



**HAL**  
open science

## Le libre accès rebat-il les cartes ?

Matthieu Noucher, Pierre Gautreau

► **To cite this version:**

Matthieu Noucher, Pierre Gautreau. Le libre accès rebat-il les cartes ?. Les Cahiers du numérique, 2013, 9 (1), p.57-83. 10.3166/LCN.9.1.57-83 . halshs-00820007

**HAL Id: halshs-00820007**

**<https://shs.hal.science/halshs-00820007v1>**

Submitted on 24 Jun 2013

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

**[VERSION AUTEUR AVANT MISE EN PAGE DE L'EDITEUR]**

**NOUCHER Matthieu et GAUTREAU Pierre (2013). Le libre accès rebat-il les cartes ? De nouvelles perspectives pour les données géographiques. *Les Cahier du Numérique*, vol. 9, 1/2013, pp. 57-83.**

# LE LIBRE ACCES REBAT-T-IL LES CARTES ?

*De nouvelles perspectives pour les données  
géographiques.*

MATTHIEU NOUCHER, PIERRE GAUTREAU

Des premières banques de données urbaines des années 70 aux portails *open data* qui fleurissent aujourd'hui, ces quarante dernières années ont donné lieu à des recompositions importantes tant au niveau des modes de production que de diffusion des données géographiques publiques. Aujourd'hui partiellement en libre accès, elles offrent un observatoire privilégié des logiques actuelles de constitutions des savoirs sur les territoires. Cet article vise à définir les concepts émergents (*crowdsourcing, open data*) qui sont associés à ces évolutions récentes et à les replacer dans le contexte de ces quarante dernières années. Cette remise en perspective des données géographiques publiques, appuyée par l'exemple des données de la biodiversité, offre ainsi une entrée par le territoire des questions de libre circulation des données publiques et permet de réinterroger leur potentiel collaboratif.

## Introduction

Les bases de données géographiques constituent un objet particulier d'étude sur les usages de l'information publique. Formées de couches d'information décrivant des objets localisés à la surface de la terre (réseaux routiers, bâti, végétation, etc.) par leur géométrie et certaines de leur propriétés (taille, fonction, etc.), les données localisées sont manipulées dans des systèmes d'information géographique<sup>1</sup>. Des premières banques de données urbaines des années 70 aux portails open data qui fleurissent aujourd'hui, ces quarante dernières années ont donné lieu à des recompositions tant au niveau des modes de production que de diffusion des données géographiques publiques.

D'une part, les données géolocalisées sont aujourd'hui les données les plus largement mises en libre accès – *open data* – par les collectivités territoriales. D'autre part, les convergences technologiques ont conduit à une expansion des usages : l'externalisation de la production par la foule – *crowdsourcing* – impacte directement le domaine de l'information géographique avec l'émergence de communautés d'internautes-cartographes et de citoyens-capteurs (Goodchild, 2007). Ces évolutions, dont nous verrons qu'elles sont à la fois institutionnelles, techniques et d'usages tendent aujourd'hui à « brouiller les cartes » en ce sens qu'il devient plus difficile de distinguer données géographiques publiques et privées. Les enjeux sont alors de savoir ce que changent effectivement ces processus. En ce qui concerne les administrations, la légitimité des autorités publiques peut-elle être remise en cause par cette multiplication des utilisateurs-réutilisateurs ? Quant aux politiques publiques, on peut se demander si l'instauration de vastes chantiers de standardisation des modes de production et de diffusion des données publiques déclenchés par l'obligation de les diffuser ne conduit-elle pas à une perte de leur richesse sémantique ?

Aussi, en mobilisant les résultats d'une enquête auprès d'experts francophones et en nous appuyant sur l'exemple des données de la biodiversité, nous analysons en quoi le libre accès rebat les cartes de la production et de la diffusion des données géographiques publiques. Pour commencer nous pointerons les différences de moins en moins nettes entre données géographiques publiques et données géographiques privées à l'heure du *crowdsourcing* et de l'*open data*. Ces changements seront remis en perspective au regard des évolutions des quatre dernières décennies. Ces dernières nous permettront alors d'identifier les débats et controverses qui émergent et

---

<sup>1</sup> Un SIG peut se définir comme système informatique de matériels, de logiciels et de processus conçus pour permettre la collecte, la gestion, la manipulation (traitement, analyse) et l'affichage de données à référence spatiale.

d'envisager quelques perspectives de recherche pour une analyse des usages des données géographiques diffusées.

## **Peut-on définir une donnée géographique « publique » ?**

### *Les nouveaux modes de production des données géographiques.*

Ces quarante dernières années, la production et la diffusion des données géographiques publiques ont connu d'importantes transformations. Ces dernières se sont accélérées avec les évolutions récentes du cadre légal à l'instar de la Directive INSPIRE<sup>2</sup> et la montée en puissance des initiatives communautaires comme *Open Street Map*<sup>3</sup>, qui conduisent aujourd'hui à une reconfiguration majeure du paysage géomatique<sup>4</sup> institutionnel. L'analyse de ces évolutions doit permettre d'identifier quelques enjeux importants pour l'information publique en général. L'une des questions clés est, notamment, d'articuler les approches expertes des autorités publiques avec les représentations citoyennes qui émergent des nouvelles pratiques. Pour ce faire, nous proposons d'analyser le double mouvement (des institutions vers les citoyens à travers l'*open data* et des citoyens vers les institutions à travers le *crowdsourcing*) à partir d'une enquête réalisée au premier trimestre 2012, auprès d'un panel d'experts francophones. L'objectif de cette enquête était de définir les concepts associés à l'*open data* et au *crowdsourcing* qui ont la particularité d'être deux mots valises (Flichy, 1995) anglais souvent non traduits. Par ailleurs, il s'agit également d'évaluer les liens entre ces deux dynamiques.

Du 30 janvier au 9 mars 2012, 32 experts francophones ont été sollicités pour participer à une enquête par Internet. En provenance de Suisse, du Canada et de France, ces experts étaient issus du monde de l'information géographique ou de celui de l'*open data*, la combinaison des deux profils n'étant pas exclue. D'un point de vue méthodologique, l'enquête de type

---

<sup>2</sup> Au niveau européen, la Directive du 14 mars 2007 dite *Directive INSPIRE* vise à « fixer les règles générales destinées à établir l'infrastructure d'information géographique dans la Communauté européenne, aux fins des politiques environnementales communautaires et des politiques ou des activités de la Communauté susceptibles d'avoir une incidence sur l'environnement » (article 1). Cette directive oblige les administrations et collectivités de tous les pays de l'Union Européenne à mettre à disposition en consultation leurs métadonnées et les données utilisées ayant un caractère environnemental.

<sup>3</sup> <http://www.openstreetmap.org/>

<sup>4</sup> La géomatique regroupe l'ensemble des outils et méthodes permettant d'acquérir, de représenter, d'analyser et d'intégrer des données géographiques.

DELPHI<sup>5</sup> a mobilisé des questionnements construits via les techniques de cadrage / recadrage<sup>6</sup>.

Les trois tours d'enquête ont permis de confronter, d'affiner puis de finaliser les définitions relatives à ces deux concepts. Ainsi, l'open data est défini comme « *des données libres d'accès et de réutilisation, disponibles sans coût dans un format ouvert* ». Les experts insistent sur les possibilités d'exploitation des données, l'accessibilité ne suffit pas, il faut pouvoir les réutiliser (tant d'un point de vue technique – notion de format ouvert, qu'économique et juridique). Après débat, ils ont rejeté certaines précisions de participants du type « *données brutes* » (qui pourraient laisser croire que les données dérivées n'ont pas vocation à être diffusées, or les données des agences de statistiques nationales agrègent les données pour préserver la confidentialité), « *données publiées sur le web* » (jugée trop exclusive), « *données issues des autorités publiques* » (car le secteur privé peut être concerné). Par ailleurs, l'ensemble des experts s'est accordé à considérer qu'il existe une complémentarité forte entre le cadre légal en cours de constitution et le mouvement d'ouverture de données. Ainsi, la Directive INSPIRE et les démarches *Open Data* (OD) permettent de consolider ou d'initier des stratégies générales de diffusion des données. Elles ne concernent pas les mêmes cibles, ne s'inscrivent pas dans la même démarche (approche 'normalisante' pour INSPIRE, approche 'brute' pour l'OD) et n'ont pas la même temporalité mais *in fine*, elles favorisent la diffusion et la réutilisation de données publiques. Par ailleurs, elles peuvent s'alimenter l'une et l'autre.

---

<sup>5</sup> La méthode DELPHI a pour but de mettre en évidence des convergences d'opinion et de dégager certains consensus sur des sujets précis, grâce à l'interrogation d'experts, à l'aide de questionnaires successifs (Linstone et Turrof, 1975). Dans le cas présent, trois itérations ont été réalisées pour initier, affiner puis finaliser les réponses. Entre chaque tour, les résultats sont synthétisés et rendus anonymes puis transmis à nouveau aux experts pour valider et enrichir les positions moyennes ou contrer et argumenter pour faire valoir d'autres positionnements. Les 32 experts ont participé au moins à 2 tours et 27 d'entre eux ont participé à l'intégralité des trois tours.

<sup>6</sup> La formulation des questions par les techniques de cadrage-recadrage (Watzlawick, 1978) a permis d'enrichir les réponses apportées. Cette technique permet d'observer si les acteurs concernés cadrent la situation de la même façon. Cette approche mobilisée essentiellement pour des entretiens *de visu* a été testée pour la première fois, dans le cadre d'une enquête DELPHI. Il s'agissait de proposer aux experts des alternatives extrêmes (caricaturales) qui les obligent à se justifier. En mettant le répondant en position de justification, on l'oblige à externaliser et argumenter son positionnement.

Le *crowdsourcing* se définit, selon le panel d'experts, comme : « *un appel à la "masse" d'utilisateurs ou d'amateurs pour créer, diffuser ou améliorer un produit ou service. Il s'agit d'une démarche volontaire. Dans le cas des bases de données (géographiques), le crowdsourcing consiste à mettre à contribution des utilisateurs pour produire, mettre à jour et enrichir des données* ». Des précisions ont été apportées pour qualifier les dispositifs de contribution volontaire car les situations peuvent être très diverses. Ainsi, la moitié des participants considère qu'il est important de rappeler que les bénéficiaires peuvent être variés (de la société privée qui revend les données à la communauté locale qui redistribue gratuitement la contribution)<sup>7</sup>. La donnée n'est, en effet, plus créée uniquement par des spécialistes au sein d'une ou plusieurs organisations, mais elle est alimentée par un collectif potentiellement dispersé sur le territoire et incluant des profils divers. Ce collectif peut être un très grand groupe bien défini (par exemple, les membres des clubs d'ornithologie du pays), voire formé ou accrédité à cet effet. Enfin, 80% des participants jugent utiles de rappeler que le *crowdsourcing* existait bien avant le développement du Web 2.0<sup>8</sup>. Ces précisions permettent de nuancer les définitions parfois un peu rapides et caricaturales et de pointer toute la diversité voire la complexité des situations. L'enquête par itération n'est pas sans rapport avec cet affinage progressif.

Au terme des trois tours d'enquête, il apparaît que la question du lien entre la diffusion de l'information et la contribution volontaire des Internautes à la saisie ou l'actualisation des données ne peut se résumer à la confrontation d'une approche ascendante (*crowdsourcing*) et d'une approche descendante (*open data*). Ce schéma simpliste ne traduit en effet que partiellement la réalité du terrain et les observations des experts mettent en évidence la porosité des situations. Ainsi, l'ouverture des données n'est le pré carré ni des autorités publiques ni des communautés locales, comme en témoigne le cas des données naturalistes.

### *L'exemple des données géographiques naturalistes.*

L'analyse de la place des données géographiques publiques en accès libre dans la diffusion des savoirs naturalistes en France permet d'éclairer certains

---

<sup>7</sup> 70% des répondants considèrent qu'il est important de prendre en compte les facteurs incitatifs (qui eux-aussi peuvent être variés) ; 75% des répondants considèrent important de rajouter des précisions sur la variété des profils des utilisateurs.

<sup>8</sup> En 1999, suite au naufrage de l'Erika au large de la Bretagne, le suivi des nappes de pétrole (400 km de côtes souillées) est effectué à partir d'une cartographie issue des données fournies par les volontaires et qui se révèlent rapidement plus efficaces que les relevés aéroportés pour évaluer, voire pour anticiper la pollution.

aspects de la façon dont ces données sont produites et de la fonction symbolique et « politique » qui leur est attribuée par les différents acteurs qui les manipulent. Les données sur la biodiversité (faune, flore, paysages) présentent la particularité d'être massivement produites par des amateurs, dans le cadre d'associations naturalistes, depuis les années 1970-80. Les raisons qui poussent les différents acteurs de la sphère naturaliste à diffuser en ligne des données géographiques sur la biodiversité, et les différentes stratégies qu'ils adoptent permettent d'interroger la place des données publiques dans ce processus de formalisation des connaissances naturalistes. Ils illustrent, par ailleurs, les liens étroits entre accès libre (*open data*) et contribution volontaire (*crowdsourcing*) évoqués par les experts sollicités dans le cadre de l'enquête DELPHI.

Une manière d'interroger la place des données géographiques publiques dans la diffusion des savoirs naturalistes est d'analyser le réseau formé par les hyperliens entre les sites web fournissant des données de biodiversité en accès libre. Ce réseau est interprété comme l'expression de liens sociaux entre les membres publics et associatifs de la sphère naturaliste, et partant comme une représentation permettant d'approcher les stratégies de diffusion de données de ces différents acteurs<sup>9</sup>. Si l'on se focalise uniquement sur les données géographiques (*figure 1*), la prééminence des sites publics est clairement marquée. Ils constituent l'essentiel des sites offrant le plus grand nombre de formes d'accès à ces données, tandis que seul un petit groupe de sites associatifs fait jeu égal<sup>10</sup>. C'est donc pour l'essentiel au travers de sites publics que les nouveaux modes de diffusion des données géographiques sont développés.

Les différentes modalités de diffusion des données entre sites publics et sites associatifs ne doivent cependant pas masquer la difficulté à distinguer clairement l'origine ou la genèse de ces données. Une des bases de données géographiques les plus importantes diffusées par un site public en France est la couche cartographique des ZNIEFF<sup>11</sup>, sur le site du *Muséum National d'Histoire Naturelle* qui valide et met à jour ce document. Celui-ci repose pour l'essentiel sur le travail d'associations mobilisées dans un cadre régional sous la direction

---

<sup>9</sup> La pertinence de l'analyse de réseaux web pour la reconnaissance de relations sociales au sein de communautés partageant des intérêts et des valeurs communes est reconnue par de nombreux auteurs (Kleinberg et al., 2001). La formalisation de ce réseau a été obtenue par collecte des liens entrants-sortants avec le logiciel *Navicrawler*, et leur spatialisation avec *Gephi*, grâce à l'algorithme Force Atlas 2 (Jacomy et al., 2012).

<sup>10</sup> Ils sont liés à deux associations : la *Ligue de Protection des Oiseaux* et *TelaBotanica*.

<sup>11</sup> Zone Naturelle d'Intérêt Ecologique Faunistique et Floristique.

des DREAL<sup>12</sup> pour la délimitation et le choix de ces zones au début des années 1980 (Coudерchet et al., 2011). Ce qui fait dire à ces auteurs que « toute la connaissance officielle de la nature en France est en définitive principalement fondée sur une construction d'amateurs » !

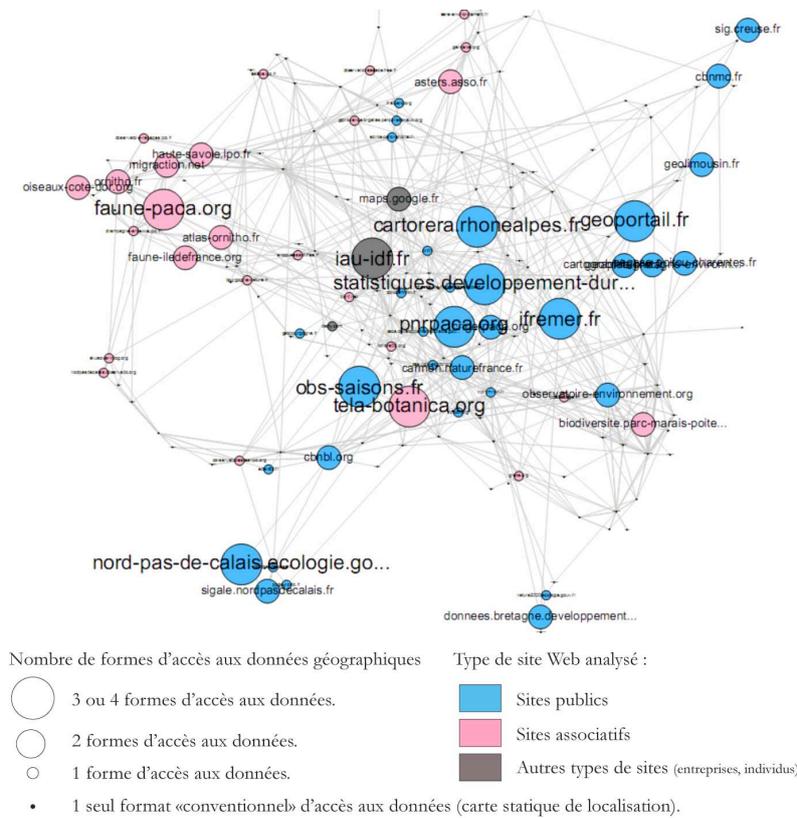


Figure 1. Sites internet offrant des données géographiques sur la biodiversité en France : formes d'accès aux données et types d'auteur<sup>13</sup>

<sup>12</sup> Directions Régionales de l'Environnement, de l'Aménagement et du Logement

<sup>13</sup> Source : inventaire représentatif de l'offre française de données en ligne sur la biodiversité (201 sites) constitué par D. Montagne dans le cadre de son master (2011).

La majeure partie des données sur la faune et la flore diffusées sous forme géographique (maillage avec fréquence d'espèces) par des sites publics provient d'associations. Inversement, nombre de données diffusées sur les sites associatifs, ou dans leurs documents d'analyse, reprennent des données des sites publics. Ce qui différencie donc fondamentalement données géographiques publiques et non publiques dans ce domaine, c'est la stratégie de diffusion dont elles font l'objet (Montagne 2011). Les institutions publiques les diffusent avant tout pour réduire le temps dédié à répondre aux demandes d'information des bureaux d'études ou de citoyens : avoir des données en téléchargement libre du temps en interne pour travailler à ses missions. Les obligations légales jouent aussi un rôle important. Dans le monde associatif, la diffusion de ce type de données joue d'abord le rôle interne de mobilisation des bénévoles, et externe d'attraction de nouveaux bénévoles : posséder un site où l'on peut voir et manipuler ce type de données est présenté fréquemment comme un élément de dynamisation de l'association par leurs dirigeants (Montagne, 2011).

Enfin, les associations utilisent la donnée géographique pour illustrer la richesse de leur base de données, dans une logique dont les enjeux symboliques ne sont pas absents, à l'heure où nombre de ces structures se professionnalisent et proposent de réaliser des tâches rémunérées similaires à celles des bureaux d'étude : laisser entrevoir sur Internet la richesse de sa base est un moyen d'obtenir des contrats. La prééminence des sites publics dans la diffusion des données géographiques sur la biodiversité exprime donc principalement une hiérarchie entre les acteurs, fondée sur des différentiels de capacité technique et financière, mais aussi sur le patient travail de normalisation imposé par les autorités publiques aux associations, souvent à leur corps défendant (Alphandéry et al., 2011).

La porosité entre données publiques et données privées, mise en évidence par l'*open data* et le *crowdsourcing*, tend donc à renforcer l'idée de frontières organisationnelles de plus en plus floues entre producteurs, utilisateurs et diffuseurs de données géographiques. Ce constat s'inscrit dans un historique d'une quarantaine d'années qui a vu se structurer la production d'information géographique puis progressivement se développer une diffusion d'abord intra- puis inter- et désormais extra- institutionnelle. Un retour sur cette évolution permet de remettre en perspective les récents changements provoqués par l'ouverture des données publiques.

## Quarante ans de diffusion de l'information géographique

### *Une première expansion à l'intérieur des organisations publiques*

#### *Les premières banques de données urbaines (années 70-80)*

Les systèmes d'information géographique sont longtemps restés une affaire de gestionnaires ou de passionnés. Les premiers utilisateurs de ce qui ne s'est pas immédiatement appelé « SIG » sont apparus à la fin des années 70 et au début des années 80. La fin des années 70 marque, ainsi, le développement du *dessin assisté par ordinateur* (DAO). Portées par cette évolution et la volonté d'automatiser les opérations de mise à jour, les premières *banques de données urbaines* (BDU) sont nées d'initiatives locales, quelquefois artisanales, souvent isolées (Roche, 2000). A l'exception des militaires qui constitue un cas particulier, les premiers à s'engager en France, furent essentiellement les services techniques des grandes villes<sup>14</sup> et les gros opérateurs de réseaux<sup>15</sup>, pour lesquels la charge financière induite par le passage au numérique pesait relativement peu par rapport aux coûts et enjeux globaux de leur activité. La voie de ce qui fut à l'origine perçu et conçu comme de la *cartographie assistée par ordinateur* (CAO) puis comme des *Systèmes d'Information Géographique* (SIG) était alors ouverte.

Cette origine *technicienne* a profondément marqué le développement des technologies de l'information géographique (TIG), non seulement en terme de domaines d'utilisation mais également dans l'ergonomie et dans la conception même des logiciels et des modèles de données. De même, l'appropriation des technologies SIG par les experts du domaine de la mesure (géomètres et informaticiens notamment) va d'emblée orienter le domaine et avoir un effet centralisateur : seuls quelques spécialistes réussissent réellement à s'approprier ces outils le plus souvent mis en œuvre dans des projets à objectifs fédérateurs. Les données produites alors sont des grands référentiels<sup>16</sup> où la précision de la

---

<sup>14</sup> La Communauté Urbaine de Bordeaux fait partie des collectivités territoriales qui, à la fin des années 1970, ont mis en œuvre les premières banques de données urbaines. Cette démarche a induit de nombreux changements dans le domaine de la cartographie. Les topographes se sont, par exemple, adaptés aux techniques SIG pour favoriser la diffusion interne des données géographiques.

<sup>15</sup> On peut citer EDF-GDF, par exemple.

<sup>16</sup> Un référentiel peut se définir comme un fonds de plan qui sert de base à la production d'autres données géographiques : canevas topographiques, plans cadastraux, photographies aériennes en sont de bons exemples.

localisation des éléments de l'espace est jugée plus importante que les informations descriptives sur ces mêmes éléments.

*Des référentiels géographiques aux données thématiques (années 90)*

Le début des années 90 coïncide avec l'arrivée de la micro-informatique, qui a, au sein des services de l'administration un impact décentralisateur en permettant aux utilisateurs de s'approprier des outils informatiques et de s'affranchir de la tutelle des informaticiens. En géomatique, c'est la génération des *SIG bureautiques* de type ArcView®, MapInfo®, GéoConcept® et autres souvent acquis de façon autonome par des services dans le cadre de projets à objectifs sectoriels. Les thématiciens<sup>17</sup> de tous horizons prennent ainsi progressivement la main sur les outils, élargissant sans cesse les champs d'intervention comme en témoigne la multiplication des néologismes qui apparaissent alors : géomarketing, géodécisionnel, géobusiness, archéomatique, géotraçabilité, géovisualisation, etc. Ce mouvement de diffusion des outils s'accompagne d'un déploiement des technologies de l'information géographique au-delà des organisations de grandes tailles.

Le secteur d'activité semble alors s'organiser principalement autour de la production, la représentation de l'information géographique, les développements informatiques mais aussi autour des thématiques de l'environnement et de l'aménagement (Mathian, 2003). En périphérie de ce noyau central historique, plusieurs autres domaines apparaissent dont la gestion des risques, l'agriculture, la foresterie, etc. tandis que des nouveaux émergent à peine : la santé, les télécommunications, l'archéologie...

Cette diffusion des technologies de l'information géographique se traduit par une augmentation exponentielle des données géographiques publiques produites. Au-delà des référentiels (levé topographique, plan photogrammétrique, etc.) qui peuplaient les premières banques de données urbaines, de plus en plus de données métier deviennent le socle informationnel de nombreux SIG thématiques<sup>18</sup>. La facilité de prise en main des outils accélère son adoption par des non spécialistes. Le déploiement des SIG se précipite

---

<sup>17</sup> Les thématiciens définissent des experts dans un domaine de l'action publique. Il est à noter que ce terme semble s'être créé en opposition au profil de « technicien » comme si la connaissance des « choses » (les « thèmes ») était indépendante des outils de formalisation de ces « choses ».

<sup>18</sup> Au sein de la Communauté Urbaine de Bordeaux (CUB) l'accès centralisé aux bases de données géographiques, est assuré par le logiciel SIG APIC® entre 1987 et 1995, puis complété par des applications métier complémentaires : assainissement, application du droit des sols, filaire de voies, consultation du PLU, etc.

alors dans les organisations sans soucis de mise en cohérence ou de fédération des outils et des données. La production cartographique individuelle est bien souvent privilégiée. Les banques de données urbaines centralisées font place à la multiplication des *couches cartographiques* produites et gérées localement. Ce que les utilisateurs gagnent en autonomie, l'organisation le perd en coordination. C'est ainsi, par exemple, que dans une ville comme Evry (50.000 habitants) on recensait début 2000, cinq logiciels SIG différents déployés dans les services<sup>19</sup>, jusqu'à quatre versions différentes du plan cadastral et une profusion de données thématiques. Cette diffusion parfois un peu brouillonne a cependant eu le mérite d'acculturer les thématiciens et parfois même les décideurs à l'information géographique.

### *Une mise en réseau entre organisations publiques*

#### *Les données géographiques à la frontière des organisations publiques*

L'émergence des technologies client/serveur va permettre une articulation entre les logiques de centralisation des débuts et les logiques d'autonomie des années 90. Les évolutions concernant les serveurs spatiaux, les outils de diffusion en environnement Internet, les architectures orientées services<sup>20</sup>, l'usage des SIG sur assistant personnel ou l'interopérabilité des logiciels aboutissent à une situation dans laquelle les outils n'imposent plus aux producteurs de données une organisation (centralisée/décentralisée) ou une logique (fédératrice/sectorielle), mais peuvent au contraire s'adapter à tous types d'objectifs et toutes formes d'organisation. Aussi, l'assemblage de référentiels géographiques et de données métier<sup>21</sup> est facilité. Ces évolutions

---

<sup>19</sup> Si la Direction des Services Techniques et la Direction de l'Urbanisme apparaissent comme les principaux producteurs et utilisateurs, d'autres directions avaient également intégrées dans leur pratique routinière les outils cartographiques. On peut ainsi citer les services en charge de la communication ou des affaires scolaires.

<sup>20</sup> L'objectif d'une architecture orientée services est de décomposer une fonctionnalité en un ensemble de fonctions basiques, appelées services. L'idée sous-jacente est de cesser de construire les systèmes d'information autour de logiciels fonctionnant en silos pour proposer des modules réutilisables dans divers contextes (implémenté la plupart du temps sous la forme de service web) et fonctionnant ainsi en réseaux.

<sup>21</sup> Les données métier concernent toutes les données produites par les utilisateurs dans la perspective de leur métier. Les thématiciens du territoire (urbaniste, conservateur du patrimoine, architecte, géologue, forestier mais aussi topographe et cartographe) partent bien souvent des référentiels géographiques auxquels ils ajoutent

technologiques permettent également aux SIG de se répandre dans des organisations de tailles inférieures<sup>22</sup>, ou encore de réduire progressivement les réticences des différentes institutions à partager leurs données dans le cadre de projets de mutualisation. Dans le domaine naturaliste par exemple, les technologies client/serveur sont un élément essentiel pour convaincre des institutions historiquement jalouses de leurs données de connecter leurs bases à des plateformes communes : les données restent déposées sur le serveur de l'institution, qui reste responsable de leur gestion et de leur mise à jour, tout en étant consultables et « moissonnables » par moteur de recherche interposé<sup>23</sup>.

La notion d'infrastructure de données géographiques (IDG) est l'une des illustrations les plus révélatrices de cette ouverture des données géographiques au-delà des frontières de leurs organisations d'origine. Classiquement, les IDG se définissent selon cinq éléments majeurs (Douglas, 2004) : des informations géoréférencées, des accords techniques et organisationnels, une documentation (métadonnées), des mécanismes pour découvrir, consulter les données et des méthodes permettant un accès aux données. Une IDG offre ainsi un ensemble de données sur une unité administrative définie (un pays, une région, un département) ou sur une thématique spécifique (l'eau, le littoral, la montagne). Ces dispositifs qui émergent dans les années 2000 s'inscrivent dans une logique institutionnelle favorisant l'organisation de la production (éviter les redondances) et de la diffusion (faciliter l'accès) des données géographiques. Elles s'orientent donc vers des principes à la fois stratégiques et économiques (politiques) et technologiques (normes) (Masser, 2005) et repose sur un cadre légal qui a été profondément renouvelé ces dernières années. Au niveau européen, la Directive INSPIRE s'inscrit pleinement dans ces objectifs. En France, ce sont essentiellement les régions qui assurent le relais entre les

---

une plus value issue de leur métier. Cette dernière peut se définir comme le reflet de l'histoire cognitive du producteur des données (Noucher, 2009).

<sup>22</sup> C'est ainsi que depuis le début des années 2000, le SIG de la CUB tend à s'ouvrir au-delà de la sphère des techniciens de l'intercommunalité. Ainsi, un Intranet donne aux communes membres un accès sécurisé à un SIG Web.

<sup>23</sup> Le consortium gérant le GBif, système d'information sur la biodiversité référençant près de 400 millions de données de présence d'espèces animales et végétales, rassure d'emblée les potentiels collaborateurs sur son site : « le GBIF offre un moteur de recherche portant sur des bases de données connectées au GBIF de manière standardisée. Les possesseurs de données peuvent connecter tout ou partie de leurs ressources au GBIF, afin de les rendre visibles et interoperables, mais restent maîtres de leurs données, qu'ils continuent à héberger et à utiliser dans le cadre de leur travail » (<http://www.gbif.fr>).

dispositifs nationaux<sup>24</sup> portés par le ministère de l'écologie, du développement durable et de l'énergie et les échelons locaux<sup>25</sup>. Ce mouvement de mise en réseau de données géographiques dépasse aujourd'hui la sphère professionnelle.

### *Vers un prolongement au-delà des organisations publiques.*

#### *L'émergence du GéoWeb ou le succès de la carte sur Internet.*

Le début des années 2000 marque le développement de véritables SIG en ligne<sup>26</sup>. Les interfaces cartographiques s'enrichissent progressivement et permettent non seulement de consulter des données géographiques mais aussi de naviguer sur les territoires (zoom, déplacement), de réaliser des requêtes (sur les attributs et/ou la géométrie des objets) et de mettre à jour à distance les données alphanumériques ou graphiques. A partir de 2005 et du lancement de *Google Maps*, on observe une véritable démultiplication des usages des cartes sur Internet. L'offre en outils cartographiques en ligne explose, avec de multiples services et l'apparition d'usages différenciés. Une organisation par l'espace de l'information sur Internet à travers un géo-référencement direct ou indirect sur la surface terrestre a fait émerger le *GéoWeb* (Fisher, 2008). Ce dernier qui permet une indexation spatiale des pages Web dans une perspective d'aide à la recherche pour les utilisateurs, se généralise à tous types de contenus (photos, vidéos, articles d'encyclopédie, dépêches d'agence de presse, etc.). La carte devient un support qui en plus de permettre de visualiser une information, la positionne dans l'espace (Joliveau, 2010).

Cette indexation spatiale massive et exponentielle de tout type d'information concerne aussi l'information publique. La directive *INSPIRE* se focalise sur les données environnementales qui par nature ne sont pas obligatoirement des

---

<sup>24</sup> Le site <http://www.geoportail.fr/> administré par l'IGN permet la consultation des données et le site <http://www.geocatalogue.fr/> administré par le BRGM permet la recherche et la consultation des métadonnées.

<sup>25</sup> Ainsi, la Communauté Urbaine de Bordeaux est membre de PIGMA (Plateforme d'Information Géographique Mutualisée d'Aquitaine) portée par la Préfecture de région et le Conseil régional et dont l'un des objectifs vise explicitement à répondre aux exigences de la Directive INSPIRE. <http://www.pigma.org/inspire>

<sup>26</sup> La CUB met à disposition des Internautes tout un panel d'applications en ligne dérivées de ses SIG métier : « Plan Local d'Urbanisme à la carte » pour la consultation du règlement d'urbanisme, « Thermographie : quelle est votre couleur ? » dont le but est de visualiser les déperditions énergétiques des toitures des habitations, « Bornes de tri » qui localise les points de collecte de tri sélectif ou encore la carte de « circulation » en temps réel qui permet d'accéder aux comptages routiers.

données géographiques. Pourtant, la directive et les spécifications qui en découlent, visent explicitement à favoriser le partage de données géographiques, renforçant *de facto* cette tendance à l'indexation de l'information par l'espace. Cette tendance forte et structurante (elle conditionne à la fois la façon dont vont se produire les données futures et la façon dont elles vont être diffusées) s'explique, en partie, par le succès du *GéoWeb* et de la carte sur Internet ; il est devenu commun de chercher des informations sur une base spatiale : quelles informations sont-elles disponibles à tel endroit ? Ces questionnements sont d'ailleurs renforcés par l'internet mobile : quelles informations/services sont-elles disponibles autour de moi ?

Dans ce contexte, les données géographiques publiques apparaissent comme singulières par rapport aux autres données publiques : elles sont à la fois des sources d'information mais également des « infrastructures » qui permettent de mettre en forme et de diffuser d'autres informations. Cette singularité explique la place centrale qu'elles occupent dans l'ouverture des données publiques.

#### *L'information géographique, vecteur d'ouverture des données publiques*

Si la diffusion des données (notamment environnementales via les sites des DIREN<sup>27</sup>) existent, en France, depuis de nombreuses années, le « *mouvement open data* » trouve un nouvel écho avec des initiatives emblématiques comme <http://data.gov> aux Etats-Unis et <http://data.gov.uk> en Grande Bretagne, dont l'enjeu est de rendre de grandes quantités de données gouvernementales disponibles aux citoyens. Aujourd'hui, partout dans le monde, les Etats et les collectivités locales<sup>28</sup> s'engagent dans l'« ouverture » de leurs données publiques. Dans ce modèle, les autorités publiques (essentiellement) produisent des données, des documents et des services ; ils les formatent selon des standards reconnus, leur donnent une adresse permanente et alertent sur la publication de nouveaux documents. A leur tour, les acteurs privés (commerciaux ou non commerciaux) exploitent ces informations pour produire des services : outils de recherche, de visualisation, *mashups*<sup>29</sup> etc. Cette stratégie mise donc sur la diffusion de données « ouvertes », c'est-à-dire de données brutes libres de réutilisation.

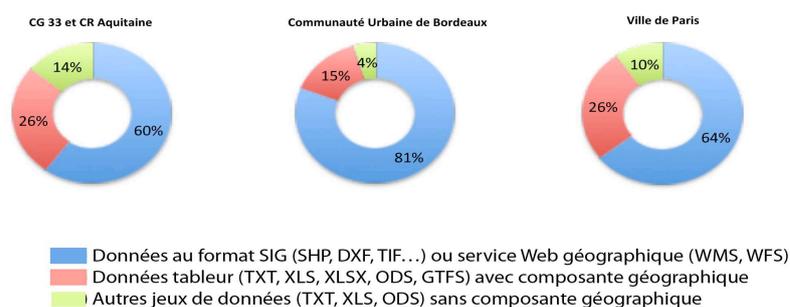
---

<sup>27</sup> Directions Régionales de l'Environnement qui sont devenues suite à la Réforme Générale des Politiques Publiques (RGPP) les Directions Régionales de l'Environnement, de l'Aménagement et du Logement (DREAL).

<sup>28</sup> En 2011, la Communauté Urbaine de Bordeaux a ouvert un portail expérimental <http://data.lacub.fr> qui met à disposition 68 jeux de données brutes.

<sup>29</sup> Un mashup (ou application composite) est une application qui combine du contenu ou du service provenant de plusieurs applications plus ou moins hétérogènes.

Une plateforme nationale <http://www.data.gouv.fr/> est complétée par les portails de plus en plus nombreux des collectivités territoriales : ville de Paris, agglomération de Rennes, Nantes, Bordeaux et Montpellier<sup>30</sup>, conseil général de Gironde et de Saône et Loire, conseil régional Aquitaine etc. Pour tous ces portails, l'information géographique occupe une place prépondérante : les données géographiques (données au format SIG ou tableur avec référence géographique) sont souvent majoritaires (fig. 2) et la maîtrise d'œuvre comme la maîtrise d'ouvrage de ces infrastructures associent souvent les directions de la communication, des systèmes d'information et de l'information géographique.



Sources :  
<http://data.lacub.fr> (consulté le 6/2/2012)  
<http://data.paris.fr> (consulté le 6/2/2012)  
<http://datalocale.fr> (consulté le 4/12/2012)

Figure 2. Part de la composante géographique dans les données libérées de trois institutions françaises (une ville, une intercommunalité, un département et une région)

Si une majorité des données qui circulent dans les systèmes d'information publique sont des données géolocalisables (*i.e.* des données ayant directement ou indirectement une composante géographique), on observe tout de même une sur-représentation de ce type de données dans les portails open data des collectivités françaises. Quatre éléments de contexte peuvent expliquer cette situation. Premièrement, les données géographiques peuvent être exploitées sous la forme de visuels (des cartes la plupart du temps) et, à ce titre, elles peuvent laisser croire qu'elles sont plus facilement appropriables par des non-spécialistes. Deuxièmement, depuis les premières banques de données urbaines des années 70-80, les données géographiques font l'objet d'un travail minutieux de structuration voire de normalisation. Des compétences nouvelles en matière d'administration de données sont apparues dans les organisations publiques

<sup>30</sup> NDLR : l'article de J.C. Plantin et J. Valentin évoque le cas de Montpellier.

pour les gérer, les contrôler, les documenter, les diffuser. Ainsi, bien souvent les données les plus facilement mobilisables lorsqu'une organisation souhaite ouvrir son patrimoine numérique et le rendre accessible sont les référentiels cartographiques, car ils sont structurés et disposent de métadonnées adéquates. On peut émettre une troisième hypothèse explicative de cette sur-représentation : les données géographiques sont parfois les données les moins sensibles... Il peut donc être plus simple de les mettre en libre accès contrairement aux données comptables, par exemple. Enfin, les données géographiques peuvent servir de support pour représenter d'autres données, elles sont à la fois sources d'informations et support pour représenter, croiser, combiner des jeux de données différents, elles ont donc un fort potentiel d'appropriation par l'utilisateur.

## Débats et controverses

Aboutissement d'une structuration lente de la production et d'une diffusion progressive des données géographiques, la mise en libre accès reste aujourd'hui un phénomène récent. Ne disposant que de peu de recul sur les nouvelles pratiques qui y sont associées, nous proposons ici de mettre en évidence les débats voire controverses qu'elles suscitent. Les questionnements et tensions sont nombreux : il s'agit notamment de savoir ce que changent ces processus à la fois pour l'administration publique et pour les politiques publiques. Pour ce faire, nous identifions six tensions et mobilisons les résultats de l'enquête DELPHI (dont on rappelle qu'elle n'avait pour vocation que de donner un aperçu des visions dominantes de l'*open data* et du *crowdsourcing* pour les données géographiques dans le monde francophone) illustrés d'exemples.

### *Tension n°1 : visibilité ou maturité ?*

Un des principaux points de débat concerne le processus de production/diffusion des données géographiques : sommes-nous dans une phase de maturité où les technologies, leurs usages, les pratiques institutionnelles associées commencent à se stabiliser ? Au contraire, sommes-nous encore dans une phase d'évolution de ces trois paramètres (technologies, usages, pratiques institutionnelles) ?

La première tension mise en évidence par l'enquête autour du couple « maturité / visibilité » interroge ici la stabilisation des usages de l'*open data* et du *crowdsourcing* pour la production et la diffusion des données géographiques publiques. La prise en compte du temps et des processus d'apprentissage est

essentielle comme le rappellent Francis Jauréguiberry et Serge Proulx (2011). Ainsi, pour affiner notre compréhension, chaque expert a été invité à positionner les deux dynamiques étudiées sur la courbe de Gartner<sup>31</sup>. Les résultats pour le *crowdsourcing* et l'*open data* sont présentés ci-dessous.

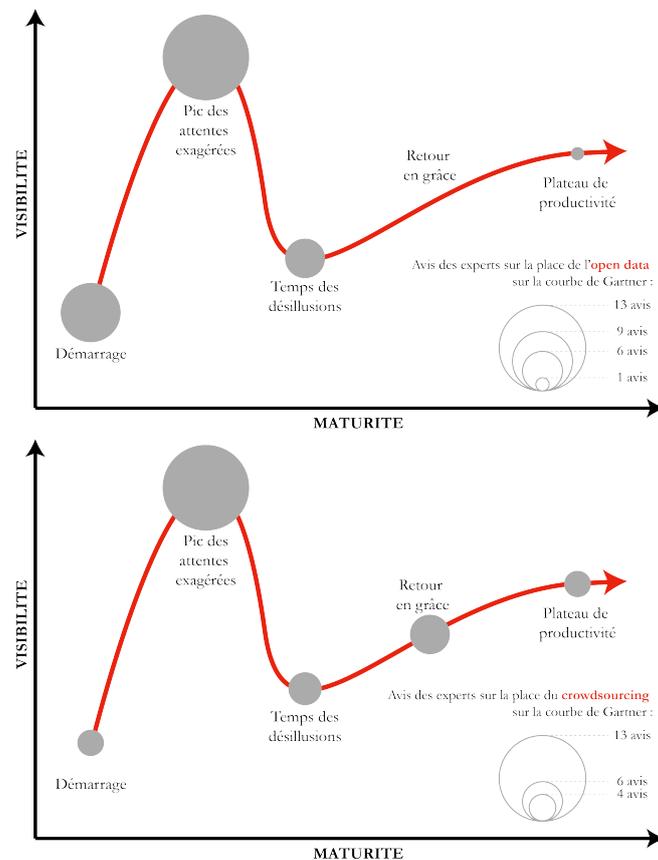


Figure 3. Evaluation de la « maturité / visibilité » de l'open data et du crowdsourcing

<sup>31</sup> La courbe de Gartner permet de suivre la visibilité et la maturité des objets techniques (Jackie et Time, 2008). Elle identifie cinq phases indiquées sur la figure 3. (1) l'attrait pour l'objet se développe, mais les applications concrètes ne sont pas encore identifiées ; (2) l'engouement pour l'objet se confirme, il est largement présent dans les médias, de multiples applications sont imaginées ; (3) l'objet se confronte à la réalité du terrain, des applications imaginées ne sont finalement pas viables ; (4) l'objet technique est désormais connu, les applications sont en phase de développement ; (5) l'objet technique est mature, des applications importantes sont établies et reconnues.

Le croisement de l'identité des experts avec les réponses permet d'identifier assez nettement les deux profils d'acteurs mis en évidence par Bédard (2012) : un *enthousiasme technologique* représenté plutôt par des jeunes experts fortement ancrés dans le milieu *open source* tend à s'opposer à une certaine *sagesse professionnelle* représentée plutôt par des experts de la géomatique ayant une certaine expérience. Ces résultats mettent aussi en évidence le paradoxe de ces dispositifs, à savoir une visibilité forte malgré un manque de maturité. La particularité des données géographiques doit ici être soulignée : contrairement à d'autres types de données publiques (textes de lois ou données comptables, par exemple), pour les données géographiques, l'*open data* dépend fortement d'innovations technologiques (formats informatiques, logiciels de gestion et traitement de ces données, etc.). Le potentiel d'évolution est donc encore important pour ce type de données ce qui rend leur production / diffusion originale.

Dans le domaine des données naturalistes et environnementales au sens large, la maturité de ces processus est liée au succès des portails offrant la possibilité de visualiser (voire d'alimenter) ce type de bases de données, portails auxquels recourent de façon routinière un nombre croissant d'acteurs. Une certaine ambiguïté demeure cependant, dans la mesure où le globe virtuel le plus populaire (*Google Earth*) est géré par une entreprise privée. Un nombre croissant d'administrations, conscientes de l'importance prise par cet outil, fournissent désormais leurs données au format « .kml » (le format Google Earth), en sus des autres formats de données consultables dans des SIG. De nombreux signes d'immaturité perdurent aussi, décelables aux débats portant sur certaines dimensions essentielles des données mises en accès libre. Ainsi, les multiples efforts des différents systèmes pour gagner en notoriété, par la mise en place de réflexions internes sur les usages et sur les moyens d'augmenter le nombre d'utilisateurs, est une preuve que l'*open data* dans ce domaine est encore loin d'être stabilisé.

### ***Tension n°2 : transparence ou confusion ?***

Le postulat selon lequel l'*open data* est un vecteur de transparence démocratique reste peu étayé par des évidences empiriques. Dans le domaine des données géographiques, ce débat est complexifié par les compétences spécifiques que supposent les usages de ces données : la connaissance n'est pas suffisante ; encore faut-il maîtriser les technologies d'analyse de ces données !

Parmi les professionnels francophones, aucune position dominante ne se dégage sur ce point. A la question : « *l'ouverture des données publiques favorise-t-elle la*

*transparence démocratique en permettant au citoyen d'accéder aux données brutes des institutions ou ne fait-elle que participer à l'infobésité ambiante en renforçant le déluge de données actuel ?* », les réponses sont très partagées. Considérant qu'il s'agit d'outils produits par des experts pour des experts, 60% des répondants n'y voient pas un renforcement du contrôle démocratique alors que, dans le même temps, 70% d'entre eux jugent ces outils porteurs d'un effort de transparence des autorités publiques. Aucune convergence d'opinion n'a pu être obtenue sur ce thème à l'issue des trois tours d'enquête. Dans un autre contexte, certaines études sur l'impact des données géographiques publiques pour résoudre des conflits environnementaux apportent un éclairage différent. Dans ces conflits affectant des territoires ruraux lors de l'arrivée de grandes entreprises agricoles, la disponibilité de données géographiques publiques ne permet pas de réduire les asymétries de pouvoir entre acteurs, dans la mesure où les capacités d'analyse de ces données sont trop inégalement partagées (Nalepa et Bauer 2012, Gautreau et Vélez 2011).

### ***Tension n°3 : centralisation ou distribution ?***

Comme évoqué plus haut, l'évolution des modes de production et de diffusion des données géographiques peut rendre délicate la distinction entre données publiques et privées. L'un des questionnements récurrents est alors de savoir si la légitimité des autorités publiques peut être remise en cause par la distribution des pratiques et la diversité des canaux de diffusion.

Dans le domaine de la biodiversité, ni les processus d'*open data* ni ceux du *crowdsourcing* ne semblent menacer des hiérarchies bien établies entre administrations publiques et associations. Les premières diffusent des données géographiques naturalistes sur leurs sites web de façon bien plus fréquente que les secondes<sup>32</sup>, du fait d'une capacité technique et budgétaire supérieure : acheter ou produire, gérer et mettre à jour, savoir interpréter ces données coûte cher et demande des ressources humaines importantes, deux éléments dont les associations sont souvent dépourvues. Tant pour des raisons de coûts que pour des raisons de répartition des fonctions entre institutions publiques et associations, les institutions publiques semblent renforcer leur prééminence

---

<sup>32</sup> Parmi les sites web publics de portée nationale, 65% fournissent des données géographiques naturalistes, tandis que les sites associatifs ne sont que 29%. Si l'on inclut également les sites de portée régionale, 58% des sites publics fournissent ce type de données, contre 26% pour les sites associatifs (source : voir figure 1).

dans la diffusion des données naturalistes grâce à leur maîtrise de la production de données géographiques.

#### *Tension n°4 : qualité interne ou externe ?*

La question de la qualité des données a également été soulevée dans de nombreux commentaires du premier tour d'enquête. Elle a donc fait l'objet d'un axe de négociation lors des deux tours suivants. La question était alors de savoir si l'ouverture des données publiques peut avoir pour effet l'amélioration de la qualité des données géographiques diffusées. On rappelle que cette dernière se définit par la notion de qualité interne (adéquation entre une donnée et ce qu'elle aurait dû être si toutes les spécifications avaient été respectées) et la notion de qualité externe (adéquation entre une donnée et les besoins de son utilisateur). L'amélioration de la qualité interne et externe des données géographiques résulterait, selon 50 à 55% des participants, d'une part d'une responsabilisation plus forte des producteurs (dès lors qu'une donnée est diffusée, plus d'attention est portée à sa bonne structuration) et, d'autre part, à la démultiplication des possibilités de contrôle par les utilisateurs eux-mêmes. 45 à 50% des répondants ne partagent cependant pas cet avis.

Dans le domaine de la biodiversité, de nombreux gestionnaires publics sont interpellés sur la qualité des données de leur portail. Tandis que certains systèmes assument ouvertement le fait qu'ils donnent accès à des données produites de façon décentralisées et dont les auteurs sont seuls responsables<sup>33</sup>, d'autres engagent des efforts importants pour donner à l'utilisateur des moyens d'évaluer si les données correspondent à la qualité qu'il cherche<sup>34</sup>.

---

<sup>33</sup> C'est par exemple la position assumée de l'Infrastructure Nationale de Données Spatiales du Brésil (Py H. 2012. *The Brazilian National Spatial Data infrastructure*. Communication au séminaire «*Sharing Environmental information. Issues of Open Environmental Data in Latin America*». Porto Alegre, Brésil, 23 et 24 août 2012. En ligne : <http://baguala.hypotheses.org/seminar-porto-alegre-2012>).

<sup>34</sup> Le GBif (voir note 23) a ainsi engagé un vaste travail de conception et mise à disposition d'outils résumant les grandes caractéristiques des jeux de données téléchargés par les usagers, dans une tentative de dépasser son image de simple dépôt de données de biodiversité. Ne pouvant garantir la qualité de données produites et gérées par d'autre, le portail essaie malgré tout d'investir de façon active le débat sur la « qualité », notamment pour convaincre les scientifiques de l'utiliser plus fréquemment.

### *Tension n°5 : carte figée ou application composite ?*

L'une des particularités de la donnée géographique, quand elle est publique et diffusée, est qu'elle peut faire l'objet d'usages très nombreux (aussi nombreux qu'il y a de formes d'utilisation du territoire) et qu'elle permet la réutilisation d'autres types de données publiques (comptables, environnementales, sociologiques). Ainsi, au contraire des cartes qui étaient figées, les données géographiques sont désormais combinables entre elles.

Les *mashups* (ou applications composites) sont devenues les briques essentielles de l'écosystème des données numériques. Ils désignent des applications qui combinent des contenus et/ou des services provenant de plusieurs applications plus ou moins hétérogènes. En juillet 2010, plus d'un tiers des *mashups* du web se fondaient sur les interfaces cartographiques de *Google, Microsoft et Yahoo!*<sup>35</sup> Certaines collectivités<sup>36</sup> développent leur propre API pour faciliter la réutilisation des données. Ainsi, à la différence des cartes imprimées, les données géographiques sont devenues malléables : on peut leur ajouter des statistiques, les « *braconner* » (de Certeau, 1990). Elles ont donc une forte capacité à être transformées (notamment par ajout sémantique) et leur mise en partage peut avoir un effet démultiplicateur dans des champs très divers de l'interaction pouvoirs publics / citoyens.

Ceci est notable dans le domaine naturaliste : c'est grâce aux possibilités offertes par l'association des technologies de géolocalisation et de numérisation que les corpus de données sur la biodiversité (faune, flore, paysages), réunies jusqu'alors sous forme de collections ou d'inventaires au format papier, ont pu être diffusées : soit sous forme de points de présence - position exacte en latitude-longitude des relevés botaniques ou floristiques- soit sous forme de listes d'espèces associées à une unité spatiale administrative. Les référentiels géographiques (couches d'information communales par exemple) sont donc un outil essentiel de « traduction » de données conventionnelles de biodiversité en données numériques. Cette traduction géographique, élément clé de la diffusion des données publiques, participe du processus progressif de « rationalisation » des données naturalistes.

### *Tension n°6 : standardisation ou richesse sémantique ?*

---

<sup>35</sup> Selon le site Web : <http://programmableweb.com>

<sup>36</sup> Par exemple, la communauté urbaine de Bordeaux : <http://data.lacub.fr/apicub>

Face à la démultiplication des canaux de diffusion des données géographiques, un travail de rationalisation est à l'œuvre. Ainsi, la création de l'Agence Européenne de l'Environnement (AEE) vise explicitement à standardiser les données pour les rendre comparables à l'échelle de l'Union Européenne. Ces comparaisons (basées sur des outils de *reporting* et de *benchmarking* de plus en plus performant) s'inscrivent « dans une logique de contrôle et de stimulation des Etats en situant leurs performances sur des échelles normatives » (Alphandéry et al., 2012). En France, pour marquer sa volonté de remédier au caractère disparate et difficilement mobilisable des données sur la nature, le Ministère de l'Ecologie met en œuvre depuis 2007 le Système d'Information sur la Nature et les Paysages (SINP) et depuis 2009 l'Observatoire National de la Biodiversité (ONB). Ces mouvements de standardisation répondent également à une demande accrue de transparence institutionnelle (Rosanvallon, 2006), fondement de nombreux cadres réglementaires comme la convention d'Aarhus.

Ainsi l'un des paradoxes identifiés lors de l'enquête DELPHI repose sur cette opposition entre la richesse sémantique (parfois aussi géométrique et graphique) que mettent en évidence les nouveaux modes de production/diffusion des données géographiques et la montée en puissance des initiatives qui visent à canaliser la diversité des représentations. Cette vision peut être illustrée par l'observation des catalogues de données qui décrivent les données diffusées. Nombre de collectivités territoriales, à l'instar de la Ville de Paris, proposent des métadonnées non normées où la simplicité d'usage est privilégiée. A l'opposé, d'autres collectivités, comme le Grand Lyon, préconisent l'usage de métadonnées et appliquent un profil de la norme ISO19115 conforme à la directive *INSPIRE*. Sous couvert de faciliter l'interopérabilité des données et des systèmes, ces vastes chantiers de standardisation des données publiques ne conduisent-ils pas à perdre de la richesse sémantique et à une normalisation de la pensée sur le territoire ? Cette question reste ouverte.

### **Conclusion : de l'importance d'une analyse des usages.**

L'analyse des modes de production et de diffusion des informations géographiques publiques nous a permis de mettre en évidence des évolutions continues depuis une quarantaine d'années. Ces dernières se sont accélérées récemment à l'instar des logiques de *crowdsourcing* et d'*open data* qui connaissent aujourd'hui une forte visibilité médiatique.

Au lieu de considérer l'émergence de ces nouvelles approches de production et de diffusion comme une rupture majeure, cet article a cherché, par une

approche historique, à les replacer dans un contexte d'ouverture progressive (intra puis inter et enfin extra-organisations publiques) des données géographiques. Il a aussi cherché, par une enquête internationale, à sortir des injonctions et des discours incantatoires pour mettre en évidence la diversité et l'interaction forte entre ces processus. Enfin, l'exemple des données de la biodiversité nous aide à analyser l'articulation entre administrations publiques et société civile dans la production et la diffusion de données géographiques.

Si le libre accès ne rebat que partiellement les cartes, puisqu'il s'inscrit, *in fine*, dans une logique d'ouverture commencée il y a plusieurs dizaines d'années, il conduit tout de même à un accroissement exponentiel des données produites et des capacités d'accès aux données. Cette évolution engendre une diversification et une expansion des usages. Une des clés de compréhension des impacts de cette mise en circulation des données publiques se trouve alors dans l'analyse de la multiplicité d'usages qui en émergent aujourd'hui.

*Remerciements.*

*Les auteurs remercient l'ensemble des participants à l'enquête DELPHI.*

## **Bibliographie**

- Alphandéry P. et Fortier A., Les associations dans le processus de rationalisation des données naturalistes, *Natures Sciences Sociétés* (19): 22-30, 2011.
- Alphandéry P., Fortier A. et Sourdril A., Les données entre normalisation et territoire : la construction de la trame verte et bleue, *Développement durable et territoires*, 3/2, 2012.
- Bédard Y., Les SIG deviennent-ils le nouveau terrain de l'arbitraire ?, In *Rencontres SIG La Lettre*, Marne-La-Vallée, 2012.
- Certeau M. de, 1990, *L'invention du quotidien, tome 1 : les arts de faire*, Éd. Gallimard.
- Couderchet L. et Amelot X., « Faut-il brûler les Znieff ? », [Cybergéo, Revue Européenne de Géographie](http://cybergeo.revues.org/23052), En ligne : <http://cybergeo.revues.org/23052>
- Douglas N., *Developing Spatial Data Infrastructures : The SDI cookbook*, Douglas, 2004.
- Fisher F., Collaborative Mapping : How Wikinomics is Manifest in the Geo-information Economy. *GEOinformatics*, 11 (2) : 28-31, 2008.
- Flichy P., *L'innovation technique. Récents développements en sciences sociales vers une nouvelle théorie de l'innovation*, La Découverte, Paris, 1995.
- Gautreau P et Vélez E. 2011. Strategies of environmental knowledge production facing land use changes: insights from the Silvicultural Zoning Plan conflict in the

- Brazilian state of Rio Grande do Sul. [Cybergéo, Revue Européenne de Géographie](#) .  
En ligne : <http://cybergegeo.revues.org/24881>
- Goodchild M., Citizens as voluntary sensors : spatial data infrastructure in the world of web 2.0. , *International Journal of Spatial Data Infrastructures Research*, 2007.
- Jackie F. et Time M., *Understanding Gartner's Hype Cycles*, Harvard Business Press, 2008.
- Jacomy, M., Heymann, S., Venturini, T., et Bastian, M., ForceAtlas2, a graph layout algorithm for handy network visualization (forthcoming), 2012.
- Jauréguiberry F. et Proulx S., *Usages et enjeux des technologies de communication*, Toulouse, Erès, 143 p., 2011.
- Joliveau T. , Le géoweb, un nouveau défi pour les bases de données géographiques, *L'Espace géographique*, 2011/2 Tome 40, p. 154-163, 2011.
- Kleinberg J & Lawrence S, The structure of the Web, *Science*, Vol. 294 no. 5548 pp. 1849-1850 2001
- Linstone H. A. et Turrof M., *The Delphi method, techniques and applications*, Addison Wesley Publishing, 1975.
- Masser I., *GIS worlds : creating SDI*, ESRI Press, 2005.
- Mathian H., *Le métier de géomaticien*, Enquête CNIG – AFIGÉO – Portail Géomatique/GéoRézo, 2003.
- Montagne, D., *Les données localisées sur la biodiversité en France. Variation géographique de disponibilité et principales logiques de constitution*. Mémoire de master 1, Université Paris 1 Panthéon-Sorbonne, 2011.
- Nalepa R et Bauer DM. 2012. Marginal lands: the role of remote sensing in constructing landscapes for agrofuel development. [The Journal of Peasant Studies](#), Volume 39, Number 2, 1 April 2012 , pp. 403-422(20)
- Noucher M., *La donnée géographique aux frontières des organisations : approche socio-cognitive et systémique de son appropriation*, Thèse de doctorat, Ecole Polytechnique Fédérale de Lausanne, 2009.
- Roche S., *Les enjeux sociaux des systèmes d'information géographique*, L'Harmattan, 2000.
- Rosanvallon P., *La contre-démocratie. La politique à l'âge de la défiance*, Seuil, 2006.
- Watzlawick P., *La réalité de la réalité : Confusion, désinformation, communication*, Seuil, 1976.