



HAL
open science

Mise en oeuvre d'un catalogue de données géographiques

Laurent Coudercy

► **To cite this version:**

Laurent Coudercy. Mise en oeuvre d'un catalogue de données géographiques. [Rapport de recherche] Centre d'études sur les réseaux, les transports, l'urbanisme et les constructions publiques (CERTU). 1999, 34 p., tableaux, 5 références bibliographiques. hal-02165547

HAL Id: hal-02165547

<https://hal-lara.archives-ouvertes.fr/hal-02165547>

Submitted on 26 Jun 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Certu |

Mise en œuvre d'un catalogue de données géographiques



Ministère de l'Équipement,
des Transports et du Logement



Centre d'études sur les réseaux, les transports,
l'urbanisme et les constructions publiques

Collection Rapports d'étude

Publications proposant des informations inédites, analysant et explorant de nouveaux champs d'investigation.
Publications susceptibles d'être amendées ou complétées ultérieurement.

Le Certu publie aussi les collections: dossiers, références, débats.
Catalogue des publications disponible sur <http://www.certu.fr>

NOTICE ANALYTIQUE

Organisme commanditaire : CERTU : Centre d'études sur les réseaux, les transports, l'urbanisme et les constructions publiques 9, rue Juliette Récamier 69456 Lyon Tel : 0 472 74 58 00 Fax : 0 472 74 59 00			
Titre Mise en œuvre d'un catalogue de données géographiques			
Organisme auteur CERTU : Centre d'études sur les réseaux, les transports, l'urbanisme et les constructions publiques	Rédacteur Laurent Coudercy , coordonateur du pôle géomatique	Date d'achèvement Janvier 1999	
Remarques préliminaires : Ce rapport fait suite aux travaux menés dans le cadre de l'AFNOR avec le CNIG. Il prolonge la réalisation des recommandations de l'AFNOR pour l'application de la norme ENV 12567, qui a aboutit à la réalisation d'un logiciel de catalogage, REPORTS, disponible sur le site du CNIG et au CERTU.			
Résumé : Le développement des SIG induit un développement des échanges de données géographiques, soit en interne aux services (au sein d'une même DDE, par exemple), soit entre différents services sur un même territoire. Dans le cadre de ces échanges, le besoin de disposer d'un catalogue des données disponibles se fait jour. Le présent document vise à aider à la mise en œuvre d'un catalogue, dans les services déconcentrés de l'Équipement. Il ne prétend pas donner La Solution à tous les problèmes que pose la mise en œuvre d'un catalogue de données géographique (le manque de recul sur ce thème étant un élément d'excuse), mais devrait plus modestement aider dans cette démarche.			
Mots clés : SIG, catalogue de données, métadonnées, norme européenne		Diffusion : Libre ; disponible sur Internet Systématique pour les services des ministères de l'Équipement	
Nombre de pages : 34 pages	Prix : gratuit	Confidentialité : Non	Bibliographie : Oui

Sommaire

Introduction	3
Un catalogue, pour quoi faire ?	3
Un catalogue pour les besoins internes au service	3
Un catalogue pour les besoins externes au service	4
Une démarche européenne reprise par la France	4
Quelles données cataloguer ?	4
<i>Données géoréférencées, référence indirecte ...</i>	4
Qualité des données à cataloguer	5
Ne pas restreindre a priori le catalogage	5
Que met-on dans une fiche de catalogue ?	5
Découpage des données en lots	5
L'information à saisir par fiche	6
Métadonnées sensibles	6
Cas d'un produit composite	7
Cas des données en provenance de l'extérieur	7
Les thésaurus et nomenclatures, outils de classification	8
Utilisation d'un thésaurus dans les métadonnées	8
Quel thésaurus ou nomenclature utiliser ?	9
Quelle stratégie de saisie des fiches du catalogue ?	9
Quelle organisation ?	10
Réalisation des fiches	10
Mise à disposition du catalogue	10
Organisation de la diffusion des données	11
Vie du catalogue	11
Annexe 1 : définitions	13
Annexe 2 : liste des champs et conseils de saisie	15
Annexe 3 : bibliographie	31

Mise en œuvre d'un catalogue de données géographiques

Introduction

Le développement des SIG induit un développement des échanges de données géographiques, soit en interne aux services (au sein d'une même DDE, par exemple), soit entre différents services sur un même territoire. Dans le cadre de ces échanges, le besoin de disposer d'un catalogue des données disponibles se fait jour.

Le présent document vise à aider à la mise en œuvre d'un catalogue, dans les services déconcentrés de l'Équipement. Il ne prétend pas donner La Solution à tous les problèmes que pose la mise en œuvre d'un catalogue de données géographiques (le manque de recul sur ce thème étant un élément d'excuse), mais devrait plus modestement aider dans cette démarche.

Un catalogue, pour quoi faire ?

La réalisation d'un catalogue, qu'il s'agisse de celui d'un organisme de vente par correspondance ou d'un catalogue de données géographiques réalisé par une DDE, répond au souci de connaître soi-même et faire connaître à d'autres les produits dont dispose un service, et qu'il peut proposer à la diffusion, en présentant leurs principales caractéristiques. Il offre ainsi la possibilité de trouver la donnée utile parmi une offre variée.

Un catalogue pour les besoins internes au service

Un catalogue de données géographiques sera avant tout réalisé pour des besoins internes au service.

La réalisation d'un catalogue de données géographiques au sein d'un service sera une occasion pour lui de se poser des questions et de documenter ses données. Il lui faudra ainsi décrire et qualifier un minimum les données dont il dispose ou qu'il crée. Ce sera aussi l'occasion d'entamer une réflexion sur les conditions de diffusion de ces données.

Le catalogue fournira alors à cette équipe une description minimum et stabilisée des données disponibles, y compris celles accessibles par plusieurs services. Le catalogue sera

la mémoire du service. Il permettra d'informer tout nouvel utilisateur sur les ressources disponibles. Il jouera ainsi le rôle d'un résumé des principales informations utiles, permettant le renvoi vers les documents plus détaillés (conventions, documentation technique ...).

Un catalogue pour les besoins externes au service

Le catalogue permet aussi de fournir à un service tiers des informations sur les données que l'on pourrait échanger avec lui.

Un catalogue de données doit être mis en œuvre dès que plusieurs utilisateurs sont en situation de vouloir s'échanger des données géographiques, et que, pour ce faire, ils souhaitent connaître et faire connaître le patrimoine de données concerné.

Cependant, l'utilisateur externe n'a pas besoin, pour faire un premier tri parmi les données auxquelles il peut accéder, de connaître l'intégralité de l'information permettant de décrire les données concernées. Ceci représenterait en fait une information trop riche, trop complexe à mettre en œuvre pour certains producteurs, et difficilement interrogeable. Un catalogue de données géographiques ne contient donc pas toutes les spécifications techniques du produit, mais seulement les informations principales nécessaires à un tiers pour juger de l'intérêt pour lui du produit.

Une démarche européenne reprise par la France

La norme européenne et le travail mené par l'AFNOR ont permis de dégager les champs d'information justes nécessaires pour une description minimale, visant à produire un catalogue de données géographiques. Le présent texte prend en compte ces travaux. Vous trouverez en annexe les métadonnées retenues, avec la définition officielle et éventuellement des commentaires.

Un outil logiciel permettant de manipuler les métadonnées retenues par l'AFNOR, dénommé REPORTS, a été développé par le CERTU pour le compte du CNIG et de l'AFNOR. Il s'agit d'un outil minimaliste, permettant principalement la saisie, la consultation et l'import-export des fiches de métadonnées. Il est diffusé libre de droit par le CNIG.

Quelles données cataloguer ?

Une des premières questions qui se pose lorsque l'on entame la réalisation d'un catalogue de données géographiques porte sur les données à faire figurer dans le catalogue.

Données géoréférencées, à référence indirecte ...

Faut-il intégrer au catalogue uniquement les données utilisables directement sous SIG ? Ou faut-il aussi y adjoindre des données localisées de manière indirecte (telles que les données INSEE ou les données VISAGE), qui, bien que ne possédant pas de coordonnées

géographiques, peuvent être intégrées à un SIG moyennant une opération de géocodage ? Le choix dépendra en particulier de l'organisation mise en place dans le service, permettant ou non l'utilisation courante de ces données indirectement localisées dans les SIG. Cependant, il paraît souhaitable d'intégrer à terme dans un même catalogue ces deux types d'informations, souvent complémentaires. De même, on peut se poser la question d'intégrer au catalogue les cartes papier ou les cartes DXF. Ces derniers types de données présentent moins d'enjeux immédiats, et sont souvent difficilement récupérables.

Qualité des données à cataloguer

Les données disponibles dans un service sont qualitativement variées. Certaines proviennent de producteurs externes, qui doivent assurer et décrire la qualité de leurs produits (IGN, Michelin, mais aussi fichiers en provenance d'autres administrations) : leur intégration dans un catalogue va de soit, tout en tenant compte que, du fait de leur statut, le service n'en est pas propriétaire, et ne peut souvent pas les diffuser à l'extérieur. D'autres données sont produites en local : le service en est propriétaire et peut souhaiter les diffuser à l'extérieur. Leur mise en catalogue est souhaitable, sous réserve que ces données soient décrites, en particulier en ce qui concerne leur qualité propre, et que leur contenu soit validé. Enfin, il y a toutes les données produites ponctuellement, de manière souvent temporaire à l'occasion d'une étude, ou d'un travail particulier, ou les données finales d'un travail cartographique. La décision de leur catalogage devra faire l'objet d'un traitement au cas par cas, afin de ne pas alourdir le catalogue avec des données peu utilisables par d'autres que le service producteur, et afin aussi de limiter la charge de travail découlant du catalogage.

Ne pas restreindre a priori le catalogue

Cependant la mise sous catalogue ne doit pas être limitée par des considérations a priori sur l'intérêt potentiel de telle donnée pour autrui. En effet, il est difficile de juger de l'intérêt d'une donnée pour un autre service et ce dernier est finalement le seul juge valable de cette utilité. Il faut, à l'occasion de la réalisation d'un catalogue de données géographiques, viser à faciliter la circulation de l'information et donc sa valorisation par la multiplication des usages, plutôt qu'avoir une position restrictive, voire frileuse. Par contre, il faut être conscient du risque représenté par la diffusion d'une donnée jugée comme non fiable, ou qui n'a pas subi une phase de qualification suffisante.

Que met-on dans une fiche de catalogue ?

Découpage des données en lots

Un catalogue se présente comme un ensemble de fiches, chaque fiche correspondant à un lot de donnée géographique, correspondant à un contenu sémantiquement et techniquement homogène.

Ce découpage en lot est assez intuitif quand la donnée géographique concernée est assez simplement structurée ; ainsi Corine Land Cover, qui est fournie sous la forme d'une seule table de donnée, fera l'objet d'un seul lot. Ce sera le cas le plus courant des données produites par les services, car elles sont souvent mono-fichier et ne comprennent pour la plus part qu'un type d'objet.

La répartition en lots peut être plus compliquée dans le cas de bases de données complexes. La BDCARTO fournit un bon exemple : elle comporte plusieurs dizaines de fichiers, dont certains ne comportent pas d'objet géographique (fichiers d'attributs ou de relations) ; elle est répartie en 7 thèmes, mais en plus de 20 couches géographiques. Une solution simple consiste à répartir la BDCARTO en autant de lots qu'il y a de thèmes. On évite ainsi de multiplier les lots, on garde une homogénéité thématique par lot, et l'on règle le cas des thèmes de la BDCARTO qui peuvent ne pas avoir été acquis (occupation des sols, par exemple).

L'information à saisir par fiche

Chaque fiche comporte une cinquantaine de champs à remplir (voir annexe), certains pouvant être remplis plusieurs fois. Ces informations à remplir sont réparties de la manière suivante :

- Des informations permettant d'identifier le lot de donnée (nom, version)
- Des informations donnant un aperçu simplifié du lot de donnée (résumé, producteur, échelle, utilisation potentielle ...)
- Des informations sur la qualité du lot de donnée (généalogie, date de mise à jour ...)
- Des informations datant les métadonnées elles même
- Des informations sur le système de référence utilisé soit direct (le système de projection) soit indirect (les types d'objets qui servent de référence)
- Des informations sur l'emprise concernée par le lot de donnée
- Une définition plus détaillée des données, avec indication des objets et de leurs attributs, ainsi que des références à un thésaurus éventuel
- Des informations administratives, précisant les organismes ayant un rôle vis-à-vis du lot de donnée (producteur, diffuseur ...), les conditions de diffusion de ces données (propriété intellectuelle, restrictions d'usage, support et format de diffusion ...)

Seuls certains champs sont rendus obligatoires par la norme : ils doivent être remplis. Cependant une réponse du type « Information disponible auprès du contact » ou « information contenue dans le document dddd » ou encore « information non encore connue » est tout à fait valable, si elle correspond à la réalité.

Le but n'est pas de faire une description complète et exhaustive de la donnée, mais bien de proposer les principales indications à un utilisateur potentiel, afin qu'il sache si cette donnée peut lui être utile, quitte à ce qu'il interroge ensuite plus précisément le service concerné. Il ne s'agit donc pas de renseigner chaque objet et chaque attribut des données disponibles, mais de faire une sélection des objets et attributs les plus significatifs. Ce travail de sélection des objets et attributs à renseigner est important à réaliser quand la donnée est compliquée ou très riche (par exemple la Bdcarto). En général, on peut omettre les attributs utiles principalement pour la gestion de la donnée (identifiant, clefs diverses).

Métadonnées sensibles

Certaines informations ne relèvent pas du contenu technique des données, mais de problèmes d'organisation du service, et de ce fait méritent une attention particulière :

- Utilisation potentielle de la donnée : ce champ doit comporter l'indication des raisons ayant abouti à la fabrication de ces données. Ce champ permet aussi d'indiquer les limites d'usage des données. Il est important de bien réfléchir à son contenu. **En particulier dans le cas d'une diffusion externe, cette information pouvant engager le service.**
- Le chapitre sur les contacts permet d'indiquer à qui s'adresser pour plus d'information sur la donnée. Selon le cas et l'organisation mise en œuvre dans le service, une personne peut être désignée comme contact, ou plus naturellement le contact peut être le producteur local ou l'utilisateur principal. **Le choix de cette personne n'est pas anodin. Elle aura à trouver les réponses à des questions tant techniques qu'organisationnelles**, par exemple : « quelles sont précisément les conditions de diffusion de cette donnée pour mon organisme ? ».
- Le chapitre diffusion précise justement ces conditions de diffusion. **Il doit faire l'objet d'une réflexion particulière, afin de bien cerner les conditions de diffusion des données.** Certaines données ne sont évidemment pas diffusables : il s'agit principalement des données acquises auprès d'un tiers et des données nominatives. Le cas des autres données est plus compliqué, et méritera une réflexion interne au service. En attendant de cette réflexion, une réponse du type « condition non encore établie » est suffisamment explicite.

Cas d'un produit composite

Un produit composite ou dérivé est issu d'autre produit, dont il intègre toute ou partie des caractéristiques. Son catalogage suppose donc de fournir une information suffisante sur le ou les produits à l'origine de ce nouveau produit. Ceci doit évidemment apparaître dans la généalogie du produit : on doit y préciser, outre les données dont il est issu, la méthode employée pour passer de ces produits vers le produit dérivé. Si le produit dérivé intègre une part importante des caractéristiques du produit d'origine, il peut être utile de fournir la fiche de catalogue de ce dernier, pour information.

Cas des données en provenance de l'extérieur

Les données acquises à l'extérieur seront de plus en plus souvent accompagnées de leurs métadonnées. Dans ce cas, la réutilisation de ces métadonnées dans le catalogue interne du service semble une bonne chose. **Cependant certains champs devront être modifiés, afin de prendre en compte l'organisation du service et les accords passés avec le producteur.**

Ces champs sont principalement :

- les utilisations potentielles de la donnée : dans certains cas, l'accord avec le producteur limite les droits d'utilisation de ces données (dans le cadre d'un certain type d'étude par

exemple). Par ailleurs, le service peut souhaiter indiquer des utilisations particulières (par exemple, pour la BDCARTO, son usage comme référentiel géographique du 1/50.000 au 1/200.000).

- La surface du lot : pointer si possible sur un fichier comportant l'objet géographique correspondant au périmètre du lot de données.
- Le chapitre sur les organismes peut être modifié, afin d'indiquer le responsable interne de la gestion de la donnée, par exemple.
- Le chapitre sur les contacts devra parfois être modifié, en particulier si une personne est chargée en interne de faire l'assistance de premier niveau vis-à-vis de la donnée concernée.
- Le chapitre sur la diffusion devra évidemment être modifié. Nous n'avons en général aucun droit de rediffusion des données acquises à l'extérieur, et dans le cas contraire les conditions de ces droits sont fixées par contrat avec le producteur. Ce sont ces conditions qui devront désormais figurer dans ce chapitre.

Par contre, il n'est souvent pas souhaitable de modifier les champs techniques du catalogue fournies par le producteur.

Les thésaurus et nomenclatures, outils de classification

Un catalogue vise à faciliter la recherche d'information géographique utile, par une personne qui a priori n'a pas créé cette donnée. Il lui faut donc, outre avoir connaissance de l'existence de la donnée, comprendre ce qu'elle représente, ce qu'elle contient comme information. Il lui faut avoir les moyens de faire le rapprochement entre les dénominations des objets présentés dans les fiches de métadonnées et ses propres besoins, les objets qu'il recherche.

C'est la fonction principale des dictionnaires ou thésaurus : proposer une classification et des définitions pour les objets manipulés dans un domaine, une profession La norme sur les métadonnées permet d'utiliser le renvoi vers des éléments de thésaurus, afin de préciser ou rendre plus universelle la dénomination des objets et attributs présentés dans chaque fiche.

Utilisation d'un thésaurus dans les métadonnées

Chaque fiche peut faire référence à un et un seul thésaurus.

Chaque objet ou attribut peut être référencé par aucun, un ou plusieurs éléments de thésaurus. En effet un objet ou un attribut particulier peut ne correspondre à rien de référencé dans le thésaurus utilisé ; il peut aussi correspondre à un élément précis du thésaurus ; enfin, il peut arriver qu'un objet ou un attribut ne puisse pas être décrit correctement en n'utilisant qu'une référence du thésaurus, mais doit faire l'objet de plusieurs références (par exemple, une rivière navigable est à la fois un milieu naturel et

une voie de communication). En règle générale, si l'on utilise un thésaurus adapté à son métier, la plupart des objets seront référencés par un et un seul élément du thésaurus.

Enfin, dans chaque fiche de métadonnées, les éléments de thésaurus utilisés sont précisés par une définition et des synonymes. Ces précisions doivent dans la mesure du possible découler du thésaurus choisi.

Quel thésaurus ou nomenclature utiliser ?

Une des principales difficultés de l'utilisation des thésaurus est leur existence encore très limitée en matière d'information géographique.

Sont principalement connus en France :

- la nomenclature EDIGéO du CNIG (disponible auprès du CNIG, à Paris) : elle est généraliste, hiérarchisée
- le dictionnaire de données du Secrétariat d'Administration National des Données Relatives à l'Eau, appelé couramment SANDRE (disponible auprès de l'OIEAU) : elle est très thématique (eau et environnement) et très détaillée (données avec attributs, indiquant en particulier les valeurs possibles).

Il peut par ailleurs exister d'autres thésaurus ou dictionnaires thématiques spécialisés, ou des démarches locales pour aboutir à une nomenclature acceptée par une communauté particulière. En l'absence de thésaurus adapté, on peut en général utiliser les premiers niveaux de la nomenclature hiérarchique du CNIG, ce qui n'indique pas précisément l'objet concerné, mais une classe plus générique incluant cet objet (voie de communication, par exemple).

Quelle stratégie de saisie des fiches du catalogue ?

Le remplissage d'un catalogue de donnée représente un travail non-négligeable, pouvant aller jusqu'à la demi-journée de travail par fiche. Il vient en surplus du travail habituel, et les principaux acteurs n'en sont pas les bénéficiaires immédiats. Il faut donc réfléchir à la meilleure stratégie permettant de réaliser au mieux cette saisie.

Il est fort probable que, dans le cas d'un SIG un peu ancien, le nombre de données existantes dans le service soit important, et que leur mise sous catalogue soit plus ou moins attendue ou facile.

Une première approche stratégique consiste à définir un ordre de catalogage des lots de données, permettant à la fois d'étaler dans le temps le travail à réaliser et d'offrir rapidement une information utile et complète. On peut ainsi commencer par cataloguer les référentiels, les données acquises à l'extérieure, puis les données déjà documentées, pour enfin finir par les données non encore décrites.

Une autre approche consiste à s'intéresser plus à la capacité de mobilisation des agents concernés par le catalogue. Une première solution est de faire réaliser le gros du travail

par un stagiaire ou un intervenant extérieur. Cette solution limite la contribution des agents producteurs qui n'ont plus qu'à répondre à une enquête, y compris en fournissant des documents déjà disponibles, et à valider la restitution. Cette solution est très adaptée pour une saisie initiale importante.

Une autre solution consiste à commencer par les services les plus motivés, les plus moteurs dans la démarche ou les plus disponibles, en espérant que l'effort à déployer pour mobiliser les suivants sera d'autant plus faible que ceux-ci verront concrètement l'utilité de la démarche.

Enfin, une dernière solution consiste dans un premier temps à ne saisir que les champs obligatoires plus quelques champs jugés faciles à saisir ou indispensables pour en savoir plus (par exemple : qui contacter). Le travail de saisie complète des fiches peut alors être reporté à plus tard. Cette solution permet rapidement de connaître toutes les données disponibles, et laisse le temps d'étudier l'information à porter sur certains champs particulièrement sensibles ou complexes. Cependant, il faut veiller à ce que le complètement des fiches ait bien lieu. Cette reprise du travail peut en effet être assez démoralisante.

Le démarrage du catalogue rapidement après le début de la mise en place des SIG reste le cas le plus simple à gérer. Le travail de saisie est alors étalé dans le temps, et les questions sont résolues au fur et à mesure de leur apparition. Cependant, certains champs complexes à remplir, et supposant une décision non encore aboutie (conditions de diffusion, par exemple), peuvent alors n'être pas remplis dès le début.

Quelle organisation ?

La réalisation d'un catalogue de données va amener à se poser des problèmes d'organisation, car elle fait intervenir de multiples acteurs au sein du service et dans son environnement.

Réalisation des fiches

Il faut définir qui remplit les fiches du catalogue, et sous quelle responsabilité. Une solution simple et assez logique consiste à donner la responsabilité du contenu des fiches au producteur de la donnée, si elle est produite en interne, au gestionnaire ou au principal utilisateur de la donnée, si elle a été acquise à l'extérieur. Par contre, la tâche de saisie peut être réalisée par une tierce personne, ce qui en assurera une certaine homogénéité, et libérera d'autant les responsables.

Enfin, certaines rubriques d'un catalogue posent des problèmes de responsabilité pour le service : conditions de diffusion des données Il est utile de prévoir une phase de validation des fiches, au moins sur ces points-là, faite sous l'égide du chef de service concerné.

Mise à disposition du catalogue

Les fiches une fois rédigées, il faut les mettre à disposition des utilisateurs potentiels internes et externes. En interne, le catalogue gagne à être mis à disposition sur le réseau interne du service, de manière centralisée ou non. Il pourra aussi être distribué en complément sous forme papier ou disquettes, en particulier à l'externe. Une fonction de gestion technique de cette mise à disposition doit être organisée, afin d'assurer le bon fonctionnement et la pérennité de ce circuit d'information.

Organisation de la diffusion des données

La mise en circulation du catalogue de données à l'extérieur posera implicitement le problème des modalités de la diffusion de ces données en dehors du service. Un des objectifs est en effet de faciliter cette diffusion, de la rendre plus professionnelle. Rapidement la demande va croître. La simple réponse au cas par cas face aux sollicitations ne sera plus une solution satisfaisante, et des décisions d'organisation de cette fonction de diffusion de données devront alors être prises.

Vie du catalogue

Enfin, il faut faire vivre ce catalogue, c'est à dire mettre à jour les fiches déjà remplies, mais aussi veiller à son utilisation et au maintien de sa pertinence. La mise à jour des fiches relève a priori plutôt des mêmes personnes que pour la saisie initiale. Par contre, un lieu d'animation entre utilisateurs internes du catalogue sera indispensable. Il permettra de valider à l'usage l'intérêt des fiches intégrées au catalogue, de veiller à ce que des pans entiers de données géographiques ne soient pas absentes du catalogue, et de proposer des solutions à quelques problèmes ou questions qui se feront jour. A cette occasion, un minimum de retour d'information sur la diffusion, en particulier vers l'extérieur, des données inscrites au catalogue trouvera son utilité.

On le voit, un minimum d'organisation et d'animation est à assurer pour la bonne mise en œuvre d'un catalogue de données. **C'est la responsabilité éditoriale directe de la direction du service.**

ANNEXE n° 1

Définitions

Mise en œuvre d'un catalogue de données géographiques

Norme Européenne Métadonnée

Définitions

Les textes suivants sont extraits de la traduction française de la norme sur les métadonnées. Il ne s'agit que d'une partie réduite des définitions et textes précisant le sens des termes employés dans la norme ; pour plus de précision, ou pour connaître la liste complète des définitions, il faut se reporter à la norme elle-même.

Définitions

Métadonnées :

Métadonnées : données relatives à un ou plusieurs lots de données géographiques

Les métadonnées peuvent être définies très simplement comme des "données sur les données". La norme européenne spécifie les données qui doivent être utilisées pour décrire un lot de données géographiques. Cela englobe les données relatives au contenu, à la représentation, à l'emprise (tant géométrique que temporelle), au système de référence spatiale, à la qualité et à la gestion du lot de données géographiques.

La norme européenne identifie également les données obligatoires pour la description des lots de données géographiques, à savoir le lot minimum de métadonnées.

Emprise :

Espace couvert par un lot de données géographiques

NOTE Le terme "espace" est utilisé ici dans son sens euclidien mais aussi pour désigner une période de temps.

Lot de données géographiques :

Ensemble identifiable de données géographiques

Un lot de données géographiques contient un ou plusieurs objets géographiques décrits par leurs attributs et leurs relations. Des données doivent être fournies pour décrire le lot lui-même, mais peuvent l'être à un niveau plus détaillé afin de décrire l'objet en utilisant le type d'objet, le type d'attribut et le type de relation que contient le lot de données géographiques.

Informations géographiques :

Informations concernant des phénomènes directement ou indirectement associés à une localisation terrestre

Qualité :

Ensemble des caractéristiques d'un produit qui lui confère son aptitude à satisfaire des besoins exprimés et implicites

Paramètre de qualité :

Elément de qualité quantifiable décrivant la conformité d'un lot de données géographiques par rapport à son terrain nominal

Schéma spatial :

Schéma conceptuel pour les aspects spatiaux (géométriques et topologiques) des données géographiques

ANNEXE n° 2

Liste des champs et conseils de saisie

Catalogue de données géographiques

Liste des champs et conseils de saisie

(*) : « o » et « f » signifient respectivement obligatoire et facultatif.

Les listes codées en chiffre (ex. liste 3) sont indiquées afin d'aider à la saisie de certaines informations. Elles peuvent donc être complétées ou modifiées.

Les listes codées en lettre (ex. liste a) découlent de la normalisation.

Chapitre	Champs	o/f (*)	Définition de la norme	Commentaires	Liste	Champs multiples
Identification du lot de données		o				
	Titre du lot de données	o	Nom explicite du lot de données géographiques, permettant son identification par les utilisateurs.	Indiquer le nom d'usage courant (BDCARTO, plutôt que banque de données cartographique)		
	Version	f	Numéro de version ou autre description de la version du lot de données géographiques.	Ex : V2.3		
Aperçu du lot de données		o				
	Résumé	o	Brève description du lot de données géographiques résumant le contenu du lot.	Mettre ici une description suffisamment détaillée, permettant des recherches sur mots-clefs.		
	Nom du producteur	o		utiliser un nom réduit (ex DDE 55, ou IGN)		
	Type de schéma spatial	o	Type du schéma spatial du lot de données géographiques.	En général « utilisateur » ou « sans schéma »	liste (a)	

Chapitre	Champs	o/f (*)	Définition de la norme	Commentaires	Liste	Champs multiples
	Schéma spatial utilisateur	o	Informations relatives aux principaux composants et aux caractéristiques générales du schéma, s'il ne s'agit pas d'un schéma spatial prédéfini dans l'ENV 12160.	Indiquer les structures prévues dans la norme EDIGéO, soit : Mapinfo : spaghetti Arcinfo : topologique ou indiquer un autre type de schéma spatial (il peut ne pas y en avoir, dans le cas d'un positionnement indirect)	liste (b)	
	Langue	o	Langue(s) utilisée(s) pour le texte dans le lot de données géographiques et identifiée(s) par des codes définis dans l'ISO 639	Prédéfini à Français		
	Jeu de caractères	o	Codification de jeu de caractères (par exemple, ISO 8859-10) utilisée dans le lot de données géographiques	Prédéfini à ISO 8859-10		
	Objectifs de la production	f	Objectifs pour lequel un lot de données géographiques a été produit par le fournisseur. Il peut s'agir d'une description	Important pour comprendre ce qu'on peut faire avec la donnée. Y indiquer éventuellement les textes législatifs ayant déclenché leur fabrication.		
	Echelle d'application prévue	f	Description de l'échelle d'application prévue du monde réel décrit par le lot de données géographiques.	On prévoit d'indiquer l'échelle mini et l'échelle maxi.		
	Utilisation potentielle	f	Liste (sous forme de texte) d'applications pour lesquelles le lot de données géographiques peut être utilisé.	Préciser en particulier les usages non recommandés		

Chapitre	Champs	o/f (*)	Définition de la norme	Commentaires	Liste	Champs multiples
	Référence(s) de document	f	Référence(s) à une documentation déjà publiée ou disponible relative au lot de données géographiques, y compris la ou les langues de la documentation.	Permet à une personne désirant avoir plus d'information de s'y référer ; mettre en particulier les documentations techniques et de livraison, ainsi que les textes de loi définissant cette donnée ou les objets concernés.		
	Echantillon	f	Un ou plusieurs exemples tirés du lot de données géographiques et étant représentatifs de l'ensemble du lot ; par exemple graphique de survol au format vecteur ou raster. Ils sont fournis par le contenu, le format et l'emplacement du nom de fichier.	Lien vers fichiers image ou autre, permettant de montrer une sorte graphique des données, ou proposant l'accès à un petit échantillon.		
Qualité du lot de données		o				
	Généalogie	o	Description de l'historique du lot de données géographiques : production, fournisseur, objectif de production, date de production, source et processus.	C'est la méthode la plus simple pour décrire la qualité		
	Date de validité	f	Date à laquelle on considère que toutes les informations contenues dans le lot de données géographiques sont correctes.			
	Texte sur la qualité	f	Texte décrivant la qualité du lot de données géographiques.	Permet d'indiquer tout autre élément de qualité, par exemple une estimation de la précision planaire ou de l'exhaustivité.		

Chapitre	Champs	o/f (*)	Définition de la norme	Commentaires	Liste	Champs multiples
Référence des métadonnées		o				
	Date de saisie	o	Date de création des métadonnées.	A remplir impérativement		
	Date de dernière mise à jour	f	Date à laquelle les métadonnées ont été mises à jour pour la dernière fois.	A remplir impérativement s'il y a eu une mise à jour des métadonnées		
Système de référence spatial		o				
<i>Système de positionnement indirect</i>		f		Cas par exemple des données INSEE, localisées selon le code INSEE communal		
	Nom du système de positionnement indirect	o	Système de positionnement indirect dans lequel sont données les références à une position.	Par exemple, divisions administratives, adresses de rue.	Liste (1)	
	Nom des classes de localisation	o	Noms des classes de localisation utilisées dans le système de positionnement indirect.	Par exemple, pays, région, municipalité.	Liste (2)	
	Date de référence	f	Date à laquelle cette version du système de positionnement indirect renvoie.	Important, en particulier pour des systèmes pouvant varier dans le temps : parcelles cadastrales, communes, adresse postale ...		
	Administrateur du système de référence	f		Important pour pouvoir remonter à la source de l'information		
<i>Système de positionnement direct</i>		o		Cas des SIG		
	Identificateur	o		Pris dans la liste de l'annexe B d'EDIGÉO ; en général, il s'agira d'un Lambert	liste (c)	
	Nom du système	o		Correspondant à l'identifiant	déduit liste (c)	

Chapitre	Champs	o/f (*)	Définition de la norme	Commentaires	Liste	Champs multiples
Emprise		o				
<i>Actualité des données d'emprise</i>		o				
	Etat de l'emprise	o	Indique dans quelle mesure la description de l'emprise reflète l'emprise finale proposée pour le lot de données géographiques ainsi que toute variation dans la classification.	Valeur par défaut à "réelle" Autre valeur possible : « future »		
	Date de validité de l'emprise	o	Date, donnée comme date de l'emprise, ou période, donnée comme date de début et/ou de fin, à laquelle l'état et la description de l'emprise sont valides.	Devrait être identique à la date de la dernière mise à jour.		
<i>Emprise planaire</i>		o				
	Surface de délimitation	f	Délimitation de la surface couverte par le lot de données géographiques.	Pointe vers un fichier contenant les coordonnées du contour de l'emprise		
	Nom de la surface couverte par le lot de données	o		Permet de nommer l'emprise (communes, départements...)		
Définition des données		o				
	Description textuelle générale	o		Complète en précisant le résumé. Y indiquer en particulier l'ensemble des objets contenus dans le lot.		
<i>Description du schéma d'application</i>		o		Mettre les principaux objets du lot, en faisant un enregistrement par objet.		
<i>Objet(s)</i>		o				plusieurs
	Nom d'objet	o	Chaîne de texte décrivant l'objet.	Donner le nom en clair, pas le nom du fichier.		
	Définition de l'objet	o	Description de l'objet ou référence à une définition standard ou de norme existant.	Mettre ici une définition de l'objet		

Chapitre	Champs	o/f (*)	Définition de la norme	Commentaires	Liste	Champs multiples
	Elément de thesaurus	f	Code de l'élément de Thésaurus utilisé pour préciser l'objet	Mettre ici le nom de l'élément de thesaurus. La définition et les synonymes sont saisis plus loin		plusieurs
<i>Attribut(s)</i>		f		Mettre les principaux attributs de l'objet, en faisant un enregistrement par attribut.		plusieurs à partir d'un objet
	Nom de l'attribut	o	Chaîne textuelle décrivant l'attribut	Donner le nom en clair, pas le nom de code de l'attribut		
	Définition de l'attribut	o	Description de l'attribut ou référence à une définition standard ou de norme existante.	Mettre par exemple les principales occurrences possibles de l'attribut.		
	Elément de thesaurus	f	Code de l'élément de Thésaurus utilisé pour préciser l'attribut	Mettre ici le nom de l'élément de thesaurus. La définition et les synonymes sont saisis plus loin		
<i>Thesaurus</i>		f				
	Nom du Thesaurus	f	Nom d'un thesaurus existant standard ou spécifique d'une application.	Faire ici référence à un thesaurus ou dictionnaire métier sur lequel est basé la nomenclature des objets (EDIGÉO, SANDRE)		
	Administrateur du Thesaurus	f	Nom de l'organisation ou de la ou des personnes responsables du maintien à jour du thesaurus.			
	Date de publication	f				
	Version	f				
<i>Eléments de thesaurus</i>				o si on ne prend pas en compte le chapitre précédent		plusieurs
	Terme	f	Texte ou code identifiant un élément de thesaurus particulier ; par exemple, route, chaussée, transport, couverture terrestre.	Mettre ici au moins tous les termes du thesaurus qui sont utilisés dans la présentation des données		

Chapitre	Champs	o/f (*)	Définition de la norme	Commentaires	Liste	Champs multiples
	Définition	f	Sens distinct du terme. Lorsque le terme n'appartient pas à une liste de codification de norme, la définition est obligatoire.	Indiquer la définition telle qu'elle est inscrite dans le thésaurus utilisé		
	Synonyme	f	Terme dont la signification est équivalente à celle du terme principal.			
Métadonnées Administratives						
	<i>Organisme et rôle de l'organisme</i>	o		Au moins les références d'un diffuseur et de l'auteur des métadonnées		plusieurs
	Nom de l'organisme	o	Nom de l'organisme, description de son rôle général sous forme d'un texte descriptif et d'un identificateur.	Ex : Institut Géographique National, Direction Départementale de l'Équipement de l'Eure		
	Nom abrégé de l'organisme	o		Ex : IGN, DDE 56		
	Adresse de l'organisme	o	Adresse postale, numéro de téléphone, numéro de télécopie, adresse de courrier électronique.			
	Rôle de l'organisme	o	Rôle pour lequel l'organisme est responsable d'un lot de données géographiques. Un organisme doit avoir au moins un rôle.	Si un organisme joue plusieurs rôles, les indiquer.	Liste (3)	
	Page d'accueil Internet de l'organisme	f	Adresse Internet (www) de la page d'accueil de l'organisme.			
<i>Contact et rôle du contact</i>						
	Nom du contact	o	Nom et/ou titre de la personne jouant le rôle de contact.	Important pour obtenir plus d'information. Indiquer ici la personne servant de contact pour toute personne voulant un renseignement.		
	Adresse du contact	o	Adresse postale, numéro de téléphone, numéro de télécopie, adresse de courrier électronique du contact.			

Chapitre	Champs	o/f (*)	Définition de la norme	Commentaires	Liste	Champs multiples
	Rôle du contact	o	Responsabilité du point de contact par rapport au lot de données géographiques. Par exemple, administrateur technique et/ou commercial du lot de données géographiques ou auteur des métadonnées..	Un point de contact doit avoir au moins un rôle. L'auteur des métadonnées doit toujours être indiqué	liste (4)	
	<i>Diffusion</i>	f				
	Restrictions d'utilisation	o	Contraintes, autres que le copyright, réglant l'accès au lot de données géographiques et son utilisation.	A remplir avec attention. A faire valider.		
	Propriété intellectuelle	o	Organisme(s) détenant le copyright du lot de données géographiques.	A remplir avec attention. A faire valider.		
	Unité de diffusion	o	Informations relatives au découpage géographique et/ou thématique du lot de données géographiques ; indique par exemple si le lot est divisé en tuiles carrées de x par x km ou s'il est diffusé par type d'objet, etc.		liste (5)	
	Tarifs	f	Frais relatifs au lot de données géographiques, incluant le prix unitaire et les possibilités de remise.	A remplir avec attention. A faire valider.		
	Support	f	Supports matériels sur lesquels le lot de données géographiques peut être enregistré et à partir desquels il peut être récupéré	Indiquer aussi le système d'exploitation correspondant	liste (6)	plusieurs
	Formats	f	Formats dans lesquels le lot de données géographiques peut être livré.		liste (7)	plusieurs
	Accès en ligne	f	Informations relatives à l'accès en ligne au lot de données géographiques.			

Chapitre	Champs	o/f (*)	Définition de la norme	Commentaires	Liste	Champs multiples
	Modalités de commande	f	Description du service de commande et de livraison des données. Cette description inclut les instructions relatives à la commande des données et le délai prévu pour la livraison.	A remplir avec attention. A faire valider.		
	Autres services	f	Autres services de traitement du lot de données géographiques, mis à la disposition des utilisateurs du lot par l'organisation.	Ex : mise au format, extraction d'une partie des attributs, conseils		

(a) : Schémas spatiaux :

- schéma spatial de base G0
- schéma topologique avec graphe planaire complet G1
- schéma réseau linéaire avec graphe planaire G2
- schéma réseau linéaire sans graphe planaire G3
- schéma réseau linéaire sans graphe planaire avec surface G4
- schéma spaghetti G5
- schéma réseau irrégulier de triangulation bornée G6
- schéma image raster G7
- schéma grille G8
- schéma utilisateur
- sans schéma spatial

(b) : Structures utilisateur

- Spaghetti
- Réseau
- Topologique
- Raster

(c) : Annexe B EDIGÉO et nom des systèmes de positionnement direct

(1) : Type de positionnement indirect

1. unité administrative
2. adresse
3. abscisse curviligne
4. découpage cartographique
5. autre
6.

(2) : Positionnements indirects (le second chiffre en début d'item renvoie à la liste (2) précédente) (seuls les termes en gras figurent dans la liste d'aide à la saisie. Les autres informations sont là soit pour préciser le sens donné au texte en gras, soit pour indiquer le codage correspondant attendu dans les données géographiques concernées) :

1. 1 . **pays** selon norme NF EN ISO 31-66-1
2. 1 . **région** française définie par son nom complet en majuscules sans accent sans blanc codé sur au plus 26 positions (voir liste ci-jointe)
3. 1 . **département** français défini par son code INSEE sur au plus trois positions (deux positions pour les départements métropolitains (exemple: "2A" pour la Corse du sud ; trois positions pour les 4 DOM ("971", "972", "973" et "974"), Saint-Pierre-et-Miquelon ("975")) et les TOM ("984", "985", "986", "987" et "988"))
4. 1 . **commune** définie par son code INSEE sur 5 positions
5. 1 . **îlot INSEE**
6. 1 . **quartier INSEE**
7. 4 . **section et subdivision de section cadastrale**
8. 1 . **parcelle cadastrale**
9. 2 . **code postal** sur 5 positions
10. 2 . **adresse postale**
11. 2 . **adresse de visite** (numéro, nom de rue, commune)
12. 3 . **PR routier** (numéro de route + PR + abscisse)
13. 3 . **PR navigation** (VNF)
14. 5 . **point RDS-TMC**
15. 5 . **code hydrographique d'un cours d'eau** sur 8 positions conformément à la circulaire ministérielle n° 91-50 du 12 février 1991
16. 3 . **code hydrographique d'un cours d'eau et point kilométrique (PK)** sur ce cours d'eau compté en kilomètres à partir 1000 à l'embouchure du cours d'eau en décroissant
17. 5 . **code d'une zone hydrographique** sur 4 positions conformément à la circulaire ministérielle n° 91-50 du 12 février 1991
18. 4 . **coupure de carte IGN au 1 :50 000**, plus éventuellement découpage est/ouest
19. ...

(3) : Rôle des organismes :

1. producteur
2. intégrateur
3. gestionnaire
4. administrateur
5. diffuseur
6. commercial
7. utilisateur
8. auteur de la fiche
9. ...

(4) : Rôle des contacts :

1. producteur
2. intégrateur

3. gestionnaire
4. administrateur
5. diffuseur
6. commercial
7. utilisateur
8. auteur de la fiche
9. gestionnaire du thésaurus
9. ...

(5) : Unité géographique de diffusion (seuls les termes en gras figurent dans la liste d'aide à la saisie. Les autres informations sont là soit pour préciser le sens donné au texte en gras, soit pour indiquer le codage correspondant attendu dans les données géographiques concernées) :

1. **pays** selon norme ISO 31-66
2. **région** définie par son nom complet en majuscules sans accent sans blanc codé sur au plus 26 positions (voir liste ci-jointe)
3. **département** défini par son code INSEE sur au plus trois positions (deux positions pour les départements métropolitains (exemple: "2A" pour la Corse du sud ; trois positions pour les 4 DOM ("971", "972", "973" et "974"), Saint-Pierre-et-Miquelon ("975")) et les TOM ("984", "985", "986", "987" et "988"))
4. **commune** définie par son code INSEE sur 5 positions
5. **section et subdivision de section cadastrale**
6. **coupure de carte IGN au 1 :50 000**
7. **km 2**
8. ...

(6) : Support :

- | | | |
|-----|----------------|------|
| 1. | Disquette 3' ½ | PC |
| 2. | CD ROM ISO | PC |
| 3. | CD ROM non ISO | PC |
| 4. | Syquest | PC |
| 5. | QIC 150/250 | PC |
| 6. | DAT 4 mm | PC |
| 7. | DAT 8 mm | PC |
| 8. | ZIP | PC |
| 9. | JAZ | PC |
| 10. | DON | PC |
| 11. | Disquette 3' ½ | Mac |
| 12. | CD ROM ISO | Mac |
| 13. | CD ROM non ISO | Mac |
| 14. | Syquest | Mac |
| 15. | QIC 150/250 | Mac |
| 16. | DAT 4 mm | Mac |
| 17. | DAT 8 mm | Mac |
| 18. | ZIP | Mac |
| 19. | JAZ | Mac |
| 20. | DON | Mac |
| 21. | Disquette 3' ½ | UNIX |

22.	CD ROM ISO	UNIX
23.	CD ROM non ISO	UNIX
24.	Syquest	UNIX
25.	QIC 150/250	UNIX
26.	DAT 4 mm	UNIX
27.	DAT 8 mm	UNIX
28.	ZIP	UNIX
29.	JAZ	UNIX
30.	DON	UNIX
31	...	

(7) : Formats :

1. EDIGéo
2. ArcInfo export
3. MIF/MID
4. DXF
5. DGN
6. FEIV
7. Export Géoconcept
8. Export APIC
9. shape files ESRI
10. DX90
11. TIF
12. JPG
13. Géospot
14.

ANNEXE n° 3

Bibliographie

Catalogue de données géographiques Bibliographie

Norme prENV 12657 : information géographique - description des données - métadonnées ;
AFNOR , 1998

Guide de mise en œuvre de la norme expérimentale XP ENV 12567 ; AFNOR 1998

Norme EDIGÉO NF Z 52000 : Echange de données informatisées dans le domaine de
l'information géographique ; AFNOR 1999

Actes du séminaire du 27-28 août 1998 à Lausanne sur « situation et développement des
géorépertoires en Europe et en Suisse » ; EPFL, 1998

Actes de la journée de la recherche géographique « le catalogage des données » à Lyon ;
CNIG , 1997

© Ministère de l'Équipement, des Transports et du Logement
Centre d'études sur les réseaux, les transports, l'urbanisme et les constructions publiques

Toute reproduction intégrale ou partielle, faite sans le consentement du CERTU est illicite (loi du 11 mars 1957).
Cette reproduction par quelque procédé que ce soit, constituerait une contrefaçon sanctionnée par les articles
425 et suivants du code pénal.

Reprographie : CETE de Lyon ☎ 04 72 14 30 30
Achévé de reprographié : janvier 1999
Dépôt légal: 1^{er} trimestre 1999
ISSN: 1263-2570
ISRN Certu RE 99-05

CERTU
9, rue Juliette-Récamier
69456 Lyon Cedex 06
☎ 04 72 74 59 71
Mel : geomatiq@certu.fr
Internet <http://www.certu.fr>



ISSN 1263-2570
ISRN Certu RE 99-05
Gratuit

Certu

Aménagement et urbanisme

*Aménagement
et exploitation de la voirie*

Transport et mobilité

Constructions publiques

Environnement

*Technologies
et systèmes d'information*

*Service technique placé sous l'autorité
du ministre chargé de l'Équipement,
des Transports et du Logement, le CERTU
(Centre d'études sur les réseaux, les transports,
l'urbanisme et les constructions publiques
a pour mission de contribuer
au développement des connaissances
et des savoir-faire et à leur diffusion
dans tous les domaines liés aux questions
urbaines. Partenaire des collectivités locales et
des professionnels publics et privés, il est
le lieu de référence où se développent
les professionnalismes au service de la cité.*