

Rendre visible l'état du patrimoine institutionnel de données géographiques (IDG) françaises

Matthieu Noucher

▶ **To cite this version:**

Matthieu Noucher. Rendre visible l'état du patrimoine institutionnel de données géographiques (IDG) françaises. Théo Quant - 12èmes rencontres, UMR THéMA, Université de Franche Comté, May 2015, Besançon, France. halshs-01187090

HAL Id: halshs-01187090
https://shs.hal.science/halshs-01187090

Submitted on 26 Aug 2015

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.

Rendre visible l'état du patrimoine institutionnel de données géographiques. Prototypage d'un observatoire des 65 Infrastructures de Données Géographiques (IDG) françaises.

Matthieu Noucher
matthieu.noucher@cnrs.fr
Chargé de recherche au CNRS
UMR 5185 ADESS (CNRS, Université Bordeaux Montaigne, Université de Bordeaux)

Mots-clés : gouvernance informationnelle ; infrastructure de données géographiques ; métadonnées ; observatoire.

Contexte : les Infrastructures de Données Géographiques comme nouvel objet de recherche

Le besoin exprimé au niveau international de faciliter l'accès, l'utilisation et le partage des données géographiques a conduit depuis les années 1990, au développement d'Infrastructures de Données Géographiques (IDG) (Crompvoets *et al*., 2004). Ces plateformes rassemblent les données, les réseaux informatiques, les normes et standards, les accords organisationnels et les ressources humaines nécessaires pour faciliter et coordonner le partage, l'accès et la gestion des données géographiques (Rajabifard *et al*., 2002). Elles sont mises en œuvre à différents niveaux organisationnels (Masser, 2010) et la coopération entre l'ensemble des producteurs et des utilisateurs de données est indispensable (Masser, 2005 ; Nedovic-Budic et *al*., 2011 ; Vandenbroucke, 2011) pour atteindre leurs objectifs de développement économique, de stabilité sociale, de bonne gouvernance et de gestion durable de l'environnement (Williamson et *al*., 2003 ; Rodriguez-Pabon, 2005 ; Masser, 2010). En Europe, la directive 2007/2/CE INSPIRE [1] vise explicitement à appuyer la politique environnementale par une Infrastructure européenne de Données Géographiques qui repose sur les IDG des Etats membres. En France, l'Infrastructure Nationale de Données Géographiques (INDG) représentée par le géoportail/géocatalogue[2] est complétée par des infrastructures régionales, comme PIGMA[3] en Aquitaine, qui constituent des maillons essentiels dans la production et la diffusion d'informations publiques. D'un point de vue technique, les IDG peuvent être assimilées à des systèmes d'information bâtis autour d'une architecture orientée services dont la composante Web est essentielle. Leur mise en réseau, du fait de l'interopérabilité et de la normalisation des données, facilite la circulation de contenus grâce, par exemple, au *moissonnage*[4] de catalogues émanant de producteurs indépendants.

Par l'application des normes imposées, les IDG sont au cœur de structures qui rassemblent et diffusent l'information publique et visent prioritairement à en faciliter l'accès. En plus de la mise à disposition de référentiels géographiques, les communautés de pratique qui se développent au sein des IDG cherchent à coproduire des données thématiques. Les IDG apparaissent dès lors non seulement comme des points de passage qui facilitent la circulation de l'information mais aussi comme des points d'ancrage où les acteurs territoriaux se réunissent pour produire collectivement de nouveaux jeux de données. Initialement centrées sur la sphère des experts, les IDG, bousculées par la mouvance *Open Data* (Chignard, 2012), tendent aujourd'hui à favoriser la mise en libre accès de contenus et les interactions avec les non spécialistes.

---

[1] http://inspire.jrc.ec.europa.eu/
[2] http://www.geoportail.gouv.fr/ et http://www.geocatalogue.fr/
[3] http://www.pigma.org
[4] http://www.geosource.fr/docs/admin/io/harvesting/index.html.

Objectif : vers un prototype d'observatoire des Infrastructures de Données Géographiques

L'objectif du projet de recherche « *GEOBS : les Infrastructures de Données Géographiques dans la gouvernance informationnelle de l'environnement* » (2015-2017) est alors d'analyser le contenu de ces plateformes pour comprendre les logiques actuelles de formalisation et de diffusion des savoirs géographiques institutionnels. Notre hypothèse est que les IDG ne doivent pas uniquement être considérées comme un support à la recherche d'information (*i.e.* un simple moteur de recherche) mais comme un véritable objet de recherche, afin de mieux appréhender la façon dont s'opère aujourd'hui la gouvernance informationnelle des territoires (Feyt et Noucher, 2014). Il s'agit alors de dresser un panorama des ressources institutionnelles qui sont disponibles et diffusées.

Pour mener à bien ces objectifs de recherche, le projet GEOBS vise à concevoir et implémenter un prototype d'observatoire des flux d'information géographique qui transitent par les 65 Infrastructures de Données Géographiques françaises (selon le recensement établi par l'AFIGEO en juin 2014). Fondé sur l'explicitation de plusieurs indicateurs de contenus et d'usages, ce prototype d'observatoire propose sous la forme d'une application web des géovisualisations dynamiques d'indicateurs dont l'objectif est de rendre visible l'état actuel du patrimoine institutionnel français de données géographiques.

Corpus de données, chaîne de traitements et résultats intermédiaires

Une chaîne de traitement pour l'extraction automatisée des métadonnées et leur restructuration dans une base de données a été développée afin de permettre l'analyse spatio-temporelle des infrastructures de données géographiques (Noucher et Gautreau, 2014). L'extraction des fichiers XML de métadonnées est opérée à partir de scripts en Python. On remarque d'emblée une très forte disparité dans le volume des catalogues : à titre d'exemples, le géocatalogue national propose 48809 fiches de métadonnées, PIGMA diffuse plus de 3100 métadonnées sur la région Aquitaine, la plateforme IGEOCOM40 compte 41 fiches de métadonnées sur les Landes (inventaire établi en décembre 2014)[5].

Il est cependant nécessaire de rester prudent sur les conclusions de tels écarts : d'une part, ces plateformes sont récentes et encore très largement en phase de montée en charge ; d'autre part, le volume de métadonnées dépend à la fois de la distribution de la donnée et de la granularité de la documentation. Ainsi, la base de données topographiques de l'IGN en France (BD TOPO®) fait l'objet de plus de 180 fiches de métadonnées suivant des logiques de découpages à la fois administratif et thématique encore approximatives (le jeu de données est parfois documenté *via* une fiche de métadonnées, parfois documenté thème par thème soit huit fiches de métadonnées).

L'analyse des métadonnées ne peut donc se contenter d'un comptage des fiches pour essayer de comprendre les logiques informationnelles des territoires. Les contenus des champs de métadonnées apportent alors plus des précisions. Leur interrogation (par des requêtes xQuery sur les balises XML) permet de traiter en particulier de la couverture temporelle (quand ?), thématique (quoi ?), organisationnelle (qui ?) et géographique (où ?) des données. A titre d'exemple, l'analyse de l'emprise géographique des données s'opère par l'extraction des 4 coordonnées des rectangles d'emprise (un même jeu de données peut posséder plusieurs emprises si la couverture est discontinue) réalisée à partir de la récupération des balises XML

[5] Une analyse du géocatalogue hélvétique (www.geocat.ch) pourrait également permettre d'apporter quelques éléments de comparaison avec le pays hôte de cette 12 édition de ThéoQuant.

correspondantes aux emprises (<EX_GeographicBoundingBox>) en format CSV pour générer ensuite des données en format GeoJSON. Un traitement dans PostGIS/QGIS, permet enfin, de produire une carte des densités de données géographiques et de faire apparaître des dynamiques spatiales dans la constitution du patrimoine institutionnel national et local de données géographiques.

Conclusion

L'analyse des résultats, en particulier les limites de la comparaison des IDG entre elles permettent de mettre en exergue les interprétations différenciées des normes. Ainsi, même si toutes les IDG se réfèrent à la directive INSPIRE et à la norme ISO19915 (métadonnées), les géocatalogues sont encore en voie de rationalisation et d'harmonisation et leur analyse comparée soulève d'intéressants enjeux méthodologiques.

Le prototype d'observatoire qui sera présenté permet de rendre visible un patrimoine institutionnel de données géographiques jusque là morcelés et constitue ainsi un support privilégié pour approcher les questions de gouvernance informationnelle des territoires. Cette *géographie de l'information géographique* s'inscrit dans le courant de la cartographie et des SIG critiques (*critical GIScience studies*) (Schuurman, 2009) en cherchant à déconstruire ces infrastructures institutionnelles pour les analyser comme des dispositifs socio-techniques de mise en ordre du monde (Noucher, 2013). Elle révèle des répartitions hétérogènes de l'information et des discontinuités spatiales qui ne sont pas le seul fait de disparité démographique. *In fine*, travailler ainsi la visibilisation (et par extension l'invisibilisation) des représentations cartographiques institutionnelles à travers les IDG c'est tenter de comprendre comment les bases de données géographiques institutionnelles ordonnent le rapport des sociétés à l'espace et aux territoires.

Bibliographie

Chignard S. (2012). *Open data. Comprendre l'ouverture des données publiques*, Edition FYP.

Crompvoets J., Bregt A., Rajabifard A., Williamson I. (2004). Assessing the world wide developments of national spatial data clearinghouses. *IJGIS*, 18, 665-689.

Feyt G. et Noucher M. (2014). La gouvernance informationnelle, outil et enjeu stratégiques des recompositions territoriales : vers l'émergence de nouveaux référentiels géographiques ? 2ème colloque du GIS CIST "*Fronts et frontières des sciences du territoire*", Mars 2014, Paris, France. pp.191-196

Nedovic-Budic Z., Crompvoets J., Georgiadou Y. (2011). *Spatial Data Infrastructures in Context: North and South*, CRC Press, Boca Raton.

Noucher M. et Gautreau P. (2014). Faire autorité sur le web. Analyse des stratégies de mise en visibilité et de contrôle de l'information géographique par les institutions publiques en Amérique Latine. *SAGEO'14*, Grenoble.

Noucher M. (2013). Infrastructure de données géographiques et flux d'information environnementale : de l'outil à l'objet de recherche. *Netcom – Networks and communication studies,* 27 (1-2), 120-147.

Masser I. (2005). *GIS worlds: creating spatial data infrastructures*, ESRI press, Redlands.

Masser I. (2010). *Building European spatial data infrastructures*, ESRI Press, Redlands.

Rajabifard A., Feeney M.-E.F., Williamson I.P. et Masser I. (2003). Chapter 6, National SDI Initiatives, in Williamson I, Rajabifard A., Feeney M.-E.F. (eds). *Development of Spatial Data Infrastructures: from Concept to Reality*, London: Taylor & Francis, pp. 95-109.

Rodriguez-Pabon O. (2005). Cadre théorique pour l'évaluation des infrastructures d'information géospatiale, Thèse de doctorat de l'Université Laval, Québec.

Schuurman N. (2009). Critical GIScience in Canada in the new millennium. *The Canadian Geographer*, 53/2, 139–144.

Vandenbroucke D. (2011). *INSPIRE & NSDI State of Play: D3.2 - Detailed survey on use and usability of INSPIRE & NSDI*, http://inspire.jrc.ec.europa.eu/reports/stateofplay2011/INSPIRENSDISoP-D32-ResultDetailedSurveyYear2-v3.1.pdf

Williamson I. P., Rajabifard A., Feeney, M-E. (2003). *Developing spatial data infrastructures: from concept to reality*, CRC Press, Boca Raton.