée.

Comme nous l'avons déjà souligné dans le précédent rapport de l'administrateur général des données, le secret ce n'est pas la destruction de l'information. Au contraire, un secret c'est une information qui est connue de certains, et qui ne doit pas être transmise à d'autres. Le plus important, dans le cas d'un secret légal est donc de bien savoir à qui il s'oppose, et dans quelles conditions. Il convient donc, pour permettre de faire circuler la donnée, de savoir qui a **le droit d'en connaître** et d'être capable de gérer ce droit.

Gérer le droit d'en connaître, cela passe notamment par une phase d'instruction : qui est le demandeur ? quelle est sa mission ? à quelles données veut-il accéder ? En fonction de cette analyse, nous devons être capables d'offrir l'accès à des données, y compris celles protégées par des secrets. Gérer le droit d'en connaître cela signifie aussi prendre la responsabilité du traitement et en assurer la traçabilité.

Notre ambition, réaffirmée, est d'être capable de fournir **la bonne donnée à la bonne personne**, dans le respect du droit d'en connaître.

Pour rendre cette doctrine pleinement opérationnelle, nos efforts porteront sur deux niveaux. Pour les données dont la circulation doit être contrôlée : nous accompagnerons le passage à l'échelle de l'API Particulier et de FranceConnect Identité. Pour les données qui doivent être anonymisées ou pseudonymisées : nous développerons une expertise et des outils afin d'être en mesure d'accompagner les administrations dans la publication des données, en s'appuyant sur l'équipe interne de data-scientists d'Etalab.

## **3. Renforcer le réseau des administrateurs ministériels des données**

Certains ministères ou directions ministérielles ont désigné un administrateur des données, sur le modèle de l'administrateur général des données. En 2018, l'une des priorités consiste à **renforcer ce réseau** naissant pour en faire un **véritable outil** au service de la politique de la donnée.

Le renforcement du réseau passe tout d'abord par la nomination d'un administrateur ministériel des données au sein de chaque ministère. Cet administrateur aura pour tâche de porter la déclinaison de la politique de la donnée au sein de son ministère sur les quatre dimensions : **inventaire**

**et cartographie** des données existantes, **production des données** essentielles, **circulation maximale** des données, **exploitation** des données, notamment par les datasciences.

Il a aussi vocation à s'assurer de la bonne maîtrise des enjeux juridiques, en particulier en cette année 2018 où l'administration devra gérer en même temps le nouveau règlement européen sur la protection des données personnelles et la généralisation du principe d'open data par défaut.

Compte tenu du **rôle essentiel joué par les opérateurs** dans la production des données essentielles, il est souhaitable qu'ils désignent eux aussi un responsable des données en leur sein, en cohérence avec les administrateurs de leurs ministères de tutelle.

L'**administrateur général des données**, de par sa position interministérielle, est chargé d'assurer l'animation et la montée en puissance de ce réseau. La mise en commun des pratiques, des difficultés rencontrées et des solutions doit permettre de faire monter en compétence l'ensemble des administrations. L'administrateur général des données veillera aussi à s'assurer de la cohérence des actions entreprises par les ministères dans une optique de mutualisation des moyens.

#### 4. Développer un pôle de compétences en matière d'intelligence artificielle

Le champ technologique autour des données et de leurs usages est en **constante et rapide évolution**, à l'image de la diffusion rapide de l'intelligence artificielle notamment sous la forme d'algorithmes apprenants (et de l'apprentissage profond – *deep learning*). L'intelligence artificielle, son potentiel et ses risques interrogent bien sûr l'État tant dans son rôle de régulateur que d'opérateur de politiques publiques.

Sous l'égide de l'administrateur général des données, la DINSIC renforcera en 2018 son expertise et ses capacités en matière d'intelligence artificielle. Les data-scientists de la mission Etalab ont déjà mis en œuvre avec succès de telles approches, notamment dans le cadre du projet OpenSolarMap (*cf. infra*).



## OpenSolarMap : combiner intelligence humaine et intelligence artificielle

Et si nous étions capables d'indiquer, rapidement et simplement, le potentiel photovoltaïque d'un bâtiment? Cela permettrait d'évaluer l'opportunité d'installer des panneaux solaires. C'est l'objectif poursuivi par OpenSolarMap, qui combine la contribution par la foule (*crowdsourcing*) et l'apprentissage par des algorithmes (*machine learning*) pour produire de nouvelles données.

L'approche suivie par OpenSolarMap est de déduire la forme des toits par analyse d'imagerie satellitaire, grâce à la mise en open data des images des satellites Spot. L'orientation de la pente d'un toit et sa surface ainsi obtenues sont les éléments les plus importants pour estimer l'opportunité d'installation de panneaux solaire. Une première interface graphique a été développée : elle se présente sous la forme d'un jeu où chaque internaute est invité à indiquer l'orientation de la pente du toit. Ainsi, en moins d'un mois, la plateforme a collecté près de 100 000 analyses de qualité. En recoupant les différentes analyses pour un même bâtiment, environ 10 000 toits ont ainsi pu être caractérisés avec certitude.



Cet échantillon de toits déjà classifiés a permis de développer un classifieur automatique en utilisant des techniques classiques en traitement d'image (régression logistique et deep learning). L'algorithme obtenu ne se trompe que dans 20% des cas, ce qui est suffisamment peu pour l'application envisagée\*.

\* Les données calculées par l'algorithme sont publiées sur [data.gouv.fr](http://data.gouv.fr). Elles sont aussi publiées sous la forme d'une carte à l'adresse [cadastre.opensolarmap.org](http://cadastre.opensolarmap.org)

L'ambition est double : d'une part être capable d'orienter les ministères dans le recours aux technologies de l'intelligence artificielle, d'évaluer les outils et les offres existantes, de **mener des projets** sur quelques cas d'usages emblématiques. D'autre part, de rester vigilant sur la dimension d'éthique et de responsabilité de la mise en œuvre de tels traitements, *a fortiori* pour des systèmes apprenants qui ne sont pas toujours entièrement explicables.

### **Définir les conditions d'une utilisation éthique et responsable**

Les enjeux de responsabilité, de transparence mais aussi de cohérence entre le droit et l'informatique (*Code is Law*) ont déjà pu être abordés en 2017 à propos d'admission post-bac et de son successeur Parcoursup<sup>1</sup>. La loi pour une République numérique a introduit des dispositions relatives à la transparence des algorithmes et à l'ouverture des codes sources.

En 2018, nous définirons, dans le cadre des engagements de la France au sein du Partenariat pour un gouvernement ouvert de l'*Open Government Partnership*, les **conditions d'une utilisation éthique et responsable des algorithmes** (apprenants ou « classiques ») pour l'action publique.

## **5. Soutenir l'écosystème des utilisateurs de données publiques**

L'écosystème des utilisateurs de données publiques est un écosystème dynamique et riche. On dénombre ainsi plus de **185 000 visiteurs uniques mensuels** sur la plateforme data.gouv.fr. Un nombre bien supérieur d'entreprises, d'associations ou de particuliers accède à des services qui sont rendus possibles grâce aux données proposées en open data et, pour certaines d'entre elles, *via* un accès contrôlé (API Entreprise, API Particulier).

Le soutien de cet écosystème est l'une des conditions de la pleine exploitation du potentiel des données. Cela passe notamment par la participation ou l'organisation d'événements publics et la reconnaissance des initiatives du secteur public, associatif ou privé les plus impactantes, voire le soutien financier à certaines d'entre elles.

En 2018, l'administrateur général des données s'efforcera de **documenter les impacts sociaux et économiques** d'une meilleure circulation des données, dans la continuité du premier rapport de l'AGD qui analysait les mécanismes de création de valeur par les données.

---

<sup>1</sup> En 2017, la mission Etalab a été mandatée pour étudier les conditions d'ouverture d'admission post-bac.

[...]



# OPEN DATA : LE MOUVEMENT PREND DE L'AMPLEUR

AU CŒUR DES PRÉOCCUPATIONS DE NOMBRE DE GÉOMATIENS, L'OPEN DATA S'INSCRIT DANS UNE LOGIQUE HISTORIQUE. DISPOSANT D'IMPORTANTES VOLUMES DE DONNÉES, LES ORGANISATIONS LES LIBÈRENT. LA RÉPONSE SANS TROP DE FORMALISME À L'OBLIGATION INITIALE FAIT DÉSORMAIS PLACE À UNE STRATÉGIE PLUS RAISONNÉE. ALTRUISTE, LE MOUVEMENT POURRAIT AUSSI NE PAS ÊTRE DÉPOURVU D'INTÉRÊTS.

ENQUÊTE RÉALISÉE PAR MICHEL BERNARD

**J**e vous parle d'un temps que les géomaticiens de moins de 30 ans ne peuvent pas imaginer ! Il n'est pourtant pas si lointain lorsque les rares informations disponibles pour commencer à mettre en place un SIG ou à mener une étude sur le territoire étaient dérivées de la Digital Chart of the World (DCW) au 1:000.000<sup>ème</sup> issue des cartes de l'armée américaine. Quant aux quelques autres bases de données vectorielles un peu plus précises, elles n'étaient qu'issues de digitalisations plus ou moins licites de cartes existantes. Côté imagerie, il fallait se contenter d'images déclassifiées dont la résolution n'excédait pas les 30 m par pixel ! Ne cherchez pas à évaluer le temps perdu durant ces années de données maigres, à numériser planches cadastrales, plans et documents d'urbanisme. D'autant,

qu'il fallait reconduire les opérations plusieurs fois sur le territoire concerné, sans garantie de qualité, de précision et le plus souvent sans respect des droits des auteurs et créateurs.

Heureusement ce passé paraît déjà lointain, nous sommes rentrés dans l'ère de l'infobésité, avec l'avalanche de données que nous proposent satellites, scanners, capteurs et autres technologies d'acquisition. Alors sous l'impulsion des associations et des groupements d'utilisateurs, l'État a pris la mesure des retards accumulés par des décennies de privation et privatisation des données pourtant essentielles au développement de la société et à l'exercice de la démocratie. Il a même annoncé la fin des redevances pour les données de référence de l'IGN

en 2022. Il faut aussi rappeler en filigrane l'aventure d'Open Street Map qui a fortement influencé notre perception de la création des données géographiques par son approche ouverte et collaborative.

## UNE OUVERTURE TANT ATTENDUE

Si certains experts font remonter le principe de la libération des données à la Déclaration des Droits de l'Homme et du Citoyen, tous s'accordent à reconnaître que les grands principes fondateurs de l'Open Data ont été établis en décembre 2007, lors d'une rencontre dans les locaux de l'éditeur O'Reilly à Sebastopol en Californie. En réalité, d'autres dates et étapes clé ont permis de progresser pour arriver à la situation actuelle, en particulier la Directive Inspire. Elle a été une expérience particulièrement

formatrice, la géomatique ayant anticipé et expérimenté les grandes questions posées par l'accessibilité et l'ouverture des données bien avant les autres domaines.

Douze ans plus tard, la France est le premier pays à avoir inscrit dans sa législation un principe d'Open Data par défaut avec la loi pour une République numérique de 2016, dite Loi Lemaire. Celle-ci instaure une obligation d'ouverture des données pour les institutions, les collectivités locales de plus de 3.500 habitants et 50 agents. Tous les acteurs investis d'une mission de service public sont aussi concernés. Mais cette nouvelle compétence pose bien des questions aux collectivités qui doivent en assumer la responsabilité et les charges. Les difficultés avancées sont que le travail est souvent sous-évalué et qu'il vient se rajouter aux missions des agents. Par ailleurs les données ne trouvent pas toujours leur public, faute de vraie réflexion et accompagnement pour leur valorisation.

Dans le cadre de l'Observatoire piloté par OpenDataFrance, une carte référence les acteurs territoriaux qui produisent et publient des données publiques ouvertes. Le dernier recensement, au mois de mars 2019, dessine malgré ces difficultés un maillage assez cohérent du territoire. Avec une concentration dans les régions qui ont également une forte culture de partage de données géographiques : PACA, Bretagne,

#### Ci-contre

En moins de deux semaines, plus d'un million de personnes ont visité l'application DVF. Leur but : obtenir les vraies valeurs foncières dans un quartier ou une rue selon les transactions enregistrées ces 5 dernières années.



Occitanie, etc. Ce n'est évidemment pas un hasard.

Ainsi, en quelques années, le panorama de l'accessibilité aux données géographiques a été complètement bouleversé. Si la mise à disposition de données urbaines par les collectivités, notamment sous l'impulsion du groupe Topo de l'Association des Ingénieurs Territoriaux (AITF), est relativement ancienne, ce n'est que depuis 2017 que l'État a modifié en profondeur les conditions d'accès aux données essentielles que sont le Cadastre, les PLU et très récemment, fin avril 2019, les données DVF.

Cette base des Demandes de Valeurs Foncières contient des informations sur près de 15 millions de transactions immobilières qui ont eu lieu en France (à l'exception de l'Alsace-Moselle et de Mayotte) au cours des cinq dernières années. Ces données sont produites par la

direction générale des finances publiques (Bercy), qui a décidé de les rendre publiques pour aider les acheteurs à bien s'informer sur les prix constatés avant qu'eux même passent à l'acte notarié. Moins de deux semaines après son ouverture, le site totalisait plus d'un million de visites. La raison de cet intérêt est que ces données permettent d'avoir enfin une vision objective du marché immobilier avec des données qui n'étaient jusqu'à présent accessibles qu'aux administrations et collectivités. Nul doute que ces données vont venir renforcer les applications déjà nombreuses autour de la prospection immobilière.

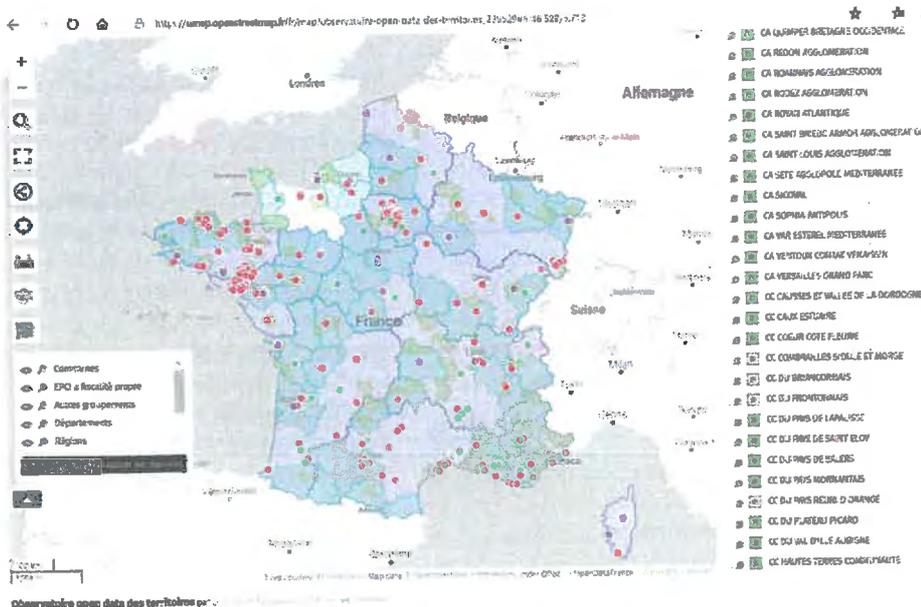
#### Ci-dessous

Les données détaillées ainsi qu'une synthèse de l'enquête sur les données ouvertes peuvent être téléchargées sur le site.

### À QUOI SERT L'OPEN DATA ?

C'est évident, les données sont devenues une infrastructure essentielle au même titre que les réseaux de transport ou d'énergie, la transformation numérique de nos sociétés a fait éclore de nouveaux modèles économiques basés sur leur exploitation. Tous les acteurs économiques doivent être en mesure de les exploiter. Les citoyens doivent aussi se les (ré)approprier. Il est donc impératif d'en assurer l'accès le plus libre et le plus ouvert possible.

L'Open Data répond donc à un ensemble d'enjeux tant politiques qu'économiques. De l'ouverture des données, on attend des bénéfices démocratiques (meilleure transparence de l'action publique, participation citoyenne, etc.), mais aussi la création de valeur économique par le développement de nouvelles activités et services. Si le premier point est



**Base Adresse Nationale (BAN)**  
 La Base Adresse Nationale est une base de données qui a pour but de référencer l'intégralité des adresses du territoire

2 30 5 France Point d'intérêt

**Base Sirene des entreprises et de leurs établissements (SIREN, SIRET) (fin le 30 avril 2019) (SIREN)**  
 Suite à plusieurs demandes émanant d'utilisateurs des fichiers Sirene, Insee a décidé de prolonger de trois mois la mise à

684 43 66 France

**Code Officiel Géographique (COG)**  
 Le code officiel géographique rassemble les codes et libellés des communes, des cantons, des arrondissements, des

21 8 20 France Commune française

**Plan Cadastre Informatisé (PCI)**  
 Le plan cadastral est un assemblage d'environ 600 000 feuilles ou planches représentant chacune une section ou une

3 5 54 France

**Registre parcellaire graphique (RPG) : contours des parcelles et toits cultureux et leur groupe de cultures majoritaire**  
 Le registre parcellaire graphique est une base de données géographiques servant de référence à l'instruction des aides de la

16 2 7 2013-2017 France

**Référentiel de l'organisation administrative de l'Etat**  
 Le décret du 14 mars 2017 a institué le Service Public de la Donnée. Celui-ci met à la disposition du public 9 jeux de données

6 4 14 2016-2026 France

**Référentiel à grande échelle (RGE)**  
 L'Etat a confié à l'IGN le développement du référentiel à grande échelle (RGE). Pour ce faire, il fait appel à ses moyens propres

12 2 8 France

**Répertoire National des Associations (RNA)**  
 Le Répertoire National des Associations (RNA) contient l'ensemble des associations relevant de la loi 1901, à savoir toutes les

28 8 22 1901-2017 France Point

**Répertoire Opérationnel des Métiers et des Emplois (ROME)**  
 Dans un contexte marqué par de fortes mutations de l'environnement économique et social, le ROME (Répertoire Opérationnel

5 2 8

■ ■ ■ désormais affirmé par l'État et encadré par l'obligation faite aux collectivités de libérer leurs données numériques, le second reste largement à développer pour qu'elles soient effectivement réutilisées et valorisées. Comme le rappelle Henri Verdier, administrateur général des données (AGD) pour l'État : « On devient plate-forme parce qu'on fait naître des usages, si on montre son efficacité et en envisageant la plateforme en fonction d'usages réels ». Dans un rapport publié en 2017, il précise en outre que « le fait de penser la donnée comme une infrastructure est un enjeu majeur de souveraineté économique. Les données sont un terrain, dans lequel on « plante » des services ». Ce rapport liste d'ailleurs les neuf données de référence jugées essentielles, qui sont majoritairement à caractère géographique ou géolocalisé ! Voilà une belle reconnaissance de la valeur de l'information géographique et de son caractère d'infrastructure déjà largement démontré par les travaux menés dans le cadre de la Directive Inspire par les nombreuses plateformes régionales et locales.

## GÉOMATIQUE ET OPEN DATA, LES ATOUTS DE L'EXPÉRIENCE

L'Open Data a fait émerger une culture nouvelle qui ne porte plus sur des domaines d'information spécialisés mais sur les compétences de gestion et de traitement de la donnée pour offrir son accès, son usabilité, sa pérennisation, et finalement son

### Ci-dessus

Le portail de data.gouv.fr propose à ce jour, 9 jeux de données de références qui couvrent les thématiques de données considérées comme des « données de référence ».

intérêt. Les métadonnées sont ainsi un enjeu majeur car elles permettent la contextualisation et l'interprétation des jeux de données, essentielles pour comprendre l'information. Mais cette dimension documentaire n'est généralement pas prise en charge par les services informatiques qui considèrent qu'elle n'entre pas dans leur cœur de métier.

Or les questions de qualité, interopérabilité, accessibilité ont été largement étudiées et mises en œuvre par les géomaticiens pour répondre à la mise en place de Portails ou Infrastructures de Données Géographiques, selon des méthodologies et des normes bien codifiées, en particulier pour le respect de la Directive Inspire. Il n'est donc pas étonnant que les géomaticiens participent fréquemment à la mise en place et au pilotage des projets Open Data. En outre, des Infrastructures existantes évoluent et récupèrent ou agrègent ces nouvelles compétences, par exemple DataSud en région PACA.

Même avec ces acquis et le capital de données accumulées dans les SIG, mettre en ligne et ouvrir ne se fait pas sans précaution. Comme le rappelle Emma Schoumaker, Chef de Service SIG et Territoire Connecté à la Métropole de Toulon Provence Méditerranée : « Nous avons commencé par les données règlementaires, pour nous conformer à la loi. Ensuite il n'est pas question de mettre en ligne des données pour l'une de nos 12 communes, si les autres ne disposent pas des mêmes ressources. C'est tout ou rien. De

plus, nous devons faire un travail sur les nomenclatures de nos bases pour les rendre compréhensibles. Nous utilisons parfois des acronymes très techniques. S'ils nous paraissent évidents, il faut documenter ou compléter »... Autant d'enjeux et de travaux qui restent à mener pour la grande majorité.

## COMMENT ABORDER L'OPEN DATA ?

Un grand nombre d'associations et d'organismes professionnels ont réalisé des études et diffusent des guides très utiles pour appréhender les enjeux de l'Open Data, que ce soit sur les aspects de la donnée elle-même, du volet juridique ou encore d'un point de vue organisationnel. Le site OpenDataFrance propose aux collectivités un référentiel pour mettre en place leur plateforme : le SCDL ou Socle Commun des Données Locales fournit les spécifications visant à homogénéiser la publication de données essentielles produites par des acteurs territoriaux (collectivités locales ou partenaires). Il pilote également l'initiative ValiData qui offre la possibilité de vérifier la conformité des données par rapport aux modèles définis dans le cadre du SCDL.

De son côté, la Fédération nationale des collectivités concédantes et régies (FNCCR) diffuse pour sa part des guides méthodologiques très détaillés sur la collecte et la gestion des données pour la mise œuvre de services et usages numériques. L'intérêt est de proposer une vision large, englobant l'Open Data : territoire-numérique.org.

Mais comme toujours, la mise en œuvre à l'échelle d'une collectivité se heurte à des choix ou des contraintes qui peuvent être politiques, techniques, juridiques ou organisationnels. « Certains services sont perturbés car la mise à disposition de leurs données n'est pas prévue, ni dans leur emploi du temps, ni dans les outils utilisés, rappelle Emma Schoumaker. Extraire les données métier pour l'Open Data peut souvent signifier une double saisie. Or, les interfaces ne sont pas encore prévues pour que ce soit un produit dérivé du travail des agents. » Les éditeurs de solutions métier vont sans

## DE L'OPEN DATA QUI NE MANQUE PAS D'AIR

La qualité de l'air est une préoccupation majeure pour tous les citoyens et les collectivités territoriales doivent disposer d'informations fiables pour prendre les mesures nécessaires dans de nombreux domaines, à commencer bien entendu par les transports.

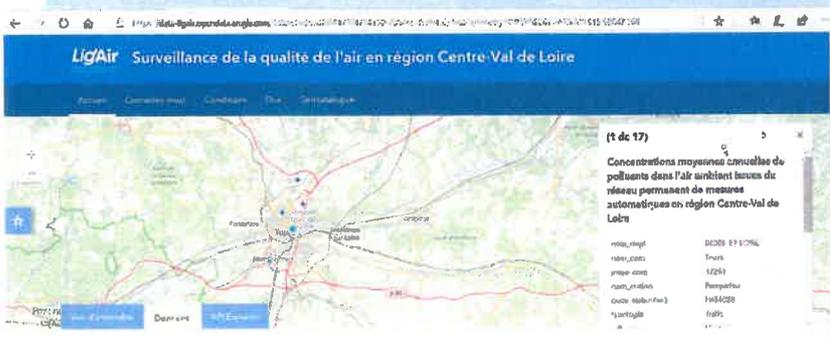
Pour les aider dans leurs prises de décision, les 18 Associations Agréées de Surveillance de la Qualité de l'Air (AASQA) ont ouvert un portail open data proposant des jeux de données harmonisés à l'échelle nationale. Ces données mesurées par les stations de surveillance sont transmises au fil de l'eau au Laboratoire Central de Surveillance de la Qualité de l'Air (LCSQA) pour être intégrées dans la base nationale de données de qualité de l'air appelée GEOD'AIR. Le site d'ATMO France, fédération qui a notamment pour but de coordonner l'action de ces organismes, se mue pour l'occasion en porte d'entrée vers ces différents portails qui utilisent la solution ArcGIS Hub.

ATMO France a engagé cette démarche pour proposer ses données dans des conditions optimales. « Les AASQA ont toujours été transparentes, elles ont toujours sorti les données avec les techniques dont elles disposaient, rappelle Marie-Blanche Personnaz, directrice d'ATMO Auvergne Rhône-Alpes

et référence numérique pour la fédération. Il était possible de les consulter et de les télécharger. Le souci, c'est qu'elles étaient assez peu répertoriées, géolocalisées. Le but aujourd'hui, c'est de sortir les données dans de bonnes conditions pour qu'elles puissent être réutilisées. »

C'est désormais le cas, à l'exception de certains territoires d'outre-mer dont l'offre est encore limitée, les différentes AASQA proposent un portail similaire avec au minimum onze jeux de données de référence accessibles en téléchargement et via des API. Les données sont également mises en forme dans des cartes de synthèse présentées et documentées avec les outils offerts par la plate-forme. ■

➔ [atmo-france.org/les-donnees](https://atmo-france.org/les-donnees)



Concentrations moyennes annuelles de polluants dans l'air ambiant issues du réseau permanent de mesures automatiques en région Centre-Val de Loire



doute devoir intégrer cette nouvelle préoccupation pour adapter les outils à cette nouvelle exigence.

### UNE OFFRE QUI SE STRUCTURE

Si la montée en charge des portails et Infrastructures de données géographiques a souvent été l'occasion de développer des solutions sur mesure, le « marché » de l'Open Data s'est très rapidement structuré et resserré autour d'un petit nombre d'éditeurs. Sur les 160 sites identifiés dans l'enquête OpenDataFrance citée plus haut, si la solution OpenDataSoft est bien représentée avec plus de la moitié des références, on constate qu'ArcGIS Hub arrive en seconde position et constitue une alternative

très pertinente avec des avantages techniques notables. Avec cette solution, les données ne sont pas dupliquées sur un serveur dédié et le producteur en reste maître, ce qui garantit la fraîcheur de la donnée. Les données (vecteur comme raster) sont gérées et les géométries complexes et/ou très volumineuses sont parfaitement accessibles et optimisées. Tout ceci est cohérent avec les principes fondateurs d'Inspire qui veulent que la donnée reste au plus près de ses producteurs et que l'Open Data a tout intérêt à appliquer. En accédant directement à la donnée, celle-ci peut être mise à jour par la communauté des utilisateurs en utilisant les nombreux outils et « apps » disponibles. Ces dernières permettent également de proposer

des représentations très poussées et personnalisables, comme des tableaux de Bord, Story Maps, Web Maps. Preuve de la pertinence de cette solution, début mai 2019, la plate-forme ArcGIS hébergeait déjà 222.070 jeux de données ouvertes partagées par 2.513 organisations à l'échelle mondiale.

En plus ArcGIS Hub ne se limite pas à l'exposition et diffusion des données, mais développe une approche novatrice, axée sur la collaboration et plus orientée sur des cas d'usage. Conçu pour être très facilement déployé, ArcGIS Hub relie les entreprises et les collectivités territoriales aux citoyens par l'intermédiaire de projets d'actions concrètes appelés initiatives. Tous les outils de la plateforme sont alors ■ ■ ■

■ ■ ■ reliés et activés à mesure des besoins de collecte, analyse, publication et diffusion des données. Cette vision évoluée de l'Open Data est en train de prendre corps notamment dans le domaine du transport, en témoigne cette citation de transport.data.gouv.fr : « L'idée est que le site agisse non seulement comme un point d'accès mais aussi comme un réseau social. Donc, pour chaque jeu de données, vous trouverez le fichier mais aussi la possibilité de poser des questions directement à l'opérateur de transport sur le territoire, qui saura mieux vous répondre ». La Ville de Brampton au Canada donne un bon exemple avec des initiatives portant sur la revitalisation du centre-ville par exemple, où l'on voit que tous les indicateurs sont liés à un engagement politique et à des données ouvertes !

## DES SECTEURS RESTENT TIMIDES

Si depuis déjà quelques années plusieurs projets ont été montés en collaboration entre le secteur public et les entreprises dans le domaine du transport (SNCF, RATP), les autres secteurs peinent encore à développer une politique Open Data. En 2018, la plate-forme ORE (Opérateurs de Réseaux d'Énergie opendata.agenceore.fr) s'est bien montée dans le secteur de l'énergie. Regroupant les 170 distributeurs de gaz et d'électricité en France, elle ambitionne de faciliter les démarches des acteurs du système énergétique dans les nouveaux mécanismes de marché. Elle veut aussi aider des acteurs publics et privés dans le cadre de leurs actions d'efficacité énergétique. Au-delà de ces exemples, beaucoup d'autres industries cherchent encore les conditions du partage d'informations avant de parler d'ouverture. Des obstacles sont identifiés tant en matière de perception des modalités de création de valeur que de risques concurrentiels. Ainsi, les entreprises veulent d'abord créer des écosystèmes basés sur la confiance pour ensuite pouvoir échanger leurs informations en sécurité. C'est ce qui s'est passé dans le secteur de l'agriculture, fortement impacté par l'abondance de



données. Les grandes exploitations agricoles disposent aujourd'hui d'outils d'aide à la décision qui s'appuient sur la collecte de données géolocalisées. En effet, les équipements agricoles sont connectés. Ils produisent de très nombreuses données propres au fonctionnement de l'équipement, mais aussi sur le volume voire la qualité des récoltes. L'American Farm Bureau, qui représente les intérêts des exploitants agricoles aux USA, s'est émue très tôt de la capacité des grands équipementiers (ex. John Deere) à utiliser ces données pour spéculer sur les marchés des céréales, allant ainsi à l'encontre de l'intérêt des utilisateurs de ses équipements. L'initiative AG Data Transparent vise à certifier les équipementiers sur le respect des principes définis dans la charte « Privacy and security principles for farm data ». Ces principes prévoient notamment l'ouverture des données entre fabricants d'équipements, l'interdiction d'usage à des fins contraires à l'intérêt des exploitants

### Ci-dessus

Le Portail Open Data de la ville de Brampton au Canada est une véritable porte ouverte aux administrés. Ils accèdent non seulement à des cartes et applications, mais aussi à des tableaux de bords ou au budgets de la ville. Le GeoHub sert aussi d'interface de dialogue direct avec les services municipaux.

agricoles ou bien encore un contrôle accru sur la revente de données non anonymisées à des tiers. AG Data Transparent est donc un outil pour créer les conditions de la confiance entre les acteurs d'une filière.

Pour conclure, force est de constater que l'adage « Vivons cachés, vivons heureux » devient bien difficile à appliquer en ces temps de géographie et d'imagerie pervasive et invasive. Si la société réclame en effet la transparence et l'ouverture des données du secteur public, la contrepartie et les conséquences sur l'espace privé génèrent des inquiétudes bien légitimes. Car d'évidence, les opérateurs d'analyse, l'IA et les algorithmes en tout genre vont de plus en plus impacter notre quotidien et risquent de s'insérer dans le moindre de nos faits et gestes.

Si les professionnels, start-ups ou autres, peuvent rapidement s'approprier les données et élaborer des applications et des services monétisables, il n'en va pas de même pour le grand public. Il semble donc indispensable d'envisager des organisations permettant aux citoyens et aux associations de s'emparer de ce mouvement. Et non seulement d'y contribuer, mais aussi de le réguler. Les collectivités et leurs partenaires s'attellent actuellement à ce chantier. Ils réfléchissent autour du concept de Service Public de la Donnée, devenu une mission capitale et vitale pour accompagner l'inéluctable transition numérique de notre société. ■ Michel Bernard

observatoire-opendata.fr



# Quatre axes pour une nouvelle vision stratégique

Le numérique ne tient pas seulement à l'établissement de nouvelles chaînes d'automatisation des systèmes d'information comme on l'entend souvent. Il englobe aussi une mise en place de nouveaux processus autonomes de données qui complètent ainsi la palette des outils à disposition des systèmes d'information du territoire. A l'avenir, les administrations devront d'une manière ou d'une autre contrôler, valider et/ou se porter garant (tiers garants) des géodonnées.

La (r)évolution du numérique imposant tant de changements sur tellement de plans que nous avons adopté une approche « bottom up » pour ordonner la réflexion. Partant de la donnée nous aboutirons à des propositions plus globales liées aux comportements nouveaux de la clientèle (citoyen, entreprise, etc.) notamment des citoyens contributeurs. Cette approche juxtaposant clients et innovation technologique nous permet de passer de l'un à l'autre en donnant tantôt la priorité à l'un puis à l'autre.

Nous avons donc choisi 4 axes pour poser les bases d'une réflexion stratégique nouvelle.

# 1

## “Smart geodata”

---

Les données du territoire vont changer de nature, de sens, de forme et de structure. La nature de celles-ci sera un mixte entre géodonnées structurées classiques exprimant des grandeurs mesurables des objets du territoire et des données moins structurées, qui nécessiteront un travail d'interprétation avant de pouvoir être incorporées aux bases de données classiques. La nature de ces nouvelles données dépendra fortement de l'origine de celles-ci (réseaux sociaux, blogs, Apps, etc.) mais aussi de la « partie prenantes » qui les émettra (citoyens, associations, entreprises, objets connectés, etc.).

De nouvelles questions de qualités et de valeurs de ces données vont se poser. La donnée mixte devra être validée. Comment ? Par qui ?

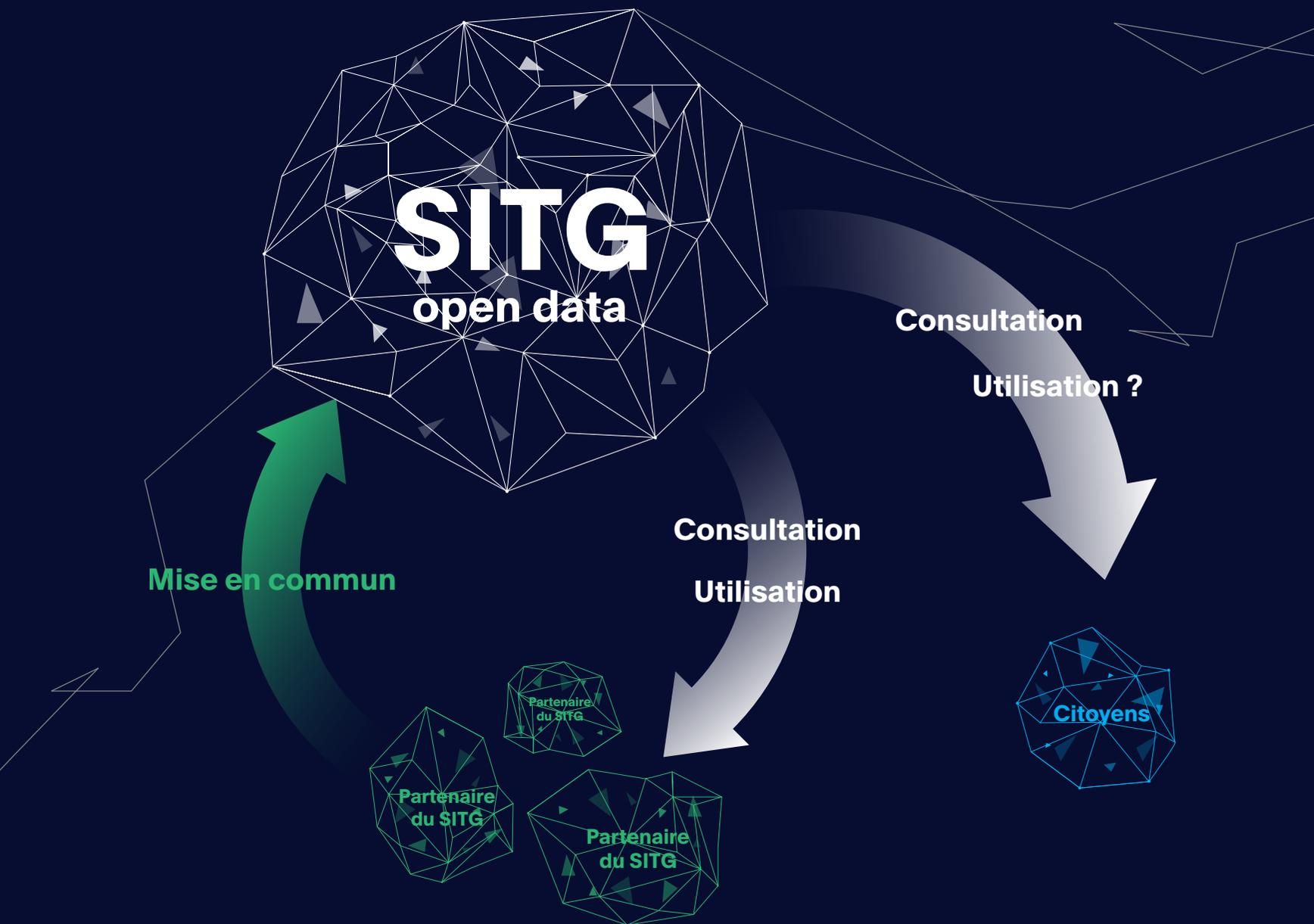
Ainsi, on peut en être sûr, les données proviendront davantage du territoire lui-même, notamment par l'Internet des objets et les nouveaux citoyens contributeurs. Les partenaires et le public fourniront de plus en plus de données localisées non-structurées, issues notamment des réseaux sociaux. L'internet des objets fournira des données en très grand nombre. Il va falloir maîtriser des nouveaux outils informatiques.

La juxtaposition d'anciennes et de nouvelles pratiques sur les données nécessite la mise en place de nouveaux outils, par exemple des capteurs de toutes sortes pour l'IoT et des logiciels de tri, de filtrage ou de réduction des géodonnées.

Les procédures de travail devront à cet effet être complètement révisées. Plus question de se limiter à collecter, stocker et restituer. Tout un travail de valorisation et d'interprétation de la donnée doit être mené d'un point de vue de la géodonnée, de ses flux et à toutes les échelles, du territoire au bâtiment. Les données des maquettes numériques des bâtiments structurées par la démarche BIM intégreront naturellement les géodonnées du territoire. Désormais, il faut donner du sens, souvent caché, aux données. Une nouvelle « science de la donnée » émerge.

# Actuel et futur

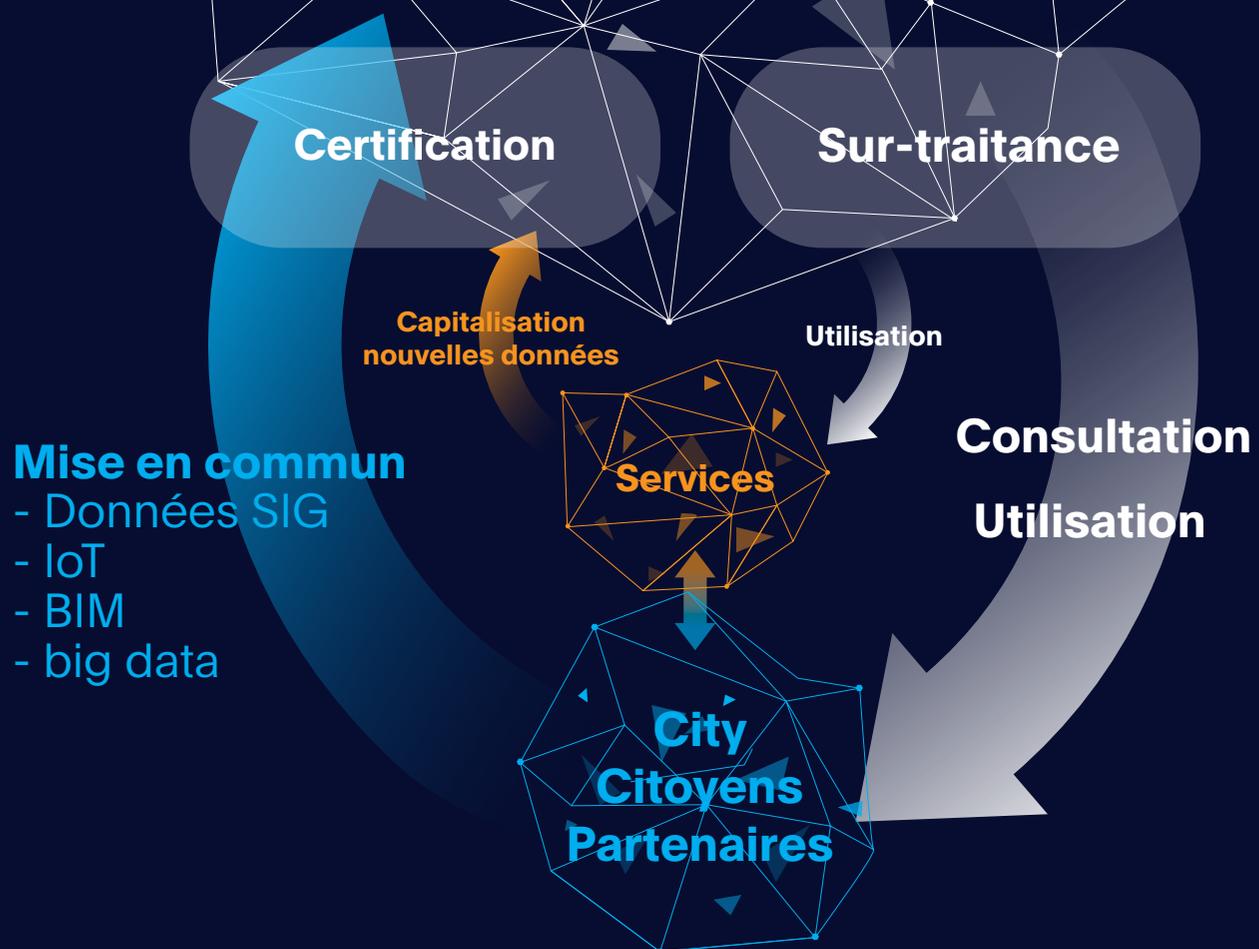
# paradigmes SITG



Les partenaires mettent leurs données en commun et en font usage. Comme ces données sont accessibles en open data, toutes autres personnes ou organisations peuvent également les utiliser. Il n'y a cependant pas de rétroaction et de capitalisation pour élargir le socle des données au-delà de celles des partenaires.

# SITG

smart geodata  
open data



Les partenaires mutualisent toujours leurs géodonnées et en font usage. L'assiette des données est considérablement élargie intégrant celles de l'IoT, du BIM et du big data. La logique de stock est étendue à une logique de flux (temps réel).

C'est la cité elle-même et les citoyens contributeurs qui peuvent produire et bénéficier des smart geodata. Pour favoriser les rétroactions, une approche de plateforme de sur-traitance est mise en place favorisant la création de services et la génération de nouvelles données qui viennent nourrir les smart geodata au service du territoire et des citoyens.

# 1a

## Les outils de management du big data

---

### **La transition numérique change la donne**

**D'un côté, les compétences « métiers » évoluent vers plus de collaboration/participation du client dans la définition de la « valeur » de l'offre.** Il s'agit, en gros, d'ajuster/créer avec les clients l'offre (co-création) et non plus d'imposer celle-ci à travers des outils tout faits.

De nouvelles pratiques d'apprentissage collectif, de discussions démocratiques, de recherches de consensus doivent être imaginées. Dès lors que le débat porte aussi sur les processus il faut changer les approches.

**Et d'un autre côté, les compétences propres aux big data que sont la « réduction » et « l'analyse prédictive » des données nécessitent de nouvelles approches.** Ces outils sont très différents de ceux plus classiques de l'analyse déductive et statistique jusqu'alors employés. Il faut absolument développer en interne la compréhension et la maîtrise de ces outils qui seront à la base des métiers de demain.

Ces deux types d'outils sont au service des professionnels de la donnée du territoire. Aujourd'hui, les outils d'analyse, de visualisation et de management des données, etc. sont développés au point qu'ils sont accessibles à tous. Les géographes, les géomaticiens et les professionnels du territoire peuvent, après une formation adéquate, parfaitement s'en servir sans faire appel à des spécialistes.

Les « démonstrateurs » (Proof of Concept) offrent cette souplesse de pouvoir réunir réellement ces compétences.

**Le big data doit être exploré en priorité selon cette approche.**

La meilleure manière de réussir cette phase de transformation est de lancer plusieurs démonstrateurs sur le territoire genevois, tout en mettant en place un « think tank » de réflexion (laboratoire d'idées) propre à enrichir le processus.

# 1b

## De nouveaux savoir faire émergents

---

Les nouveaux systèmes d'information géographique des territoires font appel à des compétences supplémentaires, comme celle des « data scientists » et évidemment aussi à d'autres experts de la gestion des données, faisant par exemple appel aux spécialistes du cloud, des plateformes ou des données IoT pour n'en citer que quelques-uns.

**De nouveaux métiers entrent en scène, de nouvelles formations se mettent en place.**

Après l'ère des géomètres puis celle des géomaticiens vient celle des « geo data scientist ». Les techniques évoluent, la science aussi. On peut dire que l'on est à un tournant pour les géographes, les géomètres et les géomaticiens. Une nouvelle discipline va désormais les accompagner sur la connaissance des territoires. Ainsi, par exemple, l'Université Catholique de Louvain offre depuis l'automne 2017 une formation d'ingénieur civil en science des données ! On observe le même mouvement dans les Hautes Ecoles en Suisse pour appréhender le big

data dans plusieurs domaines (Universités, EPF et HES). Il faut prendre très au sérieux ce changement, car la gestion et le traitement des données, qui sont à la base de tous les systèmes d'information du territoire, vont subir une mutation profonde. On parle aujourd'hui volontiers de « data driven management » pour désigner à quel point les données, géographiques en particulier, influencent la direction des affaires. Ceci est vrai pour l'économie tout comme pour les administrations.

Une ère de la maîtrise renouvelée des données s'ouvre par le big data, l'IoT et les données non-structurées. Les maîtriser, cela veut dire les mettre en forme compréhensible et donc pouvoir les retourner aux utilisateurs de manière utile. C'est en quelque sorte, les transformer en « smart data ». L'idée ici c'est de pouvoir utiliser les nouvelles géodonnées dans le cadre des « smart cities ».

Les partenaires du SITG doivent acquérir et s'adjoindre ces compétences nouvelles.

# 2

# La transition numérique et la géoinformation

---

De nouveaux acteurs majeurs arrivent avec force dans le champ des données des territoires : les sur-traitants. Par opposition à la sous-traitance, les sur-traitants coiffent la chaîne de la valeur en s'attribuant la distribution des produits via le contrôle de plateformes dominant des écosystèmes.

La sur-traitance détermine le positionnement d'une entreprise au cœur même de l'écosystème auquel elle appartient. En créant une plateforme commerciale de type Apple Store, elle va regrouper à travers celle-ci les acteurs de son écosystème. C'est ainsi qu'Apple a regroupé plusieurs centaines de milliers d'entreprises qui ont produit des Apps pour sa plateforme sur iPhone. Tous vont alors dépendre d'Apple, tant au niveau des spécifications techniques que commerciales. En quelque sorte Apple dicte le jeu et récolte les marges de sa plateforme internet de sur-traitance.

Ainsi, la sur-traitance réorganise des pans entiers de l'économie : les médias, le marketing, le commerce mais aussi la santé (digital health), la maison (domotique), l'usine (industrie 4.0), etc. et bien sûr aussi les territoires (smart cities). La sur-traitance se révèle comme la grande nouveauté économique de cette dernière décennie.

La sur-traitance se révèle comme une nouveauté économique. Tout le monde est en quelque sorte devenu sous-traitant de l'écosystème à l'exception de rares entreprises ayant encore une clientèle captive.

Les partenaires du SITG doivent s'inspirer des logiques de sur-traitance dans leur domaine de compétences. Il s'agit concrètement de favoriser la création de services à partir des smart geodata pour contribuer à une dynamique d'innovations tout en assurant la capitalisation des nouvelles données créées. L'enrichissement du socle des données et la valorisation de la contribution pour une meilleure intelligence de la ville sont les leviers de l'approche plateforme du SITG.

# 2a

## Le déploiement de l'IoT

---

Il s'agit bien ici et dans un premier temps d'acquérir des savoirs, plus que de lancer à grande échelle des projets, car l'IoT n'en est qu'à sa toute première génération technologique et les prochaines vont remettre en cause les acquis actuels. Songeons ici à l'arrivée annoncée de la 5G. Il en est de même des capteurs qui aujourd'hui en sont encore à la première génération, tant au niveau de la consommation d'énergie que celui des performances intrinsèques.

Cela ne veut cependant pas dire qu'il ne faut pas faire des « démonstratifs », bien au contraire.

Apprendre est la toute première chose à faire.

Et apprendre par la pratique est la seule manière de réellement apprendre. Alors oui, il faut lancer des « démonstrateurs » tout en sachant que ce ne sont pas des prototypes

du futur mais des essais sans véritable lendemain, car demain les projets de développement réels seront tout à fait différents.

Le numérique entraîne tellement de bouleversements qu'il faut s'appuyer sur une logique forte de mise en œuvre, basée sur la maîtrise des données d'abord, puis des capteurs ensuite pour finir avec les réseaux.

Tous ces capteurs sont géolocalisables et le SITG, en tant que réseau de partenaires, doit servir ici de principal lieu de partage de ces nouvelles géodonnées en temps réel. Une sorte de plateforme collective des apports et du potentiel de l'IoT.

# 2b

## Les nouvelles relations 2.0

---

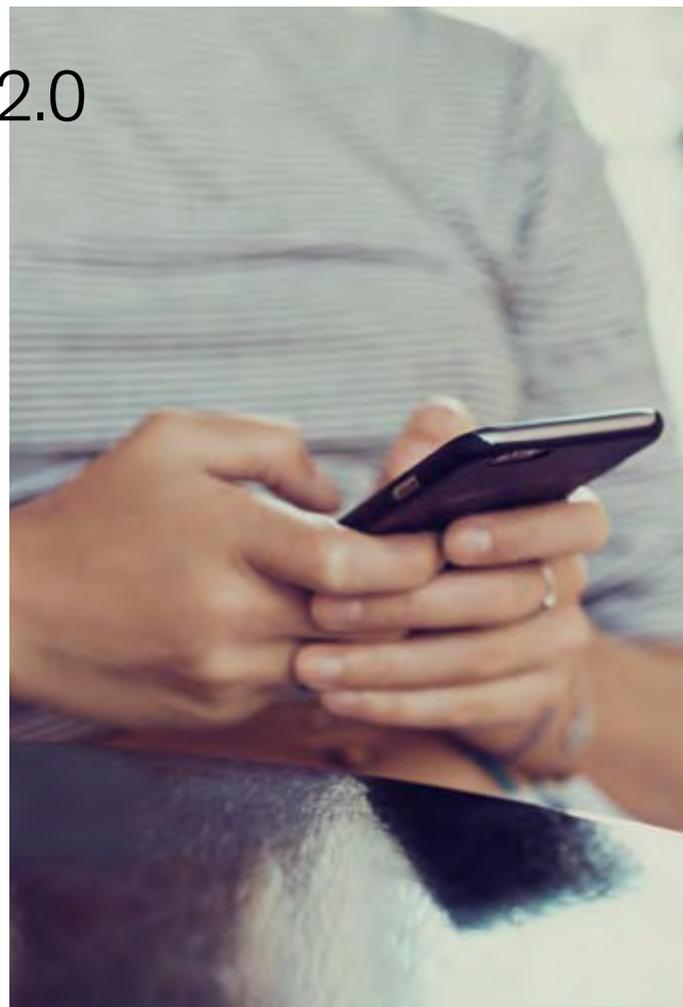
Le citoyen contributeur est sans doute le plus récent acteur sociétal et économique de cette dernière décennie. Créé en partie par le numérique, il est devenu incontournable par son action.

Au travers des réseaux sociaux, des forums de discussions, des commentaires, des chats ou même des blogs, etc. le citoyen contributeur s'exprime sur Internet.

Plus encore par ses comportements ou ses déplacements sur le territoire il signifie les nouveaux territoires : ceux d'une dynamique en temps réel. Embouteillage, trafic, attraction, concentration, etc. sont autant de paramètres relevés en temps réel et restitués aux usagers. La conduite des « acteurs », devenus des citoyens contributeurs, indique la vie des territoires. Les « objets » de celui-ci incluent désormais l'homme et notamment les objets IoT qu'il porte sur lui. Smartphones, smartwatches ou bracelets, voire lunettes connectées, informent tout le collectif-territoire de leurs comportements et attitudes. Une science du comportement géographique va naître. Les données sont en tout cas là et produites en très grand nombre et de manières très diverses.

Dès lors le citoyen contributeur intervient aussi comme producteur de géodonnées. Traduire celles-ci en « smart data » nécessitera tout de même du travail conséquent. C'est la première des tâches. Il faut nécessairement maîtriser les techniques de la donnée non-structurée et du big data.

Après l'ouverture des données publiques (open data) le SITG doit encourager des expériences de la gestion de géodonnées produites par le consommateur ou d'autres acteurs. Une structure comme le « Géofab du Grand Genève<sup>4</sup> » semble appropriée à ce type d'approche pour en favoriser l'appropriation.



# 2c

## Les blockchains

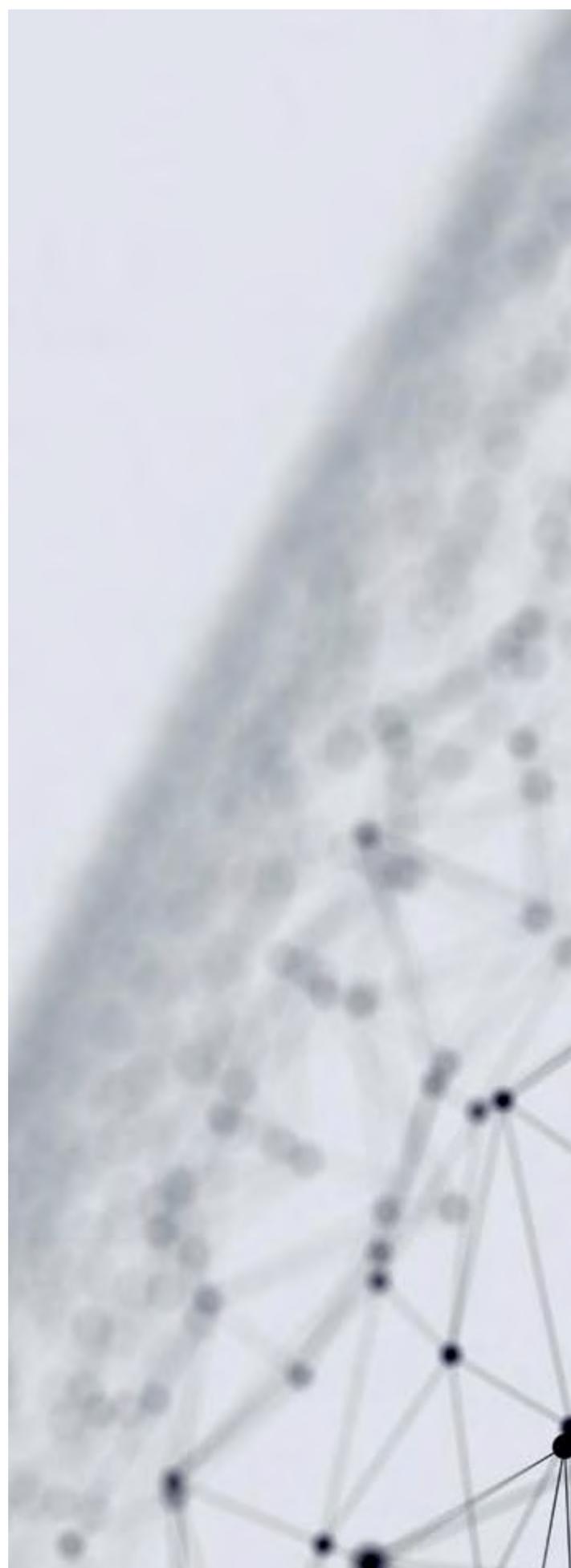
---

Les administrations utilisent pour leurs activités de nombreux registres et constituent donc une opportunité parfaite pour l'implantation de la technologie des blockchains.

Les blockchains actuellement utilisées pour la monnaie (bitcoin) posent des défis importants aux administrations. Elles s'appuient en effet sur une architecture décentralisée, échappant potentiellement au contrôle du législateur d'un pays et présentent un bilan écologique aujourd'hui défavorable, compte tenu des puissances de calcul nécessaires à l'encryptage.

De nouvelles versions préparées par les fournisseurs informatiques devraient permettre rapidement de passer à une phase de démonstration de leur valeur et de la « valeur-client ». On attend prochainement des versions plus « légères » des blockchains qui permettraient de baisser les coûts de la technologie et de son usage.

Le SITG doit encourager ses partenaires à lancer des « démonstrateurs » et partager les expériences métiers dans des domaines où les blockchains pourraient s'avérer utiles, comme par exemple le cadastre. Il ne faut pas cependant oublier que c'est avant tout les registres qui sont concernés par cette technologie et pas le stockage de toutes les données. Le SITG lui-même doit lancer des premières expérimentations en matière de certification des données par exemple.



# 3

## La gouvernance

Dans un environnement en forte transformation, dans lequel les évolutions sont de plus en plus rapides et complexes, les enjeux de gouvernance sont centraux. Ils doivent être fondés sur les principes de l'agilité et de l'adaptation permanente.

### 3a

#### Un "think tank" comme soutien au changement

---

Une réflexion continue est nécessaire pour accompagner la transition. Une organisation de séances sous la forme proposée par les « think tanks » est utile et appropriée pour le réseau des partenaires du SITG. La transformation numérique nécessite autant de réflexions que d'actions. C'est en ce sens que le concept de « think tank » s'impose. Il permet de mettre en commun les apprentissages de chacun. Le réseau est alors la source même de cette intelligence collective. Beaucoup d'entreprises du privé ont choisi cette approche, tout comme au niveau fédéral Swisstopo.

Créer les conditions de l'accompagnement, c'est aussi réfléchir ensemble !

Des séances « think tank » offrent cette souplesse par rapport au concept de commission par exemple.

La mission du SITG est ici l'échange d'idées et d'expériences pour améliorer les chances de chacun d'accomplir avec succès la transition numérique. Il s'agit d'éclairer des choix pour que les « smart geo data » aboutissent au service du territoire et des citoyens.

# 3b

Une

## démarche pédagogique/informationnelle

---

La transition numérique implique une démarche d'explications. Les partenaires et la société civile demandent souvent à être informés et formés.

Dans le domaine de la géoinformation, c'est une tâche qui pourrait incomber au SITG et à ses partenaires.

Une société de la communication impose plus de transparence mais aussi plus d'ouverture sur les savoirs. Elle impose également de savoir se raconter ou, autrement dit, de décrire les démarches entreprises.

Il faudra donc répondre à cette exigence d'une manière ou d'une autre.

Une des tâches du SITG doit être celle de développer une communication forte et proactive qui doit accompagner le développement des smart geodata. Le SITG doit aller à la rencontre des acteurs de la cité et des citoyens contributeurs.

# 3c

## La recherche et l'expérimentation socio- technologiques

---

Il est important de lier toujours plus l'activité du SITG aux centres de recherche autour du territoire. En effet, la phase de transition technologique et numérique nécessite d'être au contact des chercheurs. Ceux-ci ne font pas que trouver (innover), ils cherchent surtout. C'est cette quête qui permet de comprendre la naissance et la transformation des processus nouveaux. Perdre le contact accentue le risque de ne plus comprendre l'évolution et le progrès.

La transformation numérique comporte en effet un risque de fracture. Les expériences socio-technologiques sont tout particulièrement recommandées pour développer des liens harmonieux entre société et technologie. La rupture est possible. Il faut l'éviter.

Les partenaires du SITG doivent naturellement pouvoir conduire de nouveaux projets de partenariats avec des centres de recherche, tout en faisant remonter les informations et pratiques ainsi développées. Tout comme le fait l'Etat de Genève avec les Hautes Ecoles.

Le SITG doit veiller à l'échange de savoir entre ses membres voire élargir son partenariat et le SITG doit tisser des nouveaux liens.



# 4 “Tiers-garants”

---

Jusqu’à aujourd’hui, la bonne qualité des données du territoire (couches), des interfaces (portail) et des applications géographiques était de la responsabilité de l’État, donc garantie par les partenaires du SITG, producteurs de géodonnées.

Mais demain qu’en sera-t-il du contrôle de qualité des données provenant de sources diverses et de formats peu ou pas structurés ?

Quid des procédures de validation ? Comme dans l’exemple des « arbres remarquables » où la direction de la nature a mis dans un premier temps les données à disposition, puis a laissé les personnes compétentes se prononcer sur la base de leurs propres critères et enfin s’est assuré que la démarche a été bien respectée par les participants.

Le SITG permettra-t-il de diffuser des géodonnées ou des applications ne provenant pas des partenaires publics actuels ? Et dans ce cas, se positionne-t-il sur leur « bonne » qualité ou se limite-il à informer que la qualité n’est pas garantie ?

Il y aurait encore bien d’autres questions à aborder. Par exemple, qu’est-ce qu’une bonne donnée ou une bonne application ? Juste, à jour, pertinente, contextualisée, conforme aux valeurs sociétales ou portées par l’Etat (par exemple la protection de la vie privée) ?

L’Etat doit réglementer voire légiférer sur ces questions. Les partenaires du SITG doivent quant à eux élaborer ensemble des propositions.

C’est l’une des tâches prioritaires pour le futur des « smart geodata » et de son apport stratégique aux « smart cities ».

**I INFORMATION GÉOGRAPHIQUE ET OPEN DATA**

Le terme « **information géographique** » caractérise la nature et la structure de la donnée, tandis que le terme « **Open Data** » spécifie son accessibilité.

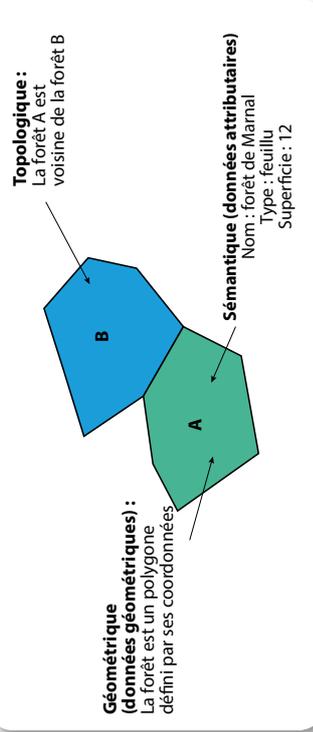
**QU'EST-CE QU'UNE DONNÉE GÉOGRAPHIQUE ?**

Selon Wikipedia<sup>6</sup>, « *La donnée géographique est la représentation d'un objet ou d'un phénomène réel ou imaginaire, présent, passé ou futur, localisé dans l'espace à un moment donné et quelles qu'en soient la dimension et l'échelle de représentation* ».

La donnée géographique est une donnée structurée qui possède 3 composantes :

- **l'information** relative à un objet décrit par sa nature, son aspect : c'est le niveau **sémantique**. L'ensemble des informations liées à l'objet forme ses attributs (ex. : le numéro d'une parcelle cadastrale et ses caractéristiques comme sa surface, son propriétaire, etc.) ;
- les **relations** éventuelles avec d'autres objets ou phénomènes, c'est le niveau **topologique** (ex. : la contiguïté entre deux communes ou l'inclusion d'une parcelle dans une commune) ;
- la **forme** et la localisation de l'objet sur la surface terrestre, exprimées dans un système de coordonnées explicite : c'est le niveau **géométrique**.

© Ministère du développement durable



**On peut ainsi distinguer 3 types de données géographiques :**



**DES USAGES ET DE LA DIFFUSION DE L'INFORMATION GÉOGRAPHIQUE**

Les données géographiques adressent l'ensemble des activités des hommes sur leur territoire.

La gestion des territoires, les volontés politiques et les usages des citoyens créent un besoin toujours croissant en données géographiques pertinentes.

Que ce soit pour des besoins aussi divers que : déclarer des travaux sur la voirie, opérer un réseau d'eau potable, orienter des visiteurs sur des sites remarquables, gérer l'éclairage public, mesurer un risque d'inondation ou distribuer le courrier, l'omniprésence d'une information géographique de qualité se vérifie au quotidien.

**QU'EST-CE QUE L'OPEN DATA ?**

Aujourd'hui, l'Open Data poursuit et complète ce mouvement en clarifiant les droits d'accès et de réutilisation des données. Le mouvement Open Data promet et valorise les données « qui peuvent être librement utilisées, réutilisées et redistribuées par quiconque - sujettes seulement, au plus, à une exigence d'attribution et de partage à l'identique<sup>8</sup> » « sans restriction juridique, financière ou technique<sup>9</sup> ». Le terme « Open Data » ne s'appuie donc pas sur des exigences de structuration, de format ou de sémantique particulières liées à la donnée elle-même, mais sur des exigences relatives à son mode de diffusion.

**LES DONNÉES GÉOGRAPHIQUES, UNE RESSOURCE CLÉ POUR LES PORTAILS OPEN DATA.**

Actuellement, près de la moitié des jeux de données ouverts en France sont des données géographiques. De plus, une très forte proportion des données peuvent être liées directement ou indirectement à une localisation et rattachées à une géométrie (entre 70% et 85% des données, par exemple pour le recensement de la population). D'après les publications d'Open Data France, déjà 45% des jeux de données recensés constituent clairement de l'information géographique.

Les systèmes d'information géographique (SIG), depuis la fin des années 90, ont évolué vers des portails, d'abord intranet puis internet, permettant le partage des données géographiques. Avec la mise en œuvre de la directive européenne INSPIRE, ces portails Web sont devenus des Infrastructures de Données Géographiques (IDG). Celles-ci se sont développées au sein des collectivités et des services de l'État, pour devenir un standard d'architecture pour la diffusion des données. Ainsi, l'accès facilité aux données qu'offrent ces infrastructures explique en grande partie la forte représentation de l'information géographique au sein des portails Open Data.



6. [https://fr.wikipedia.org/wiki/Information\\_g%C3%A9ographique#C3.891.C3.A9ments\\_de\\_d.C3.A9finition](https://fr.wikipedia.org/wiki/Information_g%C3%A9ographique#C3.891.C3.A9ments_de_d.C3.A9finition)  
7. <http://www.ign.fr/institut/activites/referentiel-a-grande-echelle>

8. <http://opendatahandbook.org/guide/fr/what-is-open-data>  
9. [https://fr.wikipedia.org/wiki/Open\\_data](https://fr.wikipedia.org/wiki/Open_data)

## II MUTUALISER ET ORGANISER LA DIFFUSION DES DONNÉES

Aujourd'hui, les plateformes territoriales de l'information géographique ne se limitent pas à de simples infrastructures de données géographiques<sup>10</sup>. Depuis plus de quinze ans, ces dispositifs partenariaux ont été créés à l'échelon régional, départemental voire métropolitain. Initialement le fruit de collaborations entre administrations territoriales et déconcentrées de l'État, elles sont de plus en plus ouvertes à tous les usagers de l'information géographique, pour lesquels elles constituent de véritables centres de ressources géomatiques<sup>11</sup>.

### COORDONNER LES POLITIQUES PUBLIQUES

Au vu de la similarité des vocations des portails Open Data et des plateformes mais aussi des problématiques de rationalisation des ressources humaines, techniques et budgétaires, la question de la mutualisation des infrastructures d'hébergement et de diffusion de données mérite d'être examinée.



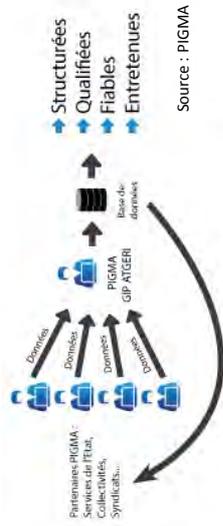
Les plateformes territoriales de l'information géographique sont aussi au cœur de l'application de la directive INSPIRE en France, pour la normalisation des données géographiques et l'amélioration de leur circulation et réutilisation. Cette mise en œuvre, qui polarise fortement leurs activités, a sans doute masqué leur expertise sur tout ce qui a trait à la donnée en général. Au-delà de cette directive, les plateformes permettent avant tout une meilleure coordination des politiques publiques. Elles facilitent le dialogue entre tous les acteurs décisionnels du domaine. Elles améliorent la mise en œuvre des actions sur le terrain.

## VERS DES INFRASTRUCTURES DE DONNÉES MUTUALISÉES

**PIGMA** : une plateforme territoriale Open et Géo en région Nouvelle-Aquitaine

PIGMA est un centre de ressources vivant de données cartographiques. Son rôle d'animation et le réseau mis en place permettent de mettre à disposition une banque de données structurées, qualifiées, fiables et entretenues.

Cette banque de données est gérée uniformément, que la donnée soit ou non géographique, sous licence ouverte, géographique et sous licence ODbL, ou à accès restreint afin de garantir le bon usage et une réutilisation optimale des données par tous ses utilisateurs. En cas d'utilisation des données lors d'une gestion de crise, par exemple, il est indispensable d'avoir des informations fiables et entretenues dans le temps.

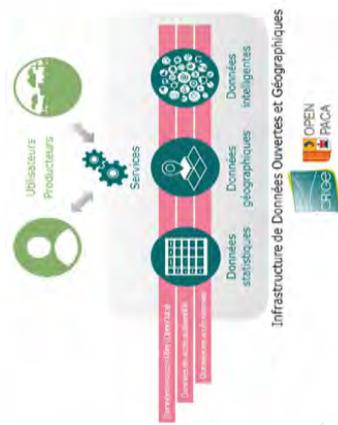


L'existence de multiples portails de données n'est pas un problème en soi si des espaces de mutualisation permettent aux utilisateurs de s'y retrouver et si ces portails sont intégrés suffisamment en amont de leur conception à l'écosystème des portails existants locaux

et nationaux, qui maillent le territoire national. Les Infrastructures de Données Géographiques sont donc un existant dont il faut tenir compte dans la mise en place d'infrastructures de Données Ouvertes.

### CRIGE PACA et OPEN PACA: vers une plateforme mutualisée de données

Les frontières entre données ouvertes, données géographiques et données intelligentes s'effacent au profit des nouveaux usages et de problématiques communes.



Sur Open PACA, plateforme opérée par le conseil régional de PACA, une partie des données ouvertes sont géographiques, alors que sur le Géoportail du CRIGE PACA, une partie des données géographiques sont ouvertes. Cette imbrication des champs d'actions des deux outils a un impact sur la lisibilité des dispositifs et la distinction des rôles vis-à-vis de partenaires en partie communs. L'analyse des évolutions législatives et technologiques et les contraintes budgétaires ont mis en évidence un rapprochement nécessaire.

En choisissant d'opérer une infrastructure mutualisée de données ouvertes, géographiques et intelligentes, sur laquelle s'appuieront des portails de valorisation et d'animation distincts, le CRIGE PACA et Open PACA font le pari de la mutualisation des moyens techniques pour simplifier la diffusion des données en région.

### III LE CATALOGAGE, UNE ÉTAPE INCONTOURNABLE

#### ➤ POURQUOI CATALOGUER SES DONNÉES ?

Pour que les données soient accessibles, les données : à l'instar des sites de commerce au plus grand nombre, les plateformes en ligne présentant leurs produits via un territoriales disposent de moyens facilitant catalogue de vente, les données géographiques leur inventaire, leur recherche et également sont également inventoriées via des fiches leur réutilisation. Le point de départ de cette descriptives, appelées « métadonnées ». volonté est la constitution d'un catalogue de

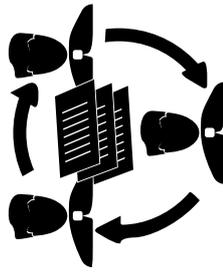
#### UN CATALOGUE

Un **catalogue** est un outil logiciel qui offre un point d'accès centralisé et structuré (d'un point de vue méthodologique et technique) de fiches descriptives sur les données mises à disposition. Cet outil est une interface ergonomique entre les gestionnaires de données et les utilisateurs.

**Cataloguer**, c'est donc faire connaître ce que l'on a produit, en faire un état sincère, pragmatique et avec recul, révélant une transparence et une démarche volontariste. C'est une pratique en partie tournée vers les autres : structures utilisatrices d'information géographique et, dans le cas de l'Open Data, vers les citoyens.

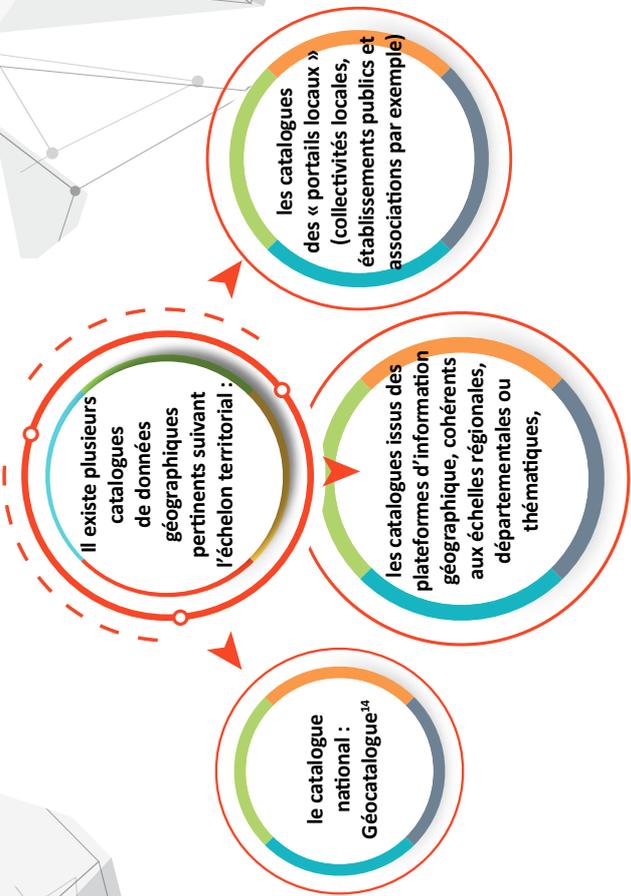
Pour celui qui produit la donnée, le catalogue doit répondre à plusieurs objectifs (tout en se conformant aux exigences de la directive européenne INSPIRE pour les données concernées par ses Annexes, et également plus récemment de la loi NOTRe<sup>12</sup>) :

- **Décrire la donnée** en spécifiant ses principales caractéristiques : titre, description, date de création, échelle, fréquence de mise à jour, contacts à prendre pour en savoir plus et qualité. Ainsi, le producteur précise toutes les conditions nécessaires à une réutilisation maîtrisée.
- **Inventorier** le patrimoine de données existantes pour son entretien et sa pérennité.
- **Permettre à tous** de prendre connaissance des données produites<sup>13</sup>.
- **Se rendre plus visible** en tant que producteur de données, en faisant connaître les données produites, par des métadonnées. Quoi de mieux qu'une donnée valorisée par d'autres partenaires pour créer un cercle vertueux de la réutilisation, allant même jusqu'à la production collaborative entre structures ?



Pour celui qui recherche un jeu de données, le catalogue doit permettre de :

- **trouver** facilement la donnée recherchée à l'aide d'un moteur de recherche et de mots-clés intuitifs présentés dans sa métadonnée,
- **disposer** d'une description détaillée de la donnée disponible afin de savoir si elle répond à son besoin,
- **connaître** les conditions (techniques et juridiques) de réutilisation de la donnée,
- **accéder** aux ressources de consommation de la donnée (services web, téléchargement...),
- **bénéficier** du contact du producteur.



Tous ces catalogues sont en fait reliés les uns aux autres pour faire « remonter/redescendre » les métadonnées et ainsi transmettre la connaissance des données, par un dispositif de « moissonnage ». Ces interconnexions assurent la synchronisation des catalogues. De par les avancées tant en termes de réglementation que d'actes d'intention vers l'ouverture des données, l'accès et la connaissance des données se trouvent facilités par les activités de catalogage.

Le catalogue constitue un vecteur pour une meilleure efficacité des pouvoirs publics. Il permet un gain de temps pour la transmission des informations et favorise leur réutilisation.

#### ➤ LES MÉTADONNÉES, UN GAGE DE QUALITÉ

Pour permettre un usage approprié et une réutilisation pertinente, les caractéristiques de production et les conditions de réutilisation d'une donnée, géographique ou non, doivent être connues des utilisateurs potentiels. Certaines données peuvent être satisfaisantes pour un usage tandis qu'elles se révéleront parfaitement inexploitable pour d'autres : on parle de « qualification » nécessaire de la donnée pour évoquer sa « qualité ».

Par exemple, des données d'études préalables pour la construction d'une infrastructure routière ne sont pas produites et ré-utilisables à la même échelle (décamétrique) que celle de données relevées pour sa construction (inframétrique). Idéalement, la description de l'échelle d'utilisation doit être reportée dans la métadonnée.

Si l'on compare les données à des produits de consommation courante, on n'imagine pas acheter une bouteille de vin, sans connaître son nom (titre), sa provenance (producteur), ni même ses cépages ou son terroir (généalogie), ses conseils de dégustation (descriptions de qualité) ou son millésime (année de production).

Appliquées aux données, cela revient à donner un titre, une date de création, un résumé et des descriptions annexes (format, producteur, distributeur) les plus pertinents possibles. Ainsi, un résumé permet de connaître les conditions de réutilisation d'un point de vue réglementaire (notamment par la désignation d'une licence d'utilisation, comme la Licence Ouverte d'Etablab<sup>15</sup>).

12. <http://cdliq.gouv.fr/?p=11652>

13. Ces trois premiers alignés correspondent aux obligations de la directive européenne INSPIRE et également à celles de la loi NOTRE.

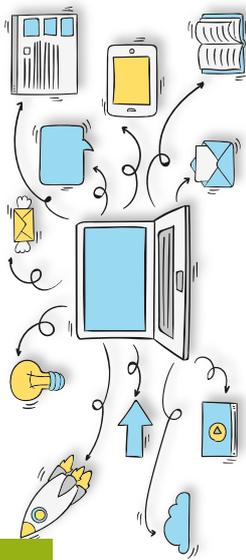
14. <http://www.geocatalogue.fr/>

15. <https://www.etablab.gouv.fr/licence-ouverte-open-licence>

#### IV HARMONISER POUR RÉUTILISER

Dans le domaine de la **géomatique** et pour répondre à des besoins d'échange de données et de métadonnées entre organisations, les métadonnées (ISO 19115), les Web services d'accès (OGC<sup>16</sup>) et les modèles de données (CNIG<sup>17</sup>, COVADIS<sup>18</sup>) sont standardisés et largement utilisés.

Grâce à l'utilisation de ces standards, en France, l'interopérabilité entre les centaines de catalogues de données géographiques et la mise à disposition harmonisée de dizaines de milliers de jeux de données sont aujourd'hui effectifs.



#### L'utilisation des services Web standardisés permet notamment :

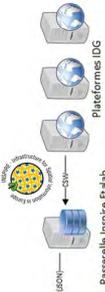
Une **duplication** intelligente (avec gestion des doublons) des métadonnées sur le catalogue national et dans data.gouv.fr

Une **diffusion et une intégration** facilitées de la dernière version des référentiels géographiques dans les outils de consultation de données. Par exemple, les Web services de l'IGN (photographie aérienne, SCAN...) permettent aux utilisateurs de disposer de ces fonds de plan dans leurs outils cartographiques sans avoir à télécharger et intégrer les données brutes correspondantes.

Aujourd'hui, l'émergence de l'Open Data offre l'opportunité d'adapter ces normes aux usages récents de l'Internet<sup>19</sup>. Ainsi, l'IGN a pris en compte le système de projection utilisé par Google pour la diffusion de ses services Web cartographiques (EPSG:3857 ou Pseudo-Mercator).

#### La Passerelle INSPIRE – Data.Gouv.FR

En outre, concernant le partage de métadonnées, des connexions entre catalogues existent déjà : développée par Etalab au sein du SGMAP (Secrétariat Général pour la Modernisation de l'Action Publique), la passerelle INSPIRE<sup>20</sup> moissonne plusieurs dizaines de catalogues de données issus des plateformes territoriales, et les valorise sur « data.gouv.fr », auprès d'un public renouvelé, susceptible de développer de nouveaux usages au-delà de la sphère professionnelle traditionnellement utilisatrice de ces géoportails. En 2016, cette passerelle recense près de 10 000 jeux de données géographiques ouvertes et 80 organismes partenaires.



Source : Passerelle INSPIRE ©Etalab

L'exploitation des données – qu'elles soient géographiques ou non – via des API<sup>21</sup>. Web est de plus en plus courante. Elle sera sans doute à l'avenir, le mode de diffusion naturelle des données afin d'en assurer un accès immédiat et à jour, à l'image des API<sup>22</sup> mises en place par Etalab pour faciliter la consultation et l'intégration de données publiques métier (cadastre, urbanisme, risques...).

#### Pour mener à bien l'objectif de favoriser l'émergence d'API pertinentes,

Il semble inévitable de promouvoir un niveau minimal d'harmonisation des productions de données répondant à des besoins identiques. En effet, l'harmonisation des données facilite l'intégration des données par les développeurs d'API leur permettant ainsi de se focaliser sur la plus-value des services qu'ils mettent en place plutôt que sur la qualité des données qu'ils doivent exploiter.

Bien sûr, l'harmonisation, lorsqu'elle est pratiquée à outrance, peut réduire la richesse des données voire les dénaturer. Elle doit donc se concentrer sur les informations dont l'exploitation à grande échelle ne serait pas possible sans cela. Notons que le fait de mettre en place une API Web unique pour un ensemble de jeux de données non harmonisés constitue en soi une opération d'harmonisation visant à fournir un service homogène aux ré-utilisateurs des données.

Différentes initiatives de premier plan montrent parfaitement que l'harmonisation à l'échelle nationale de jeux de données est un enjeu majeur actuel : le **Géoportail de l'urbanisme**<sup>23</sup> et le Service public de la donnée<sup>24</sup> en sont des exemples marquants.

Ces démarches montrent à quel point il est important d'harmoniser pour toucher un large public d'utilisateurs, qu'ils soient intégrateurs, ré-utilisateurs, développeurs d'applications internet ou de simples usagers. Cette démarche commencée de longue date pour les données géographiques doit être poursuivie et adaptée pour les données ouvertes, notamment pour l'échange de données<sup>25</sup> produites par les collectivités territoriales. C'est d'ailleurs ce qui est préconisé dans le rapport sur les dispositifs d'accompagnement des collectivités locales à l'ouverture des données publiques rédigé par l'association OpenData France (sans toutefois y préciser les modalités de réalisation).

16. Open Geospatial Consortium <http://www.opengeospatial.org/docs/ls>

17. [http://cnig.gouv.fr/?page\\_id=858](http://cnig.gouv.fr/?page_id=858)

18. COVADIS – Publication des standards de données validés par la COVADIS

19. Synthèse de l'atelier « Quand les standards du Web viennent enrichir les bonnes pratiques d'interopérabilité 1 », AFIGEO, 2014

20. <https://inspire.data.gouv.fr/>

21. API (Application Programming Interface / interface de programmation applicative)

22. <https://api.gouv.fr/>

23. <https://www.geoportail-urbanisme.gouv.fr/>

24. <https://www.etalab.gouv.fr/tae/service-public-de-la-donnee>

25. A l'instar de ce que propose la loi sur la Géoinformation en Suisse en matière de modèles de données minimales

Les données géographiques sont largement représentées sur les plateformes Open Data, ce n'est donc pas un hasard si de nombreuses collectivités demandent aux géomaticiens de s'impliquer dans la préparation et l'ouverture de leurs données.

La montée en puissance, la généralisation et la mise en cohérence des approches Open Data dans les territoires, est indispensable. Elles passent nécessairement par la structuration des données, l'organisation des échanges et la mutualisation entre acteurs. Ainsi, l'ensemble des compétences et connaissances des géomaticiens doivent pouvoir être utilisées et valorisées à bon escient par les porteurs de projets Open Data.

**Le mouvement Open Data dans son approche organisationnelle, par son agilité et son innovation, a bousculé le monde de la géomatique, attaché à ses règles et pratiques. Il a apporté son lot de nouveautés qui permettent désormais à tous de réfléchir à :**

plus d'ouverture :  
nouveaux usages,  
nouveaux utilisateurs,  
nouveaux outils ;

plus de souplesse : en  
termes d'animation, de  
mise à disposition de  
données, de prise en  
compte des remontées  
utilisateurs ;

plus de collaboration  
entre toutes les parties  
prenantes des territoires  
comme les entreprises  
privées ou la société civile.

Ces évolutions ont conduit à des changements, une et de gain en efficacité, il semble donc indispensable nouvelle vision et des rapprochements. Le partage que chaque communauté professionnelle propose des compétences entre ces deux communautés des règles correspondantes à ses propres usages, professionnelles est une opportunité à ne pas manquer. Leurs différences, désormais moins partagées au niveau national. Enfin, il apparaît nécessaire de rationaliser les démarches, représentent une complémentarité de portails de données mutualisés répondant à des enjeux de transparence et d'accessibilité des données. Cependant, il est difficile de définir des données auxquelles qu'elles soient. C'est le défi des dix prochaines années : sachons le relever !

## LE GÉOMATICIEN ET L'OPEN DATA

Le terme générique de géomaticien, regroupe les professionnels agissant dans le domaine de l'information géographique numérique, dont le cœur de métier recouvre diverses activités qui leur permettent de connaître :

- le patrimoine de données disponibles au sein de sa structure,
- leurs caractéristiques techniques,
- les modalités de leur production, gestion et de leur mise à jour,
- les contraintes juridiques les concernant.

La nécessité d'interconnecter le SIG à d'autres systèmes d'information transversaux ou métiers conduit les géomaticiens à maîtriser :

- les outils et concepts de la « qualité » des données (métadonnées, réutilisation dans d'autres applications informatiques...),
- les modalités techniques de diffusion (dont les normes et standards tels que les Web services de l'Open Geospatial Consortium (OGC)).

Pour assurer ses missions, le géomaticien s'appuie sur une formation scientifique et technique comprenant dans des proportions variables des bases en géographie, informatique et topographie ainsi que des compétences thématiques.

Il assure également des fonctions d'accompagnement, d'animation et de gestion reposant sur une dimension humaine forte, consistant par exemple à former des utilisateurs à l'usage de l'information géographique, contractualiser des partenariats avec des organismes extérieurs et animer un réseau d'acteurs s'ouvrant, de plus en plus, au grand public.

Référence : « Petit guide pratique de la géomatique à destination des employeurs, des candidats et des formateurs », AFIGEO – 2013

## RESSOURCES AFIGEO

Ressources du Groupe de travail Open Data de l'AFIGEO

Données géographiques ouvertes : bilan et perspectives de la passerelle INSPIRE / Etalab, AFIGEO-Etalab, sept. 2016  
Synthèse du colloque « Quand l'information géographique et Open Data se rencontrent », AFIGEO, 2012

Synthèse de l'atelier « Quand les standards du Web viennent enrichir les bonnes pratiques d'interopérabilité 1 », AFIGEO, 2014  
Synthèse de l'atelier « La nécessaire interopérabilité Inspire et Open Data ? », AFIGEO, 2015

Synthèse du Hackathon INSPIRE – Open Data, AFIGEO-Etalab, 2015

Tables rondes « L'Open Data et l'information géographique : synergie, usages et impacts » et « Politiques publiques et numériques : quelle place pour les plateformes d'information géographique ? » (synthèse des 8<sup>es</sup> Rencontres en Information géographique), AFIGEO, 2014  
Catalogue des IDG, AFIGEO, 2014

Objectifs et missions des plateformes territoriales, AFIGEO, 2014

Petit guide pratique de la géomatique à destination des employeurs, des candidats et des formateurs, AFIGEO, 2013

## AUTRES RESSOURCES CITÉES

Conseil fédéral, Gouvernement suisse – Loi sur la géoinformation de modèles de données minimaux COVADIS – Publication des standards de données validés par la COVADIS

isogeo (Becker M. et Halbout H.) – Livre blanc sur la gouvernance des données géographiques, 2016

Libert7C – Vers la fin du baby blues de l'open data ?, 2013

Mission de l'information géographique (Ministère de l'Environnement) – La directive Inspire pour les néophytes, V4, 2015

OpenData France – Rapport sur les dispositifs d'accompagnement des collectivités locales à l'ouverture des données publiques, 2016

## AUTRES LIENS UTILES

[https://fr.wikipedia.org/wiki/Information\\_g%C3%A9ographique#C3.89.C3.A9iments\\_de\\_d.C3.A9inition](https://fr.wikipedia.org/wiki/Information_g%C3%A9ographique#C3.89.C3.A9iments_de_d.C3.A9inition)

<http://www.ign.fr/institut/activites/referentiel-a-grande-echelle>

<http://opendatahandbook.org/guide/fr/what-is-open-data>

[https://fr.wikipedia.org/wiki/Open\\_data](https://fr.wikipedia.org/wiki/Open_data)

<http://cnig.gouv.fr/?p=11652>

<http://www.geocatalogue.fr/>

<https://www.etalab.gouv.fr/licence-ouverte-open-licence>

<http://www.opengeospatial.org/docs/ls>

[http://cnig.gouv.fr/?page\\_id=858](http://cnig.gouv.fr/?page_id=858)

<https://inspire.data.gouv.fr/>

<https://api.gouv.fr/>

<https://www.geoportail-urbanisme.gouv.fr/>

<https://www.etalab.gouv.fr/tag/service-public-de-la-donnee>

## POUR ALLER PLUS LOIN

Agence Aquitaine du Numérique – Open data : Deux ans d'action publique en Aquitaine, 2012

Collectif OpenData France – Loi NOTRe : quelles données publiques ouvrir ?, 2015

Daniel Dietrich & Co. – Le Manuel de l'Open Data

Data Grand Lyon – Commaître nos licences

La FING – Guide de l'innovation centrée-usager, 2014

OpenData France, La Reine Merlin –

Glossaire de la donnée publique, 2016

Les lois régulant la donnée publique, 2016

SGMAP – Vade-mecum sur l'ouverture et le partage des données publiques, 2013

[...]

## DOCUMENT 5

« Les données géographiques souveraines » (extraits) – Valéria FAURE-MUNTIAN – *ecologique-solidaires.gouv.fr* – Juillet 2018

### INTRODUCTION

[...]

La révolution numérique confronte la puissance publique à des défis inédits. L'un d'eux concerne l'exercice de la souveraineté même, entendue au sens large comme la capacité à agir de manière indépendante et à exercer l'autorité dont les pouvoirs publics sont démocratiquement investis, face à des acteurs internationaux dont la puissance croissante réside dans la maîtrise et l'exploitation de quantités considérables de données.

La géographie illustre mieux que n'importe quel autre domaine les mutations à l'œuvre dans la société numérique. Historiquement, la donnée géographique entretient un lien très étroit avec l'exercice de la souveraineté, d'abord dans sa dimension militaire. Citons, sans qu'il soit besoin de remonter à la « carte de Cassini », première cartographie homogène du territoire initiée par l'Académie des sciences en 1666, l'établissement de la carte d'état-major en 1817 ou à la création, quatre-vingts ans plus tard, du Service géographique de l'armée devenu en 1940 Institut géographique national puis, en 2012, Institut National de l'Information Géographique et forestières après sa fusion avec l'Inventaire forestier national (IFN). De nos jours, la puissance publique acquiert, produit et mobilise quotidiennement des données géographiques et plus largement des données géolocalisées à l'appui de ses décisions et de son action, dans des domaines aussi variés que la défense nationale, la sécurité, la prévention des risques naturels et technologiques, la fiscalité, l'urbanisme, les transports, la santé...

Illustrative, la géographie l'est aussi, et surtout, parce qu'elle constitue l'un des domaines dans lesquelles le risque de « disruption » de la puissance publique est le plus patent. Compte tenu de l'intérêt stratégique évident que présente l'information géolocalisée, les « GAFAM<sup>1</sup> » ont constitué des fonds géographiques dont la richesse, la simplicité d'utilisation et la gratuité (à tout le moins pour les données de base et les usages élémentaires) en ont fait des produits de consommation courante pour les citoyens. Plus généralement, la société civile et des acteurs historiques de la géomatique jouent un rôle actif et positif dans la mutation de l'écosystème.

Le présent rapport fait donc le présupposé que la puissance publique, pour exercer souverainement ses prérogatives et accomplir ses missions d'intérêt général, doit conserver la maîtrise des données qui fondent ses décisions, au même titre qu'elle conserve celle de ses autres « infrastructures essentielles »<sup>2</sup>. Cette maîtrise conditionne non seulement l'efficacité de l'action publique, qui doit pouvoir se fonder sur des données qualifiées, mais aussi la confiance que les citoyens<sup>3</sup> placent en elle. Elle contribue également à la souveraineté nationale entendue dans son acception économique, en permettant à nos entreprises de ne pas dépendre de grandes plateformes étrangères pour accéder aux marchés et développer leur activité et en leur offrant un support à l'innovation (voiture autonome, livraison par drone, villes intelligentes...). Une telle approche ne signifie évidemment pas que la puissance publique doive se comporter en « forteresse assiégée » à l'égard des acteurs privés.

Les autorités publiques ont parfois intérêt à coopérer avec eux dans l'élaboration de communs numériques ou à collecter leurs données lorsque le législateur leur reconnaît un caractère d'intérêt général.

La lettre de mission à laquelle s'efforce de répondre le présent rapport comportait un grand nombre de questions, dont la caractérisation même d'une donnée « souveraine » n'était pas la moins épineuse. Le calendrier des travaux était par ailleurs commandé par la nécessité d'articuler les réflexions ici conduites avec l'élaboration du nouveau projet d'établissement de l'IGN, elle-même inscrite dans le processus « Action publique 2022 ». Il a été, dans ce court laps de temps, choisi d'auditionner un panel aussi large que possible d'acteurs, publics et privés, concernés par les mutations du secteur de l'information géographique.

<sup>1</sup> Google, Apple, Facebook, Amazon, Microsoft.

<sup>2</sup> Pour reprendre l'expression employée par l'administrateur général des données dans son récent rapport au Premier ministre : « La donnée comme infrastructure essentielle », 2018.

<sup>3</sup> La récente polémique suscitée par le visionnage obligatoire, via le site Youtube, d'une vidéo de présentation du prélèvement à la source avant de pouvoir déclarer ses revenus en ligne suggère que les Français sont attentifs à la neutralité de leurs relations « numériques » avec l'administration.

Les grands axes de la commande (souveraineté de la donnée, organisation optimale de la puissance publique en matière d'information géographique) sont depuis plusieurs années au cœur des réflexions de l'IGN et de ses autorités de tutelle. Comme le rappelle un rapport récemment consacré à l'institut par le Conseil général de l'environnement et du développement durable et le Conseil général de l'alimentation, de l'agriculture et des espaces ruraux<sup>4</sup>, deux des trois orientations principales de son contrat d'objectifs et de performance 2013-2016 étaient déjà consacrées à la nécessité pour l'IGN de devenir une composante de la souveraineté et de l'indépendance nationale en fournissant aux administrations publiques les données de référence permettant une description du territoire faisant autorité, et d'ancrer son rôle d'opérateur de référence en matière d'information géolocalisée en organisant un système collaboratif d'échanges et de partage de données entre les acteurs publics et en favorisant leur diffusion. Certaines questions abordées dans les développements qui suivent étaient également en germe dans le rapport Lengagne sur les *Perspectives d'évolution de l'information géographique et les conséquences pour l'IGN*, pourtant rendu il y a près de vingt ans.

La persistance de ces interrogations trahit la difficulté qu'ont pu éprouver les pouvoirs publics à adapter leur organisation et leurs modalités d'intervention dans un contexte très évolutif, marqué par des initiatives foisonnantes et dispersées, ainsi que par des progrès technologiques rapides et continus. Les travaux dont le présent rapport est issu ont rapidement fait apparaître que l'extrême variété des données géographiques intervenant dans les décisions de la puissance publique et la complexité de l'écosystème d'acteurs que fait intervenir leur production se prêtaient mal à une approche monolithique ou systématique des solutions à mettre en œuvre.

Dans ces conditions, il n'est apparu opportun ni de conférer une portée normative à la définition des données géographiques souveraines, ni de préconiser un modèle unique d'organisation et de gouvernance qui, bien qu'intellectuellement séduisant, aurait eu toutes les chances de demeurer théorique. Une production plus efficiente des données géographiques souveraines passe essentiellement par la sensibilisation des administrations aux enjeux de la maîtrise de ces données et par le développement de coopérations autour de projets communs dont l'animation nationale doit être renforcée. Le financement de cette production doit, conformément au choix des pouvoirs publics de privilégier l'*open data*, reposer sur l'impôt.

Les préconisations qui suivent se veulent avant tout réalistes et pragmatiques. Elles constituent à la fois des orientations générales pour l'action, des recommandations opérationnelles ciblées là où il est apparu que la réflexion était assez mûre pour en formuler et des suggestions d'expérimentations dont le succès conditionnera, le cas échéant, la généralisation.

---

<sup>4</sup> Rapport CGEDD-CGAAER sur l'analyse et les propositions d'évolution de l'IGN, novembre 2017.

## SYNTHÈSE DES RECOMMANDATIONS

1. **Faire réaliser et actualiser périodiquement par les administrations centrales et les établissements publics de l'État un recensement des données géographiques souveraines qu'ils produisent ou dont ils ont besoin.**
2. **Mobiliser les administrations impliquées dans la gestion de crise autour d'un projet de mutualisation visant à ce qu'un même objet géographique fasse l'objet d'une description commune à l'ensemble de leurs systèmes d'information.**
3. **Fixer dans une circulaire du Premier ministre les exigences minimales auxquelles doivent répondre les données géographiques auxquelles les administrations centrales et déconcentrées adossent leurs décisions.**
4. **Créer un point d'accès unique aux données et métadonnées géographiques souveraines placé sous la responsabilité de l'IGN.**
5. **Transférer la tenue du Géocatalogue « INSPIRE » du BRGM à l'IGN.**
6. **Fédérer les producteurs de données géographiques souveraines autour de projets d'équipement communs, notamment dans les domaines des données de transport ou de la réalisation d'un modèle numérique tridimensionnel à maille très fine.**
7. **Renforcer les synergies entre l'IGN et les collectivités territoriales dans la réalisation de la couverture orthophotographique du territoire.**
8. **Mettre en place une plateforme mutualisée d'information et de services géolocalisés pour la préparation et le déroulement des Jeux olympiques et paralympiques de Paris 2024.**
9. **Créer une géoplateforme nationale des données géographiques souveraines conçue à terme comme une infrastructure mutualisée entre les différents producteurs.**
10. **Afin de favoriser l'adhésion des acteurs, expérimenter la constitution et l'exploitation de bases de données communes entre l'IGN et les producteurs « tiers » (collectivités notamment) à partir de l'infrastructure mutualisée.**
11. **Valoriser et encourager le développement d'acteurs français et européens du stockage et du calcul.**
12. **Faire jouer au Conseil national de l'information géographique un rôle effectif de coordination et d'animation des producteurs de données géographiques souveraines.**
13. **Placer le CNIG auprès du Premier ministre, faire assurer son secrétariat par la DINSIC et adapter sa composition.**
14. **Structurer un réseau homogène de conseils régionaux de l'information géographique.**
15. **Créer dans chaque région une plateforme unique d'information géographique, commune à la région et à l'État.**

16. **Préserver les moyens dédiés à l'information géographique dans les services déconcentrés de l'État.**
17. **Repositionner l'IGN sur l'agrégation, la standardisation, la certification des données géographiques souveraines et l'accompagnement technique des producteurs.**
18. **Structurer et animer une communauté des géomaticiens publics favorisant le partage d'expériences et de bonnes pratiques et développer avec eux une réflexion sur les compétences et savoir-faire spécifiques que requièrent la production et la diffusion des données géographiques souveraines.**
19. **Adapter la composition du conseil d'administration de l'IGN en renforçant notamment la représentation des régions. Y faire siéger le DINSIC ès-qualité.**
20. **Évaluer à échéance du 1<sup>er</sup> trimestre 2019 les impacts techniques, organisationnels, juridiques et financiers d'un transfert des missions topographiques de la DGFIP à l'IGN et procéder au transfert avant la fin de la législature.**
21. **Étendre progressivement à l'ensemble des zonages réglementaires la démarche entreprise autour du Géoportail de l'Urbanisme.**
22. **Étudier, en concertation avec les associations représentatives des communes, l'opportunité de systématiser l'obligation de réaliser des plans d'adressage et de rendre l'adresse opposable.**
23. **Mandater l'IGN pour jouer un rôle plus actif dans le pilotage du déploiement du PCRS et revoir les modalités de financement du projet, soit par la création d'un fonds de mutualisation, soit par l'identification d'une enveloppe dédiée dans le cadre de la prochaine génération de CPER.**
24. **Poser un principe de gratuité de la mise à disposition des données géographiques souveraines.**
25. **Faire couvrir par la subvention pour charges de services public les dépenses de fonctionnement et de personnel consacrées par l'IGN à la production, à l'entretien, à la certification des données géographiques souveraines, ainsi qu'à l'accompagnement technique des autres acteurs de l'écosystème et à la gestion de la géoplateforme nationale.**
26. **Clarifier l'articulation des licences ouvertes de type Etalab avec les licences dites « contaminantes ».**
27. **Diffuser à terme l'ensemble des données géographiques souveraines sous licence ouverte type Etalab.**
28. **N'autoriser l'IGN à recourir à la licence payante qu'à titre transitoire et au plus tard jusqu'en 2022.**

**LORIENT** AGGLOMERATION (exemple de catalogue de données SIG INSPIRE – outil ISOGEO)

ACCUEIL CATALOGUE CARTE ACTUALITÉS PROJET CONTACTEZ-NOUS

RECHERCHE [thématique, texte...] Adresse [OK] [G] s'authentifier | créer un compte

**RÉSULTATS DE RECHERCHE : 152 JEUX DE DONNÉES**

1 à 20 sur 152 Type de données Filtre par... Nom Page 1

### CADASTRE DE LE COMMUNE DE LOCMIQUELIC

Le plan du cadastre est un document administratif qui propose l'unique plan parcellaire à grande échelle couvrant le territoire national.

Le plan du cadastre d'une commune est découpé en sections, elles-mêmes pouvant être découpées en subdivisions de sections, communément appelées « feuilles de plan ».

- La section cadastrale est une portion du territoire communal dont le périmètre est généralement constitué par des limites présentant un caractère relativement stable sur le terrain (route, chemins, cours d'eau...).
- La parcelle est l'unité cadastrale de base. C'est un terrain d'un seul tenant situé dans un même lieudit et appartenant à un même propriétaire.
- Le plan du cadastre au format vecteur est issu majoritairement de numérisation du plan cadastral papier ou raster réalisée dans le cadre de conventions avec les collectivités territoriales. Dans une moindre mesure, il a été confectionné directement au format numérique dans le cadre de la production de plans cadastraux neufs (procédure du remaniement prévue par la loi n° 78-645 du 18 juillet 1974) ou d'aménagements fonciers agricoles et forestiers (communément désignés sous le terme « remembrements »).

Les plans cadastraux au format vecteur en France métropolitaine sont actuellement géoréférencés dans le système légal (RGF93) à

[Afficher le résumé intégral](#)

[cadastre](#) [données ouvertes](#) [locmiquelic](#) [lorient agglomération](#) [mission sig](#) [référentiels : cadastre](#)

### CADASTRE DE LE COMMUNE DE LOCMIQUELIC



Échelle : 1:1 000 Type : **Données vecteur**

[cadastre](#) [données ouvertes](#) [locmiquelic](#) [lorient agglomération](#) [mission sig](#) [référentiels : cadastre](#)

#### RÉSUMÉ

Le plan du cadastre est un document administratif qui propose l'unique plan parcellaire à grande échelle couvrant le territoire national.

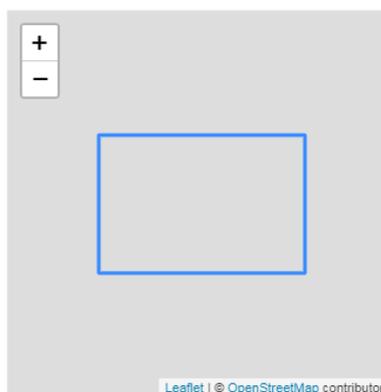
Le plan du cadastre d'une commune est découpé en sections, elles-mêmes pouvant être découpées en subdivisions de sections, communément appelées « feuilles de plan ».

- La section cadastrale est une portion du territoire communal dont le périmètre est généralement constitué par des limites présentant un caractère relativement stable sur le terrain (route, chemins, cours d'eau...).
- La parcelle est l'unité cadastrale de base. C'est un terrain d'un seul tenant situé dans un même lieudit et appartenant à un même propriétaire.
- Le plan du cadastre au format vecteur est issu majoritairement de numérisation du plan cadastral papier ou raster réalisée dans le cadre de conventions avec les collectivités territoriales. Dans une moindre mesure, il a été confectionné directement au format numérique dans le cadre de la production de plans cadastraux neufs (procédure du remaniement prévue par la loi n° 78-645 du 18 juillet 1974) ou d'aménagements fonciers agricoles et forestiers (communément désignés sous le terme « remembrements »).

Les plans cadastraux au format vecteur en France métropolitaine sont actuellement géoréférencés dans le système légal (RGF93) à l'aide des projections « coniques conformes 9 zones ». Le géoréférencement des plans vecteur a pu être obtenu :

- lors de leur confection (cas des plans cadastraux très récents) ;
- après transformation de leurs coordonnées exprimées dans la projection Lambert zones ;
- lors de leur vectorisation (cas des plans non initialement géoréférencés)

#### ZONE GÉOGRAPHIQUE



#### INFORMATIONS TECHNIQUES

**Type** : Données Vecteur  
**Nom de la couche** : Locmiquelic\_56118CAD  
**Emplacement** : \\caplorient.rcval.fr\sig-ref01\BD\_CADASTRE\DWG\Locmiquelic\_56118CAD.dwg  
**Nombre d'entités** : 19381  
**Type de géométrie** : Polygon  
**Résolution** : non renseigné  
**Échelle** : 1/1 000  
**Format de référence** : dwg 2010  
**Système de coordonnées** : RGF93 / CC48  
**Encodage des caractères** : utf-8

## HISTORIQUE DE LA DONNÉE

Date de création : 04/03/2016

Date de dernière modification : 31/07/2019

### Contexte de collecte :

Les plans cadastraux au format vecteur gérés et exportés par les services de la DGFIP ont été confectionnés à partir de levés terrain. Selon les feuilles de plan, les conditions de vectorisation et de géoréférencement sont diverses. Ces informations sont disponibles sur les fiches de chacune des communes.

### Méthode de collecte :

Les plans cadastraux au format vecteur gérés et exportés par les services de la DGFIP ont été confectionnés à partir de levés terrain. Selon les feuilles de plan, les conditions de vectorisation et de géoréférencement sont diverses. Ces informations sont disponibles sur les fiches de chacune des communes.

Les plans cadastraux sont de plusieurs types :

- plans dits « mis à jour » : ce sont des plans cadastraux mis à jour à partir de la trame des plans cadastraux napoléoniens. Ce ne sont pas des plans réguliers. Ces feuilles de plan sont identifiables par leur section à lettre unique (par exemple « section A »). Les échelles des feuilles de plan peuvent être variables et sont principalement aux échelles suivantes : 1/1250, 1/2500 et 1/5000 ;
- plans dits « réguliers » : ce sont des plans qui ont été confectionnés lors de la rénovation du cadastre sans reprendre la trame des plans napoléoniens. Parmi ces plans, l'on distingue deux catégories :
- les plans dits « renouvelés » : ce sont des plans qui n'ont pas fait l'objet lors de leur confection d'une délimitation contradictoire des propriétés sur le terrain ;
- les plans « refaits » qui ont quant à eux fait l'objet d'une délimitation contradictoire des propriétés ;

Lorsque le plan cadastral n'est plus en mesure de répondre aux besoins (notamment en raison de son échelle ou de son imprécision éventuelle), il peut être refait selon la procédure du « remaniement » prévue par la loi n° 78-645 du 18 juillet 1974.

Les plans cadastraux résultant d'opérations d'aménagements fonciers sont appelés « plans remembrés ».

L'identifiant d'une feuille de plan est de la forme : « DCCCCPPSSNN » où :

- « DD » est le numéro du département ;
- « CCC » le code INSEE de la commune ;
- « PPP » le préfixe de section. Par défaut ce préfixe est égale à « 000 » sauf dans les cas suivants :
- en cas d'absorption de commune, ce préfixe a pour valeur le code INSEE de la commune absorbée ;
- en cas de communes à arrondissements, ce préfixe contient le code de l'arrondissement (pour Paris de 101 à 120, pour Lyon de 381 à 389, pour Marseille de 201 à 216, dans le cas de la ville de Toulouse il s'agit du code de quartier prenant les valeurs de 801 à 846)
- « SS » est la désignation de la section « cadastrale » (en cas de lettre de section unique, la lettre de section est précédée du chiffre « 0 » par exemple « section 0A ») ;
- « NN » est le numéro de la feuille (« 01 » par défaut)

Fréquence de mise à jour : Annuelle

## QUALITÉ

Information non disponible

## CONTACT

Mission SIG  
France  
✉ [sig@agglo-orient.fr](mailto:sig@agglo-orient.fr)

## CONDITIONS D'ACCÈS ET D'UTILISATION

**Licence :** [Licence Utilisation Données Cadastreales](#)

## LIMITATIONS

**Directive :**

**Restriction :** Copyright

**Description :**

Pas de restriction d'accès public selon INSPIRE

le plan cadastral décrit les limites apparentes de la propriété. « En dépit des efforts et diligences mis en oeuvre pour en vérifier la fiabilité, le partenaire fournisseur n'est pas en mesure de garantir l'exactitude, la mise à jour, l'intégrité, l'exhaustivité des données et en particulier que les données sont exemptes d'erreurs, notamment de localisation, d'identification ou d'actualisation ou d'imprécisions. Les données ne sont pas fournies en vue d'une utilisation particulière et aucune garantie quant à leur aptitude à un usage particulier n'est apportée par le partenaire fournisseur. En conséquence, les utilisateurs utilisent les données sous leur responsabilité pleine et entière, à leurs risques et périls, sans recours possible contre le partenaire fournisseur dont la responsabilité ne saurait être engagée du fait d'un dommage résultant directement ou indirectement de l'utilisation de ces données. En particulier, il appartient aux utilisateurs d'apprécier, sous leur seule responsabilité : o l'opportunité d'utiliser les données ; o la compatibilité des fichiers avec leurs systèmes informatiques ; o l'adéquation des données à leurs besoins ; o qu'ils disposent de la compétence suffisante pour utiliser les données ; o l'opportunité d'utiliser la documentation ou les outils d'analyse fournis ou préconisés, en relation avec l'utilisation des données, le cas échéant. Le fournisseur partenaire n'est en aucune façon responsable des éléments extérieurs aux données et notamment des outils d'analyse, matériels, logiciels, réseaux..., utilisés pour consulter et/ou traiter les données, même s'il a préconisé ces éléments. L'utilisateur veille à vérifier que l'actualité des informations mises à disposition est compatible avec l'usage qu'il en fait. »

**Directive :**

**Restriction :** Autre

**Description :**

Usage libre sous réserve des mentions obligatoires sur tout document de diffusion : "Source : DGFIP"

**Directive :**

**Restriction :** Autre

**Description :**

Droit d'auteur / Droit moral (copyright): Restriction lié à l'exercice du droit moral

## RESSOURCES

-  [Cadastre graphique de Locmiquélic](#)

Télécharger en XML 

Télécharger en PDF 

[← fiche précédente - 4/20](#)

[fiche suivante - 6/20 →](#)

# DOCUMENT 7

« Portail Open Data – Orléans Métropole » (exemple de catalogue de données) – *outil Open Data Soft* – février 2020

## ORLÉANS METROPOLE DATA

ACCUEIL DÉMARCHÉ DONNÉES VISUALISATIONS OUTILS FORUM CONTACT

1 527 enregistrements

### Arrêts des bus et du tram du réseau TAO - Orléans Métropole

Aucun filtre actif

Rechercher...

Informations Tableau Carte Export API Commentaires (0)

Liste des arrêts des bus et du tram du réseau TAO géré par Kéolis Orléans Métropole pour le compte d'Orléans Métropole dans le cadre d'une délégation de service. Les données sont produites par Kéolis Orléans Métropole et complétées par Orléans Métropole. La dernière mise à jour date de septembre 2019.

Identifiant du jeu de données: tao\_arrets\_od

Téléchargements: 234

Thèmes: Transports, Déplacements

Mots clés: tramway, transports, keolis

Licence: Licence Ouverte v2.0 (Etalab)

Fuseau horaire: Europe/Paris

Modifié: 25 novembre 2019 15:25

Zone géographique:

Producteur: Kéolis Orléans Métropole

Suivre les mises à jour:  Suivre les mises à jour

En vous abonnant à ce jeu de données, vous pourrez recevoir des notifications par email de la part du producteur de ce jeu de données si des changements importants ont lieu.

Dernier traitement: 25 novembre 2019 15:25 (métadonnées)  
25 novembre 2019 15:25 (données)

#### Modèle de données

Cliquez pour révéler

**OBJECTID**

Aucune description n'est disponible pour ce champ.

Nom (identifiant)	objectid
Type	int
Exemple	1

**Identifiant de l'arrêt**

Aucune description n'est disponible pour ce champ.

Nom (identifiant)	cle_unique
Type	texte
Exemple	1012001

**Nom de l'arrêt**

Aucune description n'est disponible pour ce champ.

Nom (identifiant)	nom_gener1
Type	texte
Exemple	11 OCTOBRE

**Ligne(s) de passage**

Aucune description n'est disponible pour ce champ.

Nom (identifiant)	lignes_pas
Type	texte
Exemple	L02 / L03

**Directions**

Aucune description n'est disponible pour ce champ.

Nom (identifiant)	directions
Type	texte
Exemple	L02*SAINT-JEROME / L03*BEAUNEUF

**Commune**

Aucune description n'est disponible pour ce champ.

Nom (identifiant)	commune
Type	texte
Exemple	Saint-Jean-de-la-Ruelle

**Accessibilité du quai**

Aucune description n'est disponible pour ce champ.

Nom (identifiant)	access_abb
Type	texte
Exemple	oui

**Catégorie d'arrêt**

Aucune description n'est disponible pour ce champ.

Nom (identifiant)	categ_arret
Type	texte
Exemple	bus

**geo\_shape**

Aucune description n'est disponible pour ce champ.

Nom (identifiant)	geo_shape
Type	geo_shape
Exemple	{ "type": "Point", "coordinates": [ 1.682270000611401, 47.90963000020271 ] }

## geo\_point\_2d

Aucune description n'est disponible pour ce champ.

Nom (identifiant) geo\_point\_2d

Type geo\_point\_2d

Exemple [47.9856300002,1.8227800061]

## Schéma JSON

L'objet JSON ci dessous est une description standardisée du modèle de votre jeu de données. Pour en savoir plus sur JSON schema.

```
{
  "title": "tao_arrets_od",
  "type": "object",
  "oneOf": [
    {
      "$ref": "#/definitions/tao_arrets_od"
    }
  ],
  "definitions": {
    "tao_arrets_od": {
      "properties": {
        "records": {
          "type": "array",
          "items": {
            "$ref": "#/definitions/tao_arrets_od_records"
          }
        }
      }
    }
  }
}
```

## Exemple de Back office

STANDARD DCAT INSPIRE PIÈCES JOINTES

### Informations générales

Identifiant

mobilités-stationnement-des-parlons-en-tam-0-s-ree

Editer

Description

Rich text editor with toolbar (bold, italic, link, etc.) and a large text area.

Thèmes

Transports, Déplacements

Mots clés

stationnement parlons places

Afficher la date de modification

- Mettre à jour la date lors d'un changement dans les métadonnées
- Mettre à jour la date après un traitement des données
- Mettre à jour manuellement

09/05/2018 19:04:30

### Informations régionales

Langue

French

#### Fuseau horaire

Si spécifié, force l'utilisation de ce fuseau horaire pour l'affichage des dates et heures

#### Couverture géographique

**Aucune** Le contenu de ce jeu de données n'est pas associé à un territoire.

**Automatique** La couverture géographique sera calculée automatiquement à partir du contenu de ce jeu de données.

**Manuellement** Delinez une zone sur la carte

## Licence & attribution

#### Licence

La liste de licences peut être éditée dans la section configuration mentions légales

#### Producteur

Personne ou organisation qui publie les données

#### Référence

URL de la source des données

#### Attributions

Source des données à mentionner pour des raisons légales

STANDARD **DCAT** INSPIRE PIÈCES JOINTES

#### Créé

#### Publié

#### Créateur

#### Contributeur

#### Nom du contact

#### Email du contact

#### Fréquence de mise à jour

#### Emplacement

#### Période de temps

#### Granularité

#### Qualité des données

#### Type de producteur

Thème

Type

Identificateur de ressource unique

Nom du contact

Poste du contact

Adresse du contact

Email du contact

Objectif

</>  **B** *I* ~~S~~      -

Objectif

</>  **B** *I* ~~S~~      -

Description de la zone géographique

Longitude Ouest

Longitude Est

Latitude Sud

Latitude Nord

## ANNEXE 1

« L'organisation de la collectivité INGEAGGLO » – INGEAGGLO – 2020

### L'organisation de la collectivité INGEAGGLO

Le service « information géographique » de la Communauté d'Agglomération INGEAGGLO est mutualisé avec la ville centre INGEVILLE. Il est rattaché directement à la direction générale des services. La communauté d'agglomération dispose de la compétence SIG. Le service Information géographique a en charge l'information géographique, l'observatoire du territoire. Il participe activement à la mise en œuvre du projet Open Data depuis plus d'un an en collaboration avec la Direction des Systèmes d'information et la Direction de la Communication. Il apporte son savoir-faire notamment en matière d'organisation, de structuration, de qualité des données, de diffusion des données géographiques (les premières données à avoir été libérées) et en matière d'animation sur la culture de la donnée avec la Direction de la Communication.

Il est composé :

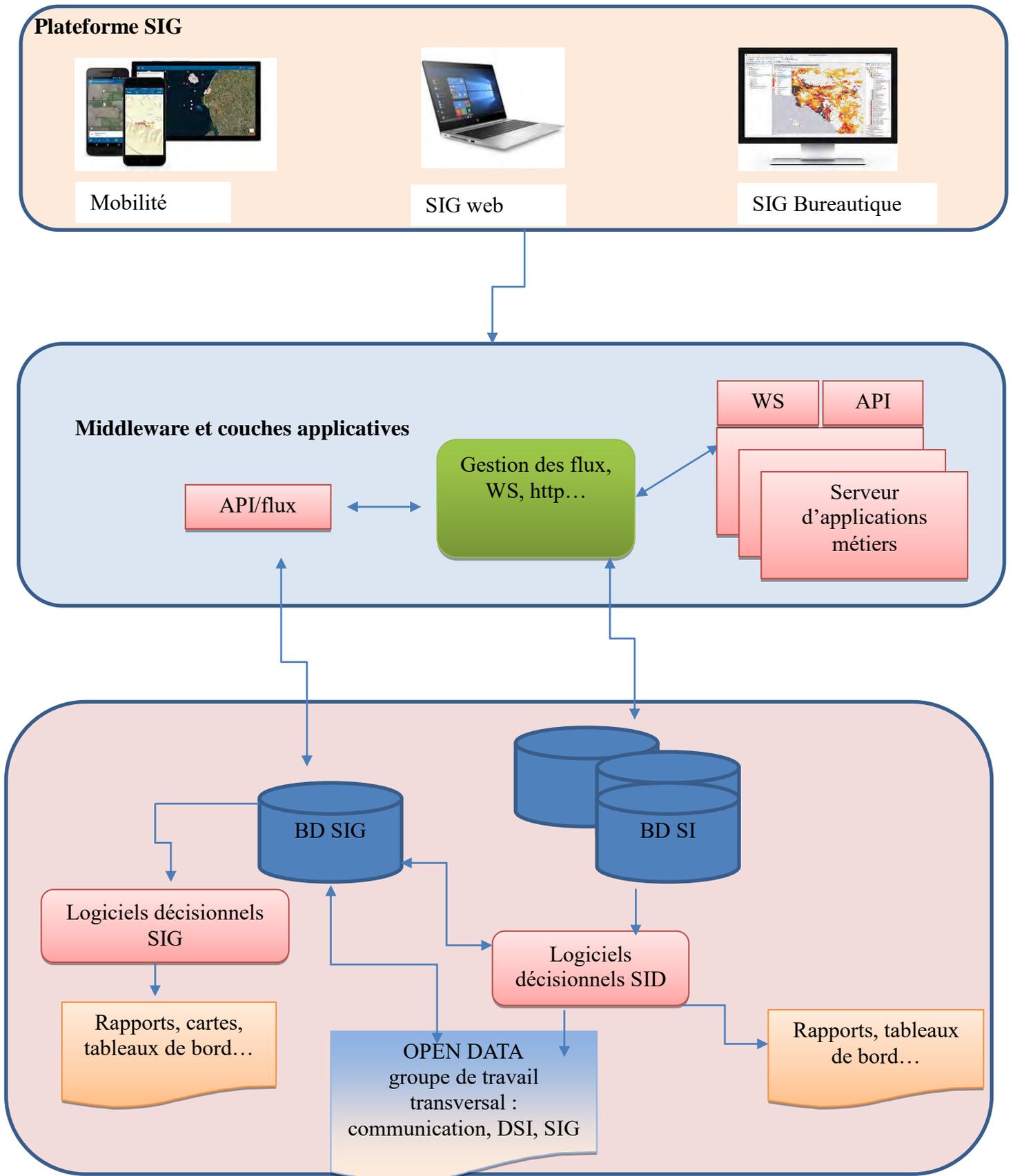
- d'un responsable de service
- de 2 chefs de projets géomatique et Open Data
- de 4 géomaticiens
- d'un topographe en charge des marchés de levés topographiques et recolement, PCRS
- d'un statisticien en charge de l'observatoire de l'Agglomération, des tableaux de bord 1
- d'un chargé de projet Open data (coordination, animation)

La Direction des Systèmes d'information gère l'infrastructure et les outils. Elle dispose d'un responsable des bases de données de la collectivité. Le service « information géographique » gère de manière autonome le SIG. Les serveurs sont mis à disposition par la Direction des Systèmes d'information et les installations et développements nécessaires sont confiés au prestataire du SIG.

Le service « information géographique » gère l'ensemble des référentiels géographiques sur son territoire : orthophotographie aérienne à 5 cm de précision mutualisée départementale, données topographiques de haute précision, cadastre, adresses, filaire de voirie, équipements publics, données d'intérêt commun (PLU, réseaux...). Des partenariats divers ont été mis en place au fil des années ( syndicats d'eau, d'énergie, gaz, plateforme régionale de l'information géographique INFOREG...). Un réseau de référents SIG et Observatoire est opérationnel depuis plusieurs années.

## ANNEXE 2

L'information géographique et l'open data d'INGEAGGLO – INGEAGGLO – 2020



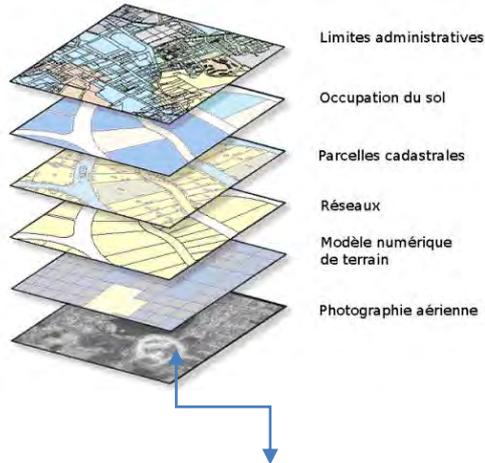
Partenaires internes

Directions transversales :  
DSI  
Communication

Directions métiers :  
Mobilité/transport  
Eau/assainissement  
Enviornnement  
Propreté  
Espace public

Réseau de référents par direction

**SIG INGEAGGLO**



Partenaires externes

Partenaires institutionnels :  
IGN  
Plateforme géographique régionale DGFIP

Autres partenaires :  
Gestionnaires de réseaux