

Commission « Règles de mise en œuvre »

Réunion du 13 mars 2018
de 10h à 15h

Membres présents

Stéphane MEVEL-VIANNAY	Région Bretagne	Fion HALLEMAN	IGN
Maria TOMANOV	Région Auvergne-Rhône-Alpes	Dominique LAURENT	IGN
Guillaume RYCKELYNCK	Région Grand-Est	Pierre VERGEZ	IGN/Appui CNIG
Seydou PAM	CD 93	Philippe MUSSI	Commission AT/INRIA
David DELHORBE	CD 93	Philippe TERME	DREAL Pays de Loire
Bérangère MOURE	CD 93	Richard MITANCHEY	CEREMA
Alix MARC	Montpellier méd. métropole	Christine LE BAS	INRA
Maël REBOUX	Rennes métropole/AITF	Bruno IRATCHET	ONF
Matthieu NOUCHER Bordeaux	CNRS / Université de	Thierry VILMUS	BRGM
Narimène DAHMANI	Université de Strasbourg	Claire CAILLEAU	BRGM
Ingrid BRUGIONI	CD 53	Jérôme DESBOEUF	
Laurent BONAREK	DDT 53	DINSIC	
Cécile JOFFRES	DGFIP	Robert RIVIERE	MTES/./CPII
Sébastien LEGER	DGFIP	Clément JAQUEMET	
			MTES/./MIG
		Benoit DAVID	MTES/./MIG
		Marc LEOBET	MTES/./MIG

Points à l'ordre du jour :

1. - Point sur les négociations européennes en cours sur INSPIRE : projets de modification des textes réglementaires européens sur le rapportage et sur l'interopérabilité des données (Marc Leobet);
2. - Nouveau tableau de bord du Géocatalogue : situation et perspectives (Thierry Vilmus, BRGM)
3. - Articulation entre le Géocatalogue et geo.data.gouv.fr (J. Desboeufs, DINSIC)
4. - Présentation de la nouvelle organisation des services en réseau dans le système de publication du ministère en charge de l'environnement, GéoIDE (Robert Rivière, MTES/CPII)
5. - Etat d'avancement du projet de recherche GÉOBS (Matthieu Noucher, CNRS – Université de Bordeaux).

6. - Présentation de l'intégration des données de l'IGN espagnol dans Openstreetmap et vice-versa, utilisation d'OSM par l'Agence européenne de l'environnement, OSM comme format INSPIRE : échanges autour de "Openstreetmap dans INSPIRE?" (Benoit David, MTES/MIG)

POINT SUR LES NÉGOCIATIONS EUROPÉENNES EN COURS SUR INSPIRE

Marc Leobet présente d'abord le projet de la Commission européenne (COM) sur le rapportage INSPIRE. Reprenant très largement le processus français basé sur l'exploitation automatique des métadonnées, le projet de la COM compense cette simplification par l'ajout de nouvelles métadonnées et mots-clés obligatoires (emprise administrative des jeux de données, jeux de données prioritaires). Ceux-ci n'ayant pas d'utilité prouvée pour les autorités publiques, la position française (consolidée par une consultation de la commission RMO) et d'autres Etats membres (EM) est de les refuser. L'adoption de cette proposition est aujourd'hui peu probable. La COM peut encore proposer une version plus acceptable d'ici le prochain comité INSPIRE.

Le processus de révision du règlement sur l'interopérabilité porte sur les modèles de données de l'annexe III. La consultation des EM sur les points à améliorer a produit des réponses peu nombreuses et complètement dispersées. Cette consultation place la COM dans l'embarras. Lors du Comité INSPIRE de décembre, une suspension de séance a même été demandée par la COM pour caler les positions de la DG ENVIR et du JRC, discordantes. Le succès de cette procédure est aujourd'hui improbable. Faute de connaître aujourd'hui quelle sera la proposition de la Commission européenne, rappelons simplement que la position française (et britannique) était de transformer les obligations d'interopérabilité de l'annexe III en simples recommandations, ce qui a été refusé.

NOUVEAU TABLEAU DE BORD DU GÉOCATALOGUE

Thierry Vilnius présente le nouveau tableau de bord du Géocatalogue (voir présentation).

Les questions portent sur la reconnaissance et la définition des données ouvertes.

Maël Reboux signale que des données obsolètes sur geo.data.gouv.fr alors qu'elles sont plutôt correctes sur le Géocatalogue. Il s'interroge sur la coexistence de trois systèmes autonomes, les deux précédents et le Géoportail de l'urbanisme (GPU), qui archive aussi et gère correctement l'historisation des données. Il souhaite des politiques d'archivage et de partages mieux établies. En particulier, la séparation des données obsolètes, des données mortes et des erreurs (données publiées par mégarde : il est difficile de faire disparaître des fiches publiées par erreur).

La Commission s'inquiète de l'impact des tests divers (Spatineo...) sur les statistiques des plateformes. Les tests Spatinéo auraient été black-listés par le Geoportail européen.

ARTICULATION ENTRE LE GÉOCATALOGUE ET GEO.DATA.GOUV.FR

Marc Leobet introduit ce point. Si le back-office du Géocatalogue est le pivot de l'architecture INSPIRE, ne serait-ce qu'en raison des besoins du rapportage, il ne permet pas de répondre de façon efficiente aux recherches des usagers. Aujourd'hui, les fonctionnalités de geo.data.gouv.fr (GDG) répondent mieux à ces besoins.

Jérôme Desboeufs rappelle le chemin parcouru : d'abord une passerelle INSPIRE-data.gouv.fr, centrée sur le triplet <open data, producteur référencé, données téléchargeables>, trois ans plus tard le site a été refondu dans une logique de pérennisation.

GDG moissonne le Géocatalogue pour obtenir des données qui sont indisponibles ailleurs. Sinon, la politique est de moissonner au plus près du producteur.

Les évolutions :

- être capable de digérer les données disponibles (géométrie) ;
- abandonner GDAL ;
- améliorer la gestion des archivages ;
- clustering par commune ;
- calcul d'une bounding box à partir de la géométrie ;
- WebGL pour des performances supérieures
- insertion de balises schema.org (d'où la nécessité de se coordonner avec les autres initiatives sur le sujet : à traiter lors de la prochaine réunion de la commission RMO)

- exploiter GeoJSON pour la visualisation, avec vocation à alimenter les autres sites en data.gouv.fr ;
- Consultation des schémas (pour les données attributaires) en cours.

Un prochain Lab (atelier d'évolution) sera organisé avant l'été. L'annonce d'une API Géo V2 aura lieu avant.

Des échanges qui suivent, il ressort que les deux plateformes, GDG et Géocatalogue, ont vocation à coexister à long terme. L'articulation entre les deux a permis de réduire l'investissement public dans le Géocatalogue. Il est noté que le dédoublement n'est pas fait en fonction de l'IRU d'origine, remplacée par l'éditeur open data sans conserver l'IRU de la source.

PRÉSENTATION DE LA NOUVELLE ORGANISATION DES SERVICES EN RÉSEAU DANS GÉOIDE

(Voir la présentation « Géo-IDE, Organisation des services en réseau »)

Robert Rivière rappelle l'exigence 52 des *Technical guidelines* : « a separate WFS endpoint for each dataset ». Pour résoudre des problèmes récurrents de performance, elle a été étendue aux services WMS et ATOM. GeoIDE a mis en place un ensemble de services (WMW, WFS, ATOM) pour chaque jeu de données alors que le guide du CNIG préconise des services « container ».

Les échanges s'enchaînent :

- Faut-il une révision des guides du CNIG ? Réponse : non : il a été de tout temps entendu que des architectures particulières pouvaient ne pas suivre certaines des recommandations du CNIG si celles-ci rendaient leurs fonctionnalités inopérantes. Les guides du CNIG visent des infrastructures de taille moyenne. Des exceptions sont possibles pour de très grosses infrastructures (ex : GeoPortail, GeoIDE)
- Geonetwork gère difficilement une volumétrie du niveau de celle de GéolIDE. Pour cette raison, Etalab ne l'utilise pas et réfléchit à déployer d'autres briques logicielles.
- quelle est le meilleur critère pour un service en mode container ? Thématique ? Mais cela produit des containers disymétriques. Par ensemble de série ? Pourquoi pas, avec un service par géostandard. Il est par exemple plus facile de développer des SLD pour des services mono-couches.

En conclusion, l'idée d'une journée dédiées à la mise en œuvre des services fait son chemin. Jérôme Desboeufs en est d'accord si on dépasse les seuls standards de l'OGC. Pour plusieurs membres, il y a bien besoin de nouvelles manières de communiquer les données tout en conservant les standards OGC pour les différentes IDG.

ÉTAT D'AVANCEMENT DU PROJET DE RECHERCHE GÉOBS

(Voir la présentation <http://www-iuem.univ-brest.fr/pops/attachments/1746>)

Plusieurs projets de recherche s'intéressent aujourd'hui à la gouvernance informationnelle des territoires. Leur objectif est de comprendre l'impact des nouveaux modes de production et de mise en circulation de l'information géographique sur les pratiques de gestion territoriale. Parmi ces projets, l'étude GÉOBS présentée par Matthieu Noucher est centrée sur les Infrastructures de Données Géographiques.

Son objectif spécifique est de décrypter :

- les stratégies des promoteurs de IDG (axe 1),
- les contenus effectivement diffusés par ces plateformes (axe 2)
- et les usages des professionnels de la géomatique qui en résultent (axe 3).

Chacun de ces axes est questionné selon quatre dimensions :

- 1) la mise en réseau des données (accessibilité) ;
- 2) la mise en réseau des outils (interopérabilité) ;
- 3) la mise en réseau des acteurs (géocollaboration) ;
- 4) la mise en réseau des territoires (égalité informationnelle).

Pour ce faire, un corpus varié (entretiens, enquêtes, littérature grise, sites web, métadonnées des géocatalogues) est mobilisé. Chaque axe a été travaillé séparément pour le moment et a donné lieu à des publications (rapports d'étape ou articles scientifiques). L'année 2018 (dernière année du projet) sera consacrée à l'analyse transversale des 3 axes pour mettre en regard discours, données et pratiques.

A ce stade de l'étude, GÉOBS constate que l'écosystème des IDG est loin d'être stabilisé. L'impact de la loi NOTRe est perceptible aussi bien dans l'évolution des discours des promoteurs, des données disponibles que des pratiques des usagers. Les évolutions des motivations construisent un bloc IDG hétérogène, des différences entre plateformes et des situations locales évolutives. Par ailleurs, l'instabilité technique des systèmes des IDG et la qualité médiocre des métadonnées ont surpris les chercheurs en informatique ayant collaboré à l'étude. Elles ont conduit l'équipe de recherche à réduire ses ambitions initiales (les travaux sur la généalogie des données ou sur les dates de publication/mise à jour n'ont pu être généralisés).

Parmi les résultats présentés, on peut relever une augmentation importante du volume de métadonnées diffusées dans les 39 géocatalogues régionaux et nationaux étudié à partir de leur service CSW : +17% de métadonnées entre juin 2016 et mai 2017 pour atteindre plus de 188.000 fiches à cette date. Matthieu Noucher souligne cependant l'importance de ne pas en rester à ces chiffres qui ne donnent qu'une vision très « macro », qui sont biaisés par de nombreuses redondances entre catalogues et qui ne disent rien sur l'accessibilité effective aux données. Ainsi, l'équipe de recherche a relevé qu'environ 20 % des données *taggées* « open data » ou « donnée(s) ouverte(s) » ont des liens invalides !

GÉOBS a également constaté un effet des dates de rapportage INSPIRE sur les dynamiques de catalogage. L'analyse de la couverture spatiale montre des contrastes importants et permet de souligner l'importance de l'échelon régional mais aussi de quelques initiatives locales. Par ailleurs, le réseau de moissonnage des géocatalogues est fortement structuré avec des relations centres (IDG nationales autour du couple géocatalogue/GéoDataGouv) / périphérie (IDG régionales et infrarégionales) très marquées, la présence d'un seul cluster thématique (autour de la mer, du littoral et de l'outre-mer). Ce réseau s'est à la fois étendu et densifié en 2017 avec l'apparition d'IDG étrangères. Matthieu Noucher, en accord avec Guillaume Ryckelynck, alerte sur le risque de sur-interprétation de ces éléments.

Du côté des actions d'animation des communautés de pratique associées aux IDG, on constate la forte dynamique des groupes de travail (pôles métier) à l'échelon régional avec des centres d'intérêt de plus en plus marqués pour des questions thématiques (plutôt que généralistes) et un travail centré sur les phases amont à la fabrication des données (co-conception de nomenclature, de modèle plutôt que co-production).

Les usagers ayant répondu à l'étude signalent l'importance majeure des IDG nationales dans leur pratique quotidienne, mais l'enquête permet également de souligner l'importance des pratiques multi-IDG : beaucoup d'usagers semblent « picorer » des données à gauche à droite...

Lors de la discussion, Maël Reboux suggère d'ajouter le nom des licences utilisées dans la sélection des données ouvertes. Matthieu Noucher explique qu'une analyse sur les types de licence avait été réalisée en 2015 : elle avait mis en évidence la piètre qualité des métadonnées sur ce sujet. Ce travail serait à reproduire maintenant si les pratiques se sont stabilisées sur ce point.

Philippe Mussi s'interroge sur l'idéal-type d'IDG que pourrait permettre d'identifier cette étude. Pour Matthieu Noucher, le projet GÉOBS n'a pas vocation à évaluer ou à définir un cadre normatif autour des IDG. L'objectif de l'équipe est d'étudier les IDG comme des dispositifs qui évoluent en fonction du contexte socio-politique et technique. C'est bien l'évolution de cet « écosystème » et la performativité de ces évolutions qui intéressent les chercheurs.

David Delhorbe souligne alors l'importance du contexte politique dans le déploiement (ou le non déploiement comme en Ile-de-France) de ce type d'initiative. GÉOBS ne peut généraliser l'étude des contextes politiques mais travaille sur des études de cas ciblées (Guyane, Nouvelle-Aquitaine) qui permettent ce type d'analyse fine.

PRÉSENTATION DE L'INTÉGRATION DES DONNÉES DE L'IGN ESPAGNOL DANS OPENSTREETMAP

(Voir la présentation *Transforming the Spanish "INSPIRE database" into OSM and vice versa* (Emilio Lopez, IGN-E) https://ies-svn.jrc.ec.europa.eu/attachments/download/2285/OpenStreetMaps+INSPIRE_databases.pptx)

Cette présentation se situe dans un contexte qui voit s'accumuler des éléments, certes d'importance variables, mais qui vont dans le sens d'un recours croissant aux données d'OSM, d'une part, et des outils d'OSM, d'autre part.

Il y a trois ans, les participants à une réunion du du groupe de maintenance technique INSPIRE (MIG-T) constataient que tous les projets européens présentés utilisaient OSM. Il y a 18 mois, Marc Leobet constatait que tous les projets transfrontaliers concernant la France abandonnaient Eurogeographics pour migrer vers OSM. En décembre, l'agence européenne de l'environnement annonçait vouloir exploiter OSM dans le cadre de la nouvelle version de Corine Land Cover. Le JRC envisage maintenant de reconnaître le format OSM dans les Technical guidelines INSPIRE. Le dernier élément en date est la présentation, durant une réunion du MIG-T, par l'IGN-E de ses travaux d'export d'une partie de ses données dans OSM et vice-versa.

Benoit David présente les diapositives produites par l'IGN-E à cette occasion. Il signale qu'Eurostat, service statistique officiel européen, utilise entre autres OSM sauf pour les parties contraires à la politique étrangère de l'UE (frontières). C'est une utilisation intéressante pour les données souveraines.

Dans les échanges qui suivent, plusieurs points sont éclaircis :

- Si le discours d'OSM reste centré sur l'apport des saisies de terrain, Matthieu Noucher signale qu'une majorité de contributeurs saisissent « depuis leur fauteuil » (*armchair mapping*) : une enquête réalisée auprès de 300 contributeurs de la communauté française en 2016¹ révèle que plus de 60 % des contributions se font par photo-interprétation, intégration de données, etc. contre moins de 40% de saisie à partir du terrain.
- il faut séparer les trois utilisations d'OSM : comme fond de plan (cas des projets européens), en tant que base de données et en tant qu'outillage informatique (séparé des données). En particulier, les premier et troisième cas n'obligent pas à passer ses données en licence ODbL.
- Plusieurs stratégies sont décrites : en Alsace, la base inter-modalité (transports) a migré de Navtech vers OSM et la mise à jour est opérée directement dans OSM, sans contrôle par l'autorité compétente. Les modifications ne sont pas historicisées.
- OSMOSE est cité comme outil de versement de données ouvertes. C'est utilisé pour les monuments historiques.
- Il existe à l'IGN un convertisseur BDTopo → OSM, mais le contraire pourrait être plus difficile. Les outils de tuilage pourraient être intéressants.
- La licence ODbL empêche d'utiliser les voies d'OSM dans la BAN
- Montpellier méditerranée métropole étudie l'import-export de ses données dans OSM. Ces dynamiques territoriales sont comme à l'accoutumée variables selon les territoires. L'idée est ainsi d'améliorer la valorisation des données des collectivités. Vu les demandes et les productions, un bon exemple serait l'occupation des sols pour en étendre les usages.
- Mettre les données publiques dans OSM pourrait être un moyen plus efficace de les valoriser que de les « pousser » dans le GéoCatalogue (et pourraient ainsi rester sous licence ouverte).
- OSM est souvent utilisé comme lanceur d'alertes pour les autorités publiques
- GoogleMap qui était en situation de force sur les sites pro des PME perd du terrain au bénéfice d'OSM

¹ <https://halshs.archives-ouvertes.fr/halshs-01371544> et <https://halshs.archives-ouvertes.fr/halshs-01672643>