

Equipements Collectifs

Comment les qualifier et quel rendu appliquer ?

Stéphane Rolle – CRIGE PACA
Mathieu Rajerison – CEREMA Med



Qualification

Noter l'importance des critères

Modèle **SCDL** (Socle Commun des Données Locales) utilisé

Sur une méthode de qualification générique, l'intérêt est d'attribuer une note aux sous-critères qui sont les plus pertinents pour effectuer les mesures

Tableau synthétique de notation par critère pour l'ensemble des données :

<https://docs.google.com/spreadsheets/d/1EejETcNnl-VXwtUR2yGj3uA-9ZnllcuscNtrbTMpJOg/edit#gid=1317631644>

Notation pour les équipco :

		Exhaustivité	Cohérence logique					Précision thématique			Précision de position		Qualité temporelle			
		Taux d'exhaustivité	Taux de conformité aux domaines de valeurs	Conflit de structure physique	Taux de connexions erronées	Nb de micro-surfaces non valides	Nb d'auto-intersections non valides	Nb de chevauchements	Matrice relative de classement erroné	Taux de valeurs d'attribut correctes	Incertitude avec seuil de signification de 95 %	Valeur moyenne des incertitudes	Classe de précision (arrêté 16/09/2003)	Exactitude avec seuil de signification de 95 %	Taux de conformité chronologique	Taux de conformité aux domaines de valeurs
Équipements Collectifs		4	4	1	NA	NA	NA	NA	4	4	/	3	3	4	2	2
Colonnes	CRIGE															d_maj
	SCDL		EQUIP_THE ME	[horaires, n° de téléphone et jours]					EQUIP_THE ME	EQUIP_THE ME				EQUIP_OUV ERTURE	EQUIP_OUV ERTURE	

→ 5 sous-critères sont importants

Proposition : si modèle existant, prendre en compte tous les sous-critères en les pondérant par sa note d'importance ?

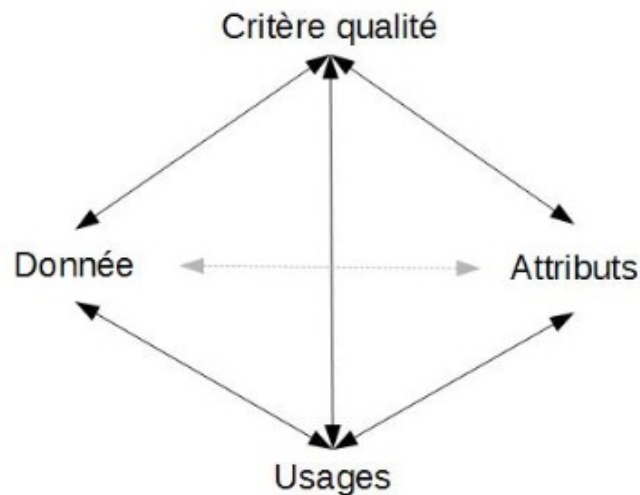
Réflexions

- Valeurs de 5 et données **critiques** relatives à la sécurité des personnes
- Des **mesures difficilement compréhensibles** nécessitant un effort pédagogique : diagrammes, illustrations nécessaires
- Des mesures qui ne s'appliquent pas à toutes les colonnes (selon leurs types)
- Différents niveaux de colonnes : celles caractéristiques du jeu de données (libelle, type d'objet), et celles décrivant les objets (source)

Au sujet des usages

- Question des **usages** centrale
- Différents usages considérés pour les EquipCo :
 - Accès sur site au public
 - Identification et visu carto
 - Stats
 - Moissonnage, flux API
- Une déclinaison nécessaire selon les usages et les colonnes

Les dimensions de la qualité



Données et critères

certains critères semblent essentiels à mesurer pour certains types de données : le critère position pour les données défibrillateur par exemple, mais pas les EquipCo

Variables et usages

certaines variables sont importantes pour certains types d'usages et pas pour d'autres : les horaires sont cruciaux pour l'accès sur site au public, mais pas pour des statistiques sur les types d'établissements.

Critères et usages

certains critères sont pertinents pour certains usages et pas pour d'autres : la précision de position pour l'accès sur site au public, mais pas pour des statistiques sur les types d'établissements.

Critères et attributs

certains critères ne s'appliquent qu'à certains types de variables : l'incertitude pour les variables quantitatives, mais pas qualitatives

Rapport détaillé

- **Exhaustivité (4)**
 - ex. oubli d'un hôpital
- **Cohérence logique**
 - Taux de conformité au domaine de valeurs (4) :
 - ex. d'un objet ne correspondant pas à la thématique EquipCo.
 - ex. des API
 - Conflit de structure physique (1) : reconnaissance possible par humain
- **Géométrie**
 - Connexions erronées, micro surfaces, etc. : Na
- **Précision thématique (4)**
 - matrice de classement et taux de valeurs d'attributs correctes : pareil que non conformité au domaine de valeurs
- **Précision de position (3)**
 - Structures de grande surface
- **Qualité temporelle**
 - Exactitude avec seuil de signification (4) : problèmes d'horaires d'accès inexacts
 - Taux de conformité chronologique (2) : jours d'ouverture, de fermeture
 - Taux de conformité au domaine de valeurs (2) :
 - date de mise à jour de l'objet vs date de création du jeu de données
 - Date de mise à jour de l'objet vs date actuelle

Questions subsidiaires

-
- EquipCo vs **ERP** ?
 - Au sujet des critères
 - Qualité temporelle et **fraîcheur**
 - Le taux de remplissage des cellules d'un tableau de données (ex. de valeurs nulles) est-il lié au conflit de structure physique ou à la conformité au domaine de valeurs ?
 - Conflit de structure physique en plusieurs catégories : format des fichiers / colonnes

Tester les données selon les 5 sous-critères identifiés

notation producteur **VS** notation calculée

=

Utiliser les balises xml contenant les notes renseignées par le producteur

VS

Calculer les résultats selon les données injectées en entrée

Quid des balises xml sur dataset multicouches?
(ex: bd carto)

Forcer le producteur à renseigner des métas que personne ne regarde ?

Restitution

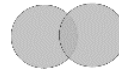
Proposition de pictogrammes qualité



Pictogrammes Cerema - fiches Qualité



Exhaustivité



Chevauchements



Conformité aux domaines de valeur



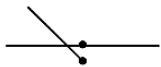
Temporel – incertitude



Conflit de structure physique



Temporel - Taux de conformité au domaine de valeurs



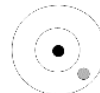
Connexions erronées



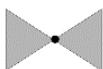
Conformité chronologique



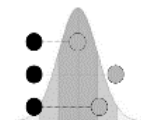
Micro-surfaces non valides



Position – valeurs moyennes des incertitudes



Auto-intersections non valides



Position – valeurs moyennes des incertitudes

Restituer le résultat de manière conviviale et compréhensible

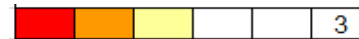
« Au regard de votre jeu données, la qualité est mesurée à 3/5 selon les critères listés ci-dessous »



- Taux d'exhaustivité 3



Exhaustivité : 3



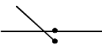
- Cohérence logique 4,6



Conformité aux domaines de valeurs : 4



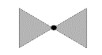
Conflit de structure physique : 5



Connexions erronées : 5



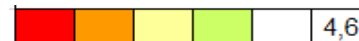
Micro-surfaces non valides : 5



Auto-intersections non valides : 5



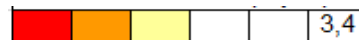
Chevauchements : 5



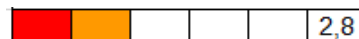
Pondération :
note du sous-critère
x
importance du critère



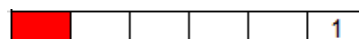
+ Précision thématique 3,4



+ Précision de position 2,8



+ Qualité temporelle 1



Restituer le résultat de manière conviviale et compréhensible

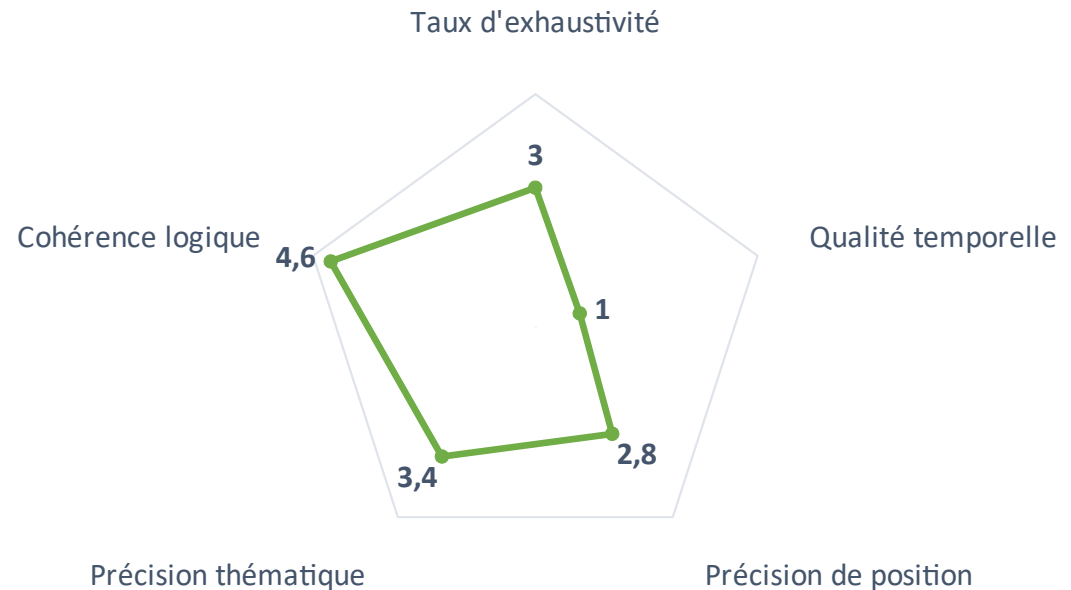
Qualité producteur



Qualité mesurée



- Balise xml 1* : 3
- Balise xml 2* : 2
- Balise xml 3* : 5
- Balise xml 4* : 5
- Balise xml 5* : 5



Conclusion

- des critères nécessitant un effort pédagogique pour être expliqués et compris
- la question des usages est essentielle et semble indissociable de la mesure qualité
- de la difficulté, donc, de noter chaque sous-critère, et donc chaque critère, et enfin, l'ensemble d'un jeu de données