

Métadonnées sur la qualité d'un lot de données

```
<mdq:specification>
  <cit:CI_Citation>
    <cit:title>
      <gco:CharacterString>Specification CNIG QUADOGEO
    </gco:CharacterString>
    </cit:title>
  </cit:CI_Citation>
</mdq:specification>
```

Objectifs

- guider et faciliter la saisie ainsi que la transmission d'informations sur la qualité de la donnée au sein des métadonnées.
- améliorer la saisie des critères qualité dans les métadonnées ISO / Inspire en préconisant de bonnes pratiques.
- Questions sous-jacentes :
 - Comment saisir des critères qualité dans les métadonnées ISO ?
=> consignes de saisie
 - Comment décrire la qualité dans les métadonnées, de façon à être facilement transmissibles et compréhensibles par l'utilisateur ?
 - Quelle part d'automatisation de la production de ces MD qualité ?
 - Comment restituer (lisiblement !) les MD qualité à l'utilisateur ?

Contexte

- GT CNIG QuaDoGéo (démarrage 2019)
 - => Réunion avec Cellule normalisation IGN, le 10 avril 2019
- Révision du guide CNIG de MD Inspire, v2,0
 - Appel à commentaires en cours, jusqu'au 21 juin
- Seulement trois éléments de description de la qualité prévus par Inspire :
 - **Généalogie**
 - **Résolution spatiale**
 - **Cohérence topologique**
- Pas l'exhaustivité, ni les autres critères..
- Mais ISO 19115 « permet la définition de plusieurs ensembles d'info. de qualité (...) pouvant fournir des informations de généalogie et des résultats d'évaluation de la qualité » ([page 99](#) du guide MD du CNIG)



Solution(s) ?

- (pour les humains) adjoindre aux MD un document spécifique à la qualification des données et/ou au résultat d'un contrôle qualité.
- (pour les machines) identifier les mécanismes ISO19115.1 permettant la transmission des éléments de qualification des données...
- ... en s'appuyant sur des registres :
 - [ISO 19135-1:2015](#) « spécifie les procédures à suivre pour établir, mettre à jour et publier les registres d'identifiants uniques, non ambigus et permanents, ainsi que les significations qui sont attribuées aux éléments de l'information géographique.
Afin d'atteindre cet objectif, cette norme spécifie les éléments qui sont nécessaires à leur enregistrement. »
=> référencer dans un registre les mesures retenues dans les [fiches méthodologiques](#) du Cerema.
=> Exemple : [Re3gistry](#), outil open source pourrait être utilisé.

Exemple d'implémentation xml (1/2)

```
1 <?xml version="1.0" encoding="UTF-8"?>
2 <!-- exemple de rapport qualité
3 2.2-Précision absolue (ou externe): Valeur moyenne des incertitudes
4
5 <mdq:report>
6   <mdq:DQ AbsoluteExternalPositionalAccuracy>
7     <mdq:measure>
8       <mdq:DQ_MeasureReference>
9         <mdq:measureIdentification>
10          <mcc:MD_Identifier> ← définir un namespace, ex : qfr
11            <mcc:code>
12              <gco:CharacterString>
13                <!-- ici, mettre l'url de la mesure dans le registre du CNIG -->
14                http://registry.cnig.fr/qualityMeasures/Valeur-moyenne-des-incertitudes
15              </gco:CharacterString>
16            </mcc:code>
17          </mcc:MD_Identifier>
18        </mdq:measureIdentification>
19        <mdq:measureDescription>
20          <gco:CharacterString>Valeur moyenne des incertitudes de position pour un ensemble de positions
21          </gco:CharacterString>
22        </mdq:measureDescription>
23      </mdq:DQ_MeasureReference>
24    </mdq:measure>
```

1) Identification et description de la mesure.

définir un namespace, ex : qfr

ici, mettre l'url de la mesure dans le registre du CNIG -->

Valeur moyenne des incertitudes de position pour un ensemble de positions

Exemple d'implémentation xml (2/2)

```
25 <mdq:result>
26   <mdq:DQ_ConformanceResult>
27     <mdq:specification>
28       <cit:CI_Citation>
29         <cit:title>
30           <gco:CharacterString>Specification CNIG QUADOGEO
31         </gco:CharacterString>
32       </cit:title>
33       <cit:onlineResource>
34         <cit:CI_OnlineResource>
35           <cit:linkage>
36             <gco:CharacterString>
37               http://specification.cnig.fr/url-de-la-spec-quadogeo</gco:CharacterString>
38           </cit:linkage>
39         </cit:CI_OnlineResource>
40       </cit:onlineResource>
41     </cit:CI_Citation>
42   </mdq:specification>
43   <mdq:pass>
44     <gco:Boolean>true</gco:Boolean>
45   </mdq:pass>
46 </mdq:DQ_ConformanceResult>
47 </mdq:result>
48 </mdq:DQ_AbsoluteExternalPositionalAccuracy>
49 </mdq:report>
```

2) Résultat de la mesure de qualité

Mesures à indiquer dans le registre

- Les mesures préconisées par le guide méthodologique :

CRITERE	SOUS-CRITERE	MESURE
Exhaustivité	<i>néant</i>	Taux d'exhaustivité
Cohérence logique	cohérence au domaine de valeurs	Taux de conformité aux domaines de valeurs
	<i>cohérence conceptuelle</i>	<i>néant</i>
	cohérence de format	Conflit de structure physique
	cohérence topologique	Taux de connexions erronées
	cohérence topologique	Nb de micro-surfaces non valides
	cohérence topologique	Nb d'auto-intersections non valides
Précision thématique	Justesse de classement	Matrice relative de classement erroné
	<i>Justesse de classement</i>	<i>Coefficient Kappa</i>
	Justesse des attributs non quantitatifs	Taux de valeurs d'attribut correctes
	Exactitude des attributs quantitatifs	Incertitude de valeur avec seuil de signification de 95 %
Précision de position	précision absolue	Valeur moyenne des incertitudes de position
	précision absolue	Classe de précision (arrêté 16/09/2003)
Qualité temporelle	exactitude de la mesure temporelle	Exactitude avec seuil de signification de 95 %
	cohérence temporelle	Taux de conformité chronologique
	validité temporelle	Taux de conformité aux domaines de valeurs

Piste complémentaire...

- Le [Geospatial User Feedback](#) (GUF)
 - standardisé par l'OGC
 - implémenté dans Geonetworks... mais pas dans le [Géocatalogue](#)
 - permet à l'utilisateur de noter et émettre des avis sur les données
 - « Cette norme définit un modèle de données conceptuel de commentaires de l'utilisateur géospatial et un codage XML pratique du modèle conceptuel.
Les commentaires des utilisateurs sont des métadonnées principalement produites par les consommateurs de données géospatiales, qui les utilisent et en acquièrent de l'expérience. La norme permet de documenter les éléments de retour, tels que (...) les rapports sur la qualité.
La norme complète les conventions de métadonnées existantes selon lesquelles les documents enregistrant les caractéristiques d'un ensemble de données et les flux de production sont générés par le producteur de données.
Dans le cadre des métadonnées, le modèle de données GUF reprend certains éléments de l'[ISO 19115-1: 2014](#), mais pas la structure générale. (...) »



Cerema

Centre d'études et d'expertise sur les risques,
l'environnement, la mobilité et l'aménagement

Direction territoriale Ouest

Merci de votre attention

Arnauld GALLAIS

+33 (0)2 40 12 84 76

arnauld.gallais@cerema.fr