

QUALITE DES DONNEES ET NORMALISATION

QUALITE

Définition

- Bcp d'utilisateurs n'imaginent pas quels autres usages pourraient être faits des données qu'ils génèrent.
- Les données sont constituées pour un usage précis, immédiat. Quid lors de la réutilisation pour d'autres usages par d'autres personnes ?
- Qualité : ensemble des propriétés et caractéristiques d'un produit ou d'un service qui lui confère l'aptitude à satisfaire des besoins exprimés ou implicites.
- Vérifier que le niveau de qualité proposé est en adéquation avec l'application envisagée et les besoins des utilisateurs.
- Résulte d'un arbitrage entre la réponse à un besoin, et un coût et délai de production.

Objectifs

- La qualité des données localisées doit permettre de renforcer les démarches d'analyse spatiale, d'aide à la décision, d'archivage et de favoriser les échanges entre partenaires et la consultation des données par le public.
- Questions : données fiables, précises, à jour, exhaustives, communicables sans risque, compatibles avec les autres données en main, aux normes en vigueur ?
- Importance des métadonnées : contribuent à l'expression même de la qualité, répondent à ces questions.

Enjeux

- Echanges : politique partenariale basée sur les échanges de données localisées, rendu fructueuse si la qualité est présente.
- Organisationnel : administration rigoureuse des données et développement de l'usage des métadonnées.
 - * Nécessité de connaître la provenance des informations à traiter.
 - * Possibilité d'exploitation des données en toute sécurité.
 - * Rôle des métadonnées : description de la qualité et aide au choix par l'utilisateur d'un lot de données correspondant à ses besoins.
- Economique : maîtriser les besoins et rentabiliser les données.
 - * Le coût élevé des données lors de l'acquisition conduit à simplifier la commande pour n'acquérir que le minimum nécessaire.
 - * Réutilisation de la donnée acquise pour satisfaire plusieurs besoins pour de multiples usages.
 - * Nécessité de mise en place de normes pour que chacun y trouve sa matière première (documenter les lots de données).
 - * Coût de la qualité.
- Professionnalisation (rigueur dans les métiers) : rigueur dans l'usage des données localisées.

Description et critères de qualité

- La description de la qualité permet de mesurer la qualité des données et vérifier que les données permettront d'en faire l'usage prévu.
- Norme française EDIGéO (Echanges de Données Informatisées dans le domaine de l'Information Géographique) : qualité des données échangées est le degré de fiabilité de ces données.
- 7 critères descriptifs :
 - * Généalogie (qualitatif) : références, histoire du lot de données.
 - * Actualité (qualitatif) : référence temporelle, données à jour.
 - * Précision géométrique (quantitatif) : écart moyen entre les mesures de contrôle et les valeurs de position planimétrique et altimétrique.
 - * Exhaustivité (quantitatif) : pourcentage d'éléments du terrain réel qui ont été pris en compte.
 - * Précision sémantique (quantitatif) : nombre d'objets correctement codifiés, identification des confusions entre objets attributs ou relations.
 - * Cohérence logique : fidélité structurelle des données par rapport à la réalité.
 - * Qualité spécifique : propre à l'utilisateur, critère personnalisable.
- Intérêt des critères : mieux comprendre comment évaluer la qualité (utilisateurs), meilleure approche des points à problèmes (producteurs).

Structuration de l'information et documentation des données

- Pour répondre aux besoins actuels et futurs d'utilisateurs différents : classer les données selon leurs caractéristiques qui doivent être structurées.
- Une structuration détaillée des informations peut toujours être simplifiée automatiquement, l'inverse requiert une sélection manuelle. Problématique du niveau de détail demandé dans un thème par différents utilisateurs.
- Structuration des données constituées ou commandées : nomenclature de la norme EDIGéO, garante de la pérennité et de la qualité des échanges.
- Connaissance des données : la qualité s'obtient par la description avec une précision suffisante des processus de production. C'est la documentation des données, conforme aux règles des métadonnées.

Contrôle qualité

- Etape complexe : une base de données localisée est une représentation numérique d'une abstraction de la surface terrestre.
- Sources d'erreur : collecte (matériel imprécis ou mal utilisé), saisie (mauvais encodage, géométrie incorrecte), traitement (modélisation abusive, interpolation fautive).
- Impératif d'avoir recours aux 7 critères qualité ensemble.
- Le diffuseur de données a une responsabilité juridique envers les futurs utilisateurs.
- Méthodologie CNIG : 8 points de contrôle.

NORMALISATION

Définition et contexte

- Information doit rester utilisable à chaque étape de sa propagation.
- Normalisation, enjeu essentiel dans la valorisation de l'information, car permet ces flux d'échanges, les réemplois et les enrichissements.
- Les formats informatiques ne permettent pas de transmettre les informations de façon satisfaisante :
 - * Dépendent de logiciels, et donc pas de garantie de rigueur ni de durée.
 - * Aucun logiciel ne gère complètement les 4 caractéristiques de l'information.
 - * Pas de vocabulaire commun.

Objectifs

- Permettre les échanges de données en altérant le moins possible leur signification et leurs caractéristiques.
- Quatre éléments composant une information géographique :
 - * Coordonnées : situer les informations dans un système de référence.
 - * Nomenclature : énumération des différents objets.
 - * Structuration des objets : description des liens entre eux.
 - * Caractéristiques de qualité des données.

Approches

- Stratégies nationales de normalisation liées à l'organisation des pouvoirs publics.
- Approche différente en fonction :
 - * De la vocation juridique ou fiscale du cadastre.
 - * Du degré de centralisation de la responsabilité de production de l'information.
 - * De l'échelle retenue pour la cartographie de base.

Normes

- EDIGéO : concerne l'échange d'informations géographiques numériques sur support informatique.
 - * Objectifs : optimiser l'efficacité et le volume des échanges de données, réduire au strict minimum le coût financier des échanges, être indépendant des SIG.
 - * Effet depuis 1999, intégrée par les producteurs de données nationales, dans des modules de conversion.
 - * Ensemble de spécifications permettant de conserver la richesse d'un lot au cours d'un échange : précision géométrique, précisions sémantique, exhaustivité.
 - * Dispositif de certification.
- Norme expérimentale sur les métadonnées : série de descripteurs.

- Normes ISO 9000 : Promouvoir la qualité dans la production des biens et services pour la satisfaction de l'utilisateur final, la clarification des responsabilités, pour une gestion explicite de la recherche de la qualité.
 - * Décrire l'organisation en terme de processus de production.
- Norme ISO / SPICE : décrit les composants majeurs du management des processus de gestion du système d'information (construction du système d'information).
 - * Offre un cadre conceptuel à l'organisation d'un système d'information : processus organisationnels, processus de base, processus de support.

A l'international

- La diversité européenne a conduit à des normes plus ouvertes, différentes d'EDIGÉO sur le langage de codage.
- Comité technique n°287 (Comité européen de normalisation) : définition de 8 normes.
 - * Modèle de référence : domaine de l'information géographique.
 - * Description des données – schéma spatial : géométrie des objets, relations.
 - * Localisation – position : système géodésique, projection.
 - * Localisation – identificateurs géographiques : méthodes de documentation des identifiants géographiques.
 - * Qualité : modèle calqué sur les critères EDIGÉO.
 - * Métadonnées : liste ordonnée.
 - * Transfert : schéma de transfert des données et métadonnées.
 - * Requête et mise à jour : interface spécialisée.
- Préparation aux normes européennes : mise en place de formation, développer des profils spécifiques (déclinaison de l'emploi des normes de façon générique), promouvoir la norme métadonnée.
- Au niveau mondial :
 - * Comité technique ISO211 : norme liées aux SIG et aux services.
 - * OGC : normes sur l'interopérabilité. Si l'objectif est atteint, les normes OGC deviendront le standard du marché.