

Compte-rendu de réunion du 4 février 2021

Participants :

Nom Prénom	Organisme	Présent	Excusé
Jean-Marie ARSAC	AZIMUT	x	
Mathieu BECKER	ISO GEO	x	
Jean-Marie BOURGOGNE	OpenDataFrance / Validata		x
Jérôme BOUTET	Idéo Ternum Bourgogne-Franche-Comté		x
Romain BUCHAUT	CRIGE PACA		x
Benjamin CHARTIER	Consultant	x	
Olivier DISSARD	CGDD/DRI/MIG		x
Bénédicte DURAND	CARENE service Topo-Cartographie	x	
Arnaud GALLAIS	Cerema Ouest	x	
Stéphane GARCIA	IGN / Cellule normalisation		x
Guillaume GRECH	UMS Patrinat, OFB - MNHN	x	
Clément JAMET	Métropole de Lyon		x
Marie LAMBOIS	IGN / Cellule normalisation	x	
Mathieu Le Moal	Axes Conseil		x
Hélène LEUFROY	Géo17		x
Jocelyne MARC	IGN / Mission Qualité	x	
Suzanne NICEY	Idéo Ternum Bourgogne-Franche-Comté		x
Nicolas PY	IGN Centre Est	x	
Clémence RABEVOLO	IFREMER	x	
Mathieu RAJERISON	Cerema Méditerranée	x	
Gessica REYNAUD	Géomap - Imagis		x
Stéphane ROLLE	CRIGE PACA	x	
Benoît SEGALA	Consultant	x	
Pierre VERGEZ	IGN / Mission CNIG	x	

Ordre du jour :

- Validation du [précédent compte-rendu](#). Points d'info et d'actu
 - Bilan du GT en commission RMO (*A. Gallais*)
- Veille sur le sujet qualité et retour sur le Q-KEN (*N. Py*)
- Appel à commentaires sur le registre des mesures de la qualité (*A. Gallais*)
- Exploitation des métadonnées de qualité (*M. Becker - Isogeo*)
- Inventaire des métadonnées de qualité sur les portails open data (*N. Py*)
- Point divers

Prochaine réunion : **20 mai 9h30**

Les documents relatifs à cette réunion sont [disponibles ici](#).

1. Validation du précédent CR - Points d'info et d'actu

Le [compte-rendu](#) du précédent GT CNIG QuaDoGéo est relu et validé.

A. Gallais indique que la formation « [Data dans les territoires](#) » a bien été relancée en 2021 avec 8 modules. Les supports et vidéos sont progressivement [déposés ici](#).

Les actions suivantes ont été réalisées :

- transmission à l'AITF du "[Retour du GT CNIG QuaDoGeo concernant les pistes d'évolution des spécifications pour le format BAL](#)" suite aux travaux menés par le groupe dans le cadre du processus de qualification des données adresses ;
- présentation des travaux du GT QuaDoGéo à Eurogéographics 19 nov 2020 (cf. [§2.2](#)) ;
- inventaire des IDG présentant des informations relatives à la qualité des données (cf. [§5](#)).

Les actions suivantes sont reportées au prochain GT CNIG QuaDoGéo :

- abonder la nomenclature des usages génériques de l'information géographique ;
- étudier le GUF de GeoNetwork, et tester la restitution sur une plateforme GeoNetwork ;
- poursuivre le développement de l'outil de génération du résultat graphique conforme à la maquette et adapter le script au flux XML de la norme ISO 19157.

Bilan du GT en commission RMO

A. Gallais fait un retour de la présentation du [bilan du GT QuaDoGéo](#) à la [Commission règles de mise en œuvre](#) (RMO) du CNIG du 10 novembre 2020 (cf. [compte-rendu](#)).

- La Commission RMO validera le registre des mesures de la qualité après appel à commentaires *[qui s'est déroulé depuis, cf [§3](#)]* et elle a validé les perspectives du groupe de travail consistant à :

- rendre les métadonnées sur la qualité externe moins facultatives dans les géocatalogues français ;
- évaluer les géostandards au regard de la thématique qualité ;
- accentuer la veille sur les outils de retour des utilisateurs – utilisation du GUF ;
- poursuivre le développement de l'outil de restitution graphique de la qualification, en ciblant la réalisation de la maquette graphique ;
- en ajoutant qu'il serait intéressant de tester la formation à distance sur la qualité des données géographiques.

Débats :

M. Rajerison souhaiterait relancer le projet de formation à distance QuaDoGéo car elle n'a jamais été réalisée. A. Gallais demande si le format webinaire ne serait pas plus adapté.

L'action d'évaluer les standards au regard de la thématique qualité peut se traduire par :

- une présentation de travaux du GT QuaDoGéo à la prochaine [Comm. Données du CNIG](#)
- l'intégration de spécifications de qualité dans les géostandards se fera au fur et à mesure des nouveaux géostandards ou de leurs mises à jour.

Un premier pas pourra être mis en place avec le futur standard Start DT Elec.

- faire inclure une sélection et hiérarchisation des critères de qualité sous forme de propositions de seuils minimum dans les géostandards. Outre le critère de cohérence logique, ceux-ci participeraient à l'évaluation de la conformité d'un lot de données par rapport à un géostandard.
- définir un modèle de géostandard incluant la qualification des données (cf. ISO 19131) à l'instar de ce que proposent les géostandards COVADIS.

Décision / Actions :

- *Présenter les travaux du GT QuaDoGéo à la prochaine Commission Données (A. Gallais)*

- *Élaborer un modèle relatif à la qualité des données pour les géostandards (N. Py)*
- *Abonder la nomenclature des usages génériques de l'information géographique (M. Lambois)*
- *Étudier le GUF de GeoNetwork et tester la restitution sur une plateforme GeoNetwork (M. Lambois)*
- *Poursuivre le développement de l'outil de restitution graphique conforme à la maquette et adapter le script au flux XML de la norme ISO 19157. (M. Rajerison)*
- *Relancer la FOAD QuaDoGéo (M. Rajerison)*

2. Veille sur le sujet qualification des données géographiques et retour sur le Q-KEN

Par N. Py. La présentation est [disponible ici](#).

Veille sur le sujet qualification des données géographiques

La veille de N. Py s'apparente à une revue de presse des dernières évolutions intéressantes sur le sujet :

- Le [site du SANDRE](#) permet à l'utilisateur de faire un retour au producteur sur la documentation du catalogue d'objets via une évaluation à cinq étoiles. Il lui permet également, pourvu qu'il soit authentifié par un login, de suggérer des améliorations.
- une nouvelle fonctionnalité sur [data.gouv.fr](#) permet de rattacher un jeu de données à un schéma existant. La conformité à un schéma active certaines fonctionnalités.

N. Py remarque que le site "*favorise la diffusion de données y compris si elles contiennent des erreurs, ce qui rend l'utilisateur (averti et courageux) capable d'auditer le respect du schéma*"... approche opposée à celle du validateur du [Géoportail de l'urbanisme](#) qui au contraire bloque toutes données non strictement conforme aux géostandards d'urbanisme.

- Rappel de la transmission à l'AITF du "*Retour du GT CNIG QuaDoGéo concernant les pistes d'évolution des spécifications pour le format BAL*" (cf. §1), à l'occasion de sa [révision au format BAL v1.2](#).

- Retours d'expériences sur le déploiement de l'open-data sur les territoires [Grandlyonnais](#), et [Rennes Métropole](#). Lyon a progressé quant à l'animation du collectif et la mise en place de moyens de contrôle et de validation. Rennes s'est en particulier penchée sur l'amélioration du catalogage et la mise en place de moyens pour diffuser des données qui n'étaient pas encore ouvertes.

- Publication très intéressante de la banque des territoires : [Boîte à outils pour la gestion des données territoriales](#), assortie d'un webinaire.

- N. Py suggère de proposer un "Modèle de document de spécifications" pour les producteurs de données, [cf diapo 12](#). Il s'agirait d'une démarche positionnée très en amont par rapport au sujet de la qualification des données géographique. B. Durand s'y associe volontiers. A. Gallais leur propose de tester la démarche mais émet une réserve par rapport au mandat et aux ressources du groupe de travail.

- le "cadeau de Noël de l'IGN" : la description par l'IGN des [contrôles BDUi et BDTopo et leur validation continue au fil de la production](#). Document à exploiter au sein du GT QuaDoGéo. A. Gallais verrait le cadeau de Noël de l'IGN [principalement là](#).

Décision / Actions :

- *Proposer un "Modèle de document de spécifications" pour les producteurs de données (N.Py et B. Durand)*

Retour sur le Q-KEN

Voir [diapos 16 à 21](#).

Le Quality Knowledge Exchange Network est un forum d'échanges entre instituts cartographiques au sujet du management de la qualité et de la qualité des données.

Le 19 novembre 2020, l'IGN y a présenté les travaux du GT CNIG QuaDoGéo et sa stratégie partenariale de réutilisation des données de producteurs de référence : collectivités, opérateurs, etc.

N. Py souligne les deux interventions allemande et néerlandaise décrivant leur politique de contrôle qualité. L'[institut allemand \(BKG\) a présenté 19 principes d'orientation de l'action décisionnelle](#) décrivant en particulier une chaîne de responsabilité. Les principes exposés font écho à QuaDoGéo : l'importance du contrôle qualité, la nécessaire synchronisation entre les outils et les spécifications, etc. L'[intervention néerlandaise](#) a permis de présenter des modèles de restitution de la donnée suivant quatre critères (completeness, accuracy, currentness, reutilisability) suivant une échelle de qualification classique à 5 étoiles. On retrouve dans le domaine de la restitution graphique des préoccupations communes avec QuaDoGéo.

3. Appel à commentaires sur le registre des mesures de la qualité

Par A. Gallais, cf. [la synthèse des commentaires](#) et [le registre amélioré](#) suite à l'appel à commentaires

L'[appel à commentaires](#) concernant le registre des mesures de la qualité organisé par le CNIG s'est déroulé du 26 novembre au 31 décembre.

Il a suscité 65 commentaires provenant de 4 contributeurs : Benjamin Chartier (1 à 54), SOLURIS (55 à 61), Benoît Segala (62 à 64), CARENE (65). Chacun a été destinataire des réponses apportées à leurs commentaires et ils ont été conviés au présent GT QuaDoGéo, trois y participent. Parmi les 65 commentaires :

- 32 (49%) ont généré des corrections
- 17 (26%) ont correspondu à des besoins d'explications (question / réponse)
- 9 (14%) portaient davantage sur les fiches méthodologiques ou la norme ISO 19157
- 7 (11%) n'ont pas été retenus

Débats :

B.Chartier considère que le terme "registre" n'est pas adapté dans la mesure où il ne présente pas de liste de valeurs, et que la partie introductive mériterait d'être plus détaillée. M. Lambois rappelle que "registre" correspond aux termes anglais "register" et "registry". A. Gallais présente les modifications de la partie introductive, où - suite à la demande de B. Chartier - le registre est décrit comme le "catalogue des mesures préconisées pour la qualification des données géographique".

B. Segala souhaiterait une révision de la fiche méthodologique "[précision de position](#)" afin d'y ajouter une mesure qui sera ensuite cataloguée dans le registre. Il suggère également qu'un item "méthode d'échantillonnage" soit ajouté dans les métadonnées, mais cette proposition n'est pas retenue par le GT QuaDoGéo.

L'appel à commentaires et le registre (version janvier 2021) sont validés par le GT CNIG QuaDoGéo, sous réserve de commentaires ultimes de ses participants d'ici la diffusion du compte-rendu de réunion.

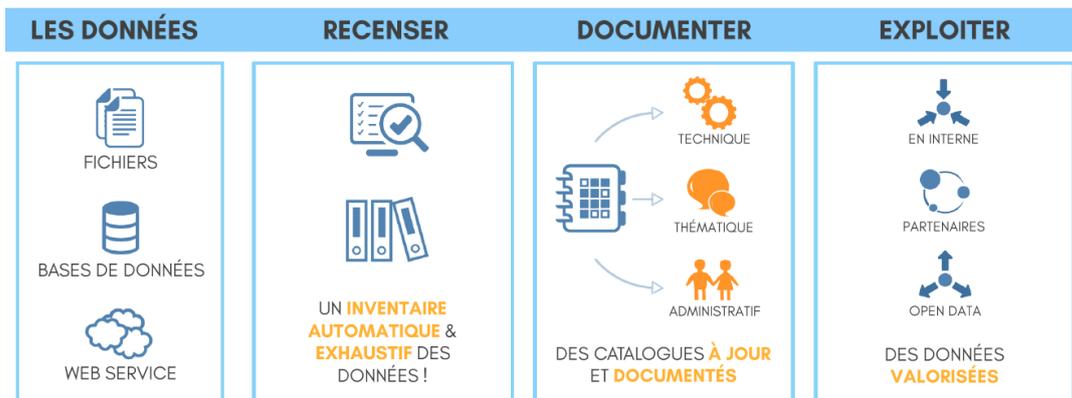
Décision / Actions :

- *Faire actualiser le registre informatique sur le geocatalogue (maintenu par le BRGM) suite aux modifications effectuées sur le document pdf (A. Gallais)*
- *Proposition de révision de la fiche méthodologique "[précision de position](#)" (B. Segala)*

4. Exploitation des métadonnées de qualité

Par M. Becker (Isogeo), la [présentation est ici](#).

Editeur de logiciel, [Isogeo](#) s'est spécialisé dans le catalogage et les métadonnées de données géographiques. M. Becker évoque la complémentarité entre les métadonnées de qualité (bien renseignées et à jour) et les métadonnées sur la qualité des données. Compatible aux normes (191115 - 19139) et basé sur une chaîne de production automatisée des métadonnées, Isogeo propose une exploitation de celles-ci au plus proche des usages, directement dans les logiciels SIG (bureautiques & web) via des plugins et widgets et innove avec des outils couplés aux métadonnées tels qu'un fil de discussion sur chaque donnée sans modération, des data-visualisations interactives et un système de notation par étoiles.



Actuellement Isogeo récupère automatiquement des métadonnées telles que : le nom du fichier, son emplacement, son format, la date de création de la métadonnée, le type et détail des modifications, le nombre d'entités, le type de géométrie, le système de coordonnées, l'enveloppe convexe, le nom et le type de chaque attribut, l'encodage des caractères, l'identifiant unique, la langue des métadonnées et le point de contact.

M. Becker considère que l'intelligence artificielle permettra d'étendre cette liste. Isogeo envisage de faire évoluer son modèle de métadonnées pour prendre en compte des critères de qualité des données, attendus par ses utilisateurs. Il pense pertinent de leur faire saisir une partie de ces critères au plus proche des usages, à savoir directement dans les plugins. Isogeo ambitionne également de faire évoluer son Scan (processus automatisé pour le parcours de répertoires et de bases de données permettant la création et la mise à jour de fiches de métadonnées par donnée rencontrée) pour calculer automatiquement certains critères de qualité comme le taux de remplissage des attributs ou les domaines de valeurs.

D'autre part, M. Becker pense que ces critères qualités devraient pouvoir être utilisés pour afficher des alertes aux utilisateurs, par exemple le dépassement d'un seuil de précision d'une donnée.

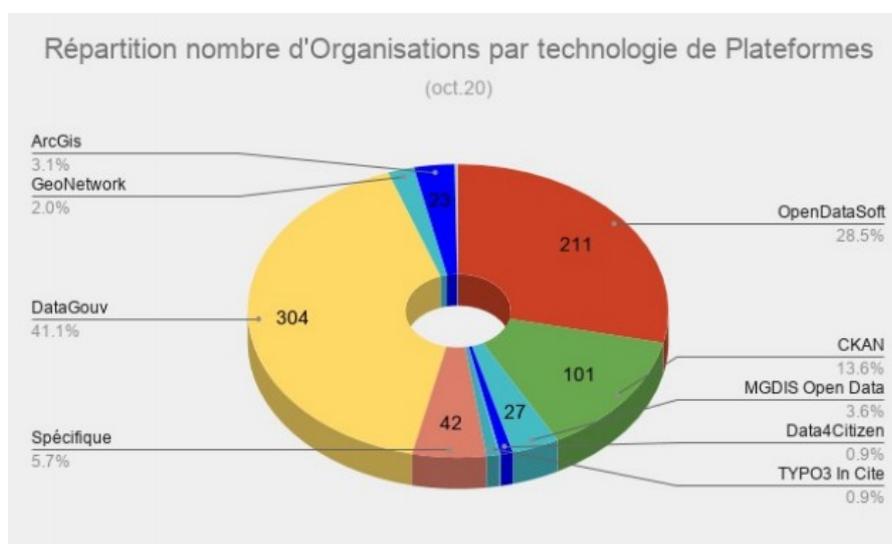
5. Inventaire des plateformes.opendata présentant des informations relatives à la qualité des données

Par N. Py (IGN), la [présentation est ici](#).

N. Py a effectué une revue des métadonnées de qualité sur les lots de données open-data exposés par les infrastructures de données géographiques (IDG) afin d'évaluer dans quelles mesures ces plateformes informent bien l'utilisateur final de la qualité des données.

L'inventaire a pris comme points de départ les [résultats annuels](#) de l'[observatoire OpenDataFrance des territoires](#) ainsi qu'une proposition de L. Haÿ. Il a examiné les données publiées par les collectivités et en adoptant le point de vue de l'internaute. L'observatoire ODF indique une multitude de jeux de données publiés, mais pas toujours de façon hétérogène ni régulièrement actualisés.

N. Py a procédé par échantillonnage de trois à quatre plateformes emblématiques par technologies de plateformes telles que inventoriées par l'observatoire ODF



Il a ainsi choisi divers exemples s'appuyant sur les technologies répertoriées ci-dessus en les évaluant suivant des critères tels que :

- description des jeux de données publiés ;
- présence des balises de métadonnées INSPIRE relatives à la qualité, avec plus ou moins de possibilités de personnalisation du modèle de présentation ;
- outils en backoffice (geosource, GeoNetwork) et standards technologiques adoptés (ex : DCAT standard du WEB alternative à ISO 19115) ;
- suggestion de réutilisations des données et possibilité donnée à l'utilisateur d'en faire un retour ;
- etc.

N.Py rappelle qu'il s'agit d'une observation ponctuelle des métadonnées réellement présentées aux potentiels réutilisateurs, indépendamment des capacités techniques des outils de catalogage utilisés. B. Chartier ajoute qu'outre les capacités techniques des outils, la présence ou non de métadonnées relatives à la qualité des données peut être grandement influencée par l'animation et la gouvernance des plateformes.

N. Py rapporte un inventaire de la Direction Territoriale Grand-Ouest de l'IGN focalisé sur l'inter-région Bretagne Pays-de-la-Loire et Normandie. Les principales plateformes y ont été étudiées, soit une cinquantaine de plateformes exposant 16 000 jeux de données. L'examen attentif a révélé :

- nombre de lots de données s'avèrent redondants entre les différents échelons

territoriaux ;

- le nombre de jeux de données n'augmente plus voire a diminué en 2020 (année particulière il est vrai..);
- l'information essentielle sur l'actualité des lots de données n'est pas toujours fournie, à l'exception notable de GéoBretagne.

Enfin, cet inventaire révèle une fracture importante entre les espaces urbains et ruraux pour ce qui concerne la disponibilité d'informations géographiques : tandis que les agglomérations et communautés de communes importantes mobilisent plus du tiers du nombre de jeux de données (pour une emprise d'un trentième seulement de la surface de la zone d'étude), certains départements ruraux s'avèrent pauvres voire absents en termes de diffusion de données géographiques ouvertes.

En conclusion, l'examen des plateformes open-data présentant des informations relatives à la qualité des données révèle :

- la rareté des informations de qualité externe, très majoritairement facultatives ;
- des rubriques "réutilisation des données" apparaissant comme des coquilles vides de contenus ;
- l'absence de mise à disposition par les producteurs de données de descriptifs de livraison, de contenu ou de qualité interne ;
- pour le moment 😊, aucune fourniture de métadonnées sur la qualité interne des lots de données (critères ISO 19157 : exhaustivité, précision de position, précision thématique, etc.), sauf des dates d'actualisation, et parfois des indications de fréquence d'actualisation (non respectée).

Si certaines plateformes présentent de réels exemples à suivre (CRIGE PACA, GéoBretagne, Lorient Agglomération, etc...) la plupart d'entre elles demeurent actuellement des vitrines quant à la fourniture de données géographiques ouvertes.

Décision / Actions :

- A. Gallais remercie M. Becker et N. Py pour leurs présentations et suggère à P. Vergez et M. Lambois qu'elles soient également présentées au [GT CNIG Métadonnées](#).

A ce stade, la production effective ou en cours du [GT CNIG QuaDoGéo](#) comprend :

- Les fiches méthodologiques du Cerema : sur le [site du Cerema](#) ou en accès individuels : [Introduction](#) - [généralités](#) - [contexte du contrôle qualité](#) - [éléments statistiques](#) - [méthodes d'échantillonnage](#) - [mode de représentation](#) - [cohérence logique](#) - [exhaustivité](#) - [précision thématique](#) - [précision de position](#) - [qualité temporelle](#) ;
- Le [registre français de mesures de qualité des données géographiques](#) et le [registre en ligne \(version initiale avant appel à commentaires\)](#) sur le Géocatalogue ;
- La [carte mentale](#) au sujet de la qualification des données suivant la norme ISO 19157 ;
- Deux pistes exploratoires de qualification de données (synopsis, logigramme) sur des thèmes particuliers ;
- La [méthodologie](#) de qualification de données (*à consolider...*), comprenant :
 - une nomenclature des usages génériques de l'information géographique ;
 - une maquette de restitution de la qualité des données ;
- Une [présentation du GUF](#), outil de retour utilisateur quant à la qualité des données.
- (*en cours*) Le prototype d'un [outil de restitution graphique de la qualification](#) des données géographiques