

LE BIG DATA TERRITORIAL

UN ENJEU D'AVENIR POUR LES COLLECTIVITÉS



La transition numérique et son corollaire urbain que constitue la ville intelligente interpellent les acteurs en charge de l'organisation et du fonctionnement des territoires. Les mutations accélérées auxquelles nous sommes confrontés supposent une évolution rapide de la philosophie de l'action des collectivités, de leurs missions et des métiers qui s'y exercent.

Parmi les conséquences de cette transition, le « big data » est à la fois l'une des plus prometteuses et l'une de celles qui posent le plus question. Les collectivités locales et l'ensemble des acteurs publics vont devoir faire face à de multiples enjeux et difficultés pour s'adapter puis tirer le meilleur profit d'un phénomène désormais aussi massif qu'inéluctable.

QU'EST-CE QUE LE « BIG DATA » ?

Le « big data » correspond au principe de stockage d'un très grand nombre d'informations sur une base numérique. Sa définition peut néanmoins varier selon que l'on se place du côté de l'utilisateur ou du fournisseur de services. Littéralement on peut le définir comme une « mégadonnée », une grosse donnée ou encore une donnée massive.

Il désigne un ensemble très volumineux de données qu'aucun outil classique de gestion de base de données ou de gestion de l'information ne peut traiter. Il se présente comme une solution pour permettre à tous d'accéder en temps réel à des bases de données géantes.

Le « big data » est un phénomène qui voit croître à vitesse exponentielle le volume des données disponibles, mouvement accompagné par des capacités de stockage et de traitement de plus en plus importantes. Certains caractérisent ce phénomène par la règle des « trois V » : un Volume énorme de données, une grande Variété et une grande Vitesse. Les informations et outils qui étaient auparavant disponibles uniquement par les géants du web ou par les grands laboratoires sont désormais accessibles à tous ; entreprises ou collectivités locales.

Ainsi, et dès aujourd'hui, le défi des données s'avère crucial. Si certaines métropoles s'en sont déjà saisi, beaucoup de collectivités, et en particulier les plus petites, peinent à maîtriser ne seraient-ce que les concepts. Il faut dire que ceux-ci ne sont pas immédiatement accessibles : open data (données ouvertes), big data (mégadonnées), data lake (lac de données = espace de stockage de données non traitées), CDO - Chief Data Officer (AGD : Administrateur Général de données), smart data (données pertinentes), smart city (ville intelligente) ...

Explosion du volume et de la diversité des données

1880

Les débuts de la surabondance d'information

- 8 ans seront nécessaires pour publier les données de recensement des États-Unis

1941

« Explosion d'information »

- 1^{ère} apparition de l'expression « explosion d'information » dans le journal The Lawton Constitution
- 1^{ère} tentative pour quantifier l'évolution du volume des données

1997

« Le problème du Big Data »

- 1^{ère} apparition de l'expression « Big Data » dans un article scientifique écrit par M. Cox et D. Ellsworth, chercheurs à la NASA
- L'augmentation du volume des données devient problématique pour les systèmes informatiques de l'époque

2001

Les 3 V du Big Data

- Les 3 V du Big Data pour Doug Laney de Gartner :
 - Volume
 - Vitesse
 - Variété

CHIFFRES-CLÉS EN FRANCE :

- **90 %** du stock existant de données a été produit au cours des deux dernières années.
- En 2013 seules **22%** des données numériques étaient potentiellement exploitables.
- Le marché du big data connaîtrait actuellement une croissance de **40%** par an et serait estimé à **652 millions d'euros** en 2018.
- En 2016 seules **300** communes, soit à peine **10 %** des **3 000** communes de plus de **3500** habitants, avaient ouvert leurs données.
- Selon une enquête menée en 2016, **60 %** des collectivités ont déclaré que la question du Big Data n'était pas prioritaire.

Source : IDC France

DES DONNÉES POUR LA GESTION DES TERRITOIRES

Dans cette masse de données disponibles, on ne retrouve pas uniquement des données personnelles collectées pour de la vente en ligne et du ciblage publicitaire, il existe des données massives d'intérêt général, et même d'intérêt local. Elles constituent incontestablement le carburant indispensable à la mise en œuvre du concept de la « smart city » (ville intelligente).

Le concept de « Smart city »

Il se définit par trois axes complémentaires :

- le transport et la mobilité intelligente,
- l'environnement durable,
- l'urbanisation responsable et l'habitat intelligent.

Une « smart city » est ainsi un espace urbain développé qui génère un développement économique durable, couplé à une haute qualité de vie, en excellant dans de multiples secteurs-clés : économie, mobilité, environnement, cohésion sociale et gouvernance.

Elle utilise le numérique et les technologies de l'information et de la communication pour accroître la qualité et la performance des services urbains, afin de réduire les coûts et préserver les ressources, et fait participer ses citoyens de manière plus active.

Correctement traitées par des algorithmes dédiés (processus mathématique qui vise à obtenir un résultat, par une suite logique d'opérations), ces données sont de nature à permettre l'optimisation de la gestion du nombre de services rendus par la collectivité : trafic routier, transports publics, éclairage, chauffage urbain, eau, arrosage, ...

Au-delà d'une gestion optimisée, le big data permet de disposer d'une connaissance fine des comportements de nos concitoyens et comparer celui de chacun avec des profils types déterminés par d'autres algorithmes. Le big data, comme le montre les utilisations commerciales, devient prédictif et permet d'anticiper sur les comportements.

Le champ d'utilisations potentielles devient alors immense :

- prévoir les risques sanitaires,
- l'accidentologie,
- atteindre, après les avoir identifiés, les bénéficiaires d'une politique publique,
- se passer d'études coûteuses pour identifier les besoins et attentes des citoyens, ...

Au final, les données couplées à de puissants outils informatiques permettent une connaissance multi-sectorielle et en temps réel d'un territoire donné. L'accès à une telle connaissance est de nature à bouleverser totalement le mode actuel d'organisation des collectivités territoriales.



Source : <https://smartcity.brussels/le-projet-definition>

OPEN DATA ET AMPLIFICATION DU BIG DATA

Le concept d'open data désigne une donnée numérique dont l'accès et l'usage sont laissés libres aux usagers constituant un véritable enjeu pour les politiques publiques.

Ces données massives sont de plus en plus disponibles à la faveur d'un mouvement mondial d'ouverture des données (open data), aujourd'hui inscrit dans les textes législatifs :

- Directive « Inspire » de l'Union Européenne,
- Loi « NOTRe » du 7 août 2015,
- Loi « Valter » du 28 décembre 2015,
- Loi « Lemaire » du 7 octobre 2016 pour une République Numérique (ouverture des données des collectivités de plus de 3 500 habitants avant l'automne 2018),
- Règlement Général européen sur la Protection des Données (RGPD).

Le législateur s'est donc engagé ces dernières années, dans une politique volontariste d'ouverture des données publiques ou « politique d'open data » allant jusqu'à consacrer dans le cadre de la loi République Numérique, l'obligation pour les administrations de mettre à disposition les données qu'elles détiennent.

L'article 106 de la loi NOTRe, intitulé « Transparence des données des collectivités territoriales » précise que :

« les collectivités territoriales de plus de 3 500 habitants ainsi que les EPCI à fiscalité propre auxquels elles appartiennent rendent accessibles en ligne les informations publiques mentionnées à l'article 10 de la loi n°78-753 du 17 juillet 1978 lorsque ces informations se rapportent à leur territoire et sont disponibles sous forme électronique. Ces informations publiques sont offertes à la réutilisation dans les conditions prévues au chapitre II du titre 1er de la même loi. »

QU'EST-CE QUE LA DONNÉE PUBLIQUE ?

Sauf exception prévue par le législateur, la quasi-totalité des données détenues par les administrations devront être mises à la disposition du public.

Demeurent non communicables, les documents administratifs dont la consultation ou la communication portent atteinte :

- A la sécurité publique soit en particulier, au secret des délibérations du Gouvernement, au secret de la défense nationale, à la sûreté de l'Etat et tous autres secrets protégés par la loi ;
- Aux personnes soit en particulier, à la protection de la vie privée, au secret médical ;
- Au secret en matière commerciale et industrielle.

Les nouvelles dispositions législatives obligent l'administration à mettre à disposition du public les données personnelles qu'elle a en sa possession, après qu'elles aient fait l'objet d'un traitement permettant de rendre impossible l'identification de ces personnes.

Au-delà de l'enjeu que constitue pour les collectivités concernées l'ouverture de leurs données publiques avant octobre 2018, il convient également de se rendre compte de la masse de données sur les territoires, notamment issues de l'Etat, qui sont désormais disponibles : INSEE, IGN, entreprises délégataires de services publics, Chambres consulaires, ...

De nombreuses innovations technologiques vont conduire les territoires à se couvrir de multiples capteurs pour surveiller et mesurer en temps réel les infrastructures et les flux. Ces capteurs vont eux-mêmes générer un jet ininterrompu de quantités de données géolocalisées.

Les collectivités doivent donc impérativement se pencher sur l'organisation de la gouvernance de cette masse colossale de données numériques pour assurer une gestion publique du bien commun territorial et éviter une privatisation de la gestion de la ville et de la donnée.

LES ENJEUX DU BIG DATA TERRITORIAL

Se posera très vite avec acuité la question du statut et de l'organisation du ou des « tiers de confiance » qui aura(ont) en charge la gestion et le traitement de ces données.

À très court terme, les collectivités concernées par la Loi Le-maire vont devoir rendre accessible leurs données publiques relatives à leur territoire recouvertes par la notion de « données d'intérêt général » introduite par la loi. Mais à ce jour, seule une soixantaine de collectivités (dont Rennes, pionnière en France) disposent d'un portail de données numériques et manifestent une maturité en la matière.

Au-delà des données, la transparence des algorithmes utilisés par les collectivités devra être garantie.

Aujourd'hui, force est de constater que de nombreuses collectivités sont démunies face aux échéances et à la technicité que requièrent ces sujets. Il existe donc à la fois un enjeu de mutualisation et d'apport d'expertise pour franchir cette étape et de tirer bénéfice des apports potentiels du big data territorial.

Cette phase ouvre d'ores et déjà des sujets juridiques complexes en termes de responsabilité des collectivités dans la gestion des données publiques : collecte, accès, traitement, réutilisation, archivage, partage ... Dès lors qu'il y a utilisation massive de données individuelles, se pose la question de l'encadrement juridique, et notamment du rôle de la Commission Nationale de l'Informatique et des Libertés (CNIL), mais également des questions de protection des données et donc de cyberdéfense.



Les principaux enjeux pour les collectivités territoriales :

Enjeu n°1 :

Accès
aux données



Vers la généralisation massive de la diffusion et de la mise à disposition de données publiques collectées.

Enjeu n°2 :

Collecte
des données



Difficultés pour les collectivités d'obtenir auprès des partenaires privés des informations nécessaires au fonctionnement des services publics.

Enjeu n°3 :

Traitement, réutilisation,
archivage



Incertitudes liées aux conditions de réutilisation des données publiques, notamment dans le cadre de licence de réutilisation ou encore au stade de l'archivage des données.

LA QUESTION CRUCIALE DE LA SÉCURITÉ DES DONNÉES

Il semble que les régimes de protection des données soient peu aptes à faire face aux défis posés par cette société numérisée.

Cette question de la cyberdéfense est au cœur des enjeux d'intelligence économique. Elle renvoie à la notion de souveraineté numérique qui constitue une question géopolitique majeure.

En effet, l'arrivée en Europe de la RGPD se heurte à la promulgation du « Cloud Act » aux Etats Unis. Ce dernier texte autorise l'Administration américaine à accéder aux données détenues à l'étranger par une entreprise américaine. Dès lors que les principaux acteurs de la sauvegarde et du stockage en ligne sont américains, il y a là un enjeu crucial de sécurité.

Pourtant, les solutions offertes notamment par Amazon, Microsoft ou Google sont séduisantes, souples et peu coûteuses. Elles peuvent tenter nombre de collectivités, rassurées par le fait que les data centers sont désormais en France. Ces sociétés, et notamment Amazon, conseillent aujourd'hui de crypter les données pour plus de sûreté.

Mais si le cryptage constitue une solution envisageable, la sécurité ne serait pas infaillible, en particulier face aux services de renseignement d'autres pays, et en particulier des USA.

Il y a donc lieu de s'interroger sur l'intérêt qu'il y aurait à développer en France, sous législation française, des dispositifs de stockage et de sauvegarde des données publiques. Il en va probablement de la souveraineté du Pays, mais également de la protection de ses citoyens.

Par ailleurs, pour une collectivité, la mise en place d'un unique entrepôt de données semble être la solution optimale pour gérer au mieux les informations et de croiser plus facilement les données.

Cependant, dès lors qu'il est nécessaire de faciliter l'accès à ces informations aux services de la collectivité mais aussi au grand public, cette unicité de site pose un réel problème de sécurité et de sauvegarde.

Les données sont désormais le terrain à partir duquel vont pousser de nouvelles activités économiques. Il convient dès aujourd'hui de choisir à qui appartient le terrain.



L'Agence Nationale de Sécurité des Systèmes d'Information (ANSSI) :

Autorité nationale en matière de sécurité et de défense des systèmes d'information qui assiste les administrations et les opérateurs d'importance vitale. Elle est chargée notamment de la promotion des technologies, des systèmes et des savoir-faire nationaux. Elle contribue au développement de la confiance dans le numérique.

Le rôle de l'ANSSI est de faciliter une prise en compte coordonnée, ambitieuse et volontariste des questions de cybersécurité en France.

www.ssi.gouv.fr

L'augmentation phénoménale des capacités de stockage et la puissance démultipliée des outils informatiques rendent possible un rêve fou : comprendre le monde tel qu'il est, en l'observant de façon immédiate sous toutes ses coutures plutôt qu'en formulant des hypothèses que l'on teste sur des échantillons par de longues et coûteuses démarches.

Si la loi pour une République numérique a permis de mettre le sujet des données à l'agenda des collectivités, un long chemin reste à faire pour tirer parti du « big data » dans le cadre de l'action publique et pour coordonner les initiatives des différents acteurs publics et privés.

De plus, les collectivités peinent à atteindre un niveau minimal de protection de leurs données, et sont, de fait, mal préparées aux futures révolutions induites par l'ouverture d'un portail de données, ainsi que par la ville connectée.

- Fédération nationale des collectivités concédantes et régies (FNCCR) : www.fnccr.asso.fr/missions
- *La ville connectée : La « Smart city » relie les acteurs entre eux - QSN n°62, avril 2014*
- Site IDC France : <https://idc.fr/>
- Livre blanc Défense et sécurité nationale (2013) : www.livreblancdefenseetsecurite.gouv.fr/pdf/le_livre_blanc_de_la_defense_2013.pdf

Directeur de la publication : Patrice DUNY
Réalisation et mise en page : AUCAME 2018
Illustrations : AUCAME, sauf mention contraire
Contact : paul.andre@aucame.fr

DÉPÔT LÉGAL : 3^e TRIMESTRE 2018
ISSN : 1964-5155



Agence d'urbanisme de Caen Normandie
21 rue de la Miséricorde - 14000 CAEN
Tel : 02 31 86 94 00
contact@aucame.fr
www.aucame.fr



LICENCE OUVERTE
OPEN LICENCE



Retrouvez nos publications en flashnt ce QR Code