

Vers une gouvernance de l'information géographique

MATHIEU BECKER ■ Gérant-fondateur d'Isogeo

HERVÉ HALBOUT ■ Consultant SIG et 3D (Halbout Consultants)

Quand les administrateurs SIG dépriment

Cet article, dont le titre peut paraître provocateur, vise à expliquer pourquoi, aujourd'hui, on ne trouve que peu de catalogues de données et de métadonnées géographiques, alors que le bon sens voudrait qu'ils soient légion. Normes contraignantes ? Besoins internes non identifiés ? Ressources insuffisantes ? Manque de temps ? Outils mal adaptés ? Etc. Ce sont autant de questions que nous allons traiter par la suite, en y apportant quelques réponses. Partons d'un exemple concret : celui d'une communauté d'agglomération de cent mille habitants, avec un service SIG (ou une cellule SIG) de deux personnes (un Responsable/Administrateur SIG et un Géomaticien), cinq référents internes, trois cent cinquante jeux de données SIG, quatre formats de données différents (*shp*, *postgis*, *dwg*, *ecw*). Et pour que cet exemple soit parlant, mettons-nous un instant à la place de l'administrateur SIG.

Le rêve

Conscient de la nécessité de bien connaître son patrimoine de données géographiques, il commence à télécharger et installer un outil de catalogage. Il est gratuit, profitons-en ! Après quelques heures d'installation et de paramétrages divers viennent

les premiers tests : il découvre enfin l'interface de saisie des métadonnées.

Le lendemain, motivé, il commence à saisir sa première fiche de métadonnées. Après de longues minutes, celle-ci est enfin achevée et soudain... prise de conscience brutale : il y en a encore trois cent quarante-neuf à créer ! Le besoin d'une boisson énergisante se fait soudain sentir, afin de conserver un moral suffisant pour continuer. Bien sûr, ce travail est indispensable, obligatoire, tout le monde doit le faire. De plus, cet outil *open source* est l'un des plus utilisés, c'est la référence du secteur...

La réalité

Cet administrateur SIG ne saisira probablement jamais l'intégralité des trois cent cinquante fiches, car il a de nombreuses autres tâches à gérer au quotidien. Même s'il le faisait consciencieusement, il lui faudrait reprendre régulièrement les fiches, car elles devront être mises à jour pour que le catalogue soit exploitable par différents utilisateurs.

Cette histoire n'est pas qu'une fiction : elle est le reflet du vécu de nombreux administrateurs SIG que nous avons rencontrés ces dernières années. Nous avons tiré quelques réflexions de ces expériences :

1. Les outils actuels ne sont pas des outils de catalogage de données, mais des outils de saisie de fiches de métadonnées.

La différence est importante : la fiche n'est pas liée à la donnée et si cette dernière est modifiée, la fiche, elle, ne l'est pas. Cela signifie que le « *catalogue* » n'est presque jamais l'image de la réalité, mais la simple somme de fiches saisies une à une... et qu'il faut mettre à jour régulièrement ;

2. Pour remplir une fiche, il faut compléter manuellement l'ensemble des champs, de A à Z. Ceci est assez étonnant, car l'évolution informatique devrait permettre d'automatiser – au moins en partie – cette tâche, à partir des informations déjà stockées dans les données ;

3. Les métadonnées doivent être saisies dans une interface qui reflète la norme ISO 19139.

Le respect de cette norme est important, car elle est garante de l'interopérabilité des échanges d'informations. Toutefois, respecter une norme d'échange ne doit pas se faire au détriment de la convivialité de saisie ;

4. Trouver une fiche de métadonnées (parmi plusieurs centaines) n'est pas toujours simple, même lorsqu'on les a soi-même saisies. Que dire alors des utilisateurs du catalogue de

données ? Est-ce qu'il ne serait pas possible de disposer d'un moteur de recherche simple, convivial et efficace ?

Au travers de ces expériences et réflexions, il devient plus facile de comprendre pourquoi les catalogues de données géographiques demeurent rares aujourd'hui et, quand ils existent, pourquoi leur mise à jour est si difficile. Réaliser un travail de catalogage est considéré comme utile, voire indispensable. Mais il s'agit aussi d'une tâche particulièrement laborieuse, répétitive et peu valorisante pour son administrateur. Néanmoins, c'est dans ce catalogage que réside la valeur des informations géographiques... et la valorisation des producteurs de ces données.

A propos d'Isogeo (www.isogeo.fr)

La société *Isogeo* a été créée en 2009 par Mathieu Becker pour aider les organismes, publics et privés, manipulant des informations géographiques à optimiser l'utilisation, le partage et la valorisation des données SIG. La solution innovante – architecture informatique en nuage et modules gérés en mode SaaS – permet de recenser, documenter et partager facilement les données SIG.

Alors, cataloguer des données géographiques est-il définitivement réhibitore ? Non, bien sûr ! Et vous découvrirez pourquoi dans la suite de cet article.

Un catalogue de données SIG exhaustif, à jour et documenté

La particularité des SIG de décomposer un environnement complexe en couches d'informations distinctes a pour conséquence la multiplication des informations descriptives. Le nombre de formats et l'accès simplifié aux outils d'édition alimentent de manière exponentielle les bases de données. De plus, le patrimoine de ces données

tend à se diversifier au niveau des supports, sous forme de fichiers, de base de données spatiales ou non et dernièrement de services web ; tout cela sur un fond d'échange et de partage désormais revendiqué mais non encore assumé.

L'environnement d'une donnée géographique a beaucoup évolué ces dernières années. À l'image de la mondialisation, nous vivons l'*expansion* et l'*harmonisation des liens d'interdépendance* entre les différentes structures manipulant de l'information géographique. La directive européenne *Inspire* accélère ce mouvement amorcé il y a quelques années à travers les conventions de partenariat d'échanges de données géographiques. Aujourd'hui, le risque encouru par un organisme est de se laisser déborder par cette masse d'informations.

Le défi des gestionnaires de données dans cet environnement de *libre circulation* consiste à rendre la donnée visible et accessible, à la valoriser. Les deux piliers de cette nouvelle forme de gestion sont le catalogue et la documentation.

Pour un utilisateur, le catalogue doit être l'image du patrimoine des données auxquelles il a accès. En théorie, un catalogue recense les données de différents formats (fichiers, bases de données, services web), les données de son SIG et celles de ses partenaires, ainsi que les données publiques, libres d'accès (*Open Data* notamment) : en bref, toutes les données disponibles sur un territoire. C'est une condition nécessaire en interne pour une exploitation optimale de son patrimoine, mais aussi en externe, pour favoriser les échanges, réduire les coûts et définir les règles d'une gouvernance des données qui, pour un territoire, prend tout son sens. En pratique, les difficultés sont nombreuses. La technologie doit être au service des gestionnaires de données et non l'inverse.

Les administrateurs de données géographiques (gestionnaires de bases de données) ont un rôle de plus en plus important dans ce contexte de libre circulation de l'information. L'élaboration d'un catalogue exhaustif, à jour et documenté est un vrai défi pour les prochaines années.

Repenser le catalogue pour en faire un véritable outil de gestion SIG

Cependant, le catalogue ne serait rien sans une documentation des données : les métadonnées. Un responsable d'études n'aura pas le même regard sur une donnée géographique que le géomaticien, l'administrateur des données ou le partenaire externe. Certes, les normes appliquées à la documentation des données géographiques ont un rôle utile dans la circulation et le partage d'informations, mais il est important de rappeler que les normes ont avant tout un objectif d'interopérabilité.

Toutefois, un catalogue de données géographiques exhaustif, à jour et documenté n'est pas une fin en soi. Une étude réalisée par la jeune société *Isogeo* montre qu'un organisme dispose en moyenne de deux cents à mille couches d'informations géographiques, sans compter les documents associés, études, cartes et spécifications. Cataloguer l'ensemble de ces données est essentiel, mais contre-productif si le catalogue est inexploité.

Le catalogue doit être la porte d'entrée d'une série de services complémentaires : d'une part pour le catalogue lui-même (tri, partage, interrogation, etc.) et d'autre part pour les données associées à chacune des fiches de métadonnées (visualisation, extraction, etc.). Couplé à des services intelligents, dans les mains

de l'administrateur de données SIG, le catalogue devient un véritable outil de gestion qui permet de mesurer, vérifier, planifier et piloter son SIG.

Le catalogage des données géographiques ?

Le tableau que nous avons brossé précédemment peut sembler assez sombre, au regard des efforts et tentatives – souvent infructueux – engagés par des administrateurs SIG pour cataloguer leurs données. Nous avons rencontré une certaine forme de résignation chez nos interlocuteurs, voire du fatalisme.

Pourtant, aujourd'hui, il y a des raisons d'espérer ! Une solution innovante dont le but est de « *simplifier les données géographiques* » vient de voir le jour. La société *IsoGeo*, spécialiste des SIG et porteuse de cette solution, se concentre plus particulièrement sur le catalogage, avec l'ambition de « *produire un catalogue de données géographiques exhaustif, à jour et documenté* ». Des outils logiciels à la fois simples d'utilisation tout en étant professionnels dans leur usage permettent de faire vivre les données cataloguées et de les valoriser à travers leur (re) connaissance.

Cette solution respecte bien évidemment les normes d'échanges et de connaissance que sont la norme ISO 19139 ou la directive *Inspire*. Sa véritable valeur ajoutée s'appuie sur plusieurs points :

– **Un inventaire automatique et exhaustif pour maîtriser son patrimoine** : ce concept d'inventaire est important, car le catalogage se fait habituellement sur des données choisies, ce qui signifie qu'il y a beaucoup d'autres données qui sont laissées

à l'écart, souvent par manque de connaissance. L'inventaire, qui va être réalisé la première fois en parcourant les formats de données géographiques sur des serveurs cibles, va permettre de prendre connaissance du patrimoine réel. Une analogie peut être faite avec un déménagement qui permet généralement de faire le tour de l'ensemble de ses biens... et de réfléchir au tri à faire ;

– **Des catalogues orientés usages** : c'est à partir de l'inventaire exhaustif de ses données que l'administrateur SIG va créer ses catalogues à destination de différents profils utilisateurs, en choisissant celles qu'il va y intégrer. Un catalogue contient un ensemble de données affectées à un groupe d'utilisateurs. Il lui restera alors à remplir les fiches de métadonnées.

Cette opération est facilitée, car le traitement intelligent des données (effectué pendant l'inventaire) permet de compléter automatiquement une quinzaine de champs normalisés (projection, emprise géographique, liste d'attributs...). L'inventaire est orienté administration et le(s) catalogue(s) orienté(s) usages ;

– **Une mise à jour en continu** : si l'un des intérêts du catalogage est de disposer d'une liste de données documentées, la mise à jour de ces informations est tout aussi importante. Dans un catalogue classique, les fiches de métadonnées sont déconnectées de la donnée originale. Ici, c'est le contraire : le lien entre donnée cataloguée et donnée source permet de détecter à chaque examen de serveurs s'il y a eu des changements sur la donnée source. Cela génère des notifications à l'attention de l'administrateur SIG ;

– **Un moteur de recherche performant** : les catalogues de données n'ont de sens que s'ils peuvent être interrogés, que ce soit par des géomaticiens, des

chargés de mission ou des intervenants de bureaux d'études. Il faut donc que la méthodologie de recherche soit suffisamment accessible pour des profils utilisateurs complètement différents. Des exemples grand public existent sur le *web*. Cette solution en suit les traces avec un moteur de recherche orienté usages ;

– **Un tableau de bord pour piloter son SIG** : il est souvent question de tableaux de bord dans les systèmes d'information. L'information géographique ne déroge pas à cette règle. Pourtant, l'administrateur SIG ne dispose pas d'outils adéquats. Cette solution lui en fournit pour piloter véritablement le SIG. Le rôle « *d'administration des données* » prend ainsi tout son sens ;

– **Un fonctionnement en mode SaaS pour une plus grande souplesse** : il n'est plus nécessaire d'installer un logiciel sur son PC, puisque la solution fonctionne en mode service *web* (*Software as a Service*). L'un des avantages en est la facilité d'administration, mais aussi la mise en relation simplifiée des acteurs d'un même territoire.

Conclusion

Les constats que nous avons dressés montrent bien les limites du concept actuel de catalogage de données géographiques qui est plus orienté sur l'interopérabilité que sur l'exploitation par les usagers. C'est ce qui explique que le catalogage est souvent perçu comme une nécessité, mais aussi et surtout comme une contrainte. La possibilité aujourd'hui de disposer d'un *inventaire exhaustif, à jour et documenté* de ses données devrait permettre de mesurer que le catalogage et la documentation des données ne sont pas une fin en soi, mais bien le début d'une nouvelle ère : celle de la gouvernance de l'information géographique. ■