

Le catalogage des données géographiques

Henri Pornon

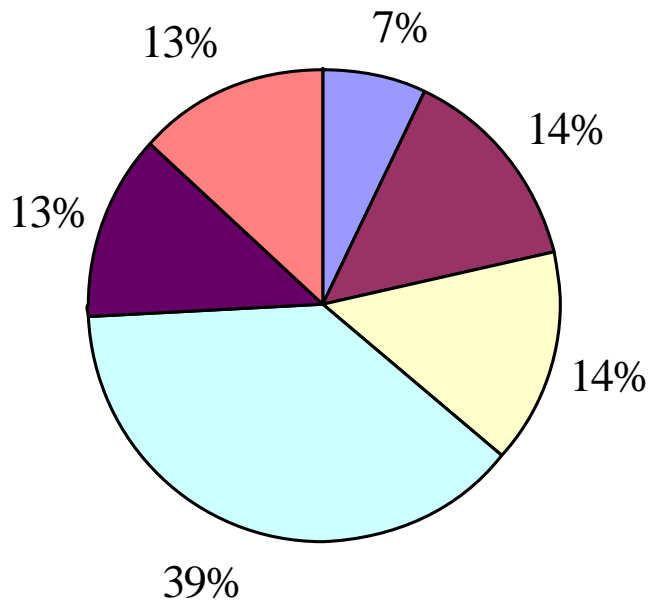
■ Plan de la présentation

- ❑ Restitution de l'état des lieux concernant les démarches de catalogage de données
- ❑ Concepts de base du catalogage des données
- ❑ Les métadonnées en 6 questions
- ❑ Comment faire ?
- ❑ Présentation du plan de formation
- ❑ Questions réponses

■ Restitution de l'état des lieux concernant les démarches de catalogage de données

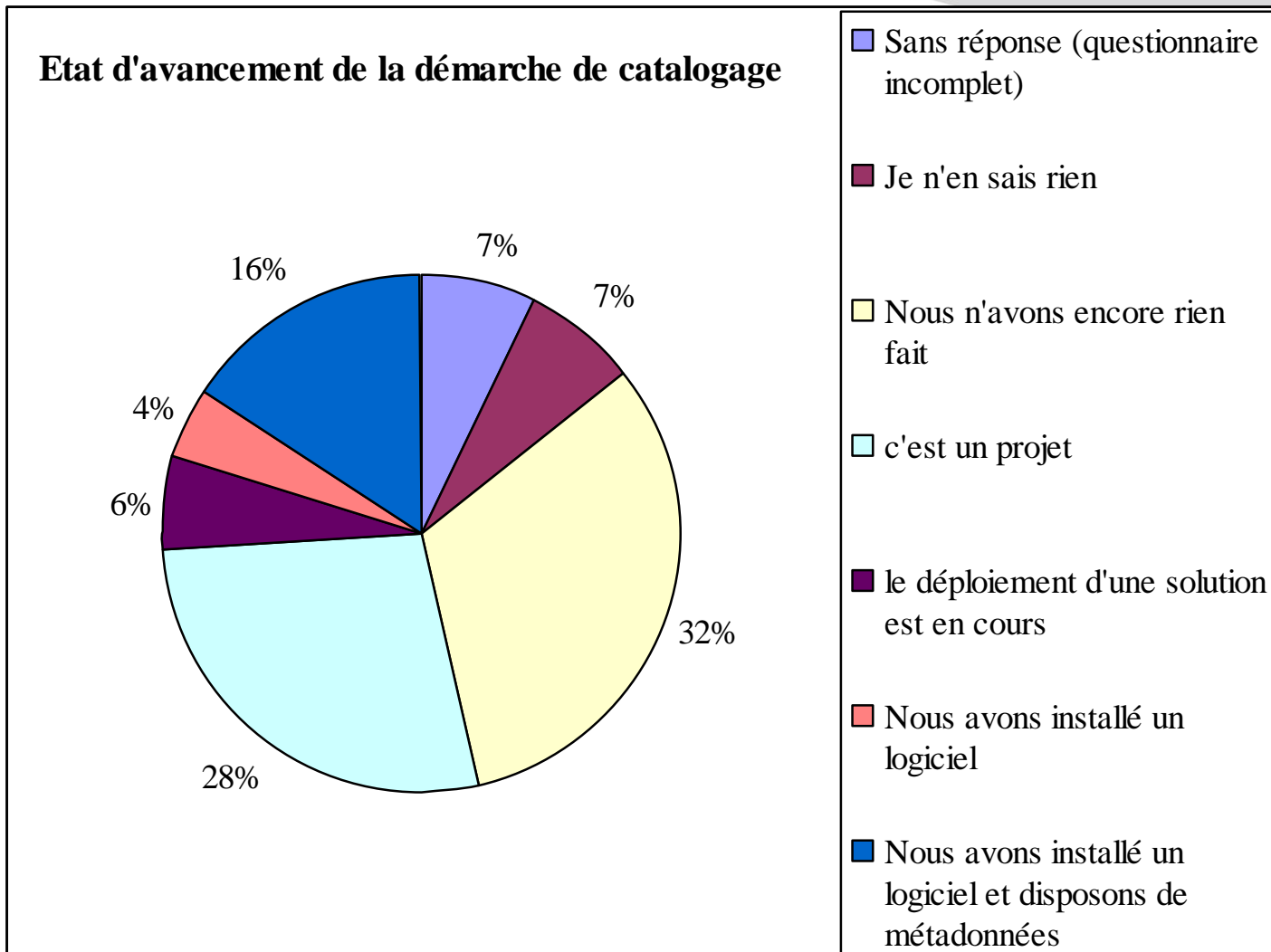
■ Connaissance du Géoportail IGN

Connaissance du Géoportail IGN



- Sans réponse (questionnaire incomplet)
- Je n'en ai jamais entendu parler
- De nom seulement
- J'en ai entendu parler
- Je connais son contenu
- Je connais son contenu et je l'ai déjà utilisé

■ État d'avancement : catalogue



■ Commentaires sur le catalogage

- 26 % des organismes ont déjà installé un outil de catalogage ou sont en train de le faire
 - 70 % des administrations, cinq collectivités et six autres organismes
- Normes évoquées
 - 21 organismes évoquent la compatibilité de leurs outils avec les normes ISO 19115 et 19139, et un avec l'ancienne norme CEN 12657
 - Deux autres standards évoqués : Dublin Core et ISO 19110

■ Inventaire des logiciels de catalogage acquis ou pressentis

- ❑ 24 organismes évoquent un logiciel de catalogage utilisé ou envisagé
- ❑ Les services déconcentrés de l'Etat utilisent ADELIE et GEOREPERTOIRE (outils de catalogage des Ministères de l'Ecologie et de l'Agriculture, cités chacun 5 fois)
- ❑ Géosource est cité 13 fois,
- ❑ Arc/Catalog d'ESRI est cité 6 fois
- ❑ Autres citations :
 - Géocatalogue par le BRGM
 - Un outil interne par l'ADT 21
 - EDITOP par la CC de Matour
 - Géonetwork par le CG21 (choix non arrêté)

■ Moyens affectés au catalogage

- ❑ Pas de moyens ou pas d'évaluation : AELB, Bourgogne Développement, CAMVAL, Chambre d'Agriculture 21, CG58, CG71, CROS, FDC71, ORS, SAFER, SDIS89, Mâcon
- ❑ Attente d'indications sur l'évaluation : CRPF, CG21
- ❑ Activité prise en compte, mais pas d'évaluation : BRGM, CUCM, Conseil Régional
- ❑ A temps perdu : Agglomération Nevers
- ❑ Evaluation réalisée : AGROSUP Dijon (0,05 ETP), CRCI (4 mois à l'initialisation, temps partiel aujourd'hui), DDT71 (2 personnes ?), DDEA58 (0,1 ETP), DDEA89 (1 ETP ?), DDAF21 (une personne à temps partiel), DIREN (une personne y consacre peu de temps), DRAAF (0,25 ETP), DRE (0,25 ETP)
- ❑ Les ressources sont en général la cellule SIG ou l'ADL
- ❑ Difficulté de distinguer les moyens alloués au catalogage et ceux alloués au SIG et aux données géographiques

■ Problèmes concernant la directive INSPIRE

- ❑ la sensibilisation insuffisante à la directive dans l'organisme
- ❑ l'identification des obligations et des impacts,
- ❑ le manque de temps et la nécessité de dégager des moyens techniques et humains, pour le catalogage et pour la diffusion des données, la difficulté d'évaluer la charge de travail requise
- ❑ les aspects concernant la mutualisation des données : méfiance des producteurs internes de données face à la démarche de catalogage, manque de sensibilisation au partage et à la diffusion des données
- ❑ la conduite de la démarche de mise en conformité
- ❑ la présence dans les services déconcentrés de deux outils de catalogage (ADELIE et GEOREPERTOIRE)
- ❑ les documents disponibles en anglais
- ❑ Les aspects relatifs à la qualité et la fiabilité des données

■ Problèmes concernant le catalogage des données

- ❑ Manque de temps et caractère chronophage de cette activité
- ❑ Faible motivation des agents, difficultés de les mobiliser pour cataloguer et mutualiser leurs données, d'où la nécessité de réaliser le catalogage avec les services producteurs constatée par une réponse
- ❑ manque de compétences en interne
- ❑ Compréhension des champs de la norme
- ❑ Que cataloguer, comment, jusqu'où aller ?
- ❑ Mauvaise adaptation de la norme à certaines données (pédologiques)
- ❑ L'intérêt et la passion que le catalogage suscite (méfiances et réticences)

■ Besoins de formation

- ❑ Sensibilisation des utilisateurs de SIG au catalogage (2 demandes)
- ❑ Formation des administrateurs au catalogage (15 demandes) et aux outils (6 demandes)
- ❑ Formation à INSPIRE (5 demandes)
- ❑ 6 organismes évoquent le besoin de formation sans le préciser

Département	Formation non précisée	Sensibilisation	Administration	Logiciels	Inspire
Autre	1	1	1	1	0
Côte d'Or	1	1	8	2	4
Nièvre	0	0	3	0	1
Saône-et-Loire	0	0	2	3	0
Yonne	0	0	1	0	0
Total	2	2	15	6	5

■ Thèmes abordés de façon plus précise

- ❑ Mise en place d'une démarche de catalogage
- ❑ Comment ça marche ?
- ❑ Quel logiciel utiliser ?
- ❑ Définition du périmètre du catalogage
- ❑ Comment renseigner les champs d'une fiche type
- ❑ Mise au point sur les normes
- ❑ Statut juridique de la donnée
- ❑ Un participant souhaiterait voir un catalogage déjà réalisé avec les différentes étapes de la réalisation

■ Besoins d'assistance

- ❑ Assistance au catalogage des premiers lots, assistance à la prise en main de l'outil de catalogage et assistance au catalogage (9 demandes)
- ❑ Aide à la mise en œuvre de la directive INSPIRE (4 demandes)
- ❑ Conseil sur le choix du logiciel (2 demandes)
- ❑ Évaluation des besoins
- ❑ 2 organismes non concernés par la formation ou l'assistance : BRGM et ONF (prise en charge par leur organisme de tutelle)
- ❑ Création d'un groupe de travail ?

■ Concepts de base du catalogage des données

■ Les données, c'est un peu comme des pots de confiture...

- Les confitures que l'on fait pour soi
 - « Pas besoin d'étiquette, je sais où sont les groseilles 2009 »
- Celles qu'on distribue aux amis
 - Il vaut mieux mettre une étiquette « groseille 2009 » sur le pot pour en faciliter l'usage
- Celles qu'on vend sur le marché ou dans les grandes surfaces
 - Il y a des mentions obligatoires : date de péremption, provenance, adresse du fabricant...

■ Qu'est-ce-que des métadonnées ?

- C'est l'étiquette sur le pot de confiture
- Le producteur des données connaît le processus de création et de mise à jour des données et n'a pas besoin de le documenter pour son propre usage
- Il doit documenter a minima ses données pour les mettre à disposition d'autres acteurs en interne dans son organisation
- Il doit les documenter en respectant la norme ISO 19115 pour les diffuser dans le cadre des obligations légales de la directive INSPIRE

■ Métadonnées

- ❑ Des données décrivant les données
- ❑ Les métadonnées décrivent :
 - Les objets géographiques
 - Leurs attributs (notamment leur géométrie)
 - Les relations entre les objets
- ❑ On peut trouver dans les métadonnées :
 - Des définitions
 - Des indications de qualité
 - Des indications sur la structuration informatique des données
 - Des informations d'ordre juridique sur le régime de propriété, le droit de diffusion
 - Les mentions des personnes à contacter
 - Et plus généralement toutes informations renseignant sur ces données

Exemple

■ Métadonnées géographiques et cartographiques

- ❑ Les **métadonnées géographiques** documentent des données géographiques (RGP, BD TOPO, Routes, aéroports de Bourgogne...)
 - Métadonnées abordées aujourd'hui et dans les formations ; des normes existent
- ❑ Les **métadonnées cartographiques** documentent des cartes (exemple : variation de population par canton en Région Bourgogne entre 1999 et 2006)
- ❑ Les **métadonnées de services** documentent les services associés aux données (consultation, géotraitement...)
 - Ces deux catégories ne seront pas abordées aujourd'hui ni dans les formations, pas de normes, descripteurs très similaires pour les métadonnées cartographiques

■ Catalogue de données, dictionnaire des données

□ Catalogue de données:

- Base de données constituée de métadonnées
- Ne contient en principe que l'état des données (par référence aux spécifications), et renvoie vers le dictionnaire de données pour les définitions.
- Dans la pratique, on met tout dans le catalogue de données

□ Dictionnaire des données

- Base de données ou document contenant la description de données en général standardisées (exemple : SANDRE pour les données sur l'eau)

□ Ontologie : concept très similaire du dictionnaire de données

■ Les métadonnées en 6 questions

- Pourquoi devons-nous créer des métadonnées ?
- Quoi ? Quelles métadonnées devons-nous constituer ?
- Comment faire ?
- Qui doit s'en occuper ?
- Quand le faire ?
- Où (dans quel outil) ?

■ *Pourquoi devons-nous créer des métadonnées ?*

- Pour nos besoins internes*
- Pour participer à GéoBourgogne*
- Pour la conformité INSPIRE*

■ Intérêt des métadonnées

- Un problème récurrent dans les organisations
 - Connaître et comprendre les données existantes
- « **Savoir** » : Mieux connaître le patrimoine de données
 - Les données que l'on récupère : données de référence, données d'études, données de partenaires, ...
 - Les données que l'on produit: garder une « traçabilité » des opérations effectuées, mieux gérer les versions, améliorer progressivement la qualité
- « **Faire savoir** » : Faciliter la diffusion des données
 - Faire connaître les données dont on dispose pour en informer les autres services de l'organisation et les autres organismes
 - Faire en sorte que les données que l'on diffuse soient mieux utilisées, car mieux connues

■ Répondre aux questions

des utilisateurs de données géographiques (1)

❑ Quoi ?

- Ce que contiennent les données (objets, attributs ; ex : cantons avec attributs donnant la variation de population entre 1999 et 2006)

❑ Où ?

- Le territoire et le découpage concernés (ex : Saône-et-Loire par canton)

❑ Qui ?

- Propriétaire des données (ex : INSEE) ?
- Autres organismes en rapport avec les données (ex : DDE71, producteur de la carte à partir des données INSEE)
- Personne à contacter pour obtenir ces données et rôle ? (ex : M X, administrateur des données localisées de la DDE71)

■ Répondre aux questions des utilisateurs de données géographiques (2)

Quand ?

- L'âge des données et la fréquence de leur mise à jour (ex : carte réalisée en 2008, données RGP 2006 et 1999)

Comment ?

- Processus de création des données

Pourquoi faire ?

- A quel type d'exploitation ces données sont-elles destinées ?

Quelles limites d'utilisation ?

- Restrictions liées à la qualité des données ?

■ Participer efficacement aux dispositifs de mutualisation

□ Objectifs de GéoBourgogne

- Acquisition de référentiels
- Mise en œuvre d'un portail pour la mutualisation d'informations géographiques
 - Pour les mutualiser, il faut les connaître !
 - 1ère étape pour chaque participant : cataloguer ses données pour les informer de ce dont il dispose
- Mise en œuvre d'actions de support géomatique
 - Formation, assistance...

■ Satisfaire aux obligations
légales (directive INSPIRE)

□ Voir intervention de ce matin

■ Quoi ? Quelles métadonnées devons-nous constituer ?

□ Objectif de documentation / diffusion

- Les données que nous souhaitons faire connaître et/ou diffuser à d'autres utilisateurs ou organismes
- Pas de contraintes particulières sur les métadonnées
- Métadonnées juridiques souhaitables pour les données diffusables

□ Objectif de mise en conformité avec INSPIRE

- Les données qui rentrent dans le champ de la directive INSPIRE
- Contraintes de format, champs obligatoires

■ Comment faire ?

Le processus de catalogage

1. Quelles données cataloguer ?
2. Où trouver les métadonnées ?
3. Comment recenser les métadonnées ?
4. Comment renseigner les métadonnées ?
5. La validation des métadonnées
6. La diffusion des métadonnées
7. La vie du catalogue

■ Etape n°1: Que cataloguer ?

- ❑ Arrêter un programme de travail et décider des types de données à cataloguer.
 - Données de référence / données métier ?
 - Données diffusables (objectifs externes) ou données partageables en interne (objectifs internes) ?
 - Données pérennes (données de gestion) / données acquises ponctuellement (données d'études) ?
 - ne pas alourdir le catalogue avec des données peu utilisables par d'autres que le service producteur ;
 - limiter la charge de travail découlant du catalogage.
- ❑ Première priorité :
 - Celles dont l'organisme est producteur et sur lesquelles il y a des contraintes liées à INSPIRE

■ Etape n°2: Où trouver les métadonnées ?

- ❑ Par votre connaissance (sujets pointus) et vos recherches.
- ❑ auprès d'interlocuteurs privilégiés (utilisateurs de SIG, équipes en charge de mesures, gestionnaires des données).
- ❑ auprès de partenaires à interviewer.
- ❑ Dans des documents qui auraient servi à constituer les données (modèle de données, cahier des charges).
 - Dans ce cas, vérifier impérativement que les données sont bien en adéquation avec les spécifications dudit document.

■ Etape n°3: Le recensement des métadonnées

- Les interlocuteurs privilégiés sont en priorité :
 - le créateur : celui qui a assuré la constitution du lot de données
 - le gestionnaire : celui qui tient les données à jour.
- Certaines questions trouveront aussi réponses :
 - auprès des utilisateurs ;
 - dans les choix politiques de l'organisme.

■ Etape n°4: Le renseignement des métadonnées

□ Deux questions ?

- Qui doit le faire ?
- Dans quel outil ?
- Voir plus loin

■ Etape n°5: La validation des métadonnées

- On distingue deux niveaux de validation
 - L'exactitude des renseignements fournis :
 - seul quelqu'un connaissant très bien la donnée en question peut l'effectuer.
 - L'actualité, la complétude ou la cohérence des métadonnées
 - Peut être fait par l'administrateur des données
- Il faut également rester pragmatique: ne pas chercher nécessairement à avoir un catalogue parfait dès le départ.

■ Etapes suivantes

- Etape n°6: Diffusion des métadonnées
 - Dans le gestionnaire du catalogue de l'organisme
 - Sur le futur portail de GéoBourgogne (moissonnage)
 - Sur d'autres portails (Géocatalogue par exemple)
- Etape n°7: La vie du catalogue
 - Il est nécessaire de faire vivre le catalogue
 - Mise à jour des fiches quand une donnée est modifiée
 - Renseigner les fiches correspondantes aux nouvelles données
 -

■ Qui doit le faire ?

- ❑ Dans certaines grandes organisations, il existe un administrateur des données disposant de compétences en catalogage
- ❑ Sinon, une personne disposant d'une compétence de base en catalogage (en pratique, c'est souvent l'administrateur SIG)
- ❑ L'administrateur doit travailler avec les personnes et services producteurs de données
- ❑ Il faut mettre en place une organisation minimale (un responsable)

■ Quand le faire ?

- ❑ Aujourd'hui, nécessité d'un rattrapage sur les données existantes à cataloguer
- ❑ Ensuite, créer si possible les métadonnées dès la publication des données

■ Où (dans quel outil) ?

□ Dans quels outils trouve-t-on des métadonnées ?

- Tableaux EXCEL, Documents WORD, bases ACCESS maison
- Gestionnaires de métadonnées associés aux outils SIG des éditeurs
 - exemple : ARC/CATALOG d'ESRI
- Gestionnaires de métadonnées indépendants des outils SIG
 - Exemple : outils Open Source GEOSOURCE, MDWEB...

■ Qualités attendues des outils de catalogage de données géographiques

□ Aspects opérationnels

- lien des métadonnées avec les bases de données
- Interopérabilité avec des bases de données diverses

□ Aspects fonctionnels

- Richesse fonctionnelle des outils du point de vue du catalogage (faciliter la saisie et les mises à jour) et de l'exploitation des métadonnées)
- Automatiser la récupération de métadonnées quand c'est possible
- Diffusion dans le respect des normes

□ Aspects documentaires

- Capacité de gérer et de diffuser des documentations accessibles aux utilisateurs et compréhensibles pour des non-informaticiens

■ Comparaison des outils

Critères	Logiciels Bureautiques	Gestionnaires éditeurs	Gestionnaires indépendants
Avantages	Pas de formation requise ni de déploiement, moyens du bord, tous les utilisateurs connaissent ces outils	Bien intégré au SIG, capacité de récupérer des métadonnées de façon automatique, liaison entre données et métadonnées	Possibilité de déployer les outils de saisie et consultation en environnement Internet pour faciliter l'accès aux métadonnées
Inconvénients	Exploitation et diffusion limitées	Pour exploiter les métadonnées en dehors des outils de l'éditeur, il faut les exporter	En général, pas de liaison avec le SIG, donc pas de récupération automatique de métadonnées
Conclusion	Bon moyen de démarrage, mais très vite limité	Bien adaptés aux usages internes à l'organisme	Bien adaptés aux dispositifs de mutualisation de données

■ Normes applicables

- ❑ Norme **ISO 19115**
- ❑ Norme internationale concernant les métadonnées géographiques publiée en mai 2003
- ❑ Fait partie d'un ensemble de normes concernant l'information géographique :
 - **ISO 19113** : principes qualité (décembre 2002)
 - **ISO 19114** : procédures pour l'évaluation de la qualité (septembre 2003)
- ❑ Norme **ISO 19139** : format des données de la norme ISO 19115
- ❑ Plusieurs outils informatiques disponibles notamment GEONETWORK, outil libre développé par la FAO et sa déclinaison française GEOSOURCE, MDWEB
- ❑ Norme adoptée par la directive INSPIRE
- ❑ Un **profil français** a été défini et validé en 2006
- ❑ INSPIRE comporte des exigences spécifiques ce qui a conduit à créer également un « **profil français INSPIRE** »

■ Présentation du plan de formation

■ Modalités pratiques

- Formation sur 2 jours consécutifs en plusieurs volets
 - Apports théoriques sur le catalogage, la norme ISO 19115
 - Formation à l'utilisation du logiciel GEOSOURCE
 - Travaux pratiques de renseignement de métadonnées dans l'outil GEOSOURCE
 - Les sessions sont réservées prioritairement aux adhérents de GéoBourgogne
 - La participation complète à une session est un préalable pour bénéficier ultérieurement d'une assistance sur site

	Lieux	1ère session	2ème session
Dijon	CRB		
Mâcon	DDE		
Auxerre	Préfecture		
Nevers	Préfecture		

■ Contenu des sessions : 1er jour

1^{er} jour MATIN

- Présentation de la session et tour de table des participants
- Concepts et mots clés du catalogage, notions de métadonnées
- Pourquoi faut-il créer des métadonnées ?
- Utilité du catalogage : à quoi servent les métadonnées ?
- Visualisation d'exemples de métadonnées
- Exercices pratiques de recherche de données dans des catalogues.

1^{er} jour APRES-MIDI

- Présentation synthétique des étapes du processus de catalogage.
- Les acteurs du catalogage
- Les outils à utiliser : présentation des logiciels de catalogage : choix du logiciel de catalogage
- Présentation de Géosource
- Normes applicables : ISO 19115 / 19139, profil français, profil français INSPIRE

■ Contenu des sessions : 2ème jour

2^{ème} jour MATIN

- Etapes du processus de catalogage
 - Que cataloguer ?
 - Où trouver les métadonnées ?
 - Comment recenser les métadonnées ?
 - Comment renseigner les métadonnées
 - Validation des métadonnées
 - Diffusion des métadonnées
 - Vie du catalogue.
- La présentation des étapes est réalisée en alternance avec un exercice pratique de catalogage

2^{ème} jour APRES-MIDI

- Suite de la présentation des étapes et de l'exercice pratique de catalogage
- Exercice de moissonnage ?
- Notions de qualité des données (présentation des critères de la norme ISO 19113)
- Que faire pour aller plus loin ?
- Bilan de la session.



I E T I
CONSULTANTS

Merci de votre attention
Des questions ?

IETI Consultants - 9 rue Lacreteille - F-71000 Mâcon

Tel : (0)3 85 21 91 91 - fax : (0)3 85 21 91 92

Email : henri.pornon@ieti.fr - Web : www.ieti.fr

■ Comparaison des outils

Critères	Logiciels Bureautiques	Gestionnaires éditeurs	Gestionnaires indépendants
Exemples	WORD, EXCEL, ACCESS	ARC/CATALOG	GEOSOURCE, MDWEB
Lien des métadonnées avec les bases de données	Aucun	Lien avec la base de données de l'éditeur	En général, pas de lien
Fonctions de création et d'actualisation	Absentes	Riches et permettant l'extraction automatique de métadonnées	Riches, mais en général sans extraction automatique
Exploitation des métadonnées par le SIG	Impossible	Facile	Dépend d'une connexion éventuelle du SIG
Exploitation des données par l'utilisateur : recherche, consultation	Limitée aux fonctions disponibles dans l'outil bureautique utilisé	Très performante, possibilité de passer des métadonnées aux données	Très performante pour la recherche de métadonnées (requêtes)
Diffusion des métadonnées (export, extraction...)	en format document, donc peu réutilisables	exploitation dans les outils de l'éditeur ou export	exploitation dans le gestionnaire ou export
Respect des normes de métadonnées (ISO 19115)	Non	à vérifier pour chaque éditeur, en général, oui	oui
Avantages	Pas de formation requise ni de déploiement, moyens du bord, tous les utilisateurs connaissent ces outils	Bien intégré au SIG, capacité de récupérer des métadonnées de façon automatique, liaison entre données et métadonnées	Possibilité de déployer les outils de saisie et consultation en environnement Internet pour faciliter l'accès aux métadonnées
Inconvénients	Exploitation et diffusion limitées	Pour exploiter les métadonnées en dehors des outils de l'éditeur, il faut les exporter	En général, pas de liaison avec le SIG, donc pas de récupération automatique de métadonnées
Conclusion	Bon moyen de démarrage, mais très vite limité	Bien adaptés aux usages internes à l'organisme	Bien adaptés aux dispositifs de mutualisation de données

■ Ressources WEB

□ Exemples de catalogues

- http://mdweb.loire-estuaire.org/mdweb_ext/
- <http://www.sigale.nordpasdecals.fr/CATALOGUE/information.asp>
- <http://adelie.application.equipement.gouv.fr/>
- <http://www.swisstopo.admin.ch/internet/swisstopo/fr/home/apps/geocat.html>

□ Autres documentations Internet

- http://www.ieti.fr/xoops/modules/xoopsfaq/index.php?cat_id=1

□ Bibliographie

- CERTU, mise en œuvre d'un catalogue de données géographiques (janvier 1999)
- CERTU, la qualité des données géographiques (février 2006)