

 <p>RÉPUBLIQUE FRANÇAISE CNIG <i>Liberté Égalité Fraternité</i></p> <p>Conseil national de l'information géolocalisée</p> <p>Commission règles & qualité</p> <p>Groupe de Travail Qualité des Données Géographiques (QuaDoGéo)</p>	<p>Groupe de Travail animé par :</p>  <p>Crige Centre de Ressources en Information Géographique Provence-Alpes-Côte d'Azur</p>  <p>IGN INSTITUT NATIONAL DE L'INFORMATION GÉOGRAPHIQUE ET FORESTIÈRE</p>
<p>COMPTE-RENDU DE REUNION (distanciel)</p>	
<p>Objet : Réunion #20 du GT QuaDoGéo du 14 octobre 2024.</p>	

Ordre du jour :

1. Accueil et rappels
2. Validation du [précédent compte-rendu](#)
3. Intervention 1 : La qualité des données géographiques au SHOM - M. Durupt (SHOM)
4. Date du prochain GT
5. Actualités diverses
6. Intervention 2 : Impact des représentations cartographiques de l'incertitude pour la prise de décision - Léa Courteille (INRAE)
7. Intervention 3 : Visualisation et incertitudes - J. Gautier (IGN Lastig)

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	X	
Muriel FOULONNEAU	DGFIP	X	
Marion LACROIX	IGN	X	
Benoît MORANDO	CNIG	X	
Matthieu LE MOAL	AXES conseil	X	
Nicolas PY	IGN	X	
Stéphane ROLLE	CRIGE Paca	X	
Benoit SÉGALA	Consultant	X	
David VIGLIETTI	DIRCO	X	
Léa COURTEILLE (invitée)	INRAE	X	
Mélanie DURUPT (invitée)	Shom	X	
Jacques GAUTIER (invité)	IGN Lastig	X	

Réunion #20 du GT QuaDoGéo du 14 octobre 2024 (distanciel).		
Mise en relecture du compte-rendu	25/10/2024	Stéphane Rolle / Nicolas Py
Relecture du compte-rendu	06/11/2024	Les participants.
Validation du compte-rendu	05/12//2024	Stéphane Rolle / Nicolas Py
Prochain rendez-vous le 13 janvier 2025 en distanciel https://webconf.numerique.gouv.fr/GTCNIGQuaDoGeo2025 .		

1. Accueil et rappels

RAS

2. Validation du précédent compte-rendu

Le compte rendu de la réunion des 14 et 15 mai 2024 (#19, 2024-2) [a été diffusé sur la page du GT \(Permalien\)](#) en omettant d'en demander une relecture collective préalable. N'hésitez pas à nous proposer vos amendements au besoin.

3. Intervention 1 : La qualité des données géographiques au SHOM - M. Durupt (Shom)

Support de la présentation [lien \(permalien\)](#).

Mélanie DURUPT (Shom ; département géomatique) présente le processus de production de l'information géographique au Shom – Service hydrographique national – en insistant sur ses aspects qualité des données géographiques.

Opérateur public, le Shom produit l'information géographique maritime et littorale de référence via ses missions d'hydrographie nationale (assurer la sécurité de la navigation), de soutien à la défense, mais aussi de soutien aux politiques publiques de la mer et du littoral (par exemple, réalisation des levés sur les zones identifiées pour la construction de champs éoliens en mer).

Parmi les produits du Shom, nous pouvons citer bien évidemment les cartes marines – sous forme papier mais aussi électronique (ENC) – qui sont normées par l'OHI (Organisation Hydrographique Internationale), par exemple sur le style, l'intégrité, le contenu – normes S-4 pour les cartes papier, S-57 et très bientôt S-100/S-101 pour les ENC. Ces produits concourent à assurer la sécurité des navigateurs (plaisancier, pêcheur, paquebot, tanker, Marine...).

La présentation expose en particulier la BDGS, pour Base de données générales du Shom , l'équivalent de la BDUi IGN. La BDGS est basée sur le modèle S-57 (<https://iho.int/fr/bulletins-de-la-s-57>, [permalien](#)) étendu d'objets et d'attributs créés par le Shom. Elle est gérée par le progiciel CARIS HPD <https://www.teledynecaris.com/en/products/hpd/> ([permalien](#)) basé sur Oracle.

L'information pérenne y est entretenue en continu. Une information urgente (par exemple des containers tombés à l'eau), est diffusée sous forme d'un avis urgent aux navigateurs, en lien avec les préfetures maritimes, mais elle n'est pas forcément intégrée en BDGS.

Contenant des données de référence pour la réalisation des produits du Shom, le maintien d'un haut niveau de qualité des données est primordial et est assuré à tout moment afin que l'information qualifiée puisse être diffusée en continu (les cartes sont corrigées toutes les semaines). Au niveau de la BDGS, CARIS HPD fournit plusieurs moyens :

- au clic de l'opérateur, dans le logiciel de saisie (les attributs obligatoires doivent être renseignés sinon une erreur remonte, seules les valeurs autorisées peuvent être saisies) ;
- par le lancement de contrôles automatisés (QC tests : certains sont natifs dans le logiciel, d'autres sont paramétrables) ;

- par le mécanisme de certification qui est une vérification croisée à 2 opérateurs. L'objet (inséré/modifié/supprimé) est flagué « under_construction » puis « not_verified » par le gestionnaire; le contrôleur étudie la modification et la passe en « verified » (disponible pour report sur les produits). Inspection régulière / supervision de ces statuts [construction ; not_verified ; verified].

Les systèmes clients de la BDGS ne peuvent utiliser que la dernière version validée des objets : seule l'information qualifiée peut être reportée sur les produits.

Les gestionnaires disposent de règles de saisie thématiques pour les aider à correctement encoder les objets et complètent ou précisent les principes énoncés dans la norme S-57 (« doit », « peut », « pourrait », « devrait »).

Enfin, un workflow externe à CARIS permet de faire circuler l'information d'un service à l'autre, de la réception au produit final, en passant par l'intégration en base de données.

Plusieurs questions ont été posées :

Une mission du Shom est de "prévoir l'évolution de l'environnement physique marin", est-ce de la cartographie prédictive ?

Il s'agit en fait plutôt des prédictions de marée, de la modélisation des états de mer (hauteur des vagues), de la vigilance vague submersion ; diffusées dans la catégorie « prévisions océanographiques » de data.shom.fr.

Il semble que le Shom soit producteur de données souveraines ("produisant un effet régalien") ; comment est-ce géré par le Shom (contentieux, juridique) ?

Le Shom opère pour le compte du SGMer le portail des limites maritimes officielles françaises (<https://limitesmaritimes.gouv.fr/>, [permalien](#)), diffusant les limites en mer, qui s'appuient sur un texte réglementaire. Il y a parfois des saisines (à tel endroit, un utilisateur a-t-il tel ou tel droit d'activité ?) mais a priori pas d'autre incidence.

Chaque signalement / donnée externe est intégrée unitairement mais n'est pas systématiquement double-vérifiée (le niveau d'investigation est fonction du niveau de confiance en le partenaire).

Le méta modèle est, de par l'adhérence à la norme s-57, embarqué (pour le périmètre s-57) dans CARIS. Cela induit qu'une bonne partie des contrôles (attributs obligatoires, liste de contrôle QC Tests) est disponible sur étagère. La cohérence entre attributs et les contrôles sur les classes/attributs spécifiques au Shom nécessitent des paramétrages ad-hoc.

La transition vers les systèmes de navigation électronique s'accélère aujourd'hui, y compris pour la marine nationale, reléguant progressivement la carte marine papier au second plan (système de secours pour certaines unités). Les cartes papier doivent être entretenues par leur propriétaire (report des corrections, parfois sous forme de patches).

Il n'y a pas de gros problèmes d'identifié sur la chaîne de production qui fonctionne depuis maintenant 10 ans. En évolution, certaines classes d'objets ne sont pas très vivantes ni entretenues, ce qui pose des questionnements d'actualité ou d'exhaustivité. Elles seront peut-être retirées de la BDGS (présumée, de par sa mise à jour et diffusion en continu, en permanence de niveau de qualité nominal) et gérées ailleurs que dans la BDGS.

4. Date du prochain GT

La prochaine réunion est proposée entre le 13 et le 24 janvier 2025 ; pas le 14, 15 ni le 16 pour **Srolle**.

Décision/actions :

Décision: La date du 13 janvier 2025 après midi est retenue.

5. Actualités diverses

ETL en voie de disparition, quelles solutions?

Les tarifs de FME explosent, en acquisition comme en maintenance ; bien que ses qualités de performance, fiabilité, capacités de lecture/écriture de formats soient reconnues.

Jusqu'à présent la société Hitachi qui avait repris l'ETL Kettle diffusait une version Open source de cet outil (Pentaho Data Integration : PDI) assez pratique pour l'intégration/transformation de données de toutes natures.

À partir de la version 10.2 de ce logiciel, Hitachi met un terme au développement de la version Open Source de PDI pour un usage de production (l'usage gratuit pour les étudiants resterait possible) et un nouveau système de licence (payante) va être mis en place.

Une licence plus légère va être proposée aux utilisateurs qui avaient plutôt tendance à utiliser la version Open Source et n'avaient pas besoin de l'intégralité des fonctionnalités, ni d'assistance.

L'outil d'ETL serait dans cette version 'légère', cependant payante.

Dans le cadre du marché de support logiciel libre au bénéfice de l'administration française piloté par la DGFIP, l'Adullact ([lien](#), [permalien](#)) dispose d'une forge ([lien](#), [permalien](#)) dans laquelle sont centralisés les rapports ([lien](#), [permalien](#)).

Début 2024, la veille « Qualité des données » a été publiée ([lien](#), [permalien](#)).

Six solutions propriétaires et cinq solutions libres sont étudiées. L'offre propriétaire apparaît dominée par Informatica Cloud Data Quality qui est une solution très complète et puissante toutefois il s'agit d'un service cloud ne pouvant répondre aux contraintes de souveraineté de certains.

L'alternative libre la plus complète est certainement Knime couvrant l'essentiel du besoin avec une interface graphique de programmation nécessitant une certaine technicité.

À noter que Talend Data Quality n'est plus libre depuis le début 2024.

Côté MTE, une étude a été réalisée en 2018 qui concluait (notamment sur des arguments budgétaires) sur l'utilisation du modèleur QGIS. L'étude est disponible pour ceux qui accèdent au RIE sous https://geoinformations.metier.e2.rie.gouv.fr/etude-amont-sur-les-etl-a3611.html?id_rub=508.

On ne lancerait pas une enquête ?

Le GT QuaDoGéo fait le constat d'une relative faible connaissance de la gestion de la qualité dans les organisations.

Indépendamment, le cas FME évoqué ci-dessus pose des questions d'outillage.

C'est donc peut être un moment où il serait souhaitable de lancer une enquête (fichier destinataire du CNIG, des CRIGE, ...), sur plusieurs types de questions :

- Est-ce que la qualité est un sujet dans votre structure ? Sur vos achat de données, sur vos productions en propre ? Ca se traduit comment ? Vous pourriez venir nous en parler ?
- Est-ce que vous auriez besoin d'outils / scripts ? Sous quel logiciel, de quelle nature ?

JMArsac est favorable à une telle enquête, il souligne que lors du dernier code sprint Qgis il y a eu des discussions et échanges sur les géotraitements de contrôle qualité.

MLeMoal y est favorable également, il annonce que nombre de ses clients et prospects basculent sur Qgis ou SGBD.

Mlacroix s'inscrit dans ces avis et annonce être disponible pour contribuer.

Décision/actions :

Décision: **SRolle** et **NPY** vont rédiger une proposition d'action mieux cadrée afin d'en faire un appel à volontaires parmi les participants du GT pour y travailler.

GT CNIG Adresse - réflexion sur la partie qualité

Sur le forum <https://forum.openstreetmap.fr/t/quelles-adresses-de-la-ban-integrer-dans-osm/24158>, les contributeurs OSM discutent de la qualité des adresses, notamment sous le prisme de ce qu'est une adresse certifiée. Cela fait suite à l'appel à commentaire du standard adresse, en cours de validation (la commission des standards du 03/10/2024 a validé le standard Adresse V1 sous réserve des reprises éditoriales), pour lequel une validation définitive et sa diffusion sont prévus en décembre 2024.

Convergence des participants sur la pertinence de rappeler, en commission, l'existence du GT QuaDoGéo et que la spécification de niveau de qualité ou la réalisation de mesures du niveau de qualité sur le thème adresse sont souhaitables !

Journée Qgis 2025 - Avignon 10,11,12 juin 2024 université d'Avignon

QGIS est un outil de saisie, mais peut être un outil de contrôle qualité ; le GT QuaDoGéo encourage évidemment l'emploi de ce logiciel libre.

10/06/2024 : ateliers

11/06/2024 : conférences

12/06/2024 : contributions, incluant code sprint, test, traduction

Exemple d'emploi, en 2024, de Qgis comme outil de contrôle qualité de l'Occupation du sol <https://conf.qgis.osgeo.fr/pretaix/qgis-french-users-days-2024/qgis-french-users-days-2024/talk/78AFNA/> ([permalien](#)).

Nouveautés sur schema.data.gouv.fr

Annonce <https://teamopendata.org/t/nouveautes-sur-schema-data-gouv-fr/4466/1> ([permalien](#)) un peu datée car de mars 2024.

L'équipe de data.gouv.fr vient d'ajouter sur schema.data.gouv.fr:

- Une page Actualités 14, qui recense toutes les modifications des schémas (nouveautés, montées de version, suppressions) : <https://schema.data.gouv.fr/actualites.html> ([permalien](#))

- Des flux RSS, pour être alerté dès qu'un schéma est mis à jour :
 - le flux global qui référence les changements de tous les schémas, accessible depuis le header du site, <https://schema.data.gouv.fr/rss/global.xml> ([permalien](#))
 - un flux par schéma (par exemple ici https://schema.data.gouv.fr/rss/gip-inclusion_data-inclusion-schema.xml ([permalien](#))) si vous souhaitez ne suivre qu'une sélection de schémas, accessibles via le bouton RSS sur chaque page de schéma
 -

A coupler, par exemple, à l'extension <https://nodetics.com/feedbro/> ([permalien](#)) pour le navigateur Firefox.

Note de lecture, « Une norme d'échange pour alimenter des référentiels et en assurer la qualité »

Article ([lien](#), [permalien](#)) notamment signé de Bertrand Dubrulle, adjoint au directeur de la Gestion de la Donnée (DGD), DSI, Caisse nationale de l'assurance vieillesse (Cnav) sur

La Cnav a défini et appliqué les principes des "normes d'échange" pour à la fois fournir une description claire et détaillée de l'échange de données et contrôler automatiquement les flux, afin de garantir la qualité de l'alimentation des référentiels.

[...]

Les données doivent être accessibles, échangées uniquement entre acteurs autorisés et seulement sur le périmètre correspondant à leur besoin d'utilisation. Les échanges sont encadrés par une convention de service fixant les objectifs et modalités (fondements juridiques, nature, modalités de transmission et de conservation des données, sécurité et traçabilité des échanges, échéances à respecter, règles de protection des données, conditions financières et modalités de suivi de la convention).

[...]

le principe de norme d'échange pour caractériser toutes les dimensions des flux de données est formalisé comme suit : une structure fonctionnelle et technique du message avec ; une cinématique de l'échange, c'est-à-dire l'enchaînement des différentes étapes de l'échange ; matérialisés par des livrables : la documentation de la norme ; des services de contrôle automatique ; un outil d' « auto-contrôle » automatique pour l'émetteur ; une gestion des versions des différents livrables : ce point est essentiel dans un contexte où la norme bouge beaucoup, notamment, en raison de contraintes réglementaires.

[...]

Les contrôles fondés sur la structure sont appelés *contrôles syntaxiques*, puisqu'ils s'appuient sur la description d'une syntaxe, à la fois pour le message et pour chaque rubrique prise isolément.

[...]

les contrôles sémantiques : cohérence inter-rubriques Cependant, les contrôles syntaxiques ne suffisent pas. Il faut assurer la cohérence interne du message. Ainsi, il existe au sein de chaque norme d'autres contrôles, qui matérialisent ce besoin de cohérence entre rubriques au sein d'un même message.

[...]

La mise en œuvre via une cinématique d'ensemble des contrôles

[...]

L'outillage, la mise en œuvre sur plusieurs années

6. Intervention 2 : Impact des représentations cartographiques de l'incertitude pour la prise de décision - Léa Courteille (INRAE)

Support de la présentation [lien \(permalien\)](#).

Léa est en 3ème année de thèse à l'INRAE, dans un laboratoire d'étude des sols sous la direction de P.Lagacherie (<https://theses.fr/083547878>).

Dans les entretiens menés pour l'étude comparative sur la meilleure manière d'apporter l'information d'incertitude en sus de l'information, Léa constate 2 faits éminemment perturbants pour les sympathisants au GT :

- la majorité des utilisateurs d'information de sol télécharge la donnée sans sa carte compagnon d'incertitude ;
- et ceux qui téléchargent la donnée et son incertitude n'utilisent in fine pas l'incertitude !

L'idéal est alors de fournir simultanément l'information et son incertitude, afin d'obliger le lecteur à considérer l'incertitude.

Il est de plus souhaitable de rechercher la bonne échelle de rendu (plus la résolution est élevée, pire est l'incertitude), où les entités sont alors de surfaces suffisantes pour être d'une incertitude maîtrisée/acceptable.

De nombreuses questions ont été posées, sur les connaissances en sciences cognitives, sur la capacité, en général, à raisonner à partir de données incertaines.

Il existe beaucoup de littérature sur le sujet, assez peu sur l'incertitude spatialisée, encore moins sur l'incertitude associée à des données environnementales. De plus, les résultats des études/cas d'étude sont souvent contradictoires.

7. Intervention 3 : Visualisation et incertitudes - J. Gautier (IGN Lastig)

Support de la présentation [lien \(permalien\)](#).

Pas de remarque notable, présentation pédagogique débouchant sur des exemples de représentation de l'incertitude, actuels et exploratoires. Rappels de la possibilité, à l'heure de la dataviz numérique, que celle-ci peut être augmentée: sons, carte animée, ...

<p>Prochaine réunion (distanciel, https://webconf.numerique.gouv.fr/GTCNIGQuaDoGeo2025) le 13 janvier 2025.</p>
--