

---

# Analyser la structure de dispositifs de mise à disposition de données géographiques

## Application aux enjeux de suivi et de gestion du trait de côte en France

Johan Desbonnet<sup>1</sup>, Françoise Gourmelon<sup>1</sup>, Christophe Claramunt<sup>2</sup>

1. LETG-Brest (UMR 6554 CNRS), Université Bretagne Occidentale IUEM, rue Dumont d'Urville, 29280 Plouzané cedex  
{johan.desbonnet@univ-brest.fr; francoise.gourmelon@univ-brest.fr}
2. Institut de Recherche de l'Ecole Navale (IRENaV) BCRM Brest, CC 600, Ecole navale, 29240 Brest, cedex 9  
{christophe.claramunt@ecole-navale.fr}

---

**RÉSUMÉ.** Dans le cadre des enjeux de suivi et de gestion du trait de côte, le besoin d'informations et de connaissances sur le milieu côtier accessibles sur le Web est décrit comme un enjeu majeur. Néanmoins sur cette thématique, la stratégie des promoteurs de dispositifs mettant à disposition de l'information géographique numérique reste encore peu connue. Cet article décrit une approche méthodologique en vue d'étudier le volet technique des possibilités de consultation de l'information géographique sur des dispositifs en ligne susceptibles d'être mobilisés dans le suivi et la gestion du trait de côte en France. L'accès aux données géographiques sur ces dispositifs est permis par la présence de liens hypertextes et de services web. L'analyse de ces deux composants structurels révèle une situation très disparate en matière de possibilités d'accès à l'information géographique.

**ABSTRACT.** In the context of shoreline management and monitoring issues, inferring information and knowledge from the Web are considered as a crucial issue. However, Spatial Data Infrastructures promoters' strategies and content are still unknown. This article presents a methodological approach to study the technical possibilities of access to the digital geographic information on online platforms often used for shoreline management and monitoring in France. Geographical data access on such platforms is granted by the presence of hyperlinks and Geospatial Web services. The analysis of this structural components (hyperlinks and Web services) shows a disparate situation for the access to geographical information.

**MOTS-CLÉS :** information géographique numérique, infrastructures de données géographiques (IDG), plateformes de données ouvertes, hypertexte, services web, suivi et gestion du trait de côte.

**KEYWORDS:** digital geographical information, Spatial Data Infrastructures (SDI), Open Data platforms, hypertext structure, Web services, shoreline management and monitoring.

---

DOI: [10.3166/riig.2019.00079](https://doi.org/10.3166/riig.2019.00079) © 2019 Lavoisier

## 1. Introduction

Face aux enjeux et aux risques associés à la dynamique du trait de côte et à l'élévation du niveau moyen de la mer, l'accès à l'information géographique numérique concernant le milieu côtier est favorisé *via* des dispositifs en ligne depuis le début des années 2000 (Bersani *et al.*, 2006 ; Georis Creuseveau, 2014). Ces technologies de l'information géographique présentes sur la toile (Web) et comprenant les infrastructures de données géographiques (IDG)<sup>1</sup>, les interfaces de programmation (API) et les plateformes de données ouvertes<sup>2</sup>, sont une des inscriptions de l'action des pouvoirs publics, des organisations non gouvernementales et des acteurs privés dans la période de transformation des modes de diffusion, de régulation et de gouvernance de l'information à l'échelle mondiale<sup>3</sup>, européenne, nationale et infranationale (Mol, 2006).

Ces dispositifs, de plus en plus nombreux sur le Web 2.0, ont pris au cours du temps, de leurs évolutions techniques, et des objectifs fixés par leurs promoteurs (services de l'État, établissements publics, collectivités territoriales, ONG) des dénominations et des formes variées : « Géoportails », « bases de données géographiques », « API et mashup cartographiques », « hypercartes Web »<sup>4</sup>. Ils visent à faciliter l'accès et les usages de l'information géographique dans certains cadres d'action publique thématiques (mers, océans et zones côtières, urbanisme, risques, conservation et gestion des ressources, etc.), territoriaux (région, département, intercommunalité) et peuvent être développés pour et par des acteurs spécifiques (Pomon, 2015 ; Georis Creuseveau *et al.*, 2016). Les travaux portant sur ces dispositifs numériques ont montré qu'ils pouvaient être influencés par des facteurs techniques, politiques, organisationnels, économiques et sociétaux (Rajabifard *et al.*, 2002 ; Noucher, 2013 ; Georis Creuseveau, 2014).

Dans les territoires côtiers français, les pouvoirs publics ont élaboré entre 2011 et 2012 une stratégie nationale de gestion intégrée du trait de côte (SNGIT) dans laquelle

---

1. Les infrastructures de données géographiques (IDG), régies par des politiques et des normes, ont l'objectif de favoriser la mise en réseau des producteurs et des utilisateurs de l'information géographique (Rajabifard, 2002 ; Georis Creuseveau, 2014 ; Noucher *et al.*, 2017).

2. Les plateformes de données ouvertes étudiées sont le résultat de la demande de la société civile et de l'investissement des pouvoirs publics dans l'ouverture de leurs données publiques. Ce sont des dispositifs qui s'appuient sur des algorithmes, des méthodes au service d'une politique publique et qui mettent à disposition ou renvoient vers des données sous une licence dite ouverte pour permettre le développement d'autres applications et d'autres usages (Colin et Verdier, 2015).

3. Convention d'Aarhus au niveau international ; directives Inspire, Commission d'accès aux documents administratifs (CADA) et sur la réutilisation des informations du secteur public (PSI) aux niveaux européen et national, loi pour une République numérique dite aussi loi Lemaire au niveau national, loi Notre au niveau régional, démarches d'ouverture des données à l'échelle locale (municipale, intercommunale).

4. Une API est une interface de programmation support de fonctionnalités spécifiques (cartographiques, statistiques et graphiques) et qui prend le plus souvent la forme de services Web (Mericksay, 2011). Une hypercarte Web est une forme d'API définie comme « une application cartographique en ligne dynamique et/ou interactive pouvant associer différents contenus hypertexte et multimédia » (Mericksay, 2011).

les questions de partage et de communication de l'information géographique numérique sur des dispositifs dédiés ont été érigées en axe prioritaire<sup>5</sup>. Les discours véhiculés par les rapports institutionnels des pouvoirs publics autour de la SNGIT nous ont conduit à nous intéresser au rôle d'un certain nombre d'IDG (en gras et soulignées dans le [tableau 1](#)) qualifiées « d'observatoires et de systèmes d'information sur le littoral (SIL) » (Suanez *et al.*, 2012). Afin d'étudier leur positionnement dans la gouvernance informationnelle numérique, nous avons tenté de comprendre comment ces dispositifs s'inséraient sur le Web d'un point de vue technique. A cet égard, nous avons analysé le parcours offert par leur trame hypertextuelle (liens hypertextes, pages Web et niveaux d'entités Web) et les possibilités de consultation de la donnée géographique auxquelles il donne accès par la présence de services web et de conventions. Afin de situer ces observatoires et SIL dans un écosystème informationnel plus large, ce corpus a été complété par un ensemble d'infrastructures thématiques, territoriales et généralistes pouvant appuyer les politiques publiques en lien avec les enjeux de suivi et de gestion du trait de côte en France ([tableau 1](#)). De l'inventaire de l'Association française pour l'information géographique (AFIGEO) réalisé en 2014, a été extrait un premier échantillon d'IDG généralistes et thématiques dites « inspiro-compatibles »<sup>6</sup> (AFIGEO, 2014). Certaines infrastructures absentes de cet inventaire, plus anciennes et liées aux activités des services de l'État et des établissements publics, ont aussi été retenues telles que Carmen, Cartélie, la Base d'observations pour le suivi des côtes (Bosco), les observatoires ministériels de cartographie thématique et statistique (Geoclip Geo Idd). Enfin, certains supports plus récents prenant la forme d'interfaces de programmation applicatives (API) telles que les hypercartes web et les mashups cartographiques développés dans le cadre de projets spécifiques (API Vigicrues, API Rivages) ont aussi été incorporés à l'analyse. En parallèle et en lien avec la mouvance « open data », nous avons choisi d'intégrer la plateforme de données ouvertes nationale qui s'est peu à peu installée dans « l'écosystème de la donnée » (Mabi, 2015) sur le Web depuis la fin des années 2000 telle que data.gouv.

Notre corpus est constitué de 43 dispositifs aux échelles supranationale (Europe), nationale, régionale et intercommunale ([tableau 1](#)).

L'analyse des composantes hypertextuelles et des services web présents sur ces dispositifs a pour objectif d'interroger les différentes possibilités de navigation et d'accès à la donnée géographique numérique. S'intéresser aux questions de consultation de l'information géographique sur une thématique donnée et sur des dispositifs numériques tels que les IDG, les API cartographiques et les plateformes de données

---

5. La stratégie nationale de gestion intégrée du trait de côte (SNGIT) est une politique publique française initiée en 2011 dédiée aux enjeux de suivi et de gestion du trait de côte (risques, aménagement du territoire, etc.) et dont l'axe A « Développer et partager la connaissance sur le trait de côte » est consacré à la question informationnelle. <http://www.geolittoral.developpement-durable.gouv.fr/strategie-nationale-de-gestion-integree-du-trait-r434.html>

6. Respectant les normes de la directive Inspire c'est-à-dire en présentant des services de consultation (catalogue, visualiseur, site éditorial, services Web) et de partage de l'information géographique (téléchargement, cartothèque).

Tableau 1. Corpus étudié (n = 43)

Type	Echelle	Nom
IDG identifiées comme « observatoires et systèmes d'information sur le littoral (SIL) » (n = 15)	Supranationale (n = 3)	Emodnet, Sedatanet, API « Coastline » (absence de site Web dédié)
	Nationale (n = 10)	API « Rivages », Anemoc, Candhis, Bosco, Geolittoral, Onml, Datashom, Sextant, Cartomer, Dynalit
	Régionale /interrégionale (n = 2)	Observatoire du littoral Bas-Normand (Oliban), Réseau d'observation du Littoral Normand et Picard (Rolnp)
Autres IDG généralistes et thématiques françaises (n = 27)	Supranationale (n = 2)	Géoportail Inspire Europe, portail de données de l'Agence européenne de l'environnement (AEE)
	Nationale (n = 17)	API Vigicrues, base de données historiques sur les inondations (BDHI), Georisques, Infoterre, EauFrance, inventaire national du patrimoine naturel (INPN), Géoportail national de l'urbanisme, Carmen Nature France, ADéLie, Cartélie, Geo-IDE ; Geocatalogue national, Géoportail national, Observatoire des territoires, portail de données publiques de Météofrance, applications Geoclip Géo (hors ligne depuis début 2018) Et Geoclip Stat du service de l'observation et des statistiques
	Régionale /interrégionale (n = 7)	Geobretagne, Bretagne Environnement, Kartenn Bretagne, Geonormandie, Geopal, Indigeo, Osuna
	Communale /intercommunale (n = 1)	Géocatalogue de Lorient agglomération
Plateforme de données ouvertes (n = 1)	Nationale (n = 1)	Datagouv

ouvertes permet de traiter des questions relatives à leurs modes de communication dans « l'écosystème numérique » qu'est le Web (Cardon, 2011). La connaissance de leurs dimensions structurelles s'appuie sur les principes de fonctionnement de l'Internet et du Web 2.0 et constitue à ce titre un enjeu de recherche pour appréhender les processus qui orientent le parcours de l'information géographique numérique et qui sous-tendent à l'accessibilité des contenus du Web dans la gouvernance informationnelle. Le Web peut être considéré comme un espace numérique ouvert façonné par des intentionnalités et des actions d'acteurs qui affichent, diffusent et partagent de l'information sur des thématiques dédiées dans un réseau interconnecté (Beaude, 2012 ; Graham et Dutton, 2014). À ce titre, Internet et la toile constituent un « objet d'investigation » (Gautreau *et al.*, 2013 ; Desbois, 2015 ; Severo et Venturini, 2016 ; Noucher, 2017) permettant de caractériser certaines pratiques et stratégies « info-communicationnelles » (Pinède, 2018) autour de la mise à disposition, de la visibilité et de la mise en réseau de l'information numérique (Gautreau *et al.*, 2013 ; Graham et Zook, 2013 ; Noucher *et al.*, 2017).

Ce travail de recherche vise donc à analyser la contribution d'un ensemble de dispositifs mettant à disposition de l'information géographique numérique sur le Web, potentiellement utilisable sur des enjeux de suivi et de gestion du trait de côte en France (risques côtiers, aménagement du territoire, protection et conservation de la biodiversité, gestion de l'eau, etc.). Pour ce faire, nous proposons d'explorer certaines caractéristiques structurelles de ces dispositifs à partir de méthodes de recherche dites numériques (Rogers, 2009, 2015 ; Plantin et Monnoyer-Smith, 2013). Ces méthodes font référence aux « recherches en sciences humaines et sociales qui portent sur des objets numériques ou numérisés » (Plantin et Monnoyer-Smith, 2013). Dans un premier temps, elles se traduisent dans notre étude par l'emploi d'un logiciel (« crawler ») qui permet d'indexer et de capturer les ressources hypertextuelles présentes sur les dispositifs étudiés (liens hypertextes, pages et entités Web renvoyées). Dans un second temps, le recensement des services web permettant de consulter de la donnée géographique d'un dispositif à l'autre est réalisé manuellement sur chacun des dispositifs et à partir des résultats d'une enquête nationale sur les IDG datant de 2015 et 2017 (<http://geobs.cnrs.fr/>). Ces composants structurels (liens hypertexte, pages et services web, entités Web) permettent de caractériser le positionnement de chacun des dispositifs en lien avec le milieu côtier par rapport à des indicateurs de dispersion statistique, au contenu des pages Web renvoyées (niveau d'information) et par rapport aux possibilités de consultation des données offertes (mise en réseau et mesures de graphe) (figure 1).

## **2. Le Web comme espace structurel d'investigation : éléments méthodologiques**

### ***2.1. L'analyse hypertextuelle : un indicateur de la diversité des ressources offertes par les dispositifs numériques***

Tandis que les acteurs publics s'inscrivent dans cette période « d'injonctions de la gouvernance de l'information » (Le Corf, 2015), la structure hypertextuelle des sites web des IDG, des plateformes de données ouvertes et des API cartographiques est une première porte d'entrée technique sur la consultation des contenus informationnels. Elle permet à l'utilisateur de parcourir et d'accéder à des ressources sur le site web de l'IDG ou de la plateforme de données ouvertes qu'il consulte, à partir de liens structurés

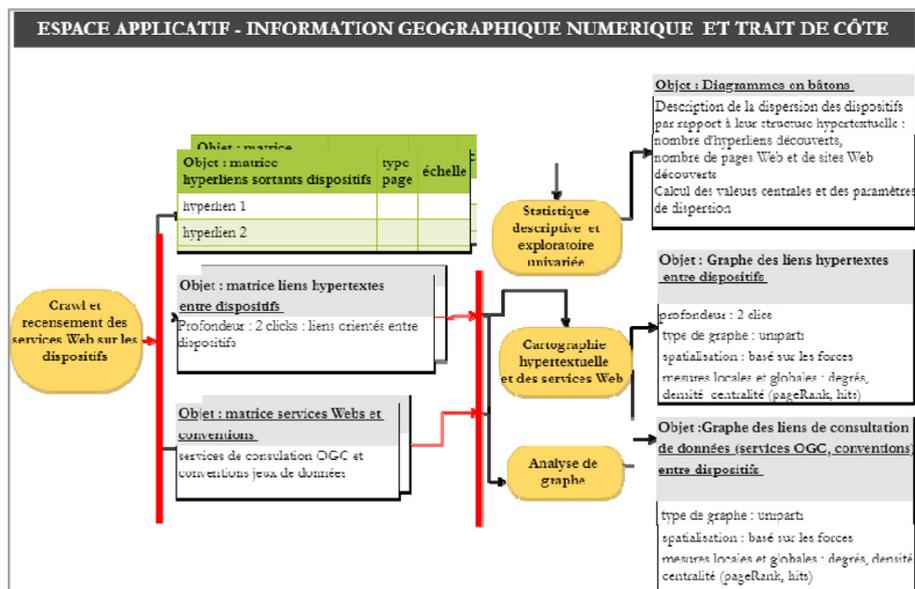


Figure 1. Présentation de la démarche d'analyse de la structure des dispositifs

renvoyant vers différents espaces de consultation sous la forme de pages (par exemple un document sur un autre site web, un espace de téléchargement, cartes, données en ligne etc.). Une telle armature hypertextuelle offre la possibilité d'observer la trame globale du dispositif et non de se concentrer sur un seul composant que serait le visualiseur cartographique, le catalogue ou bien encore l'information géographique diffusée (carte, graphique, donnée vectorielle etc.). Les hyperliens peuvent être considérés d'un point de vue du numérique (*Digital Methods Initiative*<sup>7</sup>) comme des indicateurs de richesse et de diversité informationnelle, notamment les liens sortants (Rogers, 2009). Ils mettent en évidence la redirection du site vers d'autres sources potentielles d'information auxquelles l'utilisateur n'aurait pas accès dans le dispositif visité. Les hyperliens, en permettant la mise en réseau des différents espaces informationnels, peuvent favoriser l'accès à « de nouveaux espaces de connaissances » (Pinède, 2014). Par ailleurs la « nature documentaire » du Web (Cardon, 2011) peut être caractérisée par sa structure topologique dont les liens entrants et sortants sont révélateurs de phénomènes de centralité et de hiérarchie (Kleinberg, 1999 ; Page *et al.*, 1999 ; Chakrabarti, 2002). Les liens entrants sont considérés comme des indicateurs « d'autorité » sur le Web, c'est-à-dire qu'un dispositif possède une autorité d'autant plus forte que la somme des degrés sortants du dispositif qui la mentionne est élevée. Inversement, les liens sortants mettent en évidence le phénomène de « hubs » définis comme des pages web contenant un grand nombre de liens sortants vers d'autres pages

7. Le « Digital methods initiative » ou « courtant de l'Internet » est un groupe de recherche européen qui s'intéresse aux stratégies de diffusion de l'information par les hyperliens, <https://www.digitalmethods.net/Digitalmethods/WebHome>



Figure 2. Mode de découpage des composants hypertextuels sur le Web selon le crawler Hyphe

web qui elles-mêmes possèdent une forte autorité. C'est pourquoi les liens hypertextes sont aussi définis dans la littérature comme le résultat de « politiques » (Rogers, 2009) et de stratégies de mise en visibilité (Gautreau *et al.*, 2013 ; Pinède, 2014) qui peuvent caractériser l'association ou la distance entre les acteurs porteurs des dispositifs.

Afin d'étudier la structure hypertextuelle de notre corpus, la méthode appliquée consiste à explorer, indexer et catégoriser les pages web, les liens renvoyés et les entités web découverts à partir d'un programme similaire à ceux utilisés par les moteurs de recherche. Un « crawler » est un logiciel qui permet de collecter les pages, les liens hypertextes et les entités Web renvoyées par les sites Web (figure 2) et ainsi de reconstituer « un espace de cheminement pour l'utilisateur du Web » (Rogers, 2009). Le « crawler » indexe les pages jusqu'à une certaine distance dans le site (niveau de profondeur), c'est-à-dire un nombre de clics en partant de la page du site web que l'on renseigne au « crawler » (ici  $n=2$ )<sup>8</sup>. Dans notre étude, le « crawler » utilisé est le logiciel Hyphe<sup>9</sup>. Une fois le crawl réalisé en partant de la page d'accueil de chaque dispositif (IDG, API ou plateformes de données ouvertes), des résumés statistiques et graphiques sont fournis sur le nombre de pages web, sur le nombre de liens hypertextes et d'entités web renvoyés par le dispositif à la profondeur de 2 clics. Par exemple en partant de la page d'accueil de l'IDG Geolittoral (dispositif national ressource en matière de politiques publiques sur le domaine côtier français), le logiciel Hyphe a moissonné 434 pages web internes à une distance de 2 clics de profondeur. Il a découvert et récupéré sur

8. Ce choix est motivé par le fait que la profondeur de 2 correspond à l'espace où l'on trouve les liens prioritaires dans l'environnement des sites Web (Jacomy et Ghitalla, 2007). C'est généralement à ce niveau que les pages les plus importantes se situent, ce qui peut s'expliquer par le fait que plus les pages sont profondes moins les *crawlers* des moteurs de recherche leur accordent de l'importance.

9. Hyphe est développé par le Sciences Po Médialab dans le cadre du projet Hypertext corpus initiative en 2010 (Girard *et al.*, 2012), <http://hyphe.medialab.sciences-po.fr/>

ces 434 pages web un total de 14 230 liens hypertextes, de 13 593 adresses de pages web renvoyant à 47 entités web externes. Ces valeurs extensives permettent dans un premier temps d'analyser la distribution statistique des ressources hypertextuelles offertes par chaque dispositif (calcul des valeurs centrales et des paramètres de dispersion) et d'étudier le cheminement offert par cette même structure.

Dans un second temps, nous présentons d'une part, le niveau de spécificité des pages web découvertes par le crawler (pages d'accueil versus pages dédiées) et d'autre part, la répartition spatiale des entités web externes renvoyées par les liens hypertextes présents sur les observatoires et SIL.

## **2.2. Les liens hypertextes et les services web géographiques: des indicateurs de la composante réseau des dispositifs sur la toile**

Les notions de réseau renvoient à des questions d'interopérabilité pour les dispositifs numériques sur le Web (Gasser, 2015), notamment ceux dédiés à l'information géographique (Rajabifard *et al.*, 2002 ; Noucher, 2013 ; Georis Creuseveau, 2014). En France, la mise en réseau par l'interopérabilité des IDG institutionnelles et des systèmes d'information publics de manière plus générale est promue par la directive Inspire et par la politique de la donnée de l'État<sup>10</sup>. Le principe FAIR (*Findable, Accessible, Interoperable, Reusable*) s'impose aussi pour les infrastructures de données du monde de la recherche publique (Wilkinson *et al.*, 2016). La structure en réseau de ces dispositifs peut dès lors être caractérisée par une analyse et une comparaison de graphes au niveau hypertextuel et au niveau des services de consultation de la donnée (services web, autres). Le graphe est un outil couramment employé dans la représentation et la modélisation de la topologie du Web. C'est avant tout un objet mathématique défini « comme un ensemble fini de sommets N et un ensemble fini de liens L pouvant être vide, noté  $G = (N ; L)$  » (Beauguitte, 2010).

Afin d'étudier l'existence ou non d'un réseau de relations entre les IDG, API et la plateforme de données ouvertes, un crawl a été lancé à partir du logiciel Hyphe sur les pages d'accueil des 43 dispositifs du corpus en août 2017 pour représenter les liens entre dispositifs au niveau hypertextuel. Parallèlement et en vue d'analyser la mise en réseau des dispositifs par l'intermédiaire de « services web géographiques »<sup>11</sup>, nous avons

10. La « politique de la donnée » a conduit à la création d'un poste « d'administrateur général des données » depuis 2014 et vise à « coordonner l'action des administrations en matière d'inventaire, de gouvernance, de production, de circulation et d'exploitation des données par les administrations. » (Décret n 2014-1050 du 16 septembre 2014 instituant un administrateur général des données).

11. Selon Henri Pomon (2015) les services web géographiques sont un « ensemble de technologies permettant à des applications de dialoguer *via* Internet, par l'échange de messages fondés sur des standards web d'accès aux données géographiques ». Ils permettent la consultation de données géographiques d'un dispositif à l'autre par l'intermédiaire de services web de visualisation (WMS-WMTS), de téléchargement (WFS-WCS) et de catalogue (CSW) dont les normes sont fixées par l'Open Geospatial Consortium (OGC).

réemployé les résultats de l'enquête par questionnaire du projet de recherche Geobs<sup>12</sup> sur les services web de moissonnage (standard CSW) déployés par les IDG institutionnelles françaises. Cette enquête a été complétée par le recensement des autres possibilités de consultation de données géographiques présentes sur chaque dispositif (services OGC standardisés de visualisation et de téléchargement, jeux de données d'autres institutions disponibles sous une autre forme que les services web OGC<sup>13</sup>). Ces deux matériaux (enquête du projet de recherche Geobs et recensement de l'ensemble des services) nous ont permis de représenter les possibilités de consultation de données géographiques d'un dispositif à l'autre du corpus. Les relations hypertextuelles et de service ont été cartographiées à l'aide du logiciel Gephi 0.9.1. La spatialisation des infrastructures de données, c'est-à-dire leur représentation sous forme de nœuds et de liens au niveau hypertextuel (n = 43 nœuds et 114 liens) et au niveau des relations de services de consultation (n = 40<sup>14</sup> nœuds et 152 liens), vise à faciliter la lecture du positionnement des dispositifs sur le Web à un instant donné. Cette visualisation s'appuie sur un même algorithme basé sur les forces (Fruchterman et Reingold, 1991) qui permet la comparaison à partir d'une représentation des nœuds selon des principes d'attraction (les plus liés s'attirent) et de répulsion (les moins liés sont repoussés à la périphérie). Enfin, on a appliqué des algorithmes employés dans la caractérisation de la nature topologique du Web et dans l'indexation de son contenu pour mesurer la centralité de chaque dispositif au sein du réseau des 43 dispositifs étudiés (« Autorité » et « Hub » de Kleinberg (1999) et algorithme de classement « PageRank »<sup>15</sup> de Page et al. (1999)). Le PageRank est un algorithme itératif qui va permettre d'attribuer un score sous forme de valeur à chaque nœud selon une méthode basée sur les probabilités. Ainsi chaque dispositif se voit attribuer un score proportionnel au nombre de fois où il serait visité par l'utilisateur qui en parcourant les graphes utiliserait de manière aléatoire les liens et les services présents sur le dispositif. La version actuelle du PageRank implémentée dans le moteur de recherche Google repose sur des critères plus complexes et multiples (topologiques, sémantiques, trafic) dont certains restent encore assez opaques (Cardon, 2013). Nous nous intéressons ici qu'au critère topologique du PageRank qu'on applique aux graphes de dispositifs étudiés.

---

12. <http://geobs.cnrs.fr/>

13. Certaines API cartographiques permettent la consultation de jeux de données sous d'autres formes que les géoservices WMS-WMTS, WFS, WCS. Par exemple, les observatoires territoriaux des services centraux liés à la solution technique Geoclip (Observatoire des territoires, SOES Geoclip Stat et Carto) supportent des flux de données sous le format *Table Joining Service*.

14. N = 40 dispositifs dans le réseau de services et non 43, car le géoportail Inspire européen moissonnant seulement les infrastructures nationales de chaque pays européen n'a pas été intégré dans l'analyse. Par ailleurs, les 4 dispositifs complémentaires Lorient agglomération/Lorient Kartenn grand public et Adélie/Cartélie qui présentent chacun un site web distinct ont été regroupés en deux entités.

15. Le « pagerank » est un algorithme proposé par Larry Page et Sergey Brin (1998) sur lequel repose la classification de l'information sur le moteur de recherche Google. Il peut être employé comme indicateur de centralité et s'appuie sur la notion de popularité.

### 3. Analyse de la structure hypertextuelle et de la structure en réseau des dispositifs numériques : résultats

#### 3.1. La distribution statistique des pages web, liens et entités web : une première mesure de la ressource hypertextuelle

Concernant notre corpus, les résumés statistiques des distributions hypertextuelles (liens, pages web et entités<sup>16</sup> web) montrent qu'une personne souhaitant consulter de l'information sur les dispositifs jusqu'à une distance de 2 « clics », a accès en moyenne à 22 449 pages web, à 11 791 liens hypertextuels et à 129 entités web externes. La valeur modale dans les trois distributions est assez faible, c'est-à-dire que la valeur qui revient le plus fréquemment pour les dispositifs dans les trois séries statistiques est égale à 0 pour le nombre de liens et d'entités web, et égale à 2 pour le nombre de pages web découvertes. Nous trouvons dans les trois cas des distributions dissymétriques à gauche (médiane inférieure à la moyenne), c'est-à-dire qu'une majorité des dispositifs se concentre dans des valeurs faibles en termes de nombre de pages, de liens et d'entités web découverts. Ainsi comme nous pouvons l'observer sur les diagrammes à bâtons (figure 3), certains dispositifs ne présentent aucune ou peu de pages (18 dispositifs sur 43 renvoient à un nombre de pages web inférieur à 100), peu ou aucun lien (19 dispositifs avec un nombre de liens découverts inférieur à 50) et ne renvoient qu'à peu d'entités web externes au site du dispositif (20 dispositifs sur 43 renvoyant à un nombre d'entités web inférieur à 5). Seuls quelques dispositifs thématiques et généralistes nationaux et régionaux détiennent des valeurs élevées en nombre de liens, de pages web et d'entités web (INPN, EauFrance, Datagouv, Geobretagne, Bretagne environnement).

Si l'on observe la répartition des dispositifs dans les trois distributions (pages web, liens hypertextes, entités web) selon un découpage en effectifs égaux (méthode des quantiles), on distingue un premier groupe de dispositifs caractérisés par un faible nombre de liens et de pages web (dispositifs dont la valeur est inférieure ou égale à Q1 sur la figure 3). Dans ce premier groupe, on dénombre 5 dispositifs en lien avec le suivi et la gestion du trait de côte. Le second groupe (25 % des dispositifs avec une valeur supérieure à la valeur du 3<sup>e</sup> quartile Q3) rassemble des dispositifs qui présentent un grand nombre de pages Web et de liens hypertextes tels que l'INPN, EauFrance, Geobretagne, Datagouv. Concernant les dispositifs nationaux et régionaux identifiés dans la Stratégie nationale de gestion intégrée du trait de côte française, aucun dispositif ne se situe dans les 25 % des effectifs aux valeurs les plus élevées pour l'ensemble des indicateurs (nombre de liens, de pages Web et d'entités Web externes). L'IDG européenne Emodnet, les IDG nationales Geolittoral et Sextant et l'IDG interrégionale Rolnp présentent les valeurs hypertextuelles les plus élevées dans 2 indicateurs hypertextuels sur 3 (supérieures à Q3).

L'analyse de la distribution statistique des pages web, des liens hypertextes et des entités web externes a été complétée par l'étude de deux indicateurs statistiques appliqués aux seuls dispositifs identifiés comme observatoires et systèmes

16. L'entité web est considérée comme le niveau général du site Web (domaine et sous-domaine).

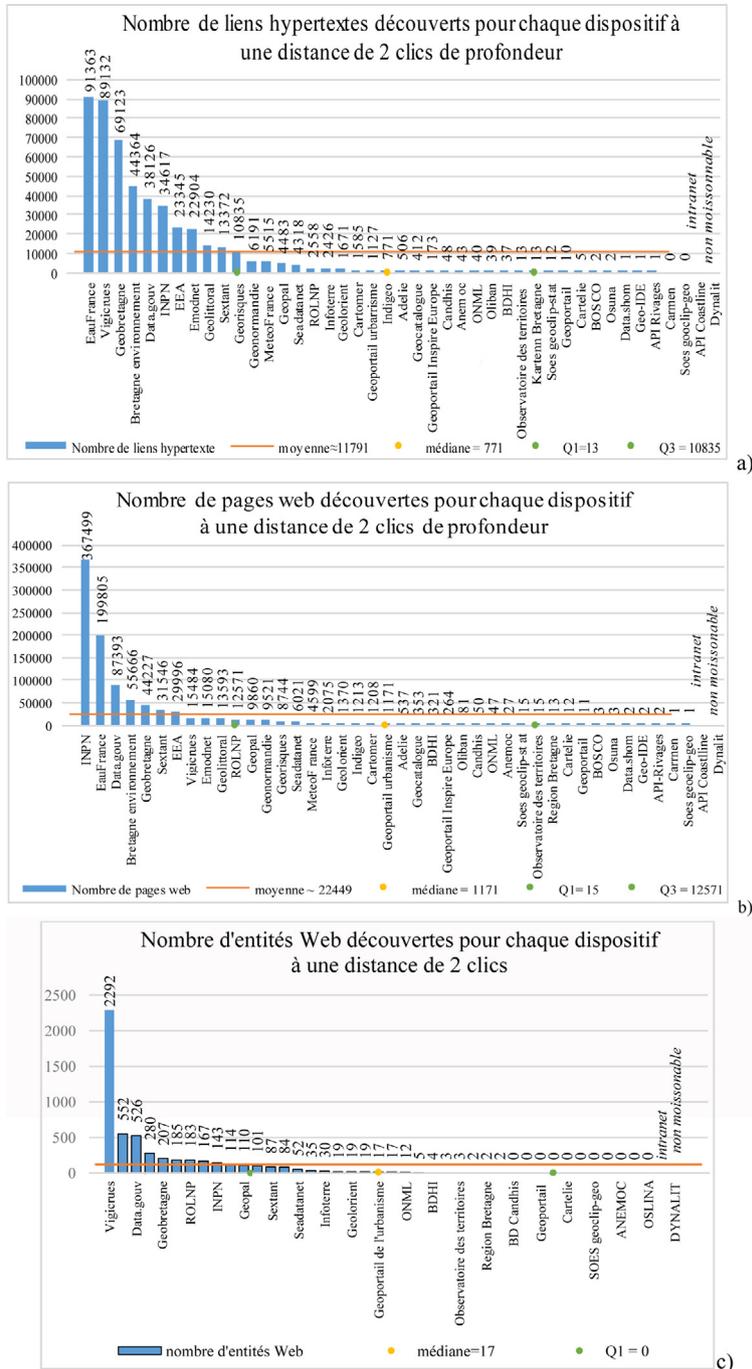
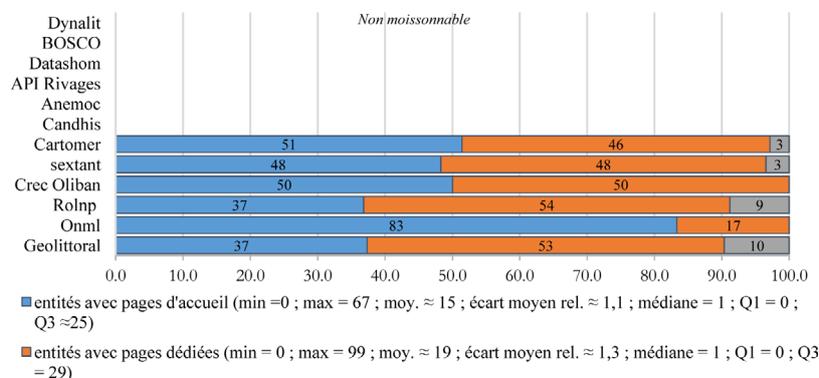


Figure 3. Les trois indicateurs de la structure hypertextuelle : a) nombre de liens découverts, b) nombre de pages web et c) nombre d'entités web (les systèmes d'information sur le littoral – observatoires sont entourés en noir)



*Figure 4. Part des entités web selon le type de pages renvoyées par les dispositifs liés à l'information géographique numérique mer/littoral et suivi du trait de côte*

d'information sur le littoral. D'une part, nous avons cherché à connaître le niveau de spécificité de la page web renvoyée par chaque dispositif sur des sites web extérieurs (figure 4). Ce niveau de spécificité peut être mesuré à partir des liens hypertextes présents sur les pages du dispositif visité. Les liens hypertextes sur lesquels l'utilisateur peut cliquer peuvent renvoyer à plusieurs types de page : une page d'accueil sur une autre entité web externe à celle crawlée (exemples : <https://data.shom.fr/> ou [www.onml.fr/accueil](http://www.onml.fr/accueil)), ou à des adresses de pages web dédiées sur l'entité web externe (exemple <https://data.shom.fr/donnees/catalogue>). Lorsque les liens sont inactifs ou cassés les pages sont introuvables.

Les résultats montrent que 5 dispositifs nationaux<sup>17</sup> moissonnés ne renvoient à aucune page sur des entité web externe (figure 4) : la base d'observation pour le suivi des côtes (Bosco), le dispositif du service hydrographique et océanographique de la Marine (Datashom), l'atlas numérique d'états de mer océanique et côtier (Anemoc), le centre d'archivage national des données de houle et l'application de suivi du trait de côte par smartphone du Cerema (API Rivages). Les liens présents sur les autres dispositifs liés à l'information géographique numérique sur le domaine côtier et le trait de côte renvoient en moyenne plus souvent à des pages web dédiées (moyenne ≈ 19 entités web avec pages dédiées) qu'à des pages web d'accueil (moyenne ≈ 15 entités Web avec pages d'accueil). On notera que l'on trouve en moyenne peu de liens hypertextes ayant renvoyé à des pages introuvables sur les entités Web externes (moyenne ≈ 3). Néanmoins, ces valeurs ne sauraient gommer des distributions statistiques encore très

17. Les dispositifs européens ont été écartés de l'analyse en raison de leur couverture informationnelle dépassant largement le cadre français en renvoyant vers un grand nombre de pages sur l'ensemble des pays européens.

dissymétriques (médiane < moyenne), ce qui témoigne des fortes variations dans la distribution des pages entre les différents dispositifs.

Concernant la répartition des pages web d'accueil et des pages web dédiées renvoyées par chaque dispositif lié à l'information géographique numérique sur le suivi et la gestion du trait de côte, on observe que seulement 2 infrastructures sur les 11 étudiées renvoient à un plus grand nombre d'entités avec des pages dédiées à une information spécifique plutôt qu'à des entités avec des pages d'accueil (figure 4). Ce sont le dispositif national ministériel dédié aux politiques publiques concernant la mer et littoral (Geolittoral) et le dispositif interrégional du réseau d'observation du littoral de Normandie et des Hauts de France (Rolnp) qui renvoient aux plus grand nombre d'entités avec des pages dédiées (figure 4). D'autres dispositifs tels que l'observatoire national de la mer et du littoral (ONML) dont la mission est de « fournir des informations sur toutes les thématiques importantes pour la mer et le littoral »<sup>18</sup> et l'IDG Cartomer<sup>19</sup> dédiée aux aires marines protégées renvoient à une majorité de pages d'accueil sur les sites Web externes auxquels ils donnent accès. Deux dispositifs possèdent une structure équilibrée entre le nombre de pages d'accueil et de pages dédiées : le dispositif national Sextant et l'application cartographique de l'observatoire régional du littoral bas-normand (Oliban).

Enfin, l'analyse a porté sur le niveau de granularité spatiale des entités web externes renvoyées par les liens hypertextes présents sur les dispositifs (figure 5). Elle fournit une première indication géographique sur le niveau d'information et d'acteur présents dans l'environnement des dispositifs dédiés au suivi et à la gestion du trait de côte. La forme des distributions statistiques des entités web renvoyées est globalement dissymétrique sur l'ensemble des échelles à l'exception de l'échelon communal et intercommunal (moyenne  $\approx$  médiane). Ces distributions témoignent du fait qu'une majorité de dispositifs donnent accès à très peu d'entités web au niveau infrarégional et qu'il existe une plus grande dispersion des entités renvoyées par les observatoires et SIL aux échelles régionales et suprarégionales.

Ainsi nous observons (figure 5) que les dispositifs liés à l'information géographique sur le domaine côtier et le suivi du trait de côte renvoient en moyenne davantage à des pages d'entités web d'emprise nationale (n = 15 en moyenne), puis à des pages de sites web d'emprise régionale (n = 9 en moyenne) et internationaux (n = 5). En fréquences cumulées, les pages de sites web régionaux, nationaux et internationaux auxquelles donnent accès les dispositifs liés à l'information géographique sur le trait de côte représentent plus de 75 % de l'effectif des pages Web externes trouvées pour chacun des six dispositifs. Certains dispositifs tels que l'observatoire national mer et littoral

---

18. L'ONML est une IDG créée en 2003 et dont l'objectif est de communiquer des informations à un public large sur les différentes politiques publiques ayant trait au domaine de la mer et du littoral, <http://www.onml.fr/qui-sommes-nous/>

19. L'IDG Cartomer est un dispositif mis en place en 2012 et coordonné par un établissement public (Agence française pour la biodiversité) dont l'objectif est de partager des données géographiques sur la protection et la conservation du milieu marin et notamment sur la thématique de la biodiversité marine (aires marines protégées).

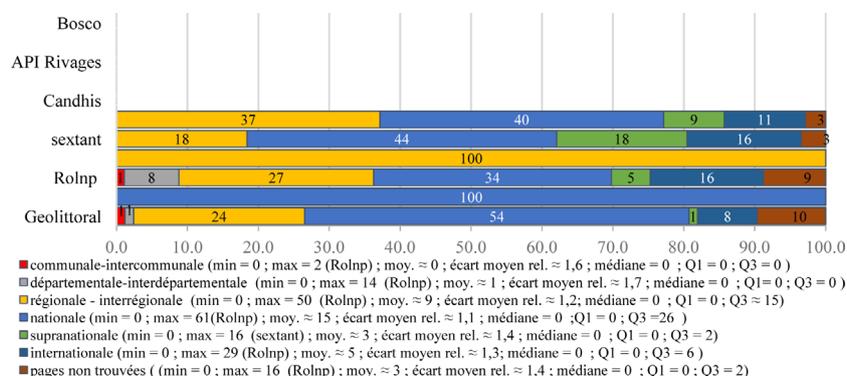


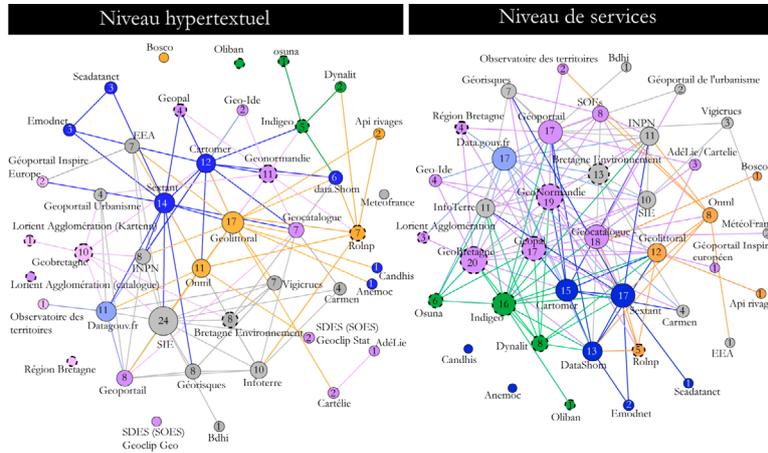
Figure 5. Fréquence des emprises spatiales des entités web renvoyées par les dispositifs liés à l'information géographique sur le domaine côtier et le trait de côte (SIL et observatoires)

(Onml) et l'application cartographique de l'observatoire du littoral bas-normand (Oliban) dirigent l'utilisateur vers des pages externes sur des sites web d'une seule échelle : nationale pour l'ONML et régionale-interrégionale pour l'Oliban. Les autres dispositifs renvoient préférentiellement vers des pages web situées sur 2 ou 3 emprises : nationale et régionale pour les dispositifs Geolittoral et Cartomer ; nationale, régionale et internationale pour les dispositifs Sextant et le Rolnp. Les adresses vers des pages sur des entités web d'échelle infrarégionale sont très minoritaires voire presque absentes des liens hypertextes renvoyés par les dispositifs en lien avec le suivi et la gestion du trait de côte à une profondeur de 2 clics. Seuls deux dispositifs sur 11 présentent des liens hypertextes renvoyant à des adresses de pages web d'échelle communale ou intercommunale (Rolnp et Geolittoral). Ces pages web situées sur des entités d'échelle communale et intercommunale représentent moins de 2 % des entités web renvoyées par les observatoires et des systèmes d'information sur le littoral.

### 3.2. Analyse de la composante réseau des dispositifs numériques : un second indicateur pour caractériser l'aspect structurel

Le second indicateur d'étude de la structure des dispositifs en lien avec le suivi et la gestion du trait de côte met en regard le niveau hypertextuel avec le niveau des services de consultation de la donnée (« services web géographiques ») par l'analyse de graphe (figure 6). Dans les deux niveaux de graphe, la densité<sup>20</sup> est relativement faible sur

20. « La densité (*density*) d'un graphe désigne le rapport entre le nombre de liens présents et le nombre de liens possibles elle varie entre 0 (graphe vide, sans aucun lien) et 1 (graphe complet). » (Beaugitte, 2013).



**Légende :**

Graphes non orientés réalisés sous Gephi 0.9.1 - Algorithme de spatialisation du graphe : Fruchterman Reingold (Paramètres : De manière successive fixer une zone de la taille du graphe de 100 puis de 100 000 puis fixer une gravité de 1000 puis de 100)

	Hypertexte	Services
Taille du graphe	43 noeuds 114 liens	40 noeuds 152 liens
densité	0,063	0,097
degré moyen	2,651	3,8
degré pondéré	2,651	4,975 (3 types de liens)

- dispositifs nationaux et supranationaux (IDG et plateforme de données)
- dispositifs infranationaux 1-20 degrés
- dispositifs informationnels identifiés dans la stratégie nationale de gestion intégrée du trait de côte
- dispositifs informationnels thématiques en lien avec l'action publique territoriale française en matière de gestion et suivi du trait de côte (risques, environnement, biodiversité, statistiques, eau)
- dispositifs informationnels ressources sur l'information géographique numérique côtière et océanique
- dispositifs informationnels généralistes nationaux et régionaux (CRIGES)
- dispositifs informationnels des structures scientifiques en lien avec les projets de recherche sur le suivi du trait de côte
- dispositif informationnel lié à la politique d'ouverture des données
- lien ayant pour source un dispositif identifié comme observatoire dans le cadre de la stratégie nationale de gestion intégrée du TC
- lien ayant pour source un dispositif informationnel ressource sur l'information géographique numérique côtière et océanique
- lien ayant pour source un dispositif informationnel de structure scientifique identifiée dans des actions de suivi du trait de côte
- lien ayant pour source un dispositif informationnel thématique en lien avec l'action publique en matière de gestion et suivi du trait de côte (risques, environnement, biodiversité, eau...)
- lien ayant pour source un dispositif informationnel généraliste national ou régional

Sources : Johan Desbonnet, à partir d'un crawl pour le niveau hypertextuel lancé en décembre 2017 et des enquêtes du programme de recherche Geobs (A. Maulpoix et al., 2015 ; 2017) complété pour le niveau de services par le recensement des différents types de services Web présents sur les dispositifs  
TC : Trait de côte

Figure 6. Mise en regard des graphes au niveau hypertextuel et au niveau des services de consultation de données

l'ensemble du réseau (< à 0,1), ce qui indique que les dispositifs du corpus n'exploitent pas le potentiel maximal de connexion. De plus, on trouve un nombre élevé de composantes<sup>21</sup> connexes, c'est-à-dire de sous-graphes (25 au niveau hypertextuel et 23 au niveau des services), auxquelles s'ajoutent des composantes faiblement connexes (7 au niveau hypertextuel et 4 au niveau des services).

21. « Un composant correspond à un sous graphe non connecté aux autres sous-graphes. » (Beauguitte, 2010).

D'un point de vue graphique, plusieurs IDG thématiques nationales sont regroupées au centre du graphe hypertextuel (figure 6, graphe de gauche), tandis que les IDG généralistes régionales et nationales se situent au centre du graphe du réseau de services (figure 6, graphe de droite). Ce positionnement des IDG s'explique par les distributions de degrés entrants et sortants des deux graphes. En effet, on observe que les dispositifs thématiques tels que Geolittoral, Sextant, l'ONML, Cartomer, l'INPN et le système d'information sur l'eau EauFrance sont situés dans le premier quartile de la distribution des degrés<sup>22</sup> au niveau hypertextuel (degrés > Q3 degrés hypertexte (= 8)). Dans le graphe du niveau de services, ces dispositifs thématiques possèdent un degré moins élevé que les dispositifs généralistes et sont pour une majorité d'entre eux en dehors du 1<sup>er</sup> quartile de la distribution des degrés du graphe des services (degrés < Q3 degrés services (= 13)).

Six dispositifs sont en situation d'isolat<sup>23</sup> au niveau hypertextuel parmi lesquels figurent Bosco et Oliban, identifiés par les pouvoirs publics dans les enjeux de suivi et de gestion du trait de côte. Au niveau du graphe de services, les deux dispositifs Candhis et Anemoc sont également en situation d'isolat. Les « systèmes d'information sur le littoral (SIL) » nationaux (Geolittoral et l'ONML) identifiés par la Stratégie nationale de gestion intégrée du trait de côte et les dispositifs d'information géographique sur le milieu marin (Sextant et Cartomer) sont les plus centraux dans les deux graphes (figure 6). Les observatoires et SIL n'ont aucun lien avec les dispositifs nationaux dédiés à l'urbanisme (Géoportail de l'urbanisme) dans les deux graphes et à la politique nationale de l'eau (inondation, gestion) au niveau du réseau de services.

D'un point de vue hypertextuel, cela se traduit par un potentiel de connexion perfectible entre les dispositifs identifiés comme observatoires et SIL et les autres dispositifs généralistes et thématiques. Les 12 observatoires et SIL partagent un total de 41 liens avec l'ensemble des autres dispositifs thématiques/généralistes sur les 114 liens recensés soit un peu plus d'un tiers des liens hypertextuels du réseau ( $\approx 36\%$  des liens). L'ensemble des dispositifs partagent entre eux 30 liens dits « mutuels »<sup>24</sup> au niveau hypertextuel soit environ 26 % des liens. Sur ces 30 liens mutuels, 8 liens concernent 4 observatoires et systèmes d'information sur le littoral.

Seuls deux dispositifs identifiés comme SIL et observatoires partagent des liens mutuels entre eux (Sextant et Seadatanet). Les SIL et observatoires partagent des liens mutuels avec un dispositif généraliste régional (Geobretagne), avec une IDG thématique (EauFrance) et avec la plateforme de données ouvertes (datagouv).

Au niveau du graphe des services web, les possibilités de consultation des données d'un dispositif en lien avec l'information géographique numérique côtière sur un autre dispositif restent limitées (43 liens avec les autres dispositifs sur 152 liens recensés soit

22. Le degré d'un dispositif est la somme des degrés entrants et sortants.

23. Un nœud ou sommet est dit en situation d'isolat lorsqu'il n'est connecté à aucun autre nœud.

24. Un lien mutuel équivaut à ce que le dispositif  $j$  envoie un lien vers le dispositif  $k$  et que le dispositif  $k$  envoie un lien vers le dispositif  $j$ .

Tableau 2. Fréquence du type de lien entre les dispositifs au niveau des services

Type de liens	Fréquences
Visualise (WMS)	≈ 52 %
Moissonne (CSW)	≈ 22 %
Moissonne et téléchargement (CSW, WFS, WCS)	≈ 10 %
Visualise et téléchargement (WMS, WMTS-WFS, WCS)	≈ 9 %
Moissonne et visualise (CSW, WMS-WMTS)	≈ 5 %
Moissonne, visualise et téléchargement (CSW, WMS-WMTS, WFS-WCS)	≈ 3 %

environ 28 % des liens). De plus les types de services web géographiques (tableau 2) présents d'un dispositif à l'autre cantonnent souvent l'utilisateur à une possibilité de consultation unique (≈ 74 % des liens permettent seulement le moissonnage ou la visualisation). Environ 52 % des liens permettent de visualiser de la donnée d'un dispositif à l'autre et 22 % des liens permettent de moissonner les données géographiques d'un dispositif à partir d'un autre dispositif (tableau 1). Les autres possibilités de consultation de la donnée d'un dispositif à l'autre représentent entre 3 à 10 % des liens sur l'ensemble du corpus.

En appliquant des métriques (algorithmes) propres à l'organisation de l'information sur le Web (« Autorités », « Hubs », « PageRank ») sur le seul réseau des 43 dispositifs étudiés, nous observons un positionnement différent et parfois stratégique des dispositifs les uns par rapport aux autres. Ces algorithmes basés sur les principes de reconnaissance et de popularité permettent d'enrichir la simple notion de centralité fournie par les degrés entrant et sortant. En effet, les IDG nationales thématiques (EauFrance, INPN) et généralistes (Géoportail, Géocatalogue) liées à la conduite de politiques publiques majeures (directive risque inondation, gestion et protection des espaces naturels ; diffusion et gestion de l'information géographique de référence) ainsi que la plateforme de données ouvertes Data.gouv possèdent de forts scores de centralité selon les indicateurs employés sur le Web (tableau 2), ce qui est moins visible sur la cartographie basée sur les forces à partir des seuls degrés (figure 6).

Parmi les dispositifs dédiés à l'information géographique numérique sur le littoral et les enjeux de suivi et de gestion du trait de côte, on retrouve ainsi 5 IDG nationales (Geolittoral, Onml, Datashom, Cartomer et Sextant) dans les plus forts scores de centralité basés sur les métriques du Web mais avec des positionnements différents que l'on peut observer en mesurant leurs scores selon l'algorithme HITS (tableau 2). Au niveau des indicateurs de centralité, seuls 2 dispositifs partagent des scores élevés d'autorité ou de hub dans les deux niveaux de réseau (en gras et soulignés dans le tableau 3 ci-après) : Datashom possède un score élevé d'autorité tandis que Cartomer

Tableau 3. Dispositifs aux scores HITS les plus élevés et les plus faibles dans le corpus des 43 dispositifs étudiés

	Hypertextuel (score HITS°)	Services (score HITS)
Plus forts scores d'autorités : 10 premiers scores (25 % de l'effectif)	EauFrance ≈ 0,41 <b>Onml ≈ 0,39</b> Geocatalogue ≈ 0,29 Inpn (Mnhn) ≈ 0,28 Geoportail ≈ 0,27 <b><u>Datashom ≈ 0,25</u></b> AEE ≈ 0,23 Datagouv ≈ 0,20 ; Vigicruces ≈ 0,20; Géorisques ≈ 0,20	<b>Sextant ≈ 0,34</b> Geoportail ≈ 0,32 Geobretagne ≈ 0,31 <b>Geolittoral ≈ 0,31</b> <b><u>Datashom ≈ 0,30</u></b> EauFrance ≈ 0,30 Inpn (Mnhn) ≈ 0,29 Geopal ≈ 0,24 Infoterre ≈ 0,24 Geonormandie ≈ 0,19
Plus forts scores de hubs : 10 premiers scores (25 % de l'effectif)	<b>Geolittoral ≈ 0,45</b> <b>Sextant ≈ 0,39</b> Eaufrance ≈ 0,38 Geonormandie ≈ 0,32 <b><u>Cartomer ≈ 0,31</u></b> Bretagne environnement ≈ 0,25 Infoterre ≈ 0,21 Geobretagne ≈ 0,20 Datagouv ≈ 0,19 <b>Rolnp ≈ 0,17</b>	Datagouv ≈ 0,38 Indigeo ≈ 0,37 Geonormandie ≈ 0,36 Bretagne environnement ≈ 0,33 <b><u>Cartomer ≈ 0,31</u></b> Geocatalogue ≈ 0,26 Geopal ≈ 0,24 <b>Dynalit ≈ 0,24</b> Geobretagne ≈ 0,22 <b>Onml ≈ 0,16</b>
Plus faibles scores d'autorités : 10 derniers scores ≈ 0 (25 % de l'effectif)	<b>Maprec (Oliban) ;</b> GEOIDD carto ; Géocatalogue Lorient Agglomération ; Kartenn Bretagne ; <b><u>Bosco ;</u></b> MeteoFrance ; Adélie ; Osuna ; Geoclip Observatoire des territoires ; Kartenn Lorient agglomération	<b>Emodnet ;</b> <b>Dynalit ;</b> Vigicruces ; <b>Seadatanet ;</b> BDHI ; <b>API suivi trait de côte (Rivages) ;</b> <b><u>Bosco ;</u></b> Géoportail Inspire européen ; <b>Candhis ;</b> <b>Anemoc</b>
Plus faibles scores de hubs : 10 derniers scores ≈ 0 (25 % de l'effectif)	<b>API suivi trait de côte (Rivages),</b> GEOIDD Stat ; Cartélie ; <b><u>Maprec (Oliban) ;</u></b> GEOIDD carto ; <b><u>Anemoc ;</u></b> Géoportail Inspire européen ; Géocatalogue Lorient Agglomération ; Geo-Ide ; Kartenn Bretagne	Géoportail ; <b>Geolittoral ;</b> EauFrance ; Carmen ; Geo-IDE ; <b><u>Maprec (Oliban) ;</u></b> MétéoFrance ; EEA ; <b>Canhis ;</b> <b><u>Anemoc</u></b>

est reconnu comme un hub dans les deux niveaux de réseaux. A l'inverse 3 infrastructures de données géographiques plus spécifiques se retrouvent dans les effectifs aux plus faibles scores d'autorités et de hubs dans les deux niveaux de réseau : la base nationale d'observation pour le suivi des côtes (Bosco), l'atlas numérique des états de mer océaniques et côtiers (Anemoc) et l'application cartographique Mapprec de l'observatoire du Littoral Bas Normand (Crec Oliban). Les autres systèmes d'information nationaux sur le littoral tels que Geolittoral et l'ONML occupent un positionnement différent sur le Web (niveau hypertextuel) et dans le Web (réseau de services). Le dispositif Geolittoral joue un rôle d'autorité au niveau du graphe de services et de hub au niveau hypertextuel tandis que l'ONML occupe une position d'autorité au niveau hypertextuel et de hub au niveau du réseau de services.

Enfin, d'après les critères topologiques de l'indicateur de centralité « PageRank » qui se base sur des notions de popularité et de reconnaissance (Cardon, 2013, 2015), les dispositifs Sextant, Datashom et Geolittoral possèdent un score élevé dans les deux réseaux tandis que l'ONML possède un score élevé essentiellement au niveau hypertextuel. Si le « pagerank » est décrit comme un algorithme « cherchant à améliorer la qualité de l'information » (Page *et al.*, 1999), Cardon (2015) rappelle que cet algorithme ne permet en aucun cas de dire si ces dispositifs sont le plus consultés ou le plus utilisés. Mais selon les principes de fonctionnement du « pagerank », les dispositifs aux plus grands scores sont ceux que les acteurs des autres dispositifs « ont choisi de reconnaître en leur adressant beaucoup de liens » et de ce fait leur conférer une plus grande visibilité (Cardon, 2015).

#### 4. Discussion

La recherche présentée dans cet article introduit un protocole méthodologique visant à utiliser des indicateurs permettant de décrire les caractéristiques structurelles de dispositifs d'accès à l'information géographique numérique sur le Web (IDG-plateforme de données ouvertes-API cartographiques). Ce protocole a été testé sur un corpus de 43 dispositifs institutionnels français, avec un focus sur ceux concernant la thématique du suivi et de la gestion du trait de côte en France. Il convient de rappeler que les méthodes numériques employées constituent une première phase d'analyse souvent exploratoire permettant de dégager des tendances globales sur la thématique étudiée et non de faire émerger des modèles généraux (Plantin et Monnoyer-Smith, 2013 ; Ash *et al.*, 2016). Elles doivent être accompagnées de méthodes complémentaires comme l'analyse de contenu et l'analyse thématique des pages renvoyées par certains dispositifs numériques.

Du point de vue des limites, notre approche méthodologique a été contrainte par une dynamique de la donnée (en volume, mise à jour, et supports) et par des logiciels évolutifs (langages, bibliothèques, systèmes d'exploitation), qui nécessitent des ressources logistiques importantes et mettent en difficulté l'emploi de suites logicielles traditionnelles. Le volume de données hypertextuelles traitées a souvent entraîné des difficultés pour stabiliser la base de données. En effet, certaines infrastructures, comme Data.gouv ou Vigicrues, contiennent près d'un million de pages à une profondeur de 2

clics, ce qui peut affecter les capacités de stockage, de traitement mais aussi d'analyse qualitative (typologies sur le contenu des pages). D'autres dispositifs au contenu dynamique de type « Content Management Systems », à l'instar de Dynalit, n'ont pas pu être moissonnés par le logiciel utilisé. Les sites Web des dispositifs numériques étudiés sont en effet assimilables à « un ensemble hétérogène de ressources techniques qui peuvent être agencées de manière très différentes » (Badouard, 2014). Pour autant, si l'utilisation du caractère exploratoire d'un logiciel de « crawl » tel qu'Hyphé doit s'inscrire en aval d'une problématique et d'hypothèses de recherche, celui-ci peut aussi contribuer à les faire évoluer comme cela a été le cas dans la sélection de notre corpus.

## 5. Conclusion

Cet article étudie le positionnement de dispositifs de mise à disposition de données géographiques mobilisables sur des enjeux de suivi et de gestion du trait de côte en France. Le recensement et l'analyse statistique et de graphe sur les composants hypertextuels (pages Web, liens hypertextes, entités Web) ainsi que sur les services de consultation de la donnée constituent un premier volet méthodologique d'une étude traitant, plus globalement, de la gouvernance informationnelle. La gouvernance informationnelle renvoie « aux processus et aux acteurs qui participent à une production de données visant à dire le territoire » (Feyt et Noucher, 2014) c'est-à-dire aux stratégies et aux formes de mise en ligne de l'information (régulations, modes de communications, interactions). L'accès à l'information géographique par le déploiement de fonctionnalités techniques qui structurent le dispositif consulté est aussi le résultat de choix stratégiques, puisque l'information est devenue un matériau important dans les mécanismes d'élaboration et d'application des politiques publiques (consultation, concertation, réglementation). A ce jour les différents éléments structurels des infrastructures de données publiques étudiées révèlent une structure très hétérogène. En effet, un premier groupe de dispositifs renvoie à un faible nombre de liens, de pages et d'entités Web parmi lesquels figurent 5 dispositifs en lien avec l'information géographique numérique sur le trait de côte ( $n = 5$  dispositifs sur les 15 dispositifs en lien avec l'information géographique numérique sur le trait de côte présentent des valeurs inférieures à Q1). Ce sont le plus souvent des dispositifs spécialisés dans une forme de consultation unique de la donnée géographique (visualisation de données). Ces infrastructures en ligne apparaissent dans les années 2005-2010 au moment de l'émergence des premiers dispositifs de cartographie grand public sur le Web (Mericksay, 2011 ; Noucher, 2017) et ont peu évolué dans la marche vers la compatibilité fixée par la directive Inspire. D'autres au contraire sont très récentes comme les applications de suivi du trait de côte par smartphone (information géographique volontaire). Il s'agit le plus souvent de dispositifs dans lesquels les possibilités de parcours sont réduites (peu d'ouverture vers des sources externes avec une faible quantité de liens et de pages Web au niveau hypertextuel). Ces dispositifs sont peu visibles dans le réseau des IDG institutionnelles (faibles scores de centralité et peu interconnecté) et le cheminement pour l'utilisateur est restreint à la fonction initiale du dispositif. Au contraire, certaines IDG nationales thématiques en lien avec les politiques

publiques environnementales (biodiversité, eau, risque inondation) ainsi que la plateforme nationale data.gouv en lien avec la politique d'ouverture des données offrent à l'utilisateur un nombre de parcours multiples au niveau hypertextuel (nombre de pages Web, liens hypertextes élevés) sans pour autant donner accès à l'ensemble des services de consultation de la donnée géographique.

Enfin, la composante réseau de ces dispositifs, deuxième indicateur structurel souvent mis en avant pour caractériser les IDG et autres plateformes de diffusion de données géographiques (Noucher *et al.*, 2017), fait apparaître une connexité<sup>25</sup> encore perfectible pour faciliter l'interopérabilité et l'accès aux différentes ressources offertes par l'ensemble des dispositifs. Ainsi, au sein des dispositifs identifiés par les pouvoirs publics dans la stratégie nationale de gestion intégrée du trait de côte (SNGIT), seuls les « systèmes d'information sur le littoral » nationaux Sextant, Datashom, Cartomer, Geolittoral et l'ONML se positionnent dans les scores les plus élevés en termes de centralités de réseaux (« Autorités », « Hubs », « PageRank »), tout en jouant des rôles différents dans les deux niveaux. Geolittoral, dispositif ministériel mis en avant dans la SNGIT et Sextant dispositif porté par un établissement public (Ifremer) présentent des profils de « hub » dans le niveau hypertextuel et d'autorité au niveau des services de données. *A contrario*, l'ONML dédié à une information plus globale (synthèses, rapports) et à un public plus large, présente un profil d'autorité au niveau hypertextuel et de « hub » au niveau des services. Ces infrastructures exhibent donc une structure sur le Web que l'on peut qualifier de niveau unique. Elles présentent une structure de « hub » ou d'« autorité » sur un seul niveau de réseau (hypertexte ou services web). Elles donnent accès à des services de consultation de la donnée entre les dispositifs limités le plus souvent à un seul type (consultation du catalogue ou visualisation). Ce constat se retrouve aussi dans la surreprésentation de pages sur des entités Web à l'échelle régionale et nationale auxquels ils donnent accès. L'autre constat qui émerge d'un point de vue structurel est la présence quasi systématique de certains dispositifs généralistes nationaux (Geocatalogue, Geoportail), régionaux (Geonormandie, Geobretagne, Geopal) et thématiques (INPN, SIE) aux indices de structure hypertextuelle et aux niveaux d'interopérabilité les plus élevés.

Ces résultats seront prochainement complétés par une analyse du contenu informationnel renseigné par les métadonnées des dispositifs. Cette seconde étape permettra d'apporter des éléments supplémentaires à l'étude du régime d'accessibilité et à la disponibilité de certaines données géographiques en termes de couverture spatiale, thématique et organisationnelle. Mis en regard des témoignages de certains promoteurs et utilisateurs, ces différents résultats permettront de fournir des éléments de réflexion à l'analyse des facteurs limitatifs à l'accès à l'information géographique numérique dans les politiques publiques en lien avec le suivi et la gestion du trait de côte<sup>26</sup>.

---

25. La connexité correspond à la mesure de la présence ou de l'absence de liens entre chaque nœud (Pumain et Saint-Julien, 2004).

26. Ce travail s'inscrit dans le cadre du travail de thèse AInfoGéo (<http://www-iuem.univ-brest.fr/pops/projects/ainfoge>).

## Bibliographie

- Ash J., Kitchin R., Leszczynski A. (2016). Digital turn, digital geographies ? *Progress in Human Geography*. vol. 42, p. 25-43, <https://doi.org/10.1177/0309132516664800>.
- Association française pour l'information géographique (AFIGEO). (2014). Catalogue des Infrastructures de données géographiques françaises, 2014, [En ligne], consulté le 04/05/2016.
- Badouard R. (2014). La mise en technologie des projets politiques. Une approche orientée design de la participation en ligne, *Participations*, n° 1, p. 31-54.
- Beaude B. (2012). *Internet, changer l'espace changer la société les logiques contemporaines de synchronization*, éditions Fyp, collection société de la connaissance.
- Beauguitte L. (2010). *Graphes, réseaux, réseaux sociaux: vocabulaire et notation: Version 1 - Septembre 2010*. Groupe f.m.r. (flux, matrices, réseaux) animé par César Ducruet et Laurent Beauguitte (CNRS, UMR,;1; 2010 <halshs-00541898>), <https://halshs.archives-ouvertes.fr/halshs-00541898/document>
- Beauguitte L. (2013). *D comme densité*, <https://groupefmr.hypotheses.org/2653>
- Bersani C., Simoni M.-L., Allain Y.-M., Ribière G., Denègre J., Planques P., Tugayé Y. (2006). *Rapport relatif au schéma d'organisation des dispositifs de recueil de données et d'observation sur le littoral*, Rapport du ministère de l'Intérieur et de l'Aménagement du territoire, ministère des Transports, de l'Équipement, du Tourisme et de la Mer, 119 p, [http://www.shom.fr/fileadmin/data/DOPS/MIP/PEP/BATHY/Litto3D/Besoins/Rapport-Bersani\\_recueil-donnees-littoral.pdf](http://www.shom.fr/fileadmin/data/DOPS/MIP/PEP/BATHY/Litto3D/Besoins/Rapport-Bersani_recueil-donnees-littoral.pdf)
- Cardon D. (2011). L'ordre du Web. *Médium*, 2011/4 vol. n° 29, p. 191-202. DOI : 10.3917/mediu.029.0191.
- Cardon D. (2013). « Dans l'esprit du PageRank » Une enquête sur l'algorithme de Google. *Réseaux*, 2013/1, n° 177, p. 63-95. DOI: 10.3917/res.177.0063
- Cardon D. (2015). *A quoi rêvent les algorithmes : Nos vies à l'heure des big data*, Le Seuil, Paris.
- Chakrabarti S. (2002). *Mining the Web: Discovering knowledge from hypertext data*, Elsevier.
- Colin N., Verdier H. (2015). *L'âge de la multitude, 2<sup>e</sup> éd., Entreprendre et gouverner après la révolution numérique*, Armand Colin.
- Desbois H. (2015). *Les mesures du territoire : aspects techniques, politiques et culturels des mutations de la carte topographique*, Presses de l'ENSSIB, Nancy.
- Feyt G., Noucher M. (2014). La gouvernance informationnelle, outil et enjeu stratégiques des recompositions territoriales : vers l'émergence de nouveaux référentiels géographiques ? *Actes du Colloque « Fronts et frontières des sciences du territoire »*, GIS CIST, 27-28 mars, Paris.
- Fruchterman T.M.J., Reingold E.M. (1991). Graph Drawing by Force Directed Placement. *Software: Practice and Experience*, vol. 21, n° 11, p. 1129-1164.
- Gautreau P., Severo M., Giraud T., Noucher M. (2013). Formes et fonctions de la « donnée » dans trois webs environnementaux sud-américains (Argentine, Bolivie, Brésil). *Networks and Communication Studies (NETCOM)*, vol. 27, n° 1-2, p. 22-59.

- Gasser U. (2015). *Interoperability in the Digital Ecosystem*, Berkman Klein Center for Internet and Society Research Publication No. 2015-13. Available at SSRN: <https://ssrn.com/abstract=2639210> or <https://dx.doi.org/10.2139/ssrn.2639210> <https://dash.harvard.edu/handle/1/28552584>
- Georis Creuseveau J. (2014). *Les Infrastructures de Données Géographiques (IDG): développement d'une méthodologie pour l'étude des usages. Le cas des acteurs côtiers et de la GIZC en France*. Thèse de Géographie, Université de Bretagne Occidentale.
- Georis Creuseveau J., Claramunt C., Gourmelon F. (2016). A modelling framework for the study of Spatial Data Infrastructures applied to coastal management and planning. *International Journal of Geographical Information Science*, vol. 31, n° 1, p. 122-138. DOI: 10.1080/13658816.2016.1188929
- Girard P., Jacomy M., Baneyx A. (2012). The Hypertext Corpus Initiative: methods and tools for Social Sciences to build corpus from the Web, *Actes du colloque Digital Humanities Luxembourg 2012 Websites as Sources*, Luxembourg.
- Graham M., Zook M. (2013). Augmented Realities and Uneven Geographies: Exploring the Geo-Linguistic Contours of the Web. *Environment and Planning, A* 45, n° 1, p. 77-99.
- Graham M., Dutton W.H. (Eds.) (2014). *Introduction, Society and the internet: How networks of information and communication are changing our lives*, OUP Oxford, p. 1-19.
- Jacomy M., Ghitalla F. (2007). *Méthodologie d'analyse de corpus en Sciences Humaines à l'aide du Navicrawler*. Rapport final, Fondation de la Maison des Sciences de l'Homme Programme TIC-Migrations, <http://webatlas.fr/wp/share/navicrawler/Guide%20m%E9thodo%20NC%202007.pdf>.
- Kleinberg J.M. (1999). Authoritative sources in a hyperlinked environment. *Journal of the ACM (JACM)*, vol. 46, n° 5, p. 604-632.
- Le Corf J. (2015). Management de l'information publique et innovation numérique de services urbains : l'intelligence territoriale en perspective. Big Data - Open Data: Quelles valeurs ? Quels enjeux, *Actes du colloque « Document numérique et société »*, Rabat (p. 97-112). Louvain-la-Neuve, Belgique: De Boeck Supérieur. doi: 10.3917/dbu.chron.2015.01.0097
- Mabi C. (2015). La plate-forme « data.gouv.fr » ou l'open data à la française ». *Informations sociales*, 2015/5, n° 191, p. 52-59.
- Mericksay B. (2011). Les Sig et la cartographie à l'ère du géoweb. Vers une nouvelle génération de Sig participatifs. *L'Espace géographique*, 2011/2 Tome 40, p. 142-153.
- Mol A.P.J. (2006). Environmental governance in the Information Age: The emergence of informational governance. *Environment and Planning C: Government and Policy*, vol. 24, n° 4, p. 497-514.
- Noucher M. (2013). Infrastructures de données géographiques et flux d'information environnementale. *Networks and Communication Studies (NETCOM)*, 27-1/2 | 2013, p. 120-147.
- Noucher M., Gourmelon F., Gautreau P., Georis-Creuseveau J., Maulpoix A., Pierson J., Rouan M. (2017). Spatial Data Sharing: A Pilot Study of French SDI. *ISPRS International Journal of Geo-Information*, vol. 6, n° 4, p. 99.

- Noucher M. (2017). *Les Petites Cartes du web. Approche critique des nouvelles fabriques cartographiques*, Editions Rue d'Ulm, collection Actes de la recherche à l'Ens no 18.
- Page L., Brin S., Motwani R., Winograd T. (1999). *The PageRank citation ranking: Bringing order to the web*, Stanford InfoLab, <http://ilpubs.stanford.edu:8090/422/1/1999-66.pdf>.
- Pinède N. (2014). Le réseau, un « entre-deux » liens ? De quelques facettes et dynamiques du lien hypertexte. *Sciences de la société*, 91, p. 94-105.
- Pinède N. (2018). Stratégies info-communicationnelles et logiques d'open data. L'exemple des infrastructures de données géographiques en France, *Actes du colloque « Repenser les cartes à l'ère du Web : Approches critiques et enjeux politiques, sociaux, économiques des données territoriales »*, Bordeaux.
- Plantin J.-C., Monnoyer-Smith L. (2013). Ouvrir la boîte à outils de la recherche numérique. *tic&société* [En ligne], vol. 7, n° 2 | 2<sup>e</sup> semestre 2013, mis en ligne le 01/06/2014, consulté le 14/12/2017. URL: <https://journals.openedition.org/ticetsociete/1527>; DOI: 10.4000/ticetsociete.1527
- Pornon H. (2015). *SIG : La dimension géographique du système d'information* (2 édition), Management des systèmes d'information, Paris, Dunod.
- Pumain D., Saint-Julien T. (2004). *L'analyse spatiale, Localisations dans l'espace*, Armand Colin.
- Rajabifard A., Feeney M.-E., Williamson I.P. (2002). Future directions for SDI development. *International Journal of Applied Earth Observation and Geoinformation*, vol. 4, n° 1, p. 11-22.
- Rogers R. (2009). Situating Digital Methods in Internet Research, *The End of the Virtual: Digital Methods*, Amsterdam University Press.
- Rogers R. (2015). Digital Methods for Web Research. *Emerging Trends in the Social and Behavioral Sciences* (eds R. A. Scott and S. M. Kosslyn). doi: 10.1002/9781118900772.etrds0076
- Severo M., Venturini T. (2016). Enjeux topologiques et topographiques de la cartographie du Web. Le cas du patrimoine culturel immatériel français. *Réseaux*, 1/2016, n° 195, p. 85-105.
- Suarez S., Garcin M., Bulteau T., Rouan M., Lagadec L., David L. (2012). Les observatoires du trait de côte en France métropolitaine et dans les DOM. *EchoGéo* [En ligne], 19 | 2012, mis en ligne le 10 février 2012, consulté le 12 décembre 2016. URL : <http://journals.openedition.org/echogeo/12942>, DOI : 10.4000/echogeo.12942.
- Wilkinson M.D., Dumontier M., Aalbersberg I.J., Appleton G., Axton M., Baak A., Mons B. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific data*, 3, 160018 doi: 10.1038/sdata.2016.18