



HAL
open science

Un cadre d'analyse des Infrastructures de Données Géographiques pour interroger la mise en réseaux des acteurs et des outils

Matthieu Noucher, Françoise Gourmelon, Antoine Laumond, Guy Melançon, Bruno Pinaud, Adeline Maulpoix, Julie Pierson, Olivier Pissot, Mathias Rouan

► To cite this version:

Matthieu Noucher, Françoise Gourmelon, Antoine Laumond, Guy Melançon, Bruno Pinaud, et al.. Un cadre d'analyse des Infrastructures de Données Géographiques pour interroger la mise en réseaux des acteurs et des outils. SAGEO : Spatial Analysis & Geomatic, Dec 2016, Nice, France. <halshs-01414234>

HAL Id: halshs-01414234

<https://shs.hal.science/halshs-01414234v1>

Submitted on 12 Dec 2016

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.



HAL Authorization

Un cadre d'analyse des Infrastructures de Données Géographiques pour interroger la mise en réseaux des acteurs et des outils

M. Noucher¹, F. Gourmelon², A. Laumond³, G. Melançon³, B. Pinault³,
A. Maulpoix², J. Pierson¹, O. Pissot¹, M. Rouan²

1. CNRS, UMR 5319 PASSAGES, site Bordeaux Universités
12 esplanade des Antilles – F-33607 Pessac Cedex
2. CNRS, UMR 6554 LETG-Brest
IUEM, rue Dumont d'Urville – 29280 Plouzané Cedex
3. Université de Bordeaux, UMR 5800 LaBRI
351 Cours de la Libération – 33405 Talence Cedex

Auteur correspondant : matthieu.noucher@cnrs.fr

RESUME. L'apparition de multiples portails cartographiques engendre une démultiplication de l'offre en données géographiques sur Internet. Notre étude propose une méthodologie mixte appliquée à 45 IDG françaises afin de comparer les objectifs affichés par leurs promoteurs et leurs contenus et services effectifs. Fondée sur l'analyse d'entretiens, de questionnaires, des sites web et des métadonnées, la méthodologie produit des résultats variés permettant de revisiter l'idée d'une mise en réseaux généralisée des outils et des acteurs. Il apparaît en effet qu'en dépit d'un discours volontariste, la mise en réseaux des outils reste confinée à des thématiques spécifiques. Par ailleurs, si la géocollaboration organisée par les IDG est très active, la mise en réseau des acteurs ne concerne que le secteur public.

ABSTRACT. The development of multiple webmapping portals generates a gear in the supply of spatial data on the Internet. Our study offers a mixed methodology applied to 45 French SDIs to compare the targets set by their promoters and their actual contents and services. Based on the analysis of interviews, survey, web sites and metadatas, methodology produces various results to visit in a different way the idea of a widespread networking tools and actors. Despite a proactive speech, the networking tools remains confined to specific themes. If geocollaboration organized by SDIs is very active, communities of practice only affects the public sector.

MOTS-CLES : Infrastructure de Données Géographiques (IDG) ; métadonnées ; interopérabilité ; géocollaboration ; analyse réseau ; France.

KEYWORDS: Spatial data infrastructure (SDI) ; metadata ; interoperability ; geocollaboration ; network analysis ; France.

1. Introduction

Par le double effet de l'interopérabilité des systèmes et des évolutions du cadre réglementaire, les données géographiques institutionnelles sont de plus en plus accessibles à travers le développement croissant d'Infrastructures de Données Géographiques (Crompvoets *et al.*, 2004). Les IDG sont ainsi devenues des plateformes web essentielles pour la diffusion des données institutionnelles (Rajabifard *et al.*, 2003). Elles visent à faciliter, à tous les échelons territoriaux, la découverte, l'accès, l'échange et le partage d'information et de services géographiques (Masser, 2010). Souvent réduites à de simples sites web permettant de consulter les données cartographiques d'un territoire à travers un géoportail et de les télécharger via un géocatalogue, les IDG sont, en fait, de véritables dispositifs sociotechniques où se joue une partie des évolutions majeures affectant le régime actuel de production et de circulation de l'information géographique.

Bien que les publications sur les IDG soient en augmentation constante depuis une décennie, Hendricks *et al.* (2012) ont montré que la plupart traitent de leurs composantes techniques, ignorant souvent les dimensions organisationnelles. Coetzee et Wolff-Piggott soulignent eux-aussi dans une revue de littérature récente (2015) et significative¹ que la plupart des recherches sont focalisées sur des développements technologiques et des preuves de concept de leurs implémentations. Considérées comme des outils, les IDG font l'objet d'analyses techniques, sur l'interopérabilité des systèmes ou la normalisation des données. Les questions organisationnelles sont parfois abordées, notamment pour évaluer la reproductibilité des systèmes de gouvernance d'un pays à l'autre (Georgiadou *et al.*, 2006).

Au niveau institutionnel, les IDG sont évaluées sur la base de critères essentiellement quantitatifs (nombre de métadonnées, nombre de services Web de consultation, nombre de téléchargements et de sources de données, etc.). Aucune analyse sur le contenu du patrimoine de données n'est généralement opérée. Par exemple, la décision de la Commission Européenne (2009/442/CE) du 5 juin 2009 inscrit le rapportage de la Directive INSPIRE avant tout dans une logique de *benchmarking*² visant à créer un référentiel cognitif commun destiné à rendre les données comparables d'un pays à l'autre, sans porter de réflexion de fond sur les types de contenus à créer ni sur leur pertinence.

Plusieurs verrous scientifiques expliquent pourquoi les études en cours sont si spécialisées. En effet, il est difficile (a) de disposer d'un panorama exhaustif de l'ensemble des données et des flux associés aux différentes plateformes décentralisées, (b) d'identifier les usagers des IDG, qu'ils soient producteurs ou consommateurs de données géographiques et (c) de décrire et qualifier les usages

¹ Coetzee & Wolff-Piggott ont analysé 237 articles dans leur revue de littérature.

² Le *benchmarking* est une technique issue du marketing pour comparer les performances des entreprises. Un *benchmark* est un indicateur chiffré de performance dans un domaine donné tiré de l'observation des résultats de l'entreprise qui a réussi le mieux dans ce domaine. Cet indicateur peut servir à définir les objectifs de l'entreprise qui cherche à rivaliser avec elle.

dont elles font l'objet. Aujourd'hui le développement des plateformes de mise à disposition d'informations géographiques sur le web contribue à la complexité de l'« univers de l'information géographique » (Ballatore, 2014). Or, la compréhension de son fonctionnement et de son évolution est un enjeu fort pour la société dans la mesure où les acteurs publics et privés à différents échelons, s'interrogent sur l'impact de ces dispositifs en termes de gouvernance et de gestion territoriale. Ces réflexions s'inscrivent pleinement dans les débats actuels menés au sein des *critical data studies* (Kitchin and Lauriault, 2014) qui concernent les contextes de production, d'analyse, de diffusion et d'usages des données, en particulier géographiques, qui circulent entre les infrastructures web.

Notre étude propose un cadre méthodologique d'analyse critique des infrastructures de données géographiques qui entend contribuer au nécessaire « travail de conceptualisation sur la base de travaux empiriques » qu'évoquent Kitchin et Lauriault (2014 : 14) pour étayer et étoffer ce champ de recherche. Une méthode mixte pour capturer des jeux de données variés et volumineux (métadonnées, pages éditoriales des sites web, résultats d'enquête en ligne et entretiens) a été mise en oeuvre pour analyser les infrastructures régionales et nationales de données géographiques françaises. L'objectif est de fournir une vue d'ensemble du partage de données géographiques en l'explorant sous quatre dimensions : l'*accessibilité des données*, la *mise en réseau des outils*, la *mise en réseau des acteurs* et l'*équité informationnelle des territoires* en distinguant la théorie (le discours des promoteurs) de la pratique (la réalité pour les usagers). L'accessibilité des données a pour but d'évaluer la capacité effective des IDG à donner accès à des contenus réutilisables en identifiant les données téléchargeables et correctement documentées permettant leur mise en exploitation dans un SIG. La mise en réseau des outils fait référence aux décloisonnements des systèmes d'information, c'est-à-dire à la capacité des IDG à favoriser les échanges entre systèmes experts pour rompre avec les logiques en silos des années 90 (démultiplication de SIG isolés par métier et par territoire sans interconnexion). La mise en réseau des acteurs sera évaluée par l'identification de dynamiques de géocollaboration. Enfin, nous questionnerons l'un des arguments récurrents de la justification de la mise en oeuvre de ces plateformes, i.e. l'argument d'égalité informationnelle des territoires. Ainsi, en abordant la déconstruction des IDG par le prisme des données, des systèmes d'information, des acteurs et des territoires, nous proposons une analyse critique d'un contexte national pour confronter les discours théoriques aux contenus effectifs.

Cet article est organisé en trois sections. La section 2 détaille le cadre exploratoire, le corpus mobilisé et la méthodologie alors que la section 3 présente les résultats de deux des quatre composantes précédemment citées : la mise en réseau des outils et la mise en réseau des acteurs. Ces résultats sont ensuite discutés et les perspectives de recherche sont présentées en section 4.

2. Corpus et méthode

2.1. Corpus

En Europe, la directive INSPIRE (2007/2/CE)³ vise au déploiement d'une infrastructure de données géographiques dans la Communauté européenne en s'appuyant sur les IDG des États membres. En France, depuis 2005 le Conseil national de l'information géographique (CNIG) assure la coordination entre l'infrastructure nationale dont le déploiement est assuré par deux acteurs historiques de la cartographie française (IGN et BRGM) et les nombreuses infrastructures thématiques (sur le sol, sur l'eau, sur le littoral, etc.) ou infranationales (départementales, régionales, transfrontalières...).

En 2014, l'AFIGEO recensait 65 dispositifs se réclamant des IDG. Ce chiffre est en progression constante depuis 2008. L'échelon régional compte à lui seul la moitié des infrastructures françaises. Relais entre les données locales et les standards nationaux, les IDG régionales ont un rôle d'animation (fédération, formation, coordination) important. Nous avons donc choisi de nous focaliser sur les infrastructures nationales et régionales qui concentrent, à elles-seules, les deux tiers des IDG recensées par l'AFIGEO. Un travail d'actualisation⁴ de l'inventaire de 2014 nous a permis de retenir 45 plateformes : 16 nationales et 29 régionales.

2.2. Méthode

Pour analyser ces 45 IDG sous l'angle de l'accessibilité des données, de la mise en réseau des outils et des acteurs et de l'égalité informationnelle des territoires, nous avons identifié onze questions de recherche (tableau 1) qui s'inscrivent dans un cadre méthodologique mixte (Johnson et Onwuegbuzie, 2004) (figure 1). L'approche qualitative est fondée sur l'analyse des composantes éditoriales des sites web et des *verbatim* d'entretiens menés avec quelques promoteurs d'IDG pour préciser leurs objectifs et stratégies. Les données d'une enquête à l'attention des promoteurs des 45 IDG et le contenu des géocatalogues en métadonnées font l'objet d'une approche quantitative à base de statistiques et d'analyses de réseaux. L'analyse spatiale des données cataloguées complète l'ensemble. Ces approches diversifiées doivent ainsi permettre de confronter les dires d'experts aux données effectivement diffusées et de s'interroger sur l'opérationnalité des objectifs assignés aux IDG.

³ <http://inspire.ec.europa.eu/>

⁴ L'actualisation de l'inventaire s'est faite à partir de l'exploration des sites web (vérification de leur maintien) et de plusieurs entretiens auprès des animateurs du réseau national des IDG. Depuis l'inventaire de 2014, certaines IDG ont disparu, fusionné ou ont été créées ce qui témoigne d'un contexte institutionnel encore instable.

TABLEAU 1. Quatre composantes déclinées en onze questions de recherche

Composante du partage de données	Question de recherche	Source
A. Accessibilité des données	A1. Quel est le discours des promoteurs sur la question de l'accessibilité des données ?	Site web et entretiens
	A2. Les données indexées dans les IDG sont-elles téléchargeables ?	Géocatalogue
	A3. Les données diffusées dans les IDG sont-elles facilement réutilisables ?	Géocatalogue
B. Mise en réseau des outils	B1. Quel est le discours des promoteurs sur la question du décloisonnement des SI ?	Site web et entretiens
	B2. Les IDG sont-elles interconnectées pour accélérer les flux de données ?	Enquête
	B3. Les IDG facilitent-elles l'hybridation thématique des données ?	Géocatalogue
C. Mise en réseau des acteurs	C1. Quel est le discours des promoteurs sur la question des dynamiques de géocollaboration ?	Site web et entretiens
	C2. Les IDG favorisent-elles des dynamiques de géocollaboration ?	Enquête
	C3. Les IDG permettent-elles la mise en réseau d'acteurs diversifiés ?	Géocatalogue
D. Équité informationnelle	D1. Quel est le discours des promoteurs sur la question de l'égalité informationnelle ?	Site web et entretiens
	D2. Les IDG couvrent-elles de manière homogène leurs territoires ?	Géocatalogue

2.2.1. L'acquisition de quatre types de données.

Les sites web

Les parties éditoriales des sites web des 45 IDG du corpus ont été explorées. Les pages relatives à la présentation de la démarche ou à ses objectifs ont été étudiées pour relever les argumentaires qui justifient le déploiement de ces plateformes.

L'enquête web

L'enquête a été lancée en mai 2015 pour une durée de cinq mois. Réalisée en ligne à partir du questionnaire d'enquête LimeSurvey, elle était destinée aux promoteurs des IDG françaises. Une série de questions fermées a permis de dresser un état des lieux des services proposés (géoportail, géocatalogue, services web, etc.) et des actions d'animation (formation, communication, coproduction, etc.) associées à chaque plateforme. Des questions ouvertes ont également invité les participants à

préciser leurs objectifs et leurs cibles. Les promoteurs de 44 des 45 IDG de notre corpus ont répondu à cette enquête.

Les entretiens.

Une série d'entretiens a été menée auprès de 12 promoteurs d'IDG nationales et régionales⁵ afin de préciser leur stratégie en matière d'objectifs, de cibles et de perspectives. A partir d'une trame semi-directive, des questions relatives aux modes de gouvernance ou aux actions d'animation et de coproduction ont également été posées pour obtenir des informations inaccessibles par l'analyse automatisée des métadonnées. Les données récoltées ont permis d'approfondir l'analyse des sites web et de l'enquête en allant plus loin dans la prise en compte de l'argumentaire des promoteurs des IDG par rapport aux quatre dimensions étudiées.

Les métadonnées.

L'extraction des métadonnées des géocatalogues est réalisée à partir des services web de catalogage CSW-2. Le contenu des balises de métadonnées jugées utiles pour répondre à nos questions de recherche est stocké dans une base de données PostGreSQL. 31 des 45 IDG du corpus disposent de services web de catalogage opérationnel qui nous ont permis d'analyser plus de 160.000 fiches de métadonnées. La base de données est exploitée par différents outils d'analyse en fonction des besoins.

2.2.2. L'analyse des données.

Des citations issues des *verbatim* de sites web et d'entretiens

Les argumentaires relevés lors de l'analyse des sites web ou lors des entretiens ont permis de récolter une série de *verbatim* qui sont mobilisés, par le biais de citations, sans traitement statistique pour conserver la richesse des points de vue.

Des statistiques pour traiter l'enquête et les métadonnées

Des traitements statistiques univariés, réalisés avec le logiciel R, ont permis de produire des éléments de synthèse facilitant l'exploitation des résultats de l'enquête et l'exploration des fiches de métadonnées.

Des analyses de réseaux pour traiter l'enquête et les métadonnées

Les traitements statistiques sur l'enquête et la base de données intégrant le contenu des balises de métadonnées sont complétés par des analyses réseaux. Elles permettent en particulier d'explorer les connexions entre IDG par des graphes sur le moissonnage des géocatalogues. D'autres analyses réseaux, réalisées avec le logiciel TULIP, permettent de lier les métadonnées en fonction de leur proximité thématique grâce au calcul d'un indicateur de similarité. Celui-ci peut se définir comme suit :

⁵ Trois entretiens dans une région métropolitaine issue de territoires nouvellement fusionnés (Aquitaine-Limousin-Poitou-Charentes) et trois entretiens dans une région ultra-marine (la Guyane) qui comptent chacune trois IDG ont été complétés par six entretiens réalisés au niveau national auprès des porteurs d'IDG qui couvrent l'ensemble du territoire.

Soit deux fiches de métadonnées **f1** et **f2** :

- On détermine un vecteur **vref** de mots-clés correspondant à l'union des mots-clés de **f1** et **f2**.
- Pour **f1** et **f2**, et pour chaque mot-clé **m** de **vref** si **m** est utilisé dans la fiche, on crée pour chacune des fiches un vecteur de valeurs binaires de la taille de **vref** déterminant si le mot-clé est présent ou non.

Ces vecteurs sont nommés **v1** et **v2** . Le score **S** de similarité est déterminé par :

$$S = \frac{\sum_{i \in [0, taille(v_{ref})[} (v_1[i] * v_2[i])}{\sqrt{\sum_{a \in v_1} a^2} * \sqrt{\sum_{b \in v_2} b^2}} \tag{1}$$

Ce score correspond au cosinus de l'angle entre les vecteurs **v1** et **v2**. Variant entre 0 et 1, il permet de déterminer l'ampleur de la proximité thématique entre deux jeux de données. On détermine ensuite un seuil arbitrairement compris dans [0,1] et si le score entre deux fiches est supérieur au seuil, on crée un lien entre les fiches.

De l'analyse spatiale appliquée aux métadonnées

La couverture territoriale des données géographiques est décrite dans les métadonnées par les quatre coordonnées de leur rectangle d'emprise via la balise <EX_GeographicBoundingBox>. Ces données sont stockées dans PostgreSQL. Un traitement par PostGIS/QGIS permet de générer des cartes de densité des données.

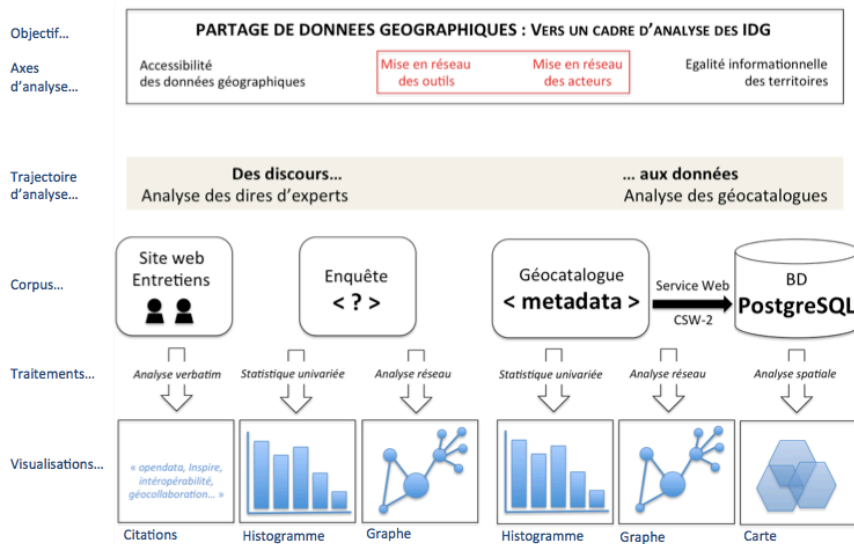


FIGURE 1. Synoptique du cadre d'analyse

Enquêtes et entretiens auprès des coordinateurs et exploration des sites web et des géocatalogues forment donc l'ossature du cadre méthodologique mis en œuvre pour évaluer l'opérationnalité des objectifs assignés aux IDG (Noucher et al., soumis). Dans la présente contribution, nous nous focalisons sur deux des quatre axes d'analyse pour nous intéresser à la question du décloisonnement des systèmes d'information (mise en réseau des SIG) et à celle des dynamiques collaboratives associées aux IDG (mise en réseau des acteurs). Ces deux axes nous permettent de mobiliser tout le spectre des corpus et outils d'analyse précédemment présentés.

3. Résultats⁶

3.1. Mise en réseau des outils

3.1.1. *Quel est le discours des acteurs sur la question du décloisonnement des SI ?*

L'un des arguments récurrents pour le déploiement d'IDG est lié aux capacités techniques dont ces plateformes disposent pour mettre en relation des bases de données stockées sur des serveurs distribués. Ainsi, 40 des 45 promoteurs d'IDG enquêtés déclarent mettre en œuvre des services web permettant d'accéder à distance à des métadonnées (CSW) ou des données (WMS, WFS) potentiellement distribuées sur différents serveurs. L'infrastructure nationale de données géographiques place ces questions d'interconnexion au cœur de sa stratégie de déploiement. La refonte en cours (2015/18) du géoportail/géocatalogue (refonte dite « geoportail v4 ») se fait, d'après les propos recueillis lors de notre enquête : « *dans un contexte d'émergence de l'État-plateforme* » et vise notamment à renforcer « *les mécanismes d'agrégation nationale des données produites régionalement ou localement* », à développer « *des géotraitements chaînés sur des sources hétérogènes* » ou encore à « *améliorer des mécanismes de moissonnage / mise en cache* ». L'État-plateforme est une stratégie nationale⁷ qui vise à favoriser l'émergence de nouveaux services publics numériques en s'appuyant sur une meilleure circulation des données. On retrouve ces préoccupations d'amélioration de la circulation de l'information au niveau régional avec des IDG qui se fixent pour objectif le « *développement de nouvelles fonctions (services web) pour faciliter les relais dans les territoires* » ou plus généralement de « *favoriser l'interopérabilité des systèmes pour réduire les dépenses publiques et faire gagner du temps* ». En fédérant les acteurs d'un même territoire ou d'une même thématique, les IDG sont censées décloisonner les SIG qui se sont déployés dans tous les métiers et à toutes les échelles depuis les années 80. Néanmoins une analyse de leurs géocatalogues permet de nuancer ces affirmations.

⁶ Des résultats détaillés sont aussi diffusés tout au long du projet sur le site www.geobs.cnrs.fr qui permet d'accéder aux publications, aux corpus de données et des visualisations interactives.

⁷ Stratégie définie par la DINSIC (Direction interministérielle du numérique, des systèmes d'information et de communication).

3.2.2. Les IDG sont-elles interconnectées pour accélérer les flux de données ?

Pour évaluer l'interconnexion des IDG, nous analysons le moissonnage des géocatalogues. Le moissonnage est un mécanisme permettant de collecter des métadonnées sur un catalogue distant et de les stocker sur le nœud local pour un accès plus rapide. Ce moissonnage est plus qu'un import de données simple car les métadonnées locales et celles du catalogue distant sont synchronisées. Par exemple, depuis un catalogue régional il est possible d'accéder à l'ensemble des métadonnées des catalogues départementaux et à une sélection de métadonnées issues d'un catalogue national. Analyser la mise en œuvre du moissonnage permet donc de mettre en lumière les relations entre IDG. Cette étude du processus de moissonnage permet en outre d'identifier par une analyse réseau les nœuds centraux (ou isolés) et d'éventuels *clusters* (à l'aide de la visualisation de sous-réseaux). En plus des connexions entre IDG du corpus, l'enquête a permis un premier recensement (sur une base uniquement déclarative) des connexions entre les géocatalogues des IDG et les géocatalogues d'observatoires, de portails *open data*, d'infrastructures infra-régionales... La figure 2 permet ainsi d'intégrer tous les liens de 2^e niveau, c'est-à-dire l'ensemble des systèmes directement en contact avec les IDG de notre corpus.

La densité⁸ de ce graphe est relativement faible ($d=0,02$). Les IDG régionales s'ouvrent vers des IDG départementales ou infra-départementales qui sont faiblement connectées au reste du réseau. La composante principale ne contient que 34 des 45 IDG du corpus initial auxquels sont associés 45 autres portails de données (d'échelon local pour la plupart). La portée de l'interopérabilité des infrastructures reste donc, encore aujourd'hui relativement limitée. Le géocatalogue du BRGM, point de relais national pour le rapportage de la directive INSPIRE, se singularise avec une connexion à 21 IDG, soit 4 fois la moyenne générale.

La figure 2 permet également de visualiser quelques circuits spécifiques d'accès à l'information :

- Deux composantes secondaires, c'est-à-dire déconnectées du réseau principal, dans le domaine de l'imagerie satellitaire (GeoSud et Theia) sont associées mais restent indépendantes du reste du réseau.
- Des composantes géographiques du réseau sont mises en évidence à travers la relation des IDG régionales aux plateformes infrarégionales. En l'occurrence, les jeux d'échelle sont très marqués à travers la structure centre-périphérie du graphe. La plupart des IDG nationales occupe une place centrale. Une première couronne périphérique donne un panorama des IDG régionales qui renvoient elles-mêmes vers une deuxième couronne périphérique composée majoritairement d'IDG initiées par des collectivités territoriales, des associations ou des syndicats départementaux ou intercommunaux.
- Des composantes thématiques du réseau apparaissent à travers les mises en relation d'IDG travaillant dans le domaine de la mer ou du côtier et

⁸ Densité d'un graphe = nombre de liens existants / nombre de liens possibles

sur des territoires ultramarins (lien Sextant – GéoMartinique ou PEIGEO) par exemple.

- Certaines infrastructures utilisant peu les techniques de moissonnage (nœuds de petite taille) occupent pourtant une place centrale car elles sont très moissonnées. C'est le cas d'EauFrance et de Géo-IDE.
- Enfin, le géocatalogue national occupe, sans surprise, le nœud central du réseau. Cependant, le portail « data.gouv.fr », bien que ne constituant pas une IDG *stricto sensu*, est lui-aussi fortement connecté au réseau.

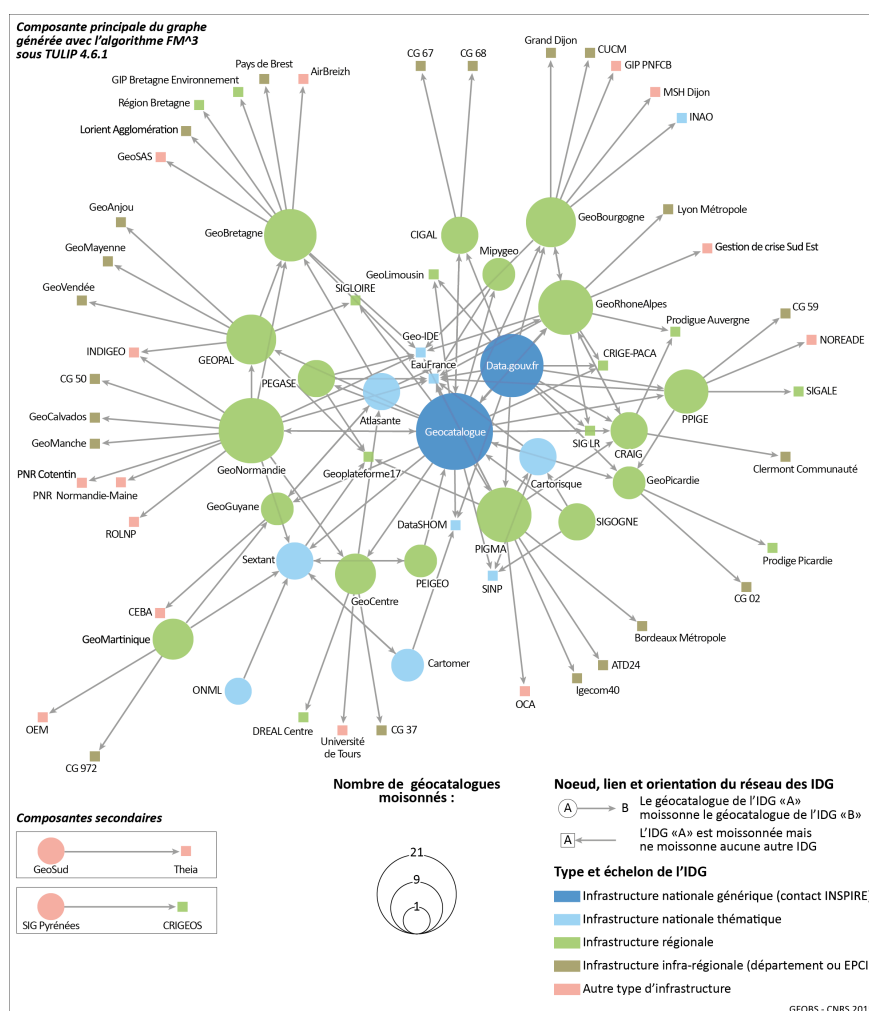


FIGURE 2. Réseau des relations inter-IDG établi par l'analyse des moissonnages entre les plateformes recensées.

3.2.4. Les IDG facilitent-elles l'hybridation de données issues de sources variées ?

Un focus sur une étude de cas régionale offre une autre lecture de la mise en réseaux des outils. Ainsi, après avoir regroupé les 2120 fiches de métadonnées issues de cinq portails de données géographiques déployés en Guyane française, nous calculons leur indice de similarité à partir des mots-clés renseignés pour chaque fiche de métadonnées. Cet indice permet la production d'un graphe (figure 3) pour visualiser les proximités thématiques de l'ensemble des différents éléments (métadonnées) du corpus. On voit alors assez nettement se reformer le découpage par IDG. Seulement trois composantes (A, B, C) mixent des métadonnées issues d'IDG différentes. Les « jonctions » observées sont thématiques (la composante A est exclusivement liée aux habitats et biotopes, la composante B ne concerne que des données relatives à l'imagerie satellite) et organisationnelle (au sein de la composante C, le pont entre les métadonnées de GéoGuyane et celles de Guyane SIG est assurée en C1 par les données du Parc Amazonien de Guyane, établissement public partenaire des deux infrastructures). De plus, aucune hybridation entre sphère scientifique et sphère institutionnelle n'est observée.

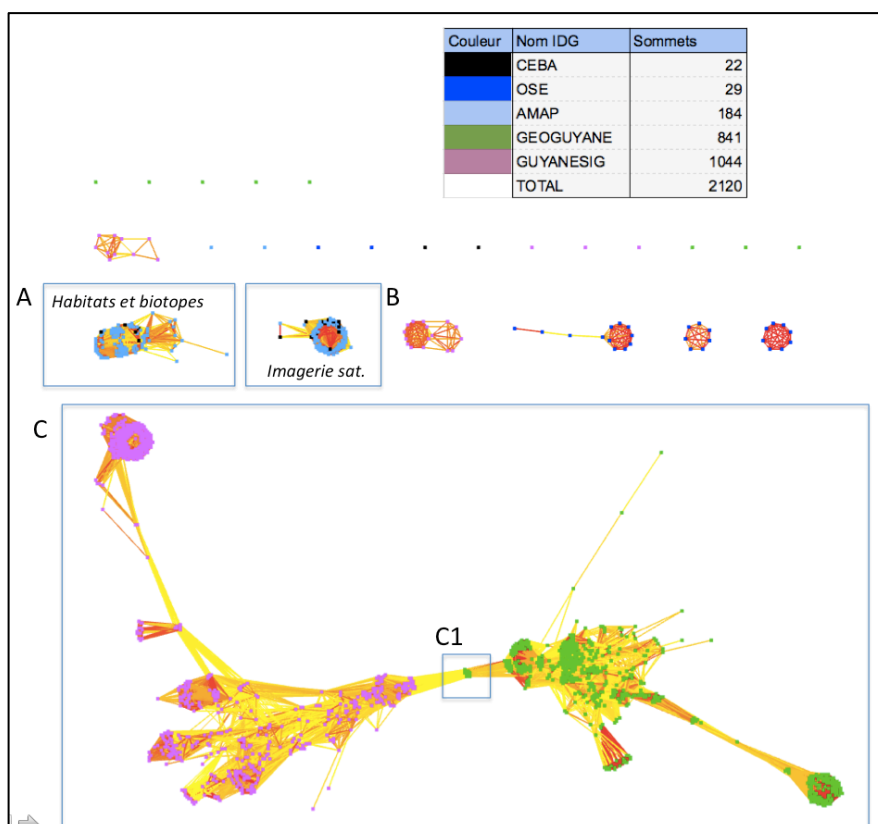


FIGURE 3. Graphe de similarité thématique des fiches de métadonnées de Guyane.

3.3. Mise en réseaux des acteurs

3.3.1. Quel est le discours des acteurs sur les dynamiques de géocollaboration ?

La plupart des IDG accompagne le déploiement de leur dispositif technique d'une série d'actions d'animation visant à favoriser l'appropriation des outils ou des données par un large spectre d'acteurs, et à impulser une dynamique autour d'un thème ou d'un territoire donné. Les formations aux outils, aux données, aux aspects réglementaires ou thématiques côtoient des actions de communication et l'animation de groupes de travail. Ce faisant, ces plateformes cherchent à initier de véritables communautés de pratique pour favoriser des dynamiques de géocollaboration et enrôler une grande diversité d'acteurs, même s'il existe des différences notables entre les échelons national (plus homogène) et régional (figure 4).

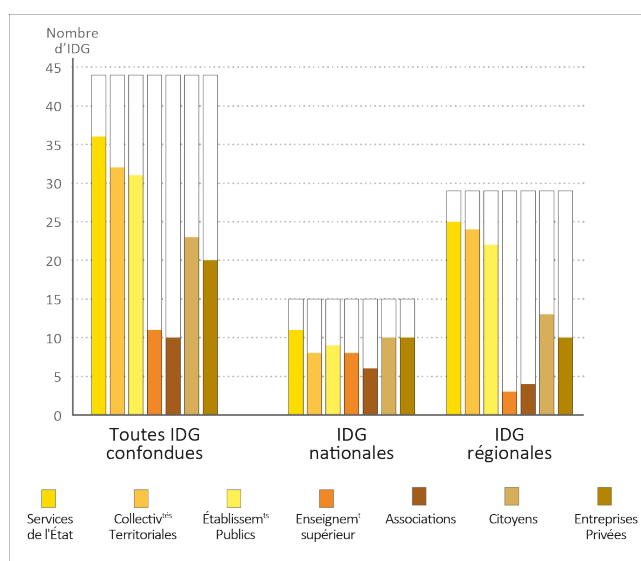


FIGURE 4. Répartition des types d'acteurs visés par les IDG régionales et nationales.

3.3.2. Les IDG favorisent-elles des dynamiques de géocollaboration ?

Environ la moitié des IDG du corpus anime un pôle métier, assimilé à un lieu d'échange dédié à une thématique et/ou à des outils pour favoriser la mise en cohérence des données et leur partage. Si en moyenne chaque IDG anime 5 à 6 pôles métier, des différences importantes sont à souligner entre les plateformes nationales (dont un tiers anime un pôle) et les plateformes régionales (dont les deux tiers animent un pôle). Les 132 pôles métier fonctionnels au sein des 45 IDG du corpus concernent un panel très diversifié de thématiques qui permet de comprendre les principaux chantiers sur lesquels travaillent actuellement les géomaticiens : occupation des sols, urbanisme, forêt, voirie, télécommunication, foncier, etc.

3.3.3. *Les IDG permettent-elles la mise en réseau d'acteurs diversifiés ?*

L'analyse des métadonnées révèle la faible diversité des contributeurs des IDG. Deux exemples peuvent illustrer ce constat. Au niveau national, le géocatalogue du BRGM accueillait, à l'été 2016, une très large majorité des données de l'État (49%) dont une part importante provient des services déconcentrés tels que les directions départementales (37%) qui gèrent le plan cadastral. La très faible présence des entreprises privées (3%) et des ONG (3%) est à souligner. Chaque IDG reste donc compartimentée dans des sphères d'acteurs spécifiques. Au niveau régional, le constat est identique. Ainsi, l'infrastructure régionale « GéoLimousin » propose essentiellement des données de l'État (93% en 2015) bien qu'elle soit la seule plateforme de la région.

4. Discussion et perspectives

4.1. *Un fonctionnement en silos qui perdure*

Si l'interopérabilité des systèmes d'information a incontestablement permis des avancées majeures pour la normalisation des services web et l'interconnexion potentielle des infrastructures de données, nos analyses concernant le moissonnage des géocatalogues tendent cependant à en limiter leur portée effective. Les IDG étudiées restent compartimentées dans des sphères d'acteurs et des thématiques spécifiques. L'analyse de l'ensemble du corpus montre des possibilités de moissonnage encore faiblement exploitées et l'étude de cas guyanais illustre une forme de fragmentation de l'information dans laquelle chaque plateforme reste dans son périmètre géographique et dans son silo thématique même si le potentiel de croisement existe.

4.2. *Des communautés actives mais peu diversifiées*

La mise en réseau des acteurs via des dynamiques de collaboration associées au déploiement des IDG semble effective. Les 132 pôles métier identifiés et l'activité importante d'animation des IDG, en particulier celles de l'échelon régional, témoignent de communautés de pratique très actives. Elles jouent un rôle clé dans la fédération des acteurs du territoire pour co-concevoir des modèles, co-produire des jeux de données et/ou les co-contrôler. Cependant, ces dynamiques de géocollaboration s'observent aujourd'hui dans un entre-soi peu diversifié : les administrations d'un côté, la recherche de l'autre, le secteur privé non représenté et le milieu associatif encore marginal.

4.3. *Perspectives*

Cette étude illustre la nature complexe, dynamique et multiforme des IDG. Nous favorisons une approche exploratoire multi-disciplinaire pour aborder leur

complexité et comparer la théorie et la pratique. Si les IDG apparaissent comme un « label » porteur pour les autorités publiques, on constate différentes formes d'acceptation du terme qui témoignent de réalités (technique et organisationnelle) variées. Le cadre méthodologique présenté est une première étape pour comparer les discours aux données effectivement partagées et aux actions de géocollaboration mises en oeuvre. Une deuxième étape consistera à mettre en parallèle les discours des promoteurs, les données et services des IDG et les usages effectifs qui y sont associés (Georis-Creuseveau et al., 2016). L'objectif sera d'évaluer l'impact du partage de données géographiques sur les pratiques de gestion territoriale en analysant finement les usages effectifs et en vérifiant leur cohérence par rapport aux objectifs initiaux des IDG étudiées.

Bibliographie

- Ballatore A. (2014). Exploring the geographic information universe : the role of search technologies. CEUR Workshop Proceedings 1273, Workshop on Geographic Information Observatories at *GIScience*, Vienna, October 2014, pp. 23–28
- Coetzee S., Wolff-Piggott B. (2015) A Review of SDI Literature: Searching for Signs of Inverse Infrastructures. *Cartography - Maps Connecting the World*. Part of the series Lecture Notes in *Geoinformation and Cartography* pp 113-127.
- Crompvoets J., Bregt A., Rajabifard A., Williamson I. (2004). Assessing the world wide developments of national spatial data clearinghouses. *International Journal of Geographical Information Science*, 18, 665-689.
- Georgiadou Y., Rodriguez-Pabon O., Lance K.T. (2006). SDI and e- Governance: A quest for appropriate evaluation approaches, *URISA Journal: Journal of the Urban and Regional Information Systems Association*, vol. 18(2).
- Georis-Creuseveau J., Claramunt C., Gourmelon F. (2016). A modeling framework for the study of spatial data infrastructures applied to coastal management and planning. *IJGIS*, <http://dx.doi.org/10.1080/13658816.2016.1188929>
- Hendriks PHJ, Dessers E., Van Hootegen G. (2012). Reconsidering the definition of spatial data infrastructure. *IJGIS*, 26(8):1479-1494
- Johnson R.B. & Onwuegbuzie AJ. (2004). Mixed Methods Research: A Research Paradigm Whose Time Has Come. *Educational Researcher*. 33:14-26.
- Kitchin R. & Lauriault T. (2014). Towards Critical Data Studies: Charting and Unpacking Data Assemblages and Their Work, in Eckert, J., Shears, A. and Thatcher, J. (eds) *Geoweb and Big Data*. University of Nebraska Press. URL: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2474112
- Masser I. (2010). *Building European spatial data infrastructures*, ESRI Press, Redlands.
- Noucher M., Gourmelon F., Gautreau P., Maulpoix A., Pierson J., Pissoat O., Rouan M. (soumis). Spatial Data Sharing: A Pilot Study of the French SDIs, *IJDE*.
- Rajabifard A., Feeney M-E., Williamson I., Masser I. (2003). «Chapter 6, National SDI Initiatives», In *Development of Spatial Data Infrastructures: from Concept to Reality*, Williamson I, Rajabifard A., Feeney M.-E.F, London: Taylor & Francis, pp. 95-109.