

Commission Règles et Qualité du CNIG

Point d'avancement du GT QuaDoGéo

Quels sont les éléments les plus importants pour déterminer la qualité d'un jeu de données ?

Déplacer les réponses selon la priorité la plus haute (1 étant l'élément le plus important)

1	Le respect à un schéma de donnée
2	La qualité de la description
3	La description des champs (colonnes)
4	La disponibilité de la ressource au cours du temps
5	La présence de tag
6	Le respect des dates de mise à jour annoncées
7	La présence de données pivot (identifiants de jeux de données référentiels, ex : n° SIREN, Code Commune, pays etc.)
8	La confiance dans le producteur des données
9	L'utilisation d'un format ouvert
10	la spécification de la licence
11	La popularité du jeu de données

Source : data.gouv.fr

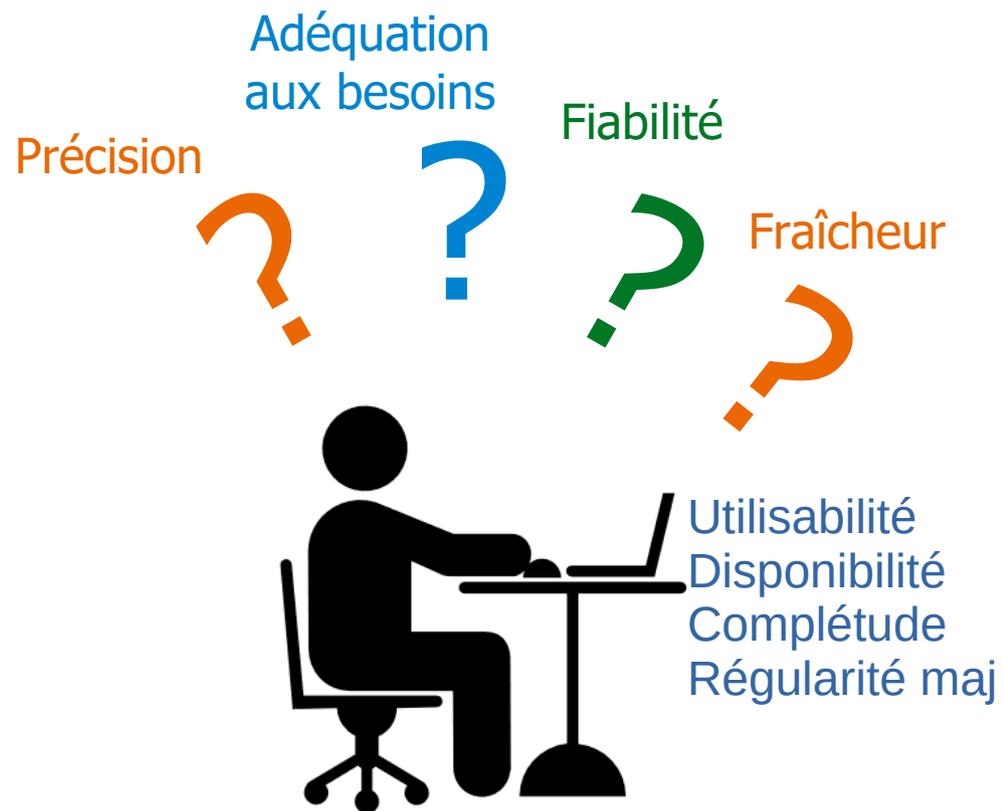
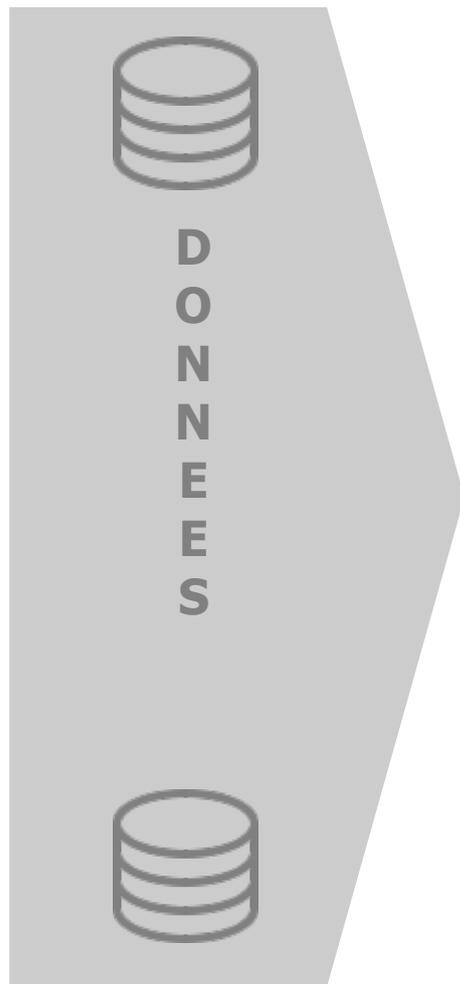


Mandat du GT CNIG QuaDoGéo

- ♦ Mission : « *aider à la réutilisation des données ouvertes grâce à une meilleure information sur leur qualité* »
- ♦ Axes de travail :
 - faire valoir les retours d'expérience
 - accompagnement des diffuseurs et des utilisateurs de données
 - prise en compte du retour utilisateur (crowdsourcing, évaluation...)
- ♦ Animation et secrétariat technique : Cerema (A. Gallais), avec le soutien du CGDD/SRI et de la Mission CNIG
- ♦ Composition du GT : collectivités, prestataires de services, consultants, IGN, Etalab, OpenData France, IFREMER, OFB, OIEau, Cerema, etc.
- ♦ 11 réunions réalisées depuis mars 2019 (bientôt 12...)

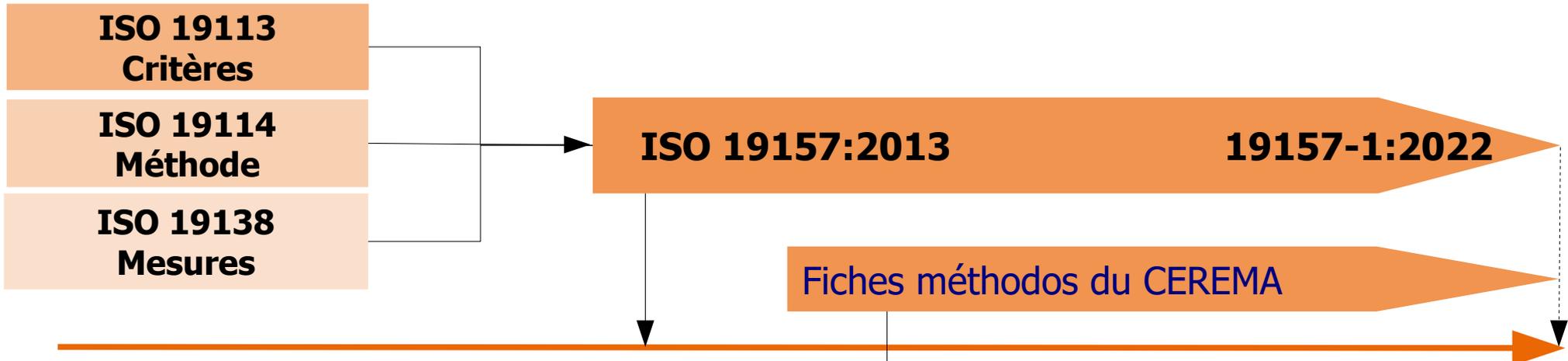
Rappel

Enjeux de la qualification des données



Rappel

Bref rappel historique



2006

2010

2013

2017

2019

2022

Expertise de G. Troipoux

GT QuaDoGéo

Bienvenue dans ce guide interactif sur la "qualité des données géographiques".



Commencer la lecture du guide par le commencement ! (Introduction, Contexte, Objet...)



Rappel

Fiches méthodologiques

Critères

19157

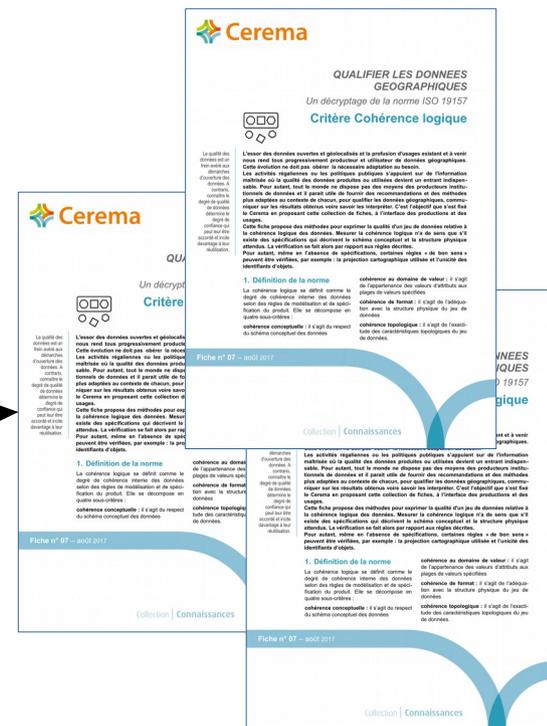
Exhaustivité
Précision thématique
Cohérence logique
Précision de position
Cohérence temporelle



Méthodes

19157

Echantillonnage
Généralités sur la qualif.
Contexte de contrôle
Éléments statistiques
Méthodes de représentation



Bilan du groupe de travail

2020 - 2021

Fiches méthododos et registre des mesures

- ◆ Validation du registre novembre 2020, après appel à commentaires
- ◆ Révision des fiches « Précision thématique » et « Précision de position » en 2021
- ◆ Actualisation du registre des mesures en conséquence (janvier 2022)

28 mesures

24 mesures ISO19157 retenues dans les fiches méthodologiques du Cerema

1 mesure hors ISO19157 mais préconisée dans les fiches Cerema :
- classe de précision arrêté 2003

3 éléments de qualité de ISO19115 :
- généalogie
- résolution spatiale
- cohérence topologique

Métadonnées relatives à la qualité des données géographiques

REGISTRE FRANÇAIS

Système de publication de registres

Registre: mesuresQuaDoGeo
URI: https://data.geocatalogue.fr/nc/_mesuresQuaDoGeo
Mesures liées à la Qualité de Données Géographiques

Registre CNIG janvier 2022

Système de publication de registres

URI: https://data.geocatalogue.fr/nc/_mesuresQuaDoGeo

Registre: mesuresQuaDoGeo
URI: https://data.geocatalogue.fr/nc/_mesuresQuaDoGeo
Mesures liées à la Qualité de Données Géographiques

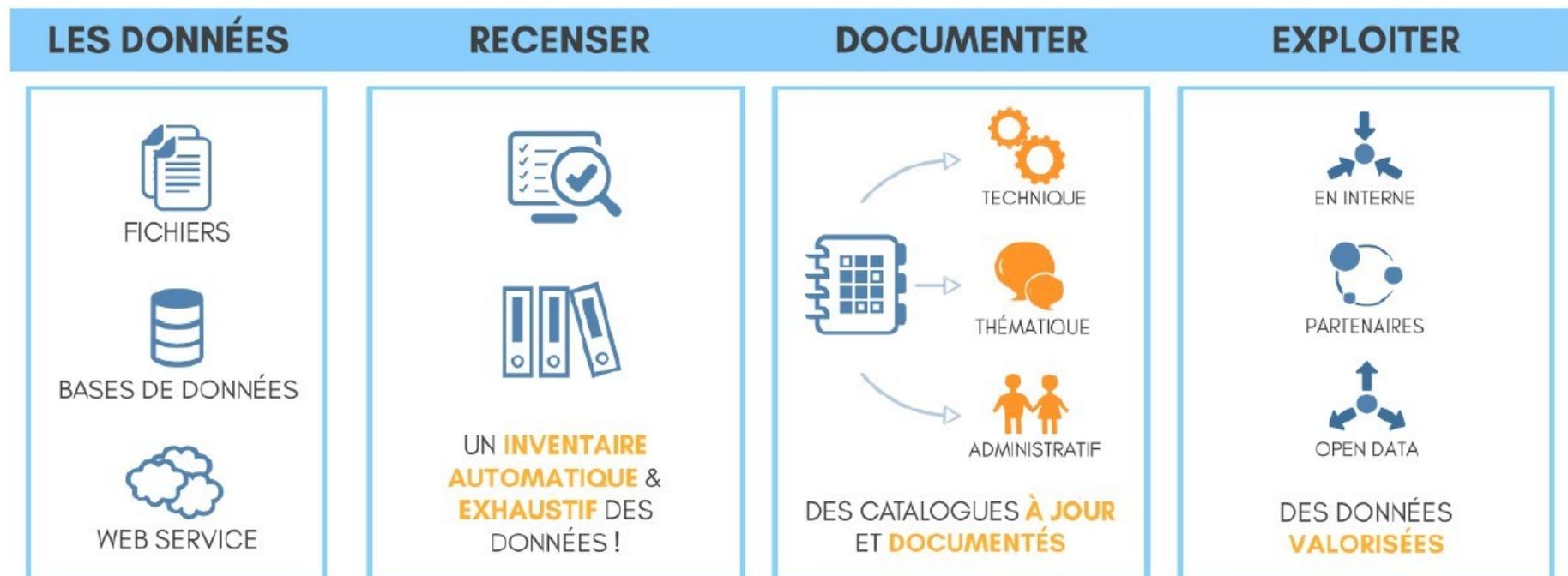
Contenu

Voir 20 lignes

Nom	Notation	Description	Types	Statut
Classe de précision au sens de l'arrêté du 10 septembre 2003	10	L'arrêté du 10 septembre 2003 porte sur les classes de précis...	dQM mesure , Concept	experimental
Coefficient Kappa	10	Matrice indiquant le nombre d'éléments de la classe (i) class...	dQM mesure , Concept	experimental
Conflit de structure physique	4	Indication que le stockage des éléments entre en conflit avec...	Concept , dQM mesure	experimental
Date de création	24	Date de création de la ressource.	dQM mesure , Concept	experimental
Date de publication	26	Date de publication de la ressource.	dQM mesure , Concept	experimental
Date de révision	25	Date de dernière révision de la ressource.	dQM mesure , Concept	experimental
Erreur circulaire avec un seuil de signification de 95 %	12	Rayon décrivant un cercle, dans lequel l'emplacement véritaibl...	dQM mesure , Concept	experimental
Erreur horizontale relative	14	Evaluation des erreurs aléatoires d'une entité dans la positi...	dQM mesure , Concept	experimental
Erreur verticale relative	13	Evaluation des erreurs aléatoires d'une entité du relief à un...	dQM mesure , Concept	experimental
Etendue temporelle	23	L'étendue temporelle définit la période de temps couverte par...	dQM mesure , Concept	experimental
Exactitude de la mesure temporelle avec un seuil de signification de 95 %	19	Demi-longueur de l'intervalle défini par une limite supérieur...	Concept , dQM mesure	experimental
Généalogie	21	La généalogie fait état de l'historique du traitement et/ou d...	dQM mesure , Concept	experimental
Incertitude de la valeur d'attribut avec un seuil de signification de 95 %	18	Demi-longueur de l'intervalle défini par une limite supérieur...	Concept , dQM mesure	experimental
Matrice relative de classement erroné (RMCM)	15	Matrice indiquant le nombre d'éléments de la classe (i) class...	Concept , dQM mesure	experimental
Nombre de micro-surfaces non valides	7	Comptage de tous les éléments du jeu de données qui constitue...	Concept , dQM mesure	experimental
Nombre d'erreurs de chevauchement	9	Comptage de tous les éléments du jeu de données qui se chevau...	Concept , dQM mesure	experimental

Retex nationaux et veille internationale

- ◆ Veille sur le sujet qualité et retour sur le Q-KEN (N. Py)
- ◆ Exploitation des métadonnées de qualité (M. Becker - Isogéo)
- ◆ Inventaire des métadonnées de qualité sur les portails open data (N. Py)
- ◆ Qualité des données dans data.gouv.fr (G. Aldebert et T. Gratier)
- ◆ Veille sur le monde de la normalisation (M. Lambois)



Source : Isogéo

Formation QuaDoGéo

- ◆ Webinaires QuadoGéo, 34 stagiaires
- ◆ 5 modules de 2h30, sur des dates resserrées, en décembre 2021
- ◆ Gratuits organisés par le CRIGE-PACA et le CEREMA
- ◆ Basés sur les fiches méthodes et la reprise de la FOAD du CEREMA
- ◆ Concepts théoriques et illustration pratique immédiate dans QGIS
- ◆ Franc succès ! => renouvellement en mai 2022, dans un format consolidé
- ◆ Une [page web](#) dédiée :



Grâce à l'opendata, les données géographiques foisonnent sur de nombreux portails. Ce foisonnement ouvre de nouvelles et nombreuses perspectives de valorisation. Il pose également la question de la confiance en la qualité des données produites et partagées au regard d'une utilisation ciblée. La qualité devient de plus en plus un sujet de premier plan pour les producteurs et les utilisateurs en quête de données fiables permettant des usages optimisés. Le CRIGE conduit depuis plusieurs années des réflexions et travaux sur la qualité des données géographiques en collaboration étroite avec le CEREMA et dans le cadre du groupe de travail ad hoc du CNIG. Il développe également des outils de qualification pour accompagner ses membres dans une démarche qualité.

Contact



Stéphane ROLLE
Chef de projet SIG
CRIGE Provence-Alpes-
Côte d'Azur
[✉ Contacter](#)

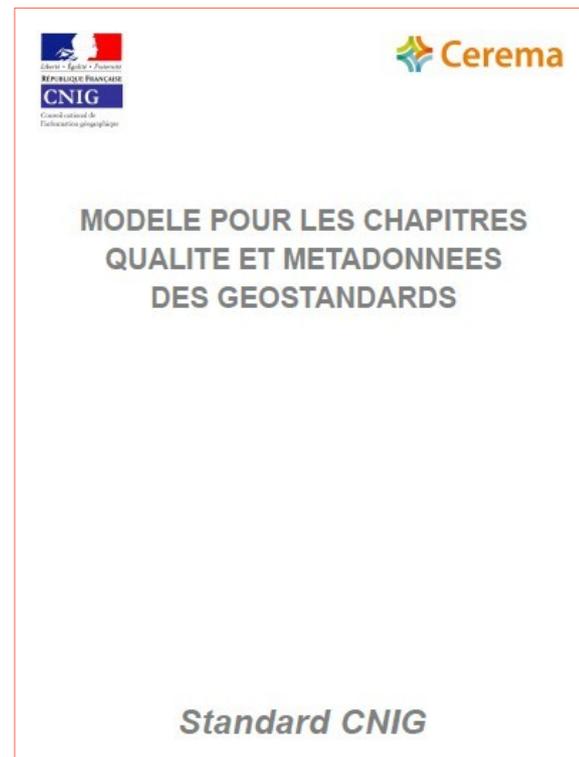
Réflexions sur les usages de l'info géo

- ◆ Deux entrées : le thème et l'usage
- ◆ Difficulté de recenser les usages de l'info géo
 - nomenclature exhaustive versus nomenclature généraliste plus resserrée
- ◆ => Deuxième option privilégiée
 - Une dizaine d'entrées « usages génériques »
 - un deuxième niveau plus détaillé, à construire
- ◆ La nomenclature alimentera l'outil « restitution de la qualité des données »

1. recenser
2. localiser
3. analyser
4. cartographier
5. adresser
6. naviguer, acheminer
7. planifier
8. gérer
9. suivre, observer, évaluer
10. servir de donnée référentielle ou et/ou de donnée pivot

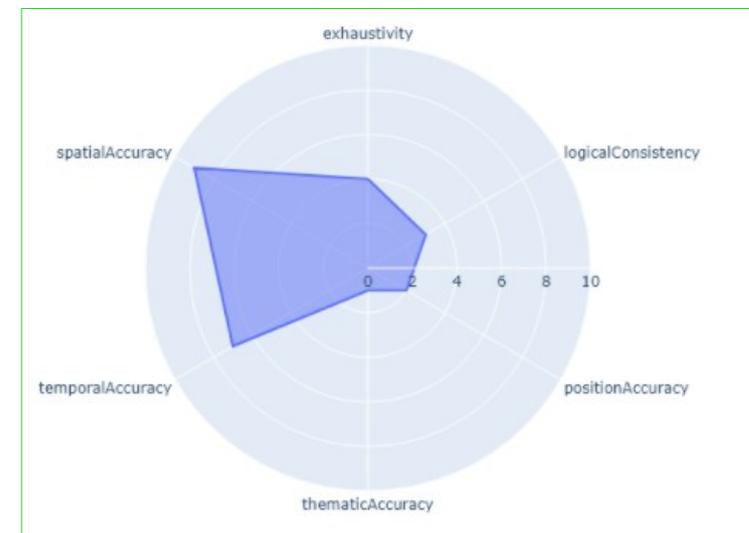
Modèle des parties qualité et métadonnées des géostandards

- ◆ Trame de rédaction, a été présentée à la Commission des Standards
- ◆ Utilisée par les standards Eclairage et Star-Elec



Restitution sur la qualité des données

- ◆ Développement d'une preuve de concept
- ◆ Basée sur la maquette graphique de restitution synthétique de la qualité
 - notes du testeur (vert), des utilisateurs (jaune), et rubrique « déposer un avis »
 - sous-notes par critères ;
 - rubrique « usages recommandés / usages à éviter »
- ◆ Liens à conforter avec
 - l'import de données de qualification ISO 19157 au format xml
 - la nomenclature des usages



Conclusions et pistes de travail

- ◆ Lien étroit avec le GT Métadonnées (participation croisée)
 - [Sujet de la description du MCD dans les métadonnées](#)
- ◆ Maintien des retours d'expériences au sein du GT QuaDoGéo
- ◆ Succès de la formation QuaDoGéo donnée au CRIGE PACA
 - [Renouvellement en mai 2022](#)
- ◆ Maintenance du registre des mesures de la qualité
- ◆ Elaboration du guide des spécifs qualité et métadonnées
- ◆ Peu d'avancée sur les outils de retour des utilisateurs
 - [RETEX en cours sur le GUF dans Géonetwork 4](#)
- ◆ industrialiser la POC de restitution graphique (*hors GT CNIG*)

MERCI DE VOTRE ATTENTION

arnauld.gallais@cerema.fr