



COMPTE RENDU DE RÉUNION	
Rédacteur : A. BRUGERON & J.B. PAROISSIEN	Entité : D3E/EVE
Projet : BDLISA 2016	Numéro : AP16D3E110
Objet : Réunion d'échanges pour l'optimisation de la production/diffusion de la BDLISA	
Date : 05/02/2016	Lieu : Orléans + visio
Participants : Régis HAUBOURG (AEAG), David VIGLIETTI (SANDRE), Philippe DANIELS (BRGM/DSI/), Robin QUIQUE (BRGM/DSI/), Olivier SEDAN-MIGEMOLLE (D3E/EVE), Jean-Baptiste PAROISSIEN (BRGM/D3E/EVE) et Alexandre BRUGERON (BRGM/D3E/EVE)	
Absents : Béatrice ULVOAS (SANDRE), Laurent BRETON (ONEMA)	
Diffusion interne : Laurence CHERY (BRGM/D3E/EVE), Laurence GOURCY (BRGM/D3E/DIR)	
Diffusion externe : Céline NOWAK, Laurent COUDERCY (ONEMA) + participants/absents externes	
En cas de diffusion externe visa et nom du responsable : Laurence CHERY	
	

RÉSUMÉ ET CONCLUSIONS
<p>1 – Support de présentation</p> <div style="text-align: center;">  <p>Reunion_Optimisation _BDLISA_05022016.pc</p> </div>
<p>2 – Partie 1 : Echanges sur la phase de production de la BDLISA</p> <p><u>Complexité des polygones</u></p> <p>Tout le monde s'accorde à dire que la complexité de certains polygones (très grosse densité de vertex pour certains) n'est pas compatible avec l'échelle de représentativité maxi de la BDLISA (1/50 000). Cette complexité pourrait s'expliquer pour les parties à l'affleurement par le fait que la délimitation des EH se base sur les cartes géologiques au 1/50 000 ? Mais ce n'est pas garanti puisque ces cartes ont probablement fait l'objet de nettoyages avant diffusion.</p> <p>Par contre, il est certain qu'une limite d'une entité située sous couverture de 20 entités ne</p>

devrait pas présenter autant de vertex. En terme « métier », en règle générale, ces limites sont plus supposées qu'avérées et donc devraient être digitalisées avec moins de vertex.

Ceci accroît très probablement le volume des données du référentiel, et donc alourdit une base déjà bien lourde. Une amélioration importante consisterait à simplifier, dans les premières étapes de la production de la BDLISA, l'ensemble des polygones. Reste que le « curseur » de simplification (un vertex tous les X m ?) devra être préalablement défini.

Proposition :

→ Dans un premier temps, le BRGM se propose d'ajouter cette étape en début de processus, à savoir sur la couche des entités de niveau 3 qui sert de couche « référence » pour l'élaboration de l'ensemble des autres couches fournies. Des échanges doivent également être menés au niveau des régions pour que cette simplification n'affecte pas la cohérence thématique du découpage.

→ Dans un second temps, pour que ce travail puisse perdurer et rester cohérent, des discussions doivent être lancées en région pour que la simplification puisse être réalisée au plus tôt dans le processus de construction de la BDLISA. Une note de recommandations techniques pourrait par exemple être rédigée.

Tests sur la conformité des attributs

Pour la V1, des contrôles topologiques ont été menés avant envoi au SANDRE (même si des erreurs subsistent toujours apparemment, cf ci-dessous).

Reste que le SANDRE mène également des analyses sur les attributs renseignés, en s'appuyant sur les nomenclatures SANDRE et les dictionnaires de données associés.

Même si des vérifications étaient déjà menées « manuellement » sur certains champs, il semble nécessaire de mener ce type de vérifications de manière plus approfondie en les automatisant et en les incluant dans le processus de génération de la BDLISA.

Proposition :

→ Des échanges BRGM/SANDRE ont été amorcés dans ce sens afin de connaître en détail les analyses menées par le SANDRE et de les transposer plus en amont (pour anticiper et abaisser le nombre de non-conformités). Le SANDRE est intéressé par cette démarche et insiste sur la nécessité d'accorder les tests (requêtes, versions des logiciels de contrôle de conformité au modèle) entre les deux structures pour faciliter les échanges.

Contrôles topologiques en amont

Plusieurs utilisateurs ont fait remonter le fait que certaines couches de la base de données présentaient un grand nombre d'erreurs topologiques :

- Philippe Daniels (administrateur de base de données côté BRGM) lors de la transposition de la BDLISA version SIG en BDLISA version base de diffusion, pour la production des services web,
- le SANDRE à l'issue de leur analyse de cohérence sur la donnée et
- Régis Haubourg (administrateur de données géographiques) de l'AEAG après qu'il ait récupéré et travaillé sur la base afin de l'intégrer dans le SI de l'agence.

Il est intéressant de noter que ces trois contributeurs ont rencontré ces problèmes de topologie lors de la transcription de la base BDLISA au format diffusé (shp, gdb ou mif/mid) en une base sous postGIS.

Les contrôles topologiques réalisés sous ArcGIS lors de la production de la BDLISA ne détectaient pas ces erreurs. Une analyse bibliographique sur ce point révèle que la nature des erreurs topologiques peut être variable d'un logiciel à l'autre et qu'il y a notamment une

différence notable dans la gestion du problème « d'auto-intersection des polygones » entre ArcGIS et PostGIS.

Philippe Daniels et Robin Quique font remarquer que pour le cas de la base de diffusion de la BDLISA, qui est sous PostGIS et qui alimente les services web de la BDLISA, il a fallu faire des corrections topologiques et un travail de simplification des polygones pour pouvoir utiliser cette base, et que par conséquent, il se peut que certaines corrections menées engendrent une légère différence en terme de contours pour quelques entités hydrogéologiques (EH) avec la donnée « référentiel » fournie via le site du SANDRE.

Olivier Sedan rajoute que la BDLISA est initialement construite sur ArcInfo. La gestion des polygones par ce logiciel lui assure une topologie très rigoureuse et il est impossible qu'il puisse générer les erreurs précédemment citées. A priori donc, les erreurs de topologie détectées par PostGIS proviennent des étapes d'éclatement de la BDLISA en plusieurs couches.

Propositions :

Plusieurs propositions sont soumises pour comprendre et solutionner les problèmes de topologie :

→ Appréhender l'origine des erreurs topologiques :

- Réaliser les tests de topologies PostGIS en intégrant directement la BDLISA (seulement NV3) en sortie d'ArcInfo vers une Gdb, un shp puis vers PostGIS.

→ Pour corriger les erreurs topologiques des différentes couches BDLISA :

- Une fois la BDLISA produite, la transposer sous postGIS,
- Réaliser les tests topologiques du SANDRE sous cette base de travail postGIS, si possible sous la même architecture que celle du SANDRE (même version postgresql/postgis)
- Corriger les erreurs détectées sous PosGIS
- Retransposer la base corrigée de postGIS à ArcGIS et contrôler aussi que les corrections apportées n'aient pas trop modifié l'aspect des EH concernées
- Faire les contrôles de topologie sous ArcGIS

Comme pour les contrôles attributaires, ces tests devront être reproductibles et les logiciels et versions utilisés devront être précisés pour faciliter les échanges avec le SANDRE. En théorie, ce processus devrait permettre de diminuer fortement le nombre de problèmes de topologie rencontrés jusqu'à présent.

Identifiant unique – clé métier pour chaque polygone de découpage des entités

Le besoin d'un champ clé unique pour chaque polygone individualisant les piles d'entités semble avéré et permettrait ainsi leur indexation.

Cette clef permettrait notamment de « pister » les erreurs topologiques et d'identifier leur origine.

3 – Partie 2 : Echanges sur la phase de livraison de la BDLISA

« Eclatement » de la BDLISA

Jusqu'à présent, la BDLISA livrée est composée de plusieurs couches distinguant les niveaux de visualisation (niveau national, régional et local), les extensions et la subdivision

selon les superpositions. La taille de la base de données complète actuellement livrée est de l'ordre de 8Go pour le format gdb, ramenée à 6Go zippée.

De plus, le processus de génération de toutes ces couches fait que plusieurs phases d'éclatement des contours et de rassemblements se succèdent, et l'on soupçonne que ceci puisse engendrer une bonne part des erreurs de topologie détectées.

Foncièrement, la donnée que l'on peut considérer « de référence » pour la BDLISA pourrait se limiter à une base beaucoup moins lourde qui serait composée :

- De la **couche des polygones correspondant au découpage des entités de niveau 3** selon les ordres de superpositions des entités entre elles ;
- du **tableau multi-échelles** reliant les EH de niveaux 1, 2 et 3 et les positionnant en âge absolu les unes par rapport aux autres.

Ce sont les deux seules données source utilisées en interne BRGM pour produire la totalité des couches de la BDLISA.

Proposition :

→ La question se pose alors s'il ne serait pas plus pertinent et optimal de :

- Restreindre le contenu du référentiel BDLISA à ces deux seules couches (avec en plus des tables des lexiques SANDRE), qui feraient dès lors l'objet des vérifications approfondies sur les attributs et la géométrie et seraient diffusées via le site du SANDRE. La taille du zip de cette base « référentiel » est estimée à moins de 2Go.
- Proposer une base de données estampillée « usages », une sorte de base « décisionnelle », contenant les produits dérivés de la base « référentiel », avec notamment les déclinaisons des extensions et ordres par niveau, mais également, pourquoi pas, des extractions pré-calculées de la BDLISA par bassin, par région....etc... Cette base s'appuierait sur la base « référentiel » pour faire ces déclinaisons mais ne ferait pas l'objet d'une analyse de validation par le SANDRE (une procédure de validation en interne similaire au SANDRE pourra être réalisée). A l'image de la base de diffusion PostGis actuelle, elle alimenterait toutes sortes de services web les actuels et ceux qui restent à développer (cf échanges partie 3).

C'est une remise en question du mode de diffusion actuel du référentiel, le sujet devra être discuté lors du prochain GT national.

Les formats d'utilisation de la BDLISA

L'enquête en ligne sur les usages de la BDLISA a pu montrer que moins de 3% des utilisateurs ayant répondu utilisent les formats Mif/Mid (pour 90% utilisant le format shp). La question se pose de la pertinence de continuer à produire ce format de moins en moins utilisé. D'autant plus que les formats Mif/Mid ont une gestion de la topologie qui présentent un risque.

A contrario, il semblerait intéressant de décliner la BDLISA aux formats libres SQLITE ou GeoPackage (.gpkg) lisibles par QGIS. Le dernier cité est d'ailleurs le seul fichier d'échange normalisé à ce jour (standards OGC).

Ce point devra être discuté lors du prochain GT national.

Possibilité de diminuer la taille des zip déposés sur le site du SANDRE

David Viglietti du SANDRE indique qu'il leur est possible de scinder chaque zip en 6 si besoin de réduire la taille des fichiers à télécharger.

Cela permettrait de ne plus subir les coupures de téléchargement au bout de 1 ou 2 Go téléchargé (car potentiellement on téléchargerait des fichiers de 200 à 400 Mo environ) mais obligerait l'utilisateur à télécharger une donnée en plusieurs clics ; et surtout, quid de l'incidence de ce tronçonnages sur la donnée contenue dans les zips ? Des tests de faisabilité devront être menés.

4 – Partie 3 : Echanges sur la phase de diffusion/valorisation de la BDLISA

Problème actuel d'accessibilité à la donnée et à ces documents explicatifs

Plusieurs retours réguliers des utilisateurs (questionnaire et contact direct) révèlent qu'ils éprouvent des difficultés à trouver les liens pour le téléchargement des couches SIG et les liens vers les documents associés.

Il est vrai que l'accès n'est pas direct, le plus simple étant de passer par l'atlas catalogue du SANDRE, rechercher la BDLISA via le moteur de recherche associé, ouvrir la fiche de métadonnées, aller sur l'onglet « ressource et téléchargement » et cliquer sur le lien « télécharger la donnée » ce qui ouvre une nouvelle page html.

Ce processus d'accès n'est pas optimal. Un accès plus direct à la donnée en téléchargement semble donc nécessaire.

Le SANDRE est conscient de ces problèmes et une réflexion commence tout juste sur l'ergonomie du site et de l'atlas catalogue. Dans cette perspective, le SANDRE souhaite mieux connaître la manière dont est utilisée la BDLISA par ses utilisateurs.

→ Pour répondre à cette attente, le BRGM enverra au SANDRE les conclusions de l'enquête sur l'utilisation de la BDLISA récemment réalisée.

De même, du fait du gros volume de données à télécharger, le temps de téléchargement est tel qu'il est difficile de télécharger certains zips (plusieurs heures nécessaires, coupures intempestives par moment). Ce problème devrait déjà être en partie résolu par la simplification des polygones, la réduction de la taille de la base et le découpage possible en plusieurs zips proposé par le SANDRE.

Aucune porte d'entrée unique pour la BDLISA

Le retour d'expérience des utilisateurs met également en évidence l'absence d'une porte d'entrée unique pour découvrir et télécharger la BDLISA. Des informations sont disponibles sur le site du SANDRE (mais peu accessibles), d'autres sur le site geotraitement.eaufrance (purement cartographique), d'autres sur ADES, le site du BRGM et les SIGES. Cette diversité éparpille l'information et l'utilisateur souhaitant comprendre la BDLISA s'y perd régulièrement.

Proposition :

Actuellement, le site geotraitement.eaufrance accueille un démonstrateur des services web et un guide technique à usage des experts pour exploiter les services web cartographiques. Ce site (**très bien référencé sur les moteurs de recherche car assez ancien**) pourrait être amélioré pour servir de portail d'accès aux données et à la documentation associée aux différents produits de la BDLISA 'usages', avec par exemple :

- Une centralisation de la documentation et des documents synthétiques et pédagogiques (vers une aide en ligne?) ;
- Une interface de saisie utilisateur avec des sélections en fonction d'une zone d'étude (certaines demandes utilisateurs pourraient être pré-calculées (par bassin, par

région, ordre 1...)

- Une BDLISA pour les administrateurs de base de données (postgreSQL/postGIS, SQLite/Spatialite) avec des modes opératoires pour le déploiement dans les services >> l'objectif est de favoriser l'exploitation de la BDLISA dans des structures aux utilisateurs potentiels.

Cette proposition devra être discutée lors du prochain GT national.

Modèle de données peu expliqué, absence de document support clair ou à jour

Le modèle de données associé au référentiel BDLISA est actuellement obsolète (dernière Maj effectuée fin 2013) ; il est prévu de le mettre à jour avec le SANDRE durant le 2nd semestre 2016.

Il semble qu'un effort doit être mené pour améliorer la description du modèle de données, dans ce document mais également dans les autres docs supports (scénario d'échange...). Un document de présentation SANDRE du modèle de données pourrait être rédigé.

Difficulté de récupérer les lexiques SANDRE et la symbologie utilisée sur les services web

Régis Haubourg remonte également la difficulté de récupérer la symbologie et les tables de lexique SANDRE associées aux caractéristiques des entités BDLISA.

Proposition :

→ La création de tables de nomenclature (.csv) reliant les codes au mnémonique, au libellé, et à la définition, pourrait être rapidement et facilement réalisée dans le cadre de la sortie de la nouvelle version du référentiel. Ce travail a par ailleurs été effectué par Régis Haubourg.

Attention à l'encodage de ce type de fichier : Latin1, UTF8

→ Concernant la question de la symbologie, les formats d'échanges des fichiers de symbologie (.lyr, .sld) ne sont soit pas intéropérables (différences formats, version et types de logiciel SIG), soit pas suffisamment riches pour les rendus souhaités. La diffusion de ces fichiers est possible mais pose encore des interrogations. Des tests vont toutefois être réalisés par Robin Quique pour permettre une récupération optimisée des symbologies actuellement diffusées sur le site géotraitement.

Action	Responsable	Délai
Mettre en place des tests sur la simplification de la BDLISA	BRGM	2016
Appréhender et mettre en place une procédure sur les tests sur les attributs et les tests de topologie sur PostGIS.	BRGM, SANDRE	2016
Envoyer un document synthétique sur les usages de la BDLISA au SANDRE	BRGM, SANDRE	2016
David Viglietti du SANDRE se renseigne sur les possibilités de scinder les fichiers en téléchargement. Des tests devront être réalisés pour appréhender l'impact de cette séparation sur l'exploitation des fichiers.	SANDRE	2016
Statuer sur les formats d'utilisation de la BDLISA (abandon Mapinfo ? création base SQLITE ?)	GT national, BRGM	2016

Statuer sur la proposition suivante : le site géotraitement comme porte d'entrée qui centraliserait les docs/supports/données de la BDLISA	GT national	2016
Mise à jour du dictionnaire de données SANDRE et rédiger un document de présentation du modèle et du dictionnaire ?	BRGM, SANDRE	S2 2016
Approfondir la question de la BDLISA référentiel et de la BDLISA 'décisionnelle', notamment via l'analyse des retours utilisateur et de l'enquête en ligne, ce qui permettra de cibler les besoins en termes d'usages de la BDLISA et les moyens à mettre en œuvre pour y répondre.	GT national, BRGM, SANDRE	2016- 2017
Intégrer un identifiant unique – clé métier sur les polygones de découpage	BRGM	2017
Création de tables de relation entre les codes BDLISA et leur signification. Ce travail s'appuiera sur le travail déjà en partie réalisée de Régis Haubourg. Ajouter une table vers les nomenclatures SANDRE.	BRGM	2017
Proposer un export de la symbologie	BRGM	2017