

Le SANDRE

Service d'administration nationale des
données et référentiels sur l'eau



Pourquoi le Sandre ?

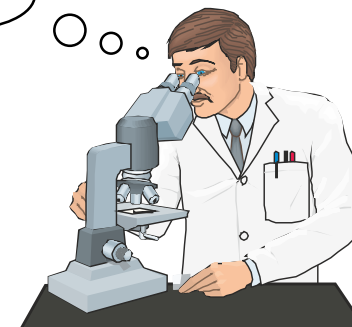
PB : Comment parler une langue commune dans un domaine aussi vaste que l'eau ?



Allô M. Dupont ? Je viens de faire le prélèvement d'eau, j'ai trouvé 18 Ctenophora près d'Abancourt

Dans de l'eau douce ? Incroyable, j'informe l'OFB. Êtes-vous en amont du barrage ?

N'y aurait-il pas confusion ?



Abancourt

⇒ dans le Nord ?
⇒ dans l'Oise ?

59001
60001

Code INSEE

Ctenophora

⇒ l'insecte ?
⇒ La diatomée ?
⇒ L'invertébré ?

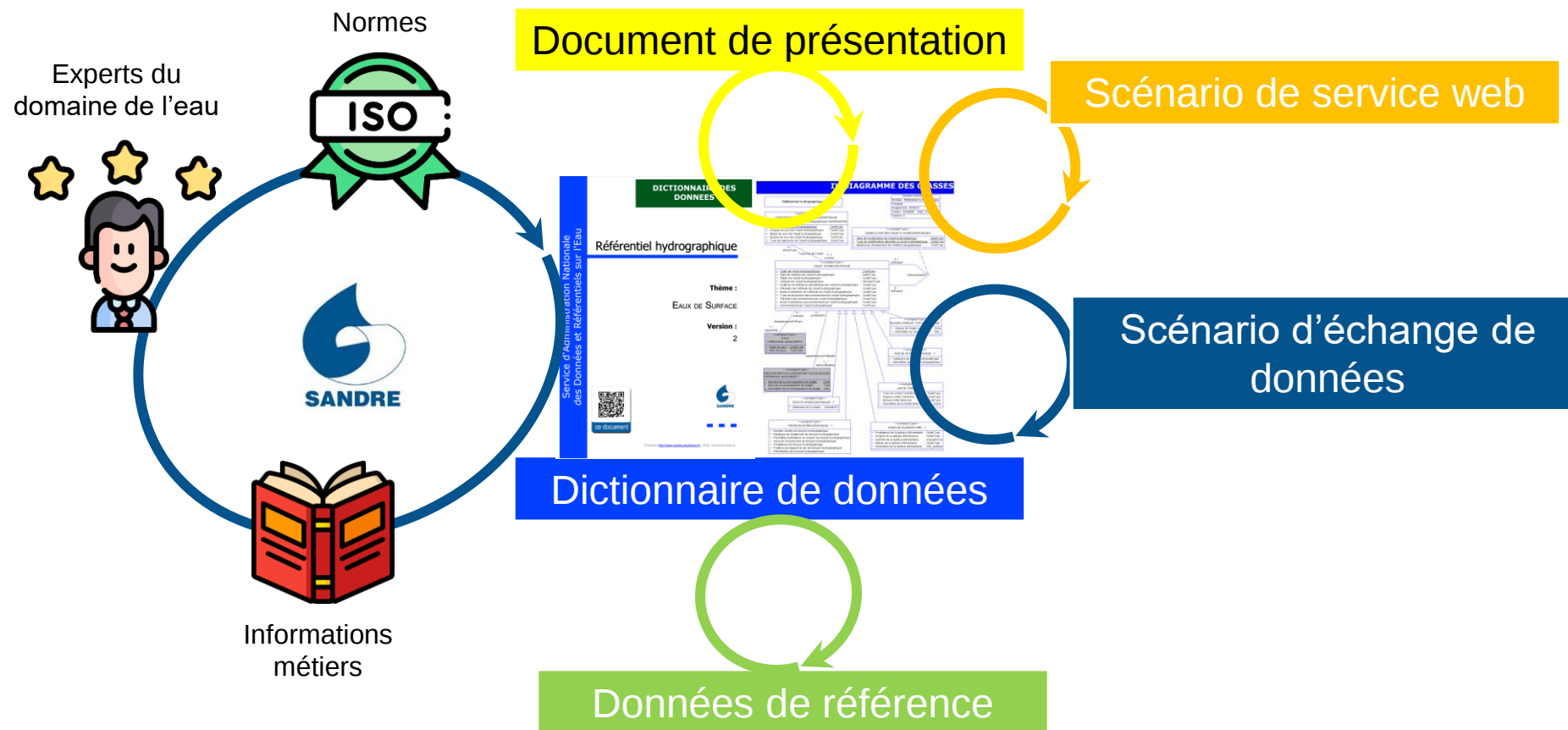
41958
9525
41957

Code Sandre



Établir le langage commun

Solution : spécifier le langage commun du SIE à travers différents documents



Les données de référence

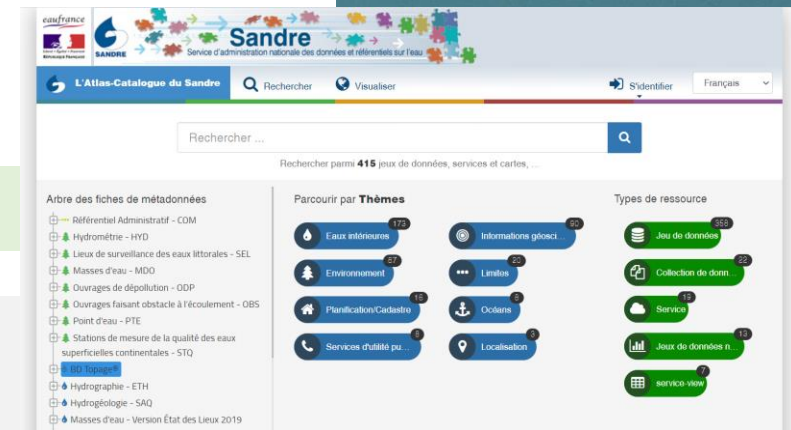
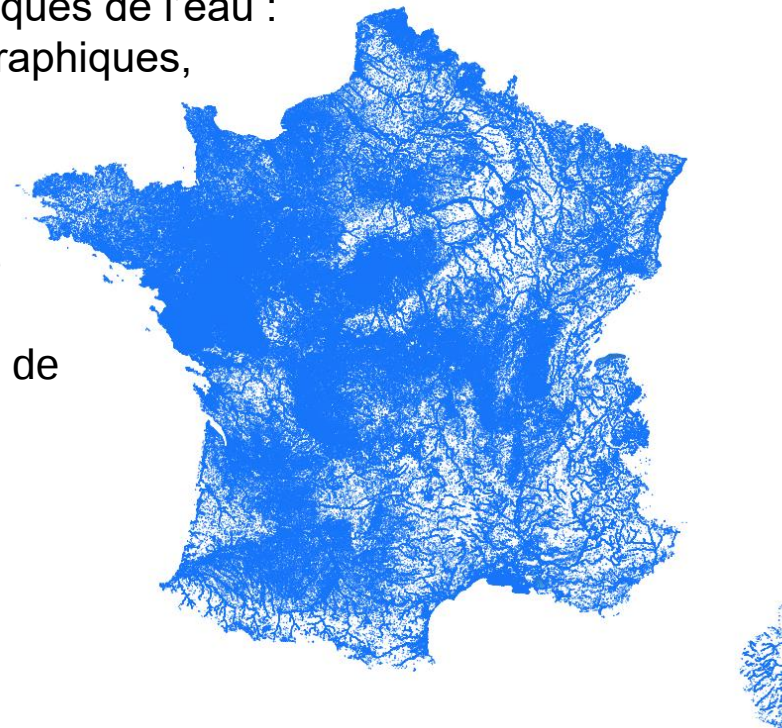
<https://www.sandre.eaufrance.fr/atlas/>

Une donnée d'un concept particulier d'un dictionnaire de données Sandre
Un code unique aussi appelé « code Sandre » est attribué à cette donnée pour être référencée de manière unique dans les systèmes d'information

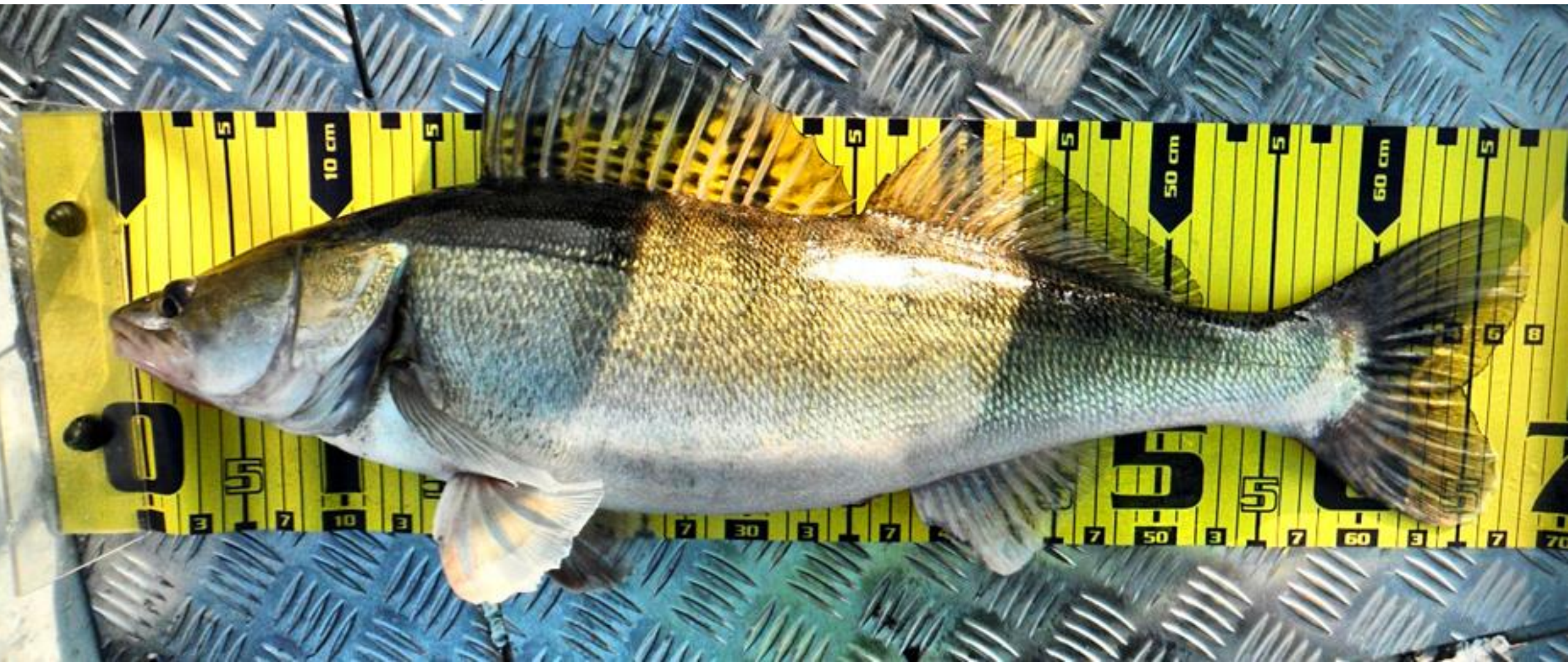
➔ Par exemple, le code Sandre 9525 du référentiel ou concept « Appellation de Taxon » désigne la diatomée Ctenophora

Les données de référence couvrent de nombreuses thématiques de l'eau :

- Hydrographie : cours d'eau, plan d'eau, tronçons hydrographiques, nœuds, etc.
- Masses d'eau
- Hydrogéologie
- Zonages : zones sensibles, zones vulnérables, zones de production ou de parage conchylicole, PGC, etc.
- Prélèvement des ressources en eau : ouvrages et points de prélèvement
- Obstacles à l'écoulement
- Stations de mesure de la qualité des eaux superficielles
- Ouvrages de dépollution
- Hydrométrie : stations, sites
- Etc.



Contrôles de conformité des jeux de données de référence





Jusqu'en 2016, le ST Sandre met les données de référence du SIE à disposition sans information relative à leur qualité. En 2017, on décide d'étendre la certification ISO 9001 du Sandre aux processus de gestion des jeux de données géographiques.

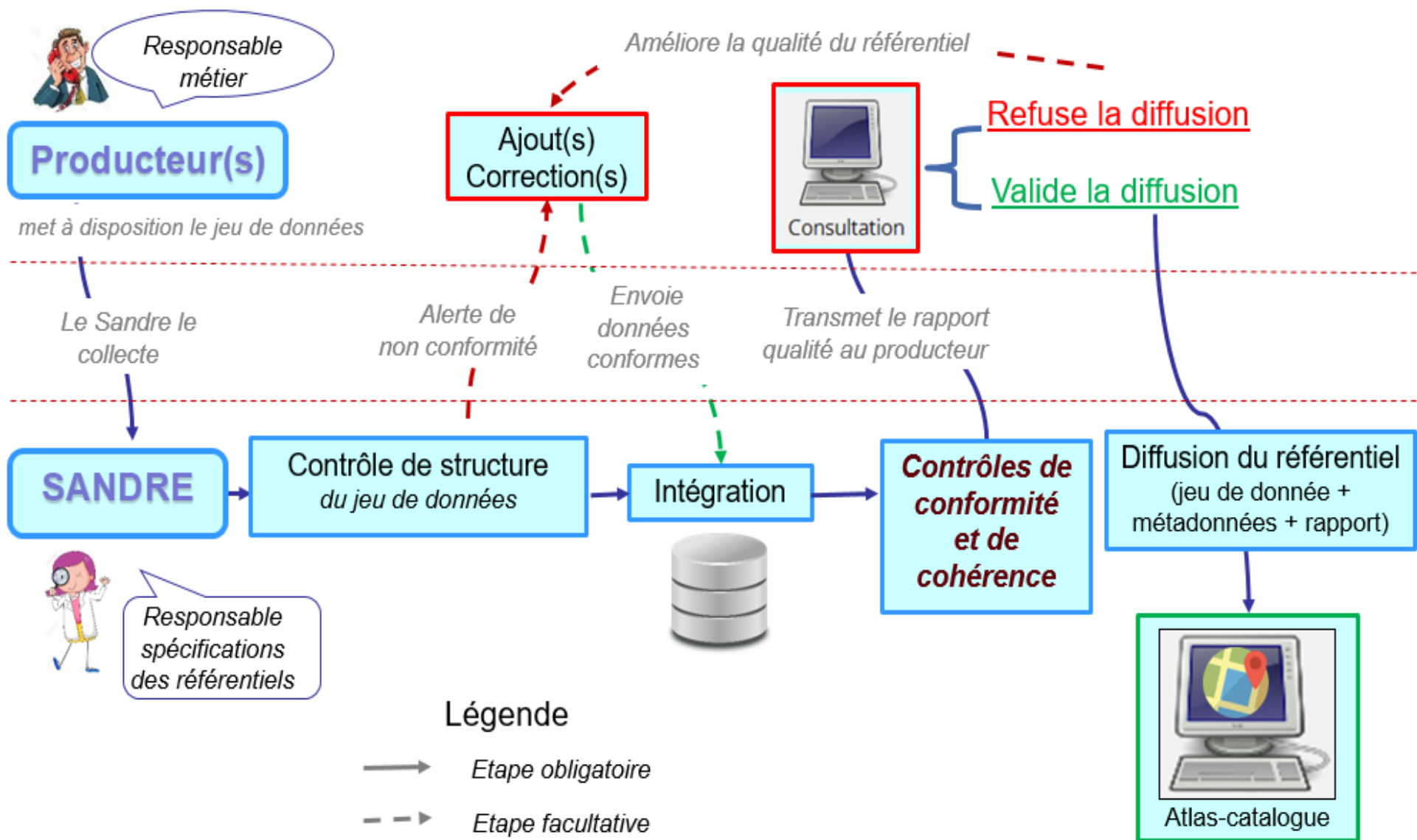
Dans ce contexte on décide de **mettre en place une interface** permettant la visualisation des résultats des contrôles réalisés, dans le but :

- **d'évaluer la qualité** des jeux de données diffusés de permettre aux utilisateurs de la données de prendre connaissance de ces informations et d'adapter leur utilisation
- de permettre **l'amélioration continue** de la donnée

En facilitant les échanges avec les producteurs ainsi que leur travail de correction



Contexte : processus général





Exemple 1 : Les AAC





Quels contrôles?

Choisis avec le producteur des données :

Contrôles de conformité Sandre

- ☐ Nom et type des champs
- ☐ Unicité des codes
- ☐ Format de attributs (longueur, règle d'écriture, ...)
- ☐ Conformité aux nomenclatures
- ☐ Conformité aux autres référentiels Sandre
- ☐ Présence des attributs obligatoires



La qualité interne → ISO 19157

Elle exprime la qualité d'un lot de données produit au regard de nécessaires spécifications, rédigées au préalable pour décrire précisément ce qui doit être produit (cahier des charges). Ces spécifications détaillent la structure et le contenu de la base de données à produire en précisant la qualité exigée.

Contrôles de cohérence

- ☐ Existence d'une géométrie
- ☐ Unicité des géométries
- ☐ Antériorité de la date de création par rapport à celle de mise à jour.
- ☐ ZAR située dans son AAC
- ☐ AAC situé dans sa circonscription administrative de bassin

Certification métiers

- ☐ Producteur identifié pour chaque AAC



Typologie de contrôles & échelle de criticité

Niveaux de criticité

1. Gravité élevée - ***Pas de diffusion***
 - Codification l'élément de référence
 - Géométrie
 - Code nomenclature
2. Gravité moyenne - ***Correction attendue***
 - Libellé nomenclature
 - Dates
 - Valeur attribut facultatif
3. Gravité faible – ***Information***
 - Attributs divers

Types de contrôles

- Unicité des géométries
- Présence des attributs obligatoires
- Respect des nomenclatures Sandre
- Respect du référentiel des intervenants
- Cohérence des dates
- Cohérence de localisation
- Respect des longueurs des attributs
- Respect des règles d'écriture
- Etc. ...

Faux positifs → Informer l'utilisateur des alertes qui concernent des cas de figure prévus par le producteur





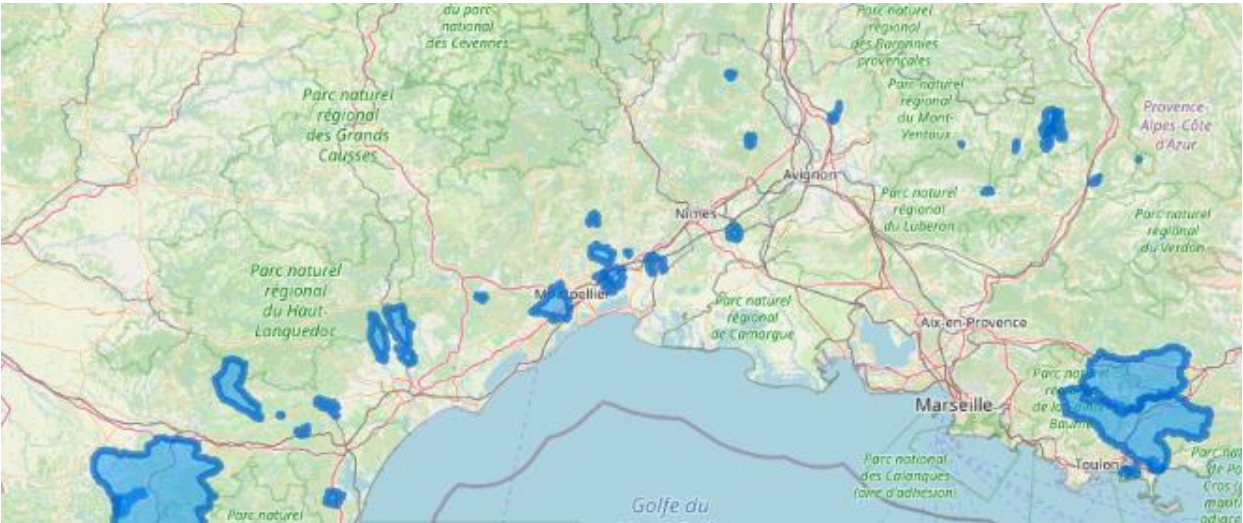
Rapport de contrôle du 04/04/2022 du référentiel Aire d'alimentation de captage sur le territoire 'France entière'

43 contrôles								
Cd	Resultat du contrôle ▼	Champ du dictionnaire	Description du contrôle	✓ 27 OK	16 échecs ❌ 1 ERREURS ⚠ 8 AVERTISSEMENTS ℹ 7 INFORMATIONS		Nb d'entités testées / Nb total	Voir
	▼	▼						
35418	❌ ERREUR	geom	Unicité de la géométrie (doublons) <i>4 entités en échec sur le test : Unicité de la géométrie (doublons)</i>				1180/1180	4 échecs
35379	⚠ AVERTISSEMENT	NumCircAdminBassin	Le code sandre du territoire administratif du bassin 'NumCircAdmin' existe dans le référentiel Sandre des Circonscriptions administratives de bassin <i>5 entités en échec sur le test : Le code sandre du territoire administratif du bassin 'NumCircAdmin' existe dans le référentiel Sandre des Circonscriptions administratives de bassin</i>				1180/1180	5 échecs
35382	⚠ AVERTISSEMENT	NumCircAdminBassin	Le périmètre d'un AAC intersecte sa circonscription administrative de bassin - intersections multiples autorisées: un AAC peut recouvrir le périmètre de plusieurs circonscriptions administratives de bassin <i>20 entités en échec sur le test : Le périmètre d'un AAC intersecte sa circonscription administrative de bassin - intersections multiples autorisées: un AAC peut recouvrir le périmètre de plusieurs circonscriptions administratives de bassin</i>				1175/1180	20 échecs
35387	⚠ AVERTISSEMENT	NumCircAdminBassin	Numéro de la circonscription administrative de bassin = 2 caractères <i>5 entités en échec sur le test : Numéro de la circonscription administrative de bassin = 2 caractères</i>				1180/1180	5 échecs
35393	⚠ AVERTISSEMENT	AuteurAAC	Auteur de l'AAC = < 50 caractères <i>85 entités en échec sur le test : Auteur de l'AAC = < 50 caractères</i>				1180/1180	85 échecs
35406	ℹ INFO	RefBiblioAAC	Valeur NULL de la référence bibliographique de l'AAC <i>1180 entités en échec sur le test : Valeur NULL de la référence bibliographique de l'AAC</i>				1180/1180	1180 échecs
35407	ℹ INFO	CdAltAAC_1	Valeur NULL du Code alternatif n°1 <i>240 entités en échec sur le test : Valeur NULL du Code alternatif n°1</i>				1180/1180	240 échecs
35408	ℹ INFO	OrCdAltAAC_1	Valeur NULL de l'Origine du code alternatif n°1 <i>236 entités en échec sur le test : Valeur NULL de l'Origine du code alternatif n°1</i>				1180/1180	236 échecs
35376	✓ OK	geom	Localisation dans une emprise (METROPOLE, GUADELOUPE, MARTINIQUE, GUYANE, REUNION, MAYOTTE) <i>Localisation dans une emprise (METROPOLE, GUADELOUPE, MARTINIQUE, GUYANE, REUNION, MAYOTTE)</i>				1180/1180	
35377	✓ OK	publie_site_aac	AAC publié sur le portail aire-captage <i>AAC publié sur le portail aire-captage</i>				1180/1180	
35378	✓ OK	valide_referent	AAC validé par le référent <i>AAC validé par le référent</i>				1180/1180	



Rapport de contrôle du 04/04/2022 du référentiel Aire d'alimentation de captage

Auteur de l'AAC =< 50 caractères



Description du contrôle: Respect d'une longueur de chaîne de caractères(paramètres => operateurs: < > <= >= !=, longueur: 0-99

Télécharger le fichier CSV des résultats.

Code sandre de l'entité ^	Commentaire	Code sandre de l'entité comparée
5498	La valeur "numérisation DREAL Bretagne d'après docs. EdM (dossier DUP)" du champ "AuteurAAC" de l'entité "AAC" de code Sandre "5498" a une longueur de 61 caractères. Elle ne doit pas être > à 50 caractères	numérisation DREAL Bretagne d'après docs. EdM (dossier DUP)
5493	La valeur "numérisation DREAL Bretagne d'après docs. EdM et DDTM56" du champ "AuteurAAC" de l'entité "AAC" de code Sandre "5493" a une longueur de 57 caractères. Elle ne doit pas être > à 50 caractères	numérisation DREAL Bretagne d'après docs. EdM et DDTM56
1673	La valeur "Agence de l'eau Rhône Méditerranée et Corse (AE RMC)" du champ "AuteurAAC" de l'entité "AAC" de code Sandre "1673" a une longueur de 55 caractères. Elle ne doit pas être > à 50 caractères	Agence de l'eau Rhône Méditerranée et Corse (AE RMC)
1658	La valeur "Agence de l'eau Rhône Méditerranée Corse (AERMC)" du champ "AuteurAAC" de l'entité "AAC" de code Sandre "1658" a une longueur de 51 caractères. Elle ne doit pas être > à 50 caractères	Agence de l'eau Rhône Méditerranée Corse (AERMC)
1656	La valeur "Agence de l'eau Rhône Méditerranée Corse (AERMC)" du champ "AuteurAAC" de l'entité "AAC" de code Sandre "1656" a une longueur de 51 caractères. Elle ne doit pas être > à 50 caractères	Agence de l'eau Rhône Méditerranée Corse (AERMC)
1636	La valeur "Agence de l'eau Rhône Méditerranée Corse (AERMC)" du champ "AuteurAAC" de l'entité "AAC" de code Sandre "1636" a une longueur de 51 caractères. Elle ne doit pas être > à 50 caractères	Agence de l'eau Rhône Méditerranée Corse (AERMC)





Spécification sandre

▼ Auteur de l'AAC

Premier organisme à avoir demandé au SANDRE l'enregistrement de l'AAC au sein du référentiel. L'attribution d'un auteur à une AAC relève de la responsabilité du SANDRE.

Identifiant: <http://id.eaufrance.fr/ddd/ZGR/1/AuteurAAC>

Balise XML: AuteurAAC

Définition:  Premier organisme à avoir demandé au SANDRE l'enregistrement de l'AAC au sein du référentiel.
L'attribution d'un auteur à une AAC relève de la responsabilité du SANDRE.

Cardinalité min: 0

Cardinalité max: 1

Type: cct:TextType

Longueur max: 50





Les 5 critères proposés par la norme ISO 19157

Cohérence logique

- la cohérence **conceptuelle**
- la cohérence des **domaines** de valeur
- la cohérence du **format**
- la cohérence **topologique** 🌐

Exhaustivité

- l'**excédent**
- l'**omission**

Précision thématique

- la justesse du **classement**
- la justesse des **attributs non quantitatifs**
- la précision des attributs **quantitatifs**

Précision de position 🌐

- la précision de **position absolue**
- la précision de position **relative**
- la précision de position de **données matricielles**

Qualité temporelle

- l'**exactitude** de la mesure temporelle
- la **cohérence** temporelle
- la **validité** temporelle.





RÉSUMÉ : PLU de Numérac

NOTE DE L'EVALUATEUR



NOTE DES UTILISATEURS (16)



VOTRE AVIS

[DÉPOSER UN AVIS >](#)

SOUS-NOTES

EXHAUSTIVITE



PRECISION THEM



PRECISION GEO



COHERENCE LOGIQUE



Usages recommandés

- ✓ Analyse statistique
- ✓ Calculs de surface

Usages proscrits

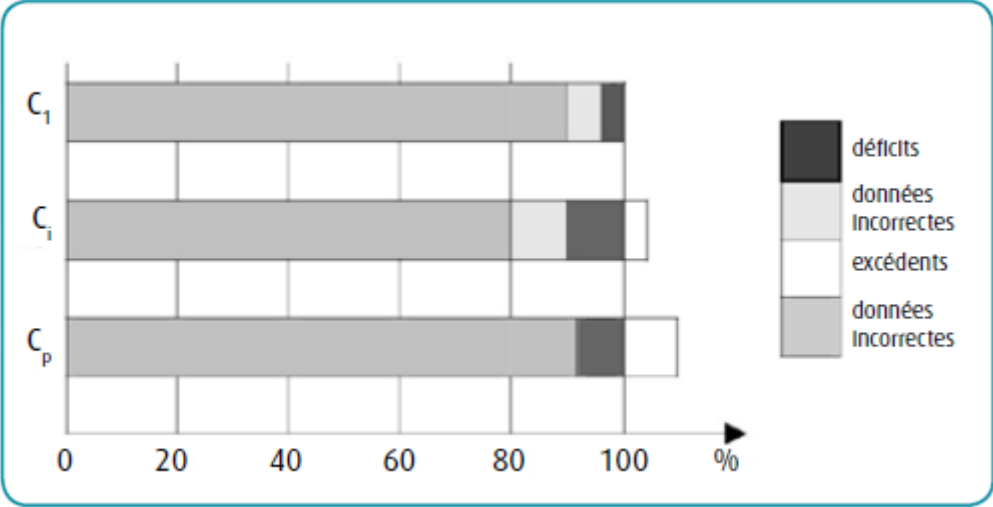
- ✗ Instruction des autorisations d'urbanisme
- ✗ Représentation cartographique





Ex: Indicateurs qualité et représentation synthétique

Représentation – Notation / critères ISO - 19157



Classification/ catégorisation	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

Dataset quality metrics

exhaustivity ★★★★★
logicalConsistency ★★★★★
positionAccuracy ★★★★★
thematicAccuracy ★★★★★
temporalAccuracy ★★★★★
spatialAccuracy ★★★★★
usability ★★★★★



Cohérence topologique taux de cohérence topologique	Note sur 5
De [99 % à 100 %]	5
De [97,5 % à 99 %[	4
De [95 % à 97,5 %[	3
De [90 % à 95 %[	2
Taux < 90 %	1

Cohérence des domaines de valeurs taux de conformité aux domaines de valeurs	Note sur 5
De [99,5 % à 100 %]	5
De [99 % à 99,5 %[	
De [97,5 % à 99 %[	
De [95 % à 97,5 %[	
Taux < 95 %	

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

Exemple 2 : Stations hydrométriques





Contrôles réalisés

Contrôles de conformité Sandre

- ❑ Nom et type des champs
- ❑ Unicité des codes de stations
- ❑ Présence des attributs obligatoires
- ❑ Respect des nomenclatures
- ❑ Respect des longueurs des attributs
- ❑ Respect du référentiel des intervenants



Contrôles de cohérence métier

- ❑ Unicité des géométries
- ❑ Antériorité de la date de mise en service par rapport à celle de fermeture
- ❑ Station dans la zone hydrographique
- ❑ Station à moins de 1 km de la commune associée
- ❑ Station à moins de 50 m du tronçon hydrographique associé
- ❑ Station à moins de 50 m du cours d'eau associé
- ❑ Station dans l'emprise de la métropole



Rapport détaillé

En ligne pour le producteur de la donnée

<https://mdm.sandre.eaufrance.fr/geo/rapportsv3/>



Rapports de contrôle des référentiels géographiques

	Date ^	Dictionnaire	Référentiel	Territoire	Etat
			Station hydrométrique		
	20/11/2019	HYD 1	Station hydrométrique	Guyane	✓ 9 OK ⚠ 4 Avertissements
	20/11/2019	HYD 2.3	Station hydrométrique	Guyane	✓ 9 OK ⚠ 4 Avertissements
	19/11/2019	HYD 1	Station hydrométrique	Guyane	✓ 5 OK ⚠ 3 Avertissements
	19/11/2019	HYD 2.3	Station hydrométrique	Guyane	✓ 5 OK ⚠ 3 Avertissements
	30/05/2017	HYD 1	Station hydrométrique	Métropole	✓ 5 OK ⚠ 11 Avertissements



Rapport détaillé

En ligne pour le producteur de la donnée

Rapport de contrôle du 30/05/2017 du référentiel Station hydrométrique sur le territoire 'Métropole'

[Retour à la liste des Rapports](#)

16 contrôles	
✓ 5 OK	11 échecs
	✖ 0 ERREURS
	⚠ 11 AVERTISSEMENTS
	i 0 INFORMATIONS

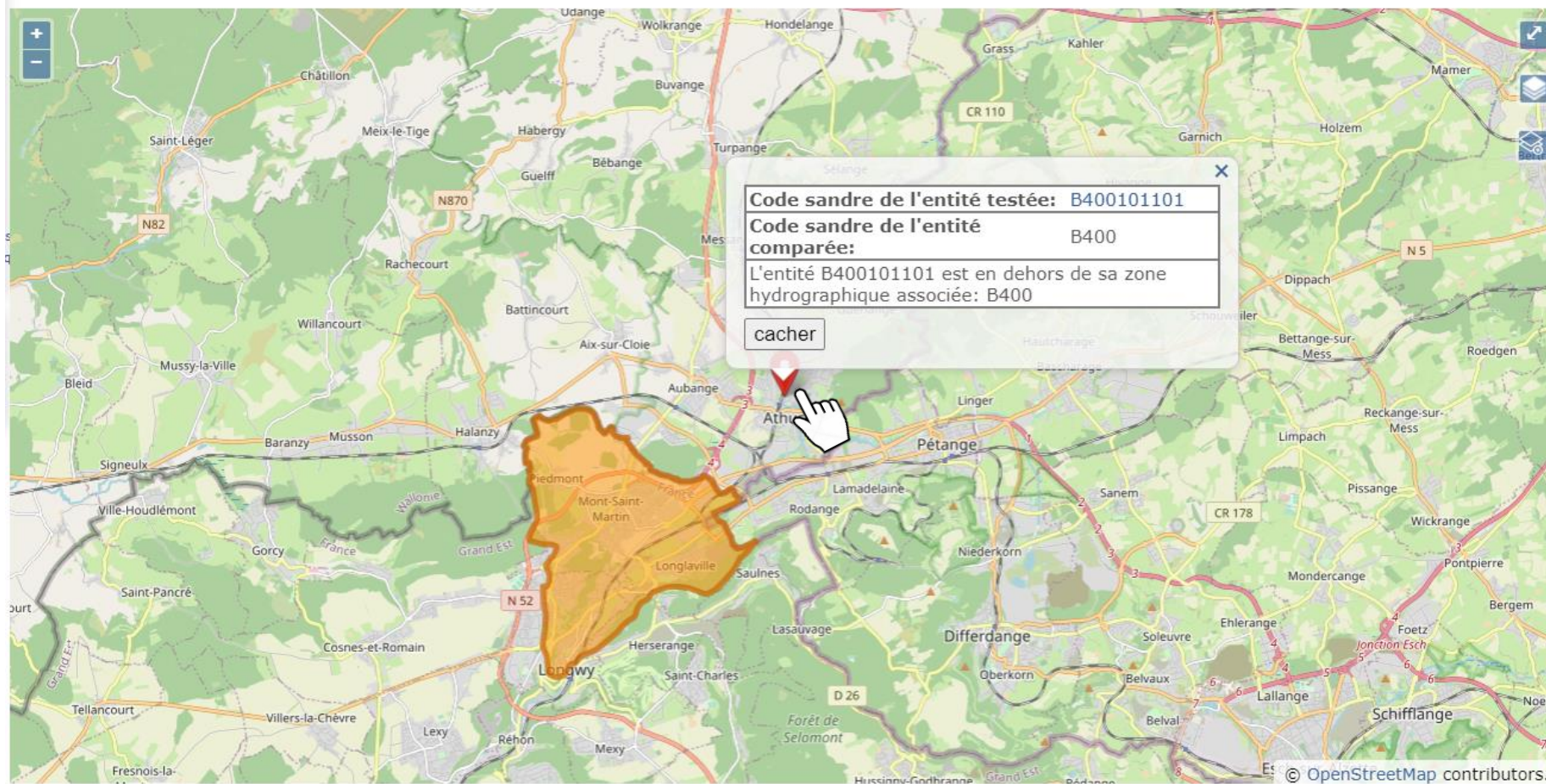
Cd	Resultat du contrôle ▼	Champ du dictionnaire	Description du contrôle	Nb d'entités testées / Nb total	Voir
	▼	▼			
1301	⚠ AVERTISSEMENT	CdIntervenant ; NomIntervenant	Respect du référentiel des intervenants pour le gestionnaire de station hydrométrique. <i>Relance du contrôle de conformité des intervenants après correction du code erroné 13001779101321.</i>		2763 échecs
1296	⚠ AVERTISSEMENT	geom	Présence de la station hydrométrique à moins de 50 m du cours d'eau associé		653 échecs
1298	⚠ AVERTISSEMENT	geom	Présence de la station hydrométrique à moins de 50 m du tronçon hydrographique associé		594 échecs
1279	⚠ AVERTISSEMENT	geom	Présence de la station hydrométrique sur le cours d'eau associé		1000 échecs

Rapport détaillé

En ligne pour le producteur de la donnée

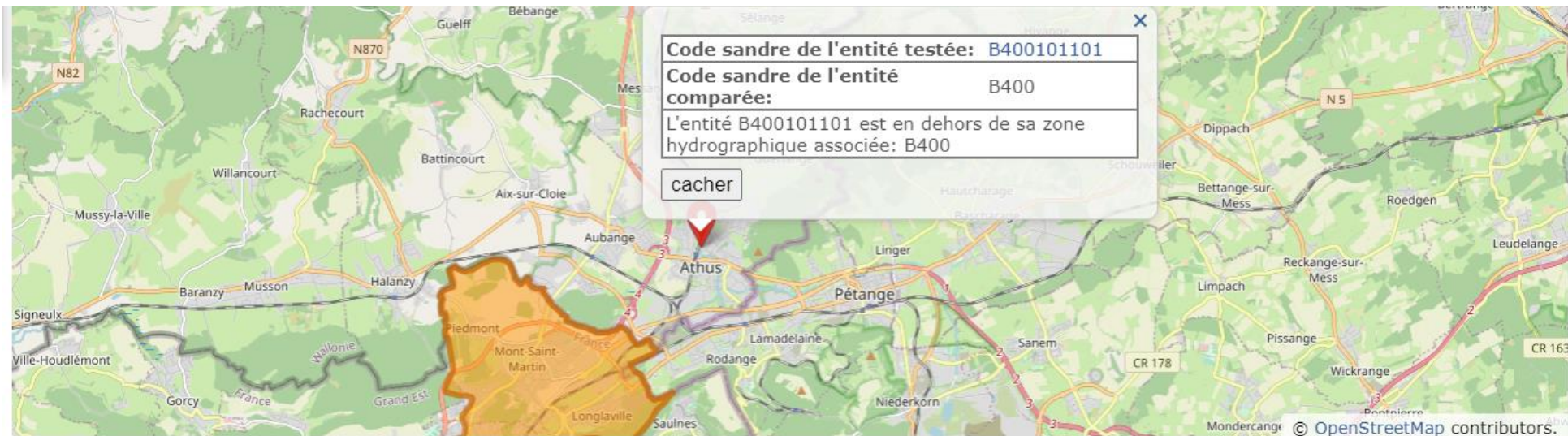
Rapport de contrôle du 30/05/2017 du référentiel Station hydrométrique sur le territoire 'Métropole'

[Retour à la liste des autres contrôles.](#)



Rapport détaillé

En ligne pour le producteur de la donnée



Description du contrôle: Cohérence de localisation

Télécharger le fichier CSV des résultats.

Code sandre de l'entité ^	Commentaire	Code sandre de l'entité comparée
Y991000101	L'entité Y991000101 est en dehors de sa zone hydrographique associée: Y991	Y991
Y911000201	L'entité Y911000201 est en dehors de sa zone hydrographique associée: Y911	Y911
Y830000101	L'entité Y830000101 est en dehors de sa zone hydrographique associée: Y830	Y830
Y060401001	L'entité Y060401001 est en dehors de sa zone hydrographique associée: Y061	Y061
Y047402001	L'entité Y047402001 est en dehors de sa zone hydrographique associée: Y047	Y047
Y047401001	L'entité Y047401001 est en dehors de sa zone hydrographique associée: Y047	Y047

1^{er} Bilan (2017)



Points forts:

- Amélioration de la qualité du processus de gestion des données
- Un outil pour la consolidation en continu des référentiels géographiques

Mais...



Des améliorations nécessaires:

- Export des erreurs
- Concertation sur le calcul des notes
- Pondération des erreurs
- Ergonomie
- Description des tests
- Validation en ligne des erreurs



Perspective: Vers un rapport simplifié accessible à tous sur le catalogue

- Définir ensemble des règles d'élaboration de ce rapport (à partir du rapport détaillé)
- Définir une note globale évaluant la qualité du jeu de données?



Note



Perspectives (2022)



- ▶ Gestion des faux positifs
- ▶ Exhaustivité des contrôles $\Rightarrow$ cf. travaux du QuaDoGéo
- ▶ Simplification de la lecture des résultats
- ▶ Rapport simplifié et évaluation par les utilisateurs $\Rightarrow$ cf. travaux du QuaDoGéo

- ▶ **Intégration des recommandations du**  **!**

