

 <b>RÉPUBLIQUE FRANÇAISE</b> <i>Liberté Égalité Fraternité</i> <b>CNIG</b> Conseil national de l'information géolocalisée <b>Commission règles &amp; qualité</b>  <b>Groupe de Travail Qualité des Données Géographiques (QuaDoGéo)</b>	Groupe de Travail animé par :   <b>Crige</b> Centre de Ressources en Information Géographique Provence-Alpes-Côte d'Azur   <b>IGN</b> INSTITUT NATIONAL DE L'INFORMATION GÉOGRAPHIQUE ET FORESTIÈRE
<p align="center"><b>COMPTE-RENDU DE REUNION (distanciel)</b></p>	
<p align="center"><b>Objet : Réunion #22 du GT QuaDoGéo du 13 mai 2025.</b></p>	

**Ordre du jour :**

1. Accueil et rappels
2. Validation du [précédent compte-rendu](#)
3. Date du prochain GT
4. Assurer la qualité des données avec ArcGIS (ESRI)
5. Retex Grand Lyon, gestion de la qualité dans le socle et rapprochement avec les données référentielles (Grand Lyon)
6. Exploitation de l'outil OSM Osmose à l'IGN
7. Avancées du sous-groupe "Questionnaire/sondage sur la pratique des ETL dans le contrôle qualité"
8. Décision de principe, page QuaDoGéo "Comment faire pour valider vos données" réutiliser dans fabrique des standards

**Liste des participants****Animateurs** : Stéphane Rolle (CRIGE PACA) et Nicolas Py (IGN)

Nom Prénom	Organisme	Présent	Excusé
Jean-Marie ARSAC	AZIMUT	X	
Mamadou-Bailo BALDE	Ineris	X	
Chantal COULOMB	IGN / Contrôle Qualité	X	
Laurène DEBRAY	Office international de l'eau / Sandre	X	
Arnauld GALLAIS	Cerema Ouest	X	
Gaëtan LAVENU	ESRI	X	
Mathieu Le MOAL	Axes Conseil	X	
Alexis LEAUTIER	CNIG	X	
Loïc MATHIS	Métropole de Lyon	X	
Benoît MORANDO	CNIG	X	
Nicolas PY	IGN Centre Est	X	
Stéphane ROLLE	CRIGE PACA	X	
David SARDIN	ESRI	X	
Benoit SEGALA	Consultant - Cabinet d'études	X	
Oscar URHY	Métropole de Lyon	X	

Réunion #22 du GT QuaDoGéo du 13 mai 2025 (distanciel).		
Mise en relecture du compte-rendu	28/05/2025	Stéphane Rolle / Nicolas Py
Relecture du compte-rendu	Octobre 2025	Les participants.
Validation du compte-rendu	Octobre 2025	Stéphane Rolle / Nicolas Py
<b>Prochain rendez-vous en octobre 2025 en distanciel ou présentiel.</b>		

## 1. Accueil et rappels

### Vie du GT

La réunion (#22, 2025-2) prévue initialement le lundi 10 mars 2025 a été décalée au mardi 13 mai. Malgré un stock de sujets/intervenants potentiels approchés et enclins à partager leur expérience, leur conversion en intervention confirmée est difficile à obtenir.

### Archives du GT

La reprise du patrimoine de comptes-rendus pour vérifier les présentations, documents, sites web liés s'est poursuivie, les réunions jusqu'à la #07 2020-3 sont traitées.

### Fabrique des standards

Nombreuses réunions de travail en ce début d'année, les premiers livrables sortent en beta, <https://cnig.gouv.fr/gt-fabrique-des-standards-a26189.html> ([permalien](#)) sur le processus (ses étapes, pourquoi, comment) mais aussi sur le standard des standards, aka le *template* de rédaction de spécifications.

### Décision/actions :

**Tous:** les membres du GT sont invités à s'emparer dès à présent de ces documents.

## 2. Validation du précédent compte-rendu

Aucune remarque à signaler sur le précédent compte-rendu (réunion du 13 janvier 2025) [disponible](#) ([permalien](#)) sur la page QuaDoGéo.

## 3. Date du prochain GT

La prochaine réunion est envisagée en octobre 2025, entre le 06 et le 10 octobre ; choix via un sondage envoyé aux membres du GT.

Répondre à <https://framadate.org/hY8ddCGUBxVSxmfg> avec la possibilité d'une réunion présentielle 😊 dont le lieu sera alors à déterminer.

Un rappel sera effectué le 11/06/2025.

Le sondage sera clos le 18/06/2025.

## 4. Assurer la qualité des données avec ArcGIS

[Présentation](#) ([permalien](#)) de David SARDIN et Gaëtan LAVENU (ESRI france) sur les fonctionnalités liées aux qualités des données et métadonnées dans l'univers des produits Esri (bureautiques, web, etc.).

La géodatabase est la première fonctionnalité pouvant être citée. Il s'agit d'une base de donnée spatiale disponible en plusieurs implémentations (fichier, extension de SGBDR, SQLite) dont certaines caractéristiques dépendent de l'implémentation (caractère mono vs multi utilisateurs notamment) [[lien/permalien](#)].

Au sein de la géodatabase, le concepteur de la donnée peut spécifier :

- des domaines de valeur associés à chaque attribut - intervalles, listes - (exemple, plage de valeurs autorisées sur un attribut nombre) [[lien/permalien](#)]

- des sous-types, aka des règles applicables à une partie des objets de la table [\[lien/permalien\]](#)
- 
- Des valeurs conditionnelles, qui dépendent de la valeur d'un attribut (attributs en cascade) [\[lien/permalien\]](#)
- des relations, afin de vérifier des cardinalités [\[lien/permalien\]](#)
- des règles de topologie (ex., x doit être inclus dans y) [\[lien/permalien\]](#)
- des règles attributaires, aka des scripts permettant de spécifier des contraintes plus complexes [\[lien/permalien\]](#)

Ces règles/contraintes restreignent la saisie et/ou sont vérifiables après éditions.

Les géodatabases associées à un SGBDR disposent en outre de fonctionnalités de versionnement, permettant de créer des branches de travail puis offrant des outils de réconciliation et gestion des conflits de la réintégration dans le socle de ces données temporairement traitées à part dans une branche [\[lien/permalien\]](#).

Ainsi, à l'aide de ces outils et fonctionnalités de conception et qualité by design, il est possible d'implémenter des modèles de données métiers. Cela s'applique particulièrement aux réseaux de transport, hydrographiques, d'assainissement, et même à des modèles plus complexes dit *utility network* [\[lien/permalien\]](#) comme le réseau électrique. En France, ce sont notamment des propositions du programme Arcopole [\[lien/permalien\]](#).

Bien que la maîtrise de la qualité des données lors de la saisie s'effectue le plus efficacement via le client lourd ArcGIS Pro, les outils web ArcGIS (Map Viewer, Web Editor, etc.) embarquent de plus en plus de fonctionnalités dédiées à la garantie de la qualité des données. En leur sein, les outils/formulaires de saisie s'adaptent aux règles ; mais également aux outils d'accroche (*snapping*).

ArcGIS Pro dispose aussi d'inspecteurs de topologie et d'erreurs (objets non conformes aux règles, à la topologie) [\[lien/permalien\]](#).

Possibilité également de *formulaires intelligents* (quels champs sont éditables, visibles, calculés, listes déroulantes, valeurs dépendantes). Le comportement des formulaires est programmé en langage de script Arcade (langage propriétaire, mélange de javascript et python). Ces formulaires sont utilisables dans les clients (bureautique, web, mobile, ...).

Après la saisie, le système ArcGIS propose des outils de data-engineering :

- Visualisation des champs, représentation graphique/statistiques des valeurs contenues dans chaque champ ; aka statistiques descriptives pour chaque champ directement dans la table attributaire.
- Outils de la toolbox : calculs de champs, snapping (avec spécification de tolérance) sur une grille ou entre les objets, interpolation, généraliser, multipart vers unipart,...
- Outils ETL, extension *data interoperability* intégrant FME dans ArcGIS

Le système ArcGIS dispose d'une extension spécifiquement dédiée à la qualité des données, quasi no-code, l'extension ArcGIS Data Reviewer [\[lien/permalien\]](#).

Cette extension industrialise le suivi de la qualité, via un workflow, en s'appuyant sur les outils de base (géodatabase, domaines, sous-types, règles, topologie, ...). Le workflow comprend les étapes :

1. Définition des règles de validation topologiques et attributaires ; règles de contraintes (respecter impérativement) ou pour validation (alerte, suspicion) ; avec niveaux de criticité des règles
2. Revue automatique ou semi-automatique des erreurs et écarts aux règles
3. Les objets sont alors flagués et disposent d'un statut de revue (null, corrigé, vérifié, exception, ...) [[lien/permalien](#)]
4. Outils d'édition et de visualisation dédiés au passage en revue des objets vus en erreur
5. Rapport et suivi des anomalies (et notamment statistiques sur le devenir des objets flagués en erreur et passés en revue)

Les métadonnées dans le système ArcGIS [[lien/permalien](#)] emploient une implémentation propriétaire, à base de xml, et un outil d'édition ; cet outil permet de traduire les informations de métadonnées stockées de manière « ArcGIS » dans la géodatabase vers un export au profil/norme de métadonnées quelconque (ISO, Inspire, DCAT, ...).

L'éditeur a été repensé ~2022 dans son design d'interface, notamment pour suivre le niveau d'avancement de la saisie des éléments de métadonnées, de proposer une saisie alternative à l'utilisateur entre saisie simple (métadonnées obligatoires selon le profil de métadonnées choisi) ou davantage complète. Les métadonnées sont stockées au plus près de la donnée, ex. pour un shapefile création d'un xml compagnon. La métadonnée s'enrichit de plus automatiquement en documentant les traitements appliqués à la donnée (rubrique généalogie).

Dans les perspectives d'évolution du système ArcGIS qui pourraient avoir une utilité pour la maîtrise de la qualité des données géographiques, David et Gaëtan citent l'emploi de l'IA pour détecter des anomalies ou simplifier la saisie des métadonnées, la création de davantage d'outils sur les données temps réel (ArcGIS Velocity), et des travaux autour des BDs de graphe de connaissance (ArcGIS Knowledge [lien/permalien](#)).

**NPY** se demande si les standards CNIG sont repris dans les modèles de données métier, dans le programme Arcopole. **AGallais** reformule en prenant l'exemple GT Accessibilité. Arcopole vise à faire communauté et mutualiser dans ces communautés la création de ressources, des widgets comme des modèles de données. Ces modèles sont fournis sous forme de fichiers xml\_générateur-de-gdb [[lien/permalien](#)]. Le programme Arcopole, né il y a environ 15 ans dans contexte plutôt français, est ouvert à tous, y compris aux non client ESRI qui ont certes un intérêt moindre aux ressources spécifiquement ArcGIS. Arcopole visait initialement à répondre à un besoin plutôt exprimé par les collectivités, qui consistait à disposer de modèles de structure des BDs, mais également des widgets (plugins développés à façon) davantage opérationnels à la thématique métier afin de mutualiser les besoins et les financements. Le programme a ensuite évolué vers d'autres ressources que des scripts/plugins.

A la suggestion de relier les communautés Arcopole-CNIG, il est acté une réunion de présentation mutuelle.

**NPY** s'interroge sur la faisabilité de stocker dans les métadonnées des métadonnées de qualité, au sens ISO19157 ; aka est-ce possible de documenter la précision planimétrique (code iso, méthode de mesure, valeur de la mesure) ?

**AGallais** répond que si les métadonnées Inspire sont pleinement supportées alors vraisemblablement les métadonnées de qualité le sont également.

Après vérification, la documentation le mentionne [[lien/permalien](#)], ESRI participe aux travaux de cette section ISO [[lien](#), *whitepaper de 2015*].

Décision/actions :

**SG du CNIG, Dsardin (ESRI):** monter une réunion de présentation au CNIG et notamment Fabrique des Standards du programme Arcopole (vision, existant, prochain projets)

**5. Retex Grand Lyon, gestion de la qualité dans le socle et rapprochement avec les données référentielles**

Présentation ([permalien](#)) de Loïc MATHIS.

La première partie de l'intervention est consacrée au projet « socle ». Ce projet consiste à centraliser les données utilisées par toutes les directions et services, ou d'intérêt pour tout agent du Grand Lyon, dans une même base de données socle. Cela inclut l'agrégation des données inféodées à des progiciels métiers.

D'un point de vue architecture, le SGBDR retenu est PostgreSQL/PostGIS, fourni en 5 environnements (test => prod).

L'information est organisée en « schémas de référence » (les tables brutes) complétées de « schémas de publication » contenant les vues et vues matérialisées (pour la sélection ou le filtre de donnée, la gestion des droits d'accès).

La seconde partie de l'intervention décrit les travaux du groupe de travail du Grand Lyon sur la qualité des données géographiques.

Son objectif était de rédiger un guide pour toutes les données du socle (cf. supra), quelle que soit l'unité du Grand Lyon gestionnaire ; accompagné de scripts FME & FME Flow, PostGIS, QGIS. Ces livrables (scripts et guide) ont été partagés avec les direction/services ; notamment via le réseau des géomaticiens de la métropole et ses réunions trimestrielles.

La qualité des données du socle est auditée par le service fédéral géomatique ; les contrôles opérés via FME Flow sont utilisables par tous.

**NPY** interroge Loïc sur l'emplacement des descriptions des schémas (noms et types des attributs), celles-ci ne sont pas lisibles machine mais cependant stockées dans le catalogue de données.

Il n'est pas toujours simple mais généralement faisable de vérifier la conformité 'automatisable', le respect de formalismes ; il est cependant complexe et couteux de vérifier les précisions planimétriques, thématiques, temporelles ou l'exhaustivité.

**BSegala** pose la question d'une données vérifiée après échantillonnage. Le seul cas connu de **LMathis** n'est pas au service géomatique mais au service de la voirie, où des carrés de 5\*5km sont sélectionnés et les données du service (tronçons de voirie, objets rattachés) sont auditées par contrôle photo/satellite/fond de plan (donc sans retour terrain).

La dernière partie de l'intervention cite les référentiels nationaux avec lesquels Grand Lyon se synchronise (BAL/BAN, PCRS, RNB, BDTopo, ...).

Un nouveau projet en cours d'étude de faisabilité consiste à refondre les données « Ariane » : pour les lieux et édifices un rapprochement vers la BDTopo avec l'utilisation

de l'Espace Collaboratif (<http://espacecollaboratif.ign.fr>) est envisagé. Pour le filaire routier, Grand Lyon a un intérêt à se synchroniser avec la BDTopo mais de nombreux et importants effets de bord sont encore à investiguer (logiciels, reprise des données, ...).

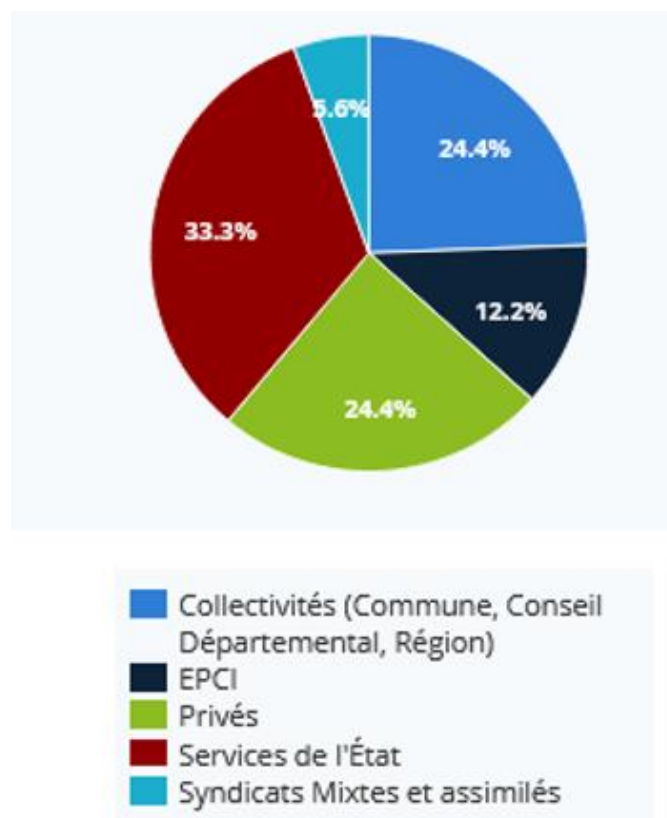
**BMorandot** est étonné de constater la multitude d'articulations/liens entre les entités, les outils.

#### 6. Avancées du sous-groupe "Questionnaire/sondage sur la pratique des ETL dans le contrôle qualité"

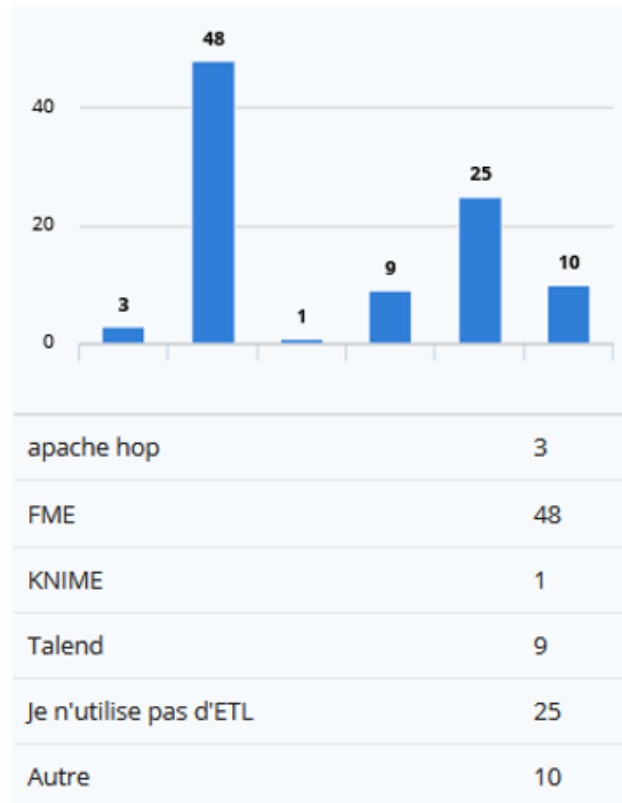
Après quelques réunions de préparation, le questionnaire a été diffusé début avril, notamment sur [Georezo](#) ([permalien](#)), newsletters (CNIG, CRIGE, OpenIG). Une brève Geotribu a été proposée (prévue le 13/06), **Srolle** espère une revue de presse avant la clôture escomptée du sondage fin juin 2025.

A date, 90 répondants avec une représentation significative des profils d'acteurs géomatique (état, collectivités, privé).

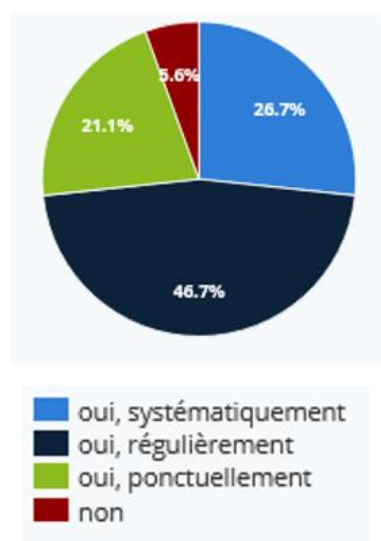
## Typologie des répondants



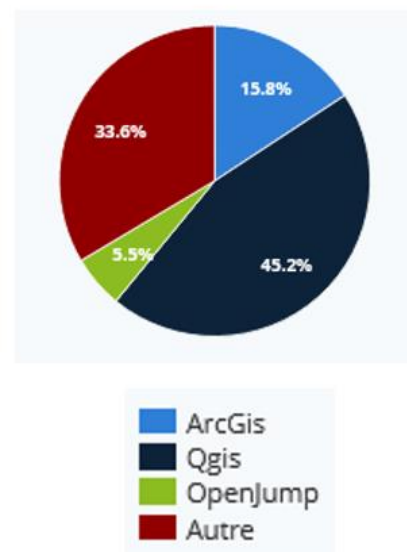
## Quel(s) ETL ?



## Contrôlez-vous les données géographiques ?

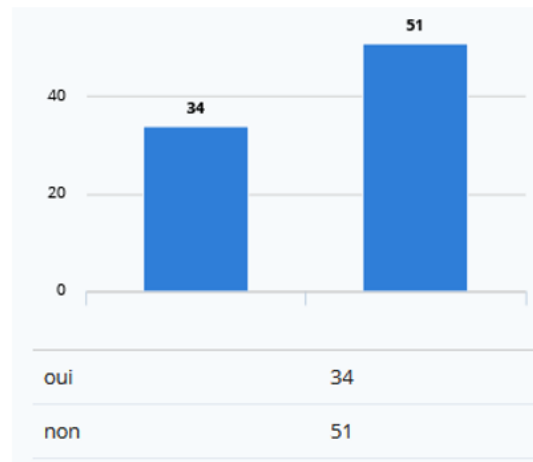


## Quel(s) logiciel(s) pour les contrôles ?





## Connaissez vous la norme ISO 19157 ?



En l'état des réponses, quelques enseignements provisoires (tirés des graphiques présentés) :

- FME est utilisé par la moitié des répondants
- La moitié des répondants contrôlent leur/des données ; de façon « machine » bien davantage qu'au sens ISO19157 (retour terrain pour acquérir le terrain nominal et constater les écarts avec le jeu de données produit)
- Les 2/3 des répondants ne connaissent pas l'ISO19157

A ce stade aucune étude approfondie des résultats n'a été faite, de plus, une coquille en fin de questionnaire a été découverte en cours de route, malgré les contrôle du sous-groupe. Cela nécessitera une petite vérification manuelle mais rien de bloquant pour l'analyse des résultats.

**BMorando** réfléchit à une page du site du CNIG pour lister et décrire les normes (Inspire, OGC, ISO19157, ...) et requiert l'assistance du GT pour l'écriture d'un paragraphe descriptif de l'ISO19157. En l'absence de volontaire, **NPY** proposera une version martyr.

### 7. Exploitation de l'outil OSM Osmose à l'IGN

[Présentation](#) ([permalien](#)) de **NPY** pour **Gabriel Bregand** (alternant IGN, excusé car en séquence scolaire).

Dans l'écosystème OpenStreetMap, l'outil Osmose ([application web](#) et son [permalien](#) ; [wiki \(fr\)](#) et son [permalien](#)) :

- compare la base OSM à des données de référence,
- génère des ponctuels descripteurs des incohérences détectées
- propose aux contributeurs OSM ces incohérences, afin qu'ils éditent/corrigent la base OSM ou rejettent l'incohérence si celle-ci leur semble hors de propos.

Les données de l'IGN gratuites et en licence ouverte telle la BDTopo et notamment son thème routier sont des données de référence ainsi régulièrement comparées à la base OSM, conduisant à la détection des différences entre ces bases de données.

L'automate développé par GBregand est basé sur l'hypothèse qu'une incohérence Osmose rejetée par un contributeur OSM l'est au motif que la base OSM reflète fidèlement la réalité, et donc que l'information IGN a cet endroit est erronée.

Le script consulte quotidiennement l'API Osmose pour récupérer ces suspicions d'erreurs IGN et les propose aux opérateurs IGN via un signalement déposé sur l'[Espace collaboratif \(permalien\)](#).

Globalement, une proportion relative des incohérences IGN/OSM détectées par Osmose est étudiée par un contributeur OSM, conduisant dans l'immense majorité des cas à une édition de la base OSM. Pour 2024, ce sont seulement 150 signalements à destination des agents IGN qui ont ainsi été générés, dont 1/3 ont été rejetés car hors spécifications ou hors de propos.

Même si ces résultats semblent limités, ils démontrent la faisabilité d'une complémentarité entre les données OSM sous licence ODBL (avec clause de repartage, exemple d'explication dans [le guide pratique « Tout savoir sur la licence ODbL » \(permalien\)](#)) et les données en licence ouverte d'un organisme d'état comme l'IGN.

Ce point de complémentarité, allant dans le sens d'une amélioration des données d'OSM comme de l'organisme d'état, est l'objectif du [projet « Public domain map » \(permalien\)](#) du chapitre américain d'OSM.

#### **8. Décision de principe, page QuaDoGéo “Comment faire pour valider vos données” réutiliser dans fabrique des standards**

Lors des travaux de la fabrique des standards, il a été suggéré de créer une page « Vérifier la qualité des données ».

Après discussion, un guide grossier pourrait être proposé, avec par exemple les étapes suivantes accompagnées chacune de quelques paragraphes descriptifs et de liens vers des ressources :

- Quel est le data-product ?
  - Se reposer les questions du pourquoi le jeu de données, de son usage, avec par exemple <https://teamopendata.org/t/traduction-et-adaptation-du-modele-de-description-des-donnees-datasheet-for-datasets/1400>
  - Retrouver ou rédiger l'ontologie/les définitions
  - Retrouver, ou rédiger, les spécifications du jeu de données, avec par exemple le standard des standards CNIG
  - Retrouver ou rédiger les spécifications de saisie
  - Retrouver, ou rédiger les spécifications de niveau de qualité
- Vérification du respect du formalisme
  - Respect du schéma de données
  - Respect du domaine de valeurs de chaque attribut
  - Respect de l'intégrité référentielle
  - Respect des contraintes (unicité, regex, temporalités, ...)
  - Vérification de la géométrie (validité OGC, non rebroussements, non superposition de sommets, dimensions minimales, ...)
- Vérification, par échantillonnage, du respect du niveau de qualité
  - Définition du terrain nominal
  - Acquisition de la données nominales

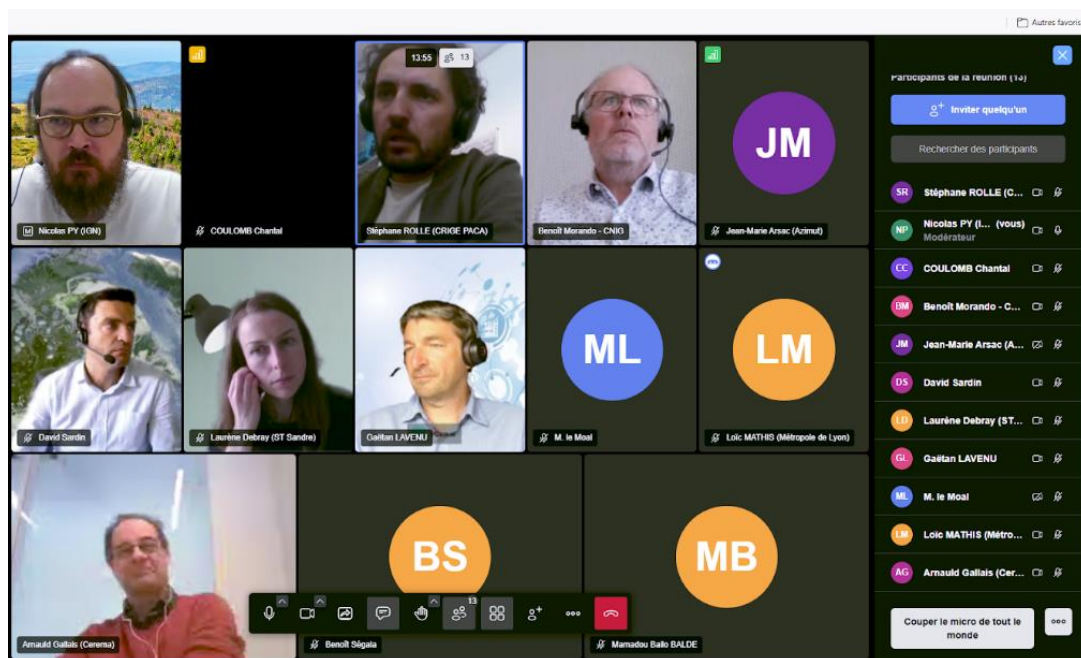
- Constatation des écarts
- Analyse
- Mise à jour de la fiche de métadonnées
  - Descriptif
  - Spécifications
  - Niveau de qualité mesuré

Parmi les ressources, AGallais rappelle celles du MTE tel le diaporama sur les [Principes généraux de qualité des données \(permalien\)](#).

Décision/actions :

**NPY**: proposer une V0 de la page pour par les membres du GT.

**Tous** : session de travail sur cette v0 lors du prochain GT.



Prochaine réunion (distanciel ou présentiel) en octobre 2025 (sondage en cours)