



Enjeux  
de la qualification

Les fiches  
Méthodologiques

Et la suite ?

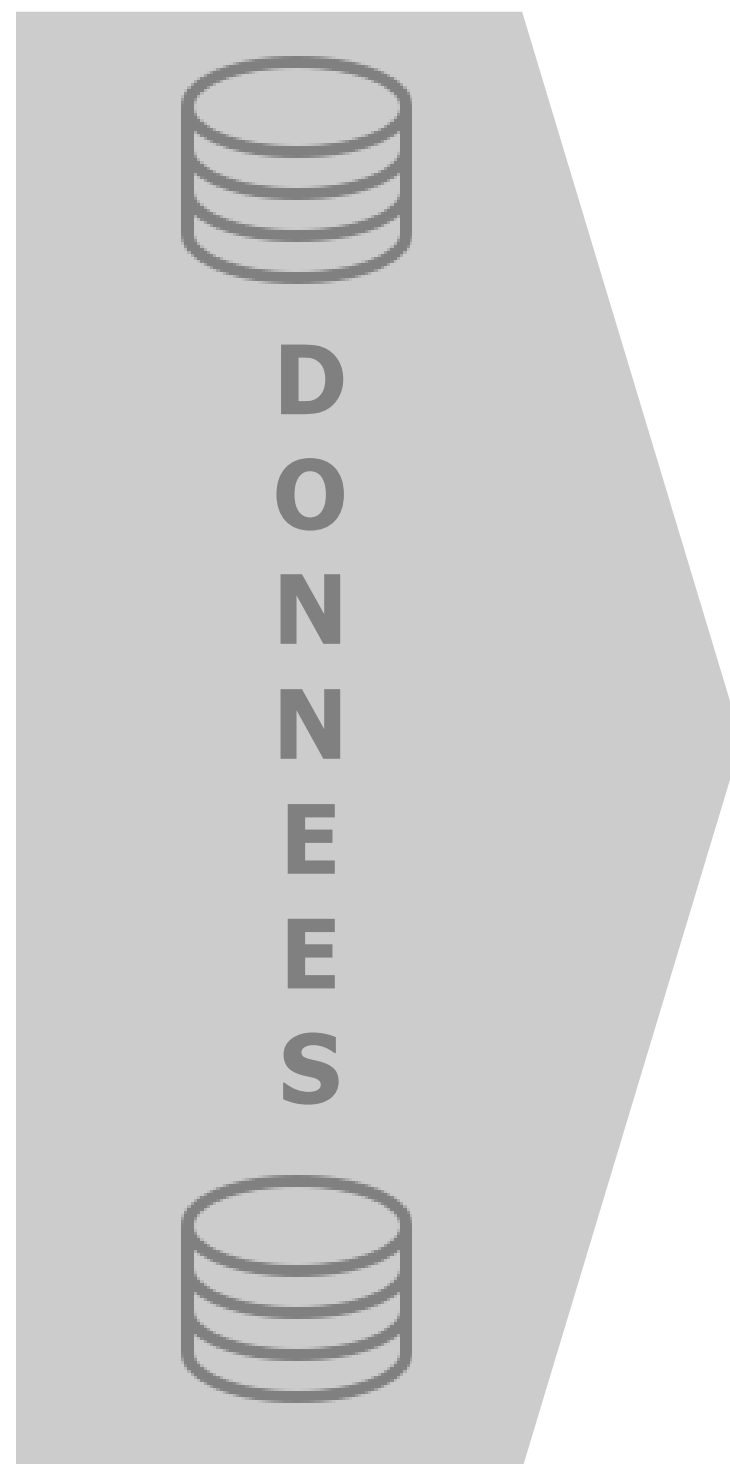
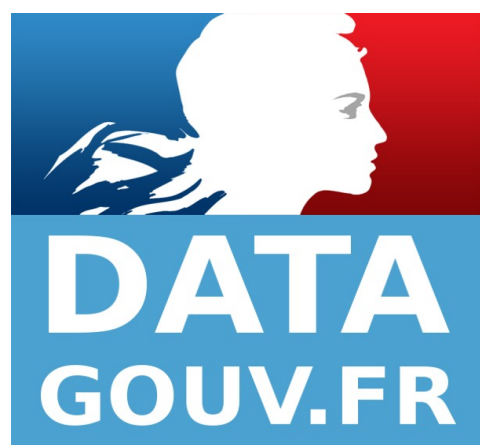
Le projet  **QuaDoGéo**  
Qualifier les données géographiques

**Stéphane Lévêque**  
Cerema Territoires et ville



#quadogeo  
@quadogeo

# Pourquoi qualifier



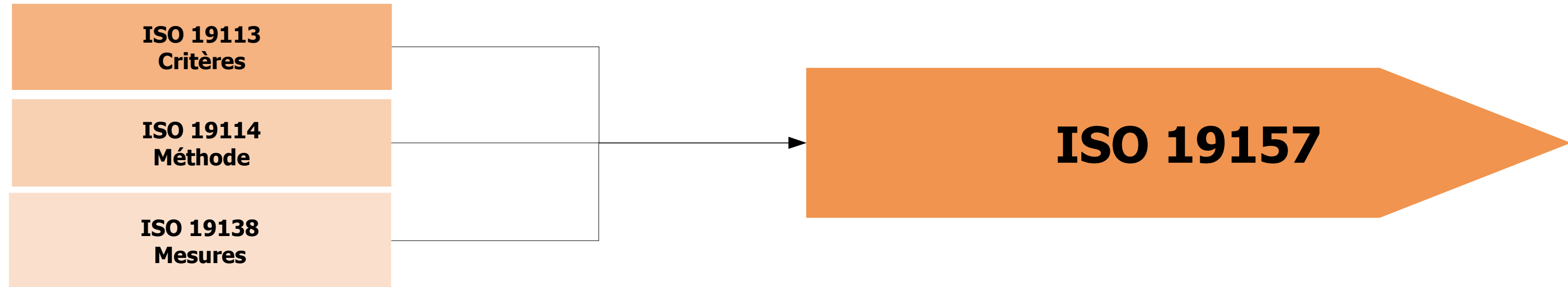
Correspondance  
aux besoins

Précision

Fiabilité



# Les travaux du Cerema



2013



2006

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



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

2010

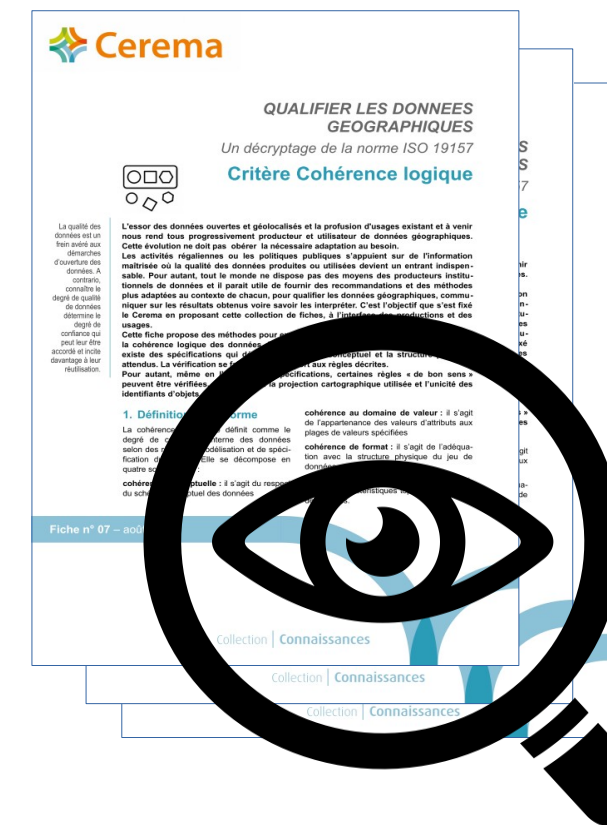


2015

La description de différents critères et sous-critères nécessaires à la description de la qualité des données géographiques

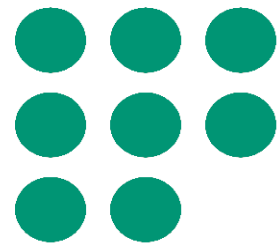


2017



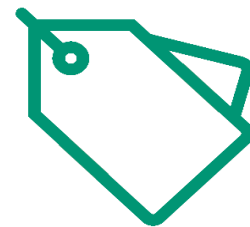
# 5 fiches critères

Exhaustivité



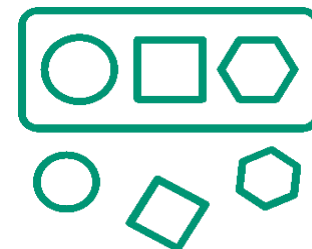
**Géométrie**  
**Attributs**

Précision  
thématique



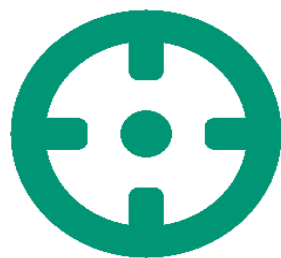
**Attributs**

Cohérence  
logique



**Géométrie**  
**Attributs**

Précision de  
position




**Géométrie**

Cohérence  
temporelle



**Attributs**

 **Cerema**

**QUALIFIER LES DONNEES GEOGRAPHIQUES**  
Un décryptage de la norme ISO 19157

**Critère d'exhaustivité**

La qualité des données est un frein avéré aux démarches d'ouverture des données. A contrario, connaître le degré de qualité de données détermine le degré de confiance qui peut leur être accordé et incite davantage à leur réutilisation.

L'essor des données ouvertes et géolocalisées et la profusion d'usages existant et à venir nous rend tous progressivement producteur et utilisateur de données géographiques. Cette évolution ne doit pas obérer la nécessaire adaptation au besoin. Les activités régaliennes ou les politiques publiques s'appuient sur de l'information maîtrisée où la qualité des données produites ou utilisées devient un entrant indispensable. Pour autant, tout le monde ne dispose pas des moyens des producteurs institutionnels de données et il paraît utile de fournir des recommandations et des méthodes plus adaptées au contexte de chacun, pour qualifier les données géographiques, communiquer sur les résultats obtenus voire savoir les interpréter. C'est l'objectif que s'est fixé le Cerema en proposant cette collection de fiches, à l'interface des productions et des usages. Cette fiche propose des éléments de méthode pour utiliser le critère d'exhaustivité ainsi que des indicateurs pour rendre compte de la mesure de cet élément de qualité.

**1. Les définitions utilisées**  
L'exhaustivité est un critère de la norme ISO 19157 qui se définit comme la présence et l'absence d'entités, de leurs attributs et de leurs relations. Il se décompose en deux sous-critères :

- l'**excédent** : données excédentaires présentes dans le jeu de données ;
- l'**omission** : données absentes d'un jeu de données.

*Remarque : ce critère ou ses deux sous-critères peuvent se mesurer pour chaque classe d'objets, attribut ou relation que l'on désire évaluer.*

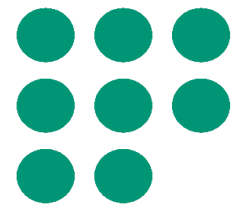
**2. Description des mesures possibles à réaliser**  
Les deux sous-critères sont abordés dans ce qui suit plus en détail, notamment par la description de mesures qui leur sont rattachées dans la norme. Les différentes définitions de ces mesures sont rappelées dans le **tableau 1** et les types de valeurs qu'elles doivent fournir sont indiquées pour chacune.

Fiche n° 09 – août 2017

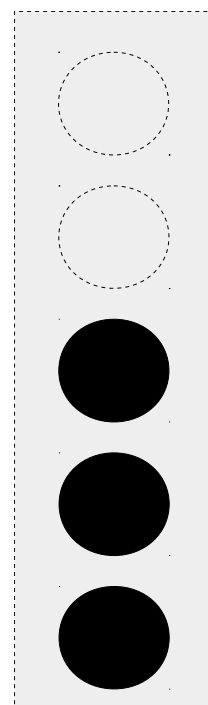
Collection | **Connaissances**

1. Les définitions utilisées
2. Description des mesures possibles à réaliser
3. Indicateurs retenus
4. Méthodes de contrôle
5. Représentation-Notation
6. Ce qu'il faut retenir

# Exemple : taux d'exhaustivité



Ce critère rassemble tout ce qui caractérise l'absence de données attendues ou leur présence à tort.



Éléments manquants  
NbO

Éléments en  
excès  
NbE

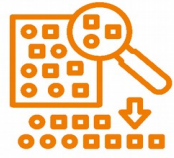
$$\text{taux d'exhaustivité} = 1 - \frac{(Nb_O + Nb_E)}{N_t}$$

Dans notre cas, le taux d'exhaustivité est de  $1 - ((2+2)/5) = 20\%$

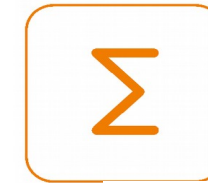
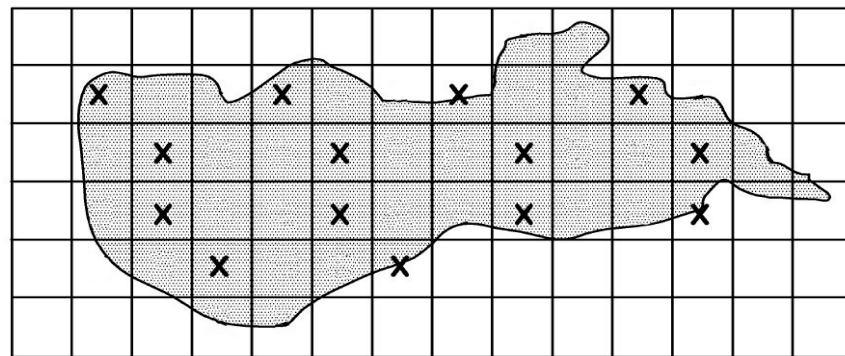
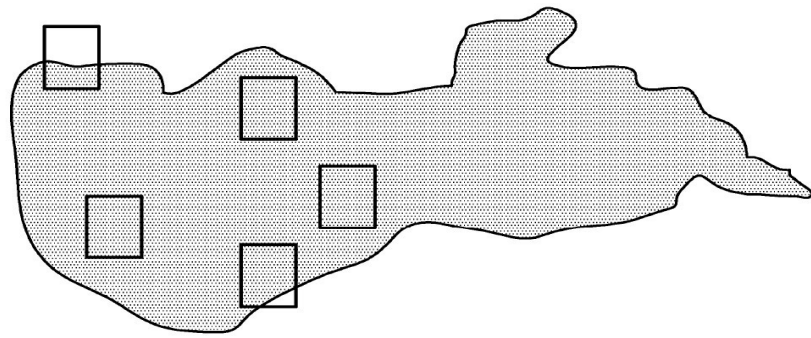
Taux d'exhaustivité	Note sur 5
De 95 % à 100 %	5
De 90 % à 95 %	4
De 75 % à 90 %	3
De 50 % à 75 %	2
Taux < 50 %	1

Dans notre cas, la note sera de 1

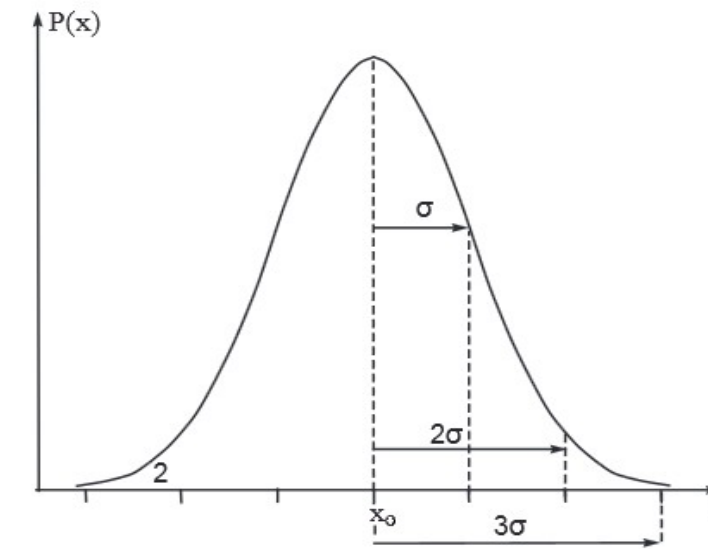
# Complément : les fiches méthodes



## Échantillonnage

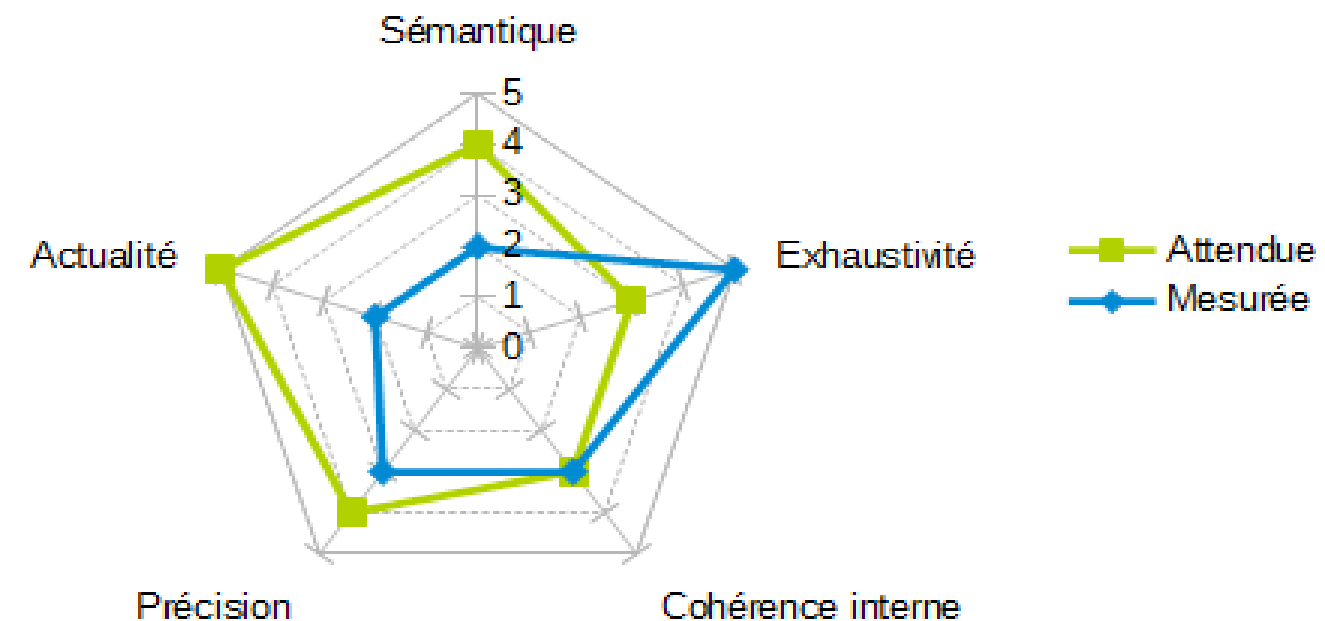


## Éléments statistiques



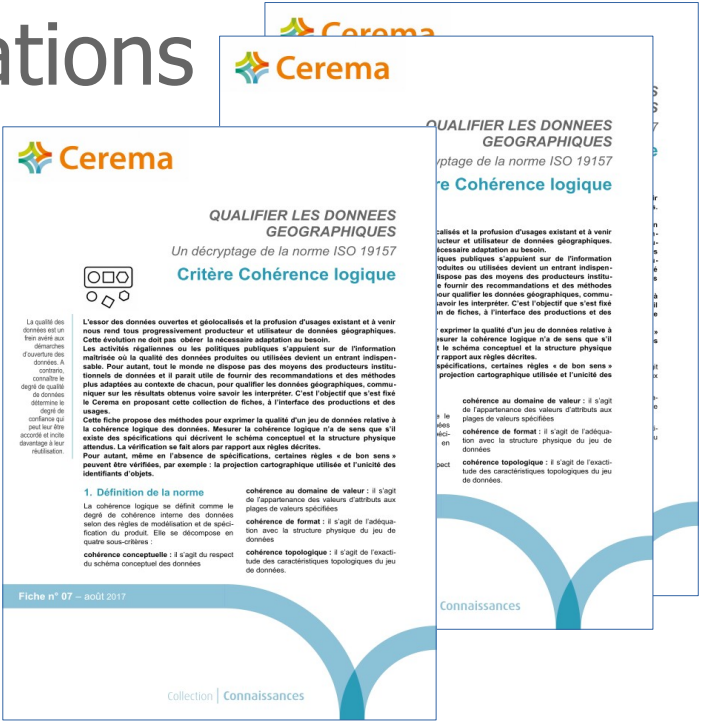
## Méthodes de représentation

Représentation simplifiée des notes attendues et mesurées d'un jeu de données

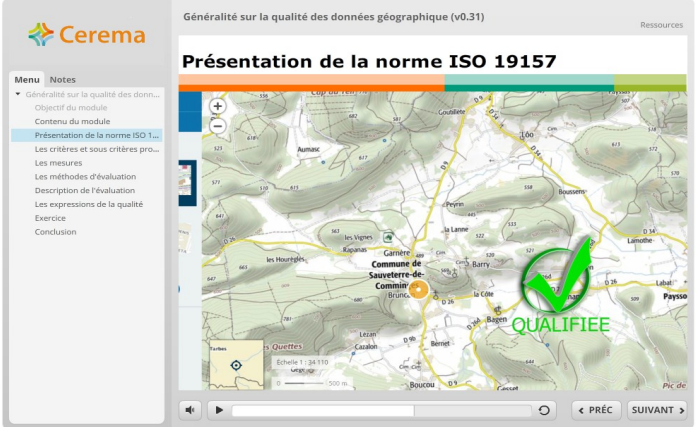


# Et maintenant ? Au Cerema...

Publications



Formation



Critères



Méthodes



Partenariats

Expérimentation

**GRAND LYON**  
la métropole

Centre Régional  
de l'Information  
Géographique  
**CRIGE**  
Provence-Alpes  
Côte d'Azur



# Et maintenant ? Les autres acteurs...

Recherche

*GdR MAGIS - CNRS*

IGN



Outillage



Animation



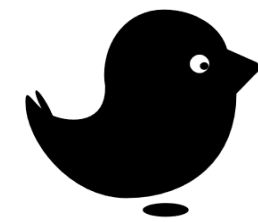
Accompagnement





En savoir plus sur  QuaDoGéo

[www.cerema.fr/fr/mots-cles/quadogeo](http://www.cerema.fr/fr/mots-cles/quadogeo)



#quadogeo @quadogeo



Crédits (images Creative Commons issues de nounproject.org)

Desktop *par marc david rapoza*

Data *par Aybige*

Fichier *par I cons*

Interface *par Aftergrind*

Outil *par Hea Poh Lin*

Utilisateur *par Edward Boatman*

Valise *par Shashank Singh*

Loupe *par Alena*

Click *par Adrien Coquet*

Valise *par Shashank Singh*

Bird *par Ker'is*

### **Visuels des critères**

Cohérence logique *par Made by Made*

Exhaustivité *par Made by Made*

Précision de position *par Gregor Cesnar*

Précision thématique *par Pictohaven*

Qualité temporelle *par Icon 54*

### **Visuels des méthodes**

Contexte de contrôle *par Yu Luck*

Échantillonnage *par Becris*

Globe *par Proletkult Graphik*

Méthode de représentation *par NOPIXEL*

Statistiques *par Burak Kucukparmaksiz*