
Protocole de préparation et de republication des données du BRGM sur les formations à silicites

Travaux réalisés dans le cadre des PCR « Réseau de lithothèques » et du GDR « SILEX » – Article de protocole

Christophe Tufféry, Vincent Delvigne, Paul Fernandes et Jérémy Garniaux



Édition électronique

URL : <https://journals.openedition.org/pm/3483>

DOI : 10.4000/pm.3483

ISSN : 2105-2565

Traduction(s) :

Preparation and re-publication protocol for the BRGM data on silicites formations - URL : <https://journals.openedition.org/pm/3549> [en]

Éditeur

Association pour la promotion de la préhistoire et de l'anthropologie méditerranéennes

Référence électronique

Christophe Tufféry, Vincent Delvigne, Paul Fernandes et Jérémy Garniaux, « Protocole de préparation et de republication des données du BRGM sur les formations à silicites », *Préhistoires Méditerranéennes* [En ligne], 10 | 2022, mis en ligne le 13 novembre 2022, consulté le 17 avril 2023. URL : <http://journals.openedition.org/pm/3483> ; DOI : <https://doi.org/10.4000/pm.3483>

Protocole de préparation et de republication des données du BRGM sur les formations à silicites

Travaux réalisés dans le cadre des PCR « Réseau de
lithothèques » et du GDR « SILEX » – *Article de protocole*

Christophe Tufféry, Vincent Delvigne, Paul Fernandes et Jérémy Garniaux

Mots-clés

Préhistoire, pétrographie,
lithothèques, base de
données, silix

Keywords

*Prehistory, petrography, lithological
collections, database, flint*

Résumé

Les données sont issues des travaux de quatre PCR : « Réseau de lithothèques en Centre-Val-de-Loire », « Réseau de lithothèques en Nouvelle Aquitaine », « Réseau de lithothèques en Auvergne Rhône-Alpes » « Les Silicites Céozoïques d’Île de France » et du GDR « SILEX ». Ces travaux s’appuient sur une méthode commune de production et de validation. Elles permettent une description harmonisée des formations à silicites par départements à partir des données de la Bd Charm-50 du BRGM (source : site Infoterre¹). Les données sont également consultables depuis un logiciel de SIG sous la forme de webservices de type WFS (ArcGIS Feature Service).

Abstract

The data is the result of four PCR: “Réseau de lithothèques en Centre-Val-de-Loire,” “Réseau de lithothèques en Nouvelle Aquitaine,” “Réseau de lithothèques en Auvergne Rhône-Alpes,” (translating to: lithological collection networks in the Centre-Val-de-Loire, Nouvelle Aquitaine, and Auvergne Rhône-Alpes regions, respectively) “Les Silicites Céozoïques d’Île de France” (which translates to: Cenozoic silicate mineral formations in Île de France) and the GDR “SILEX.” These projects and their resulting data are based on a common method of production and validation. Thus, they provide for a harmonized description of the siliceous rock formations by department, using the BRGM’s Bd Charm-50 data (source: Infoterre website²). The data can also be downloaded using GIS software in the form of WFS (ArcGIS Feature Service).

1 <https://page.hn/2ie938>

2 <https://page.hn/2ie938>

Reçu / Received 28-04-2022 ; Version révisée reçue / Received in revised form 04-07-2022 ; Accepté / Accepted 20-08-2022

Inrap – Nouvelle-Aquitaine, Nouvelle Aquitaine et Outre-Mer, 140 avenue Maréchal Leclerc, 33130 Bègles, France,

ORCID : 0000-0002-1818-9388, christophe.tuffery@inrap.fr

CNRS, UMR 8068 TEMPS, Maison René Ginouvès, 21 allée de l’université, 92023 Nanterre, France / Service de Préhistoire, Université de Liège, Place du XX août, 4000, Liège, Belgique, ORCID : 0000-0001-6477-2104, vincent.delvigne@cnrs.fr

Paléotime, 75 avenue Jean-Séraphin Achard-Picard, 38250 Villard-de-Lans, France, ORCID : 0000-0001-6253-8250, paul.fernandes@paleotime.fr

Univ Aix Marseille, CNRS, UAR3125 MMSH / UMR 7269 LAMPEA, 5 rue du Château de l’Horloge, 13097 Aix-en-Provence, France

ORCID : 0000-0002-2603-4407, jeremy.garniaux@univ-amu.fr

1. VUE D'ENSEMBLE

Correspondants du GDR SILEX

Céline Leandri - celine.leandri@culture.gouv.fr - ORCID: 0000-0001-7668-3248

Vincent Delvigne - vincent.delvigne@cnrs.fr - ORCID: 0000-0001-6477-2104

Données

Emplacement du référentiel

Site de cartographie sur AGOL : <https://bit.ly/3a4ZfmZ> (URL ouverte à tous depuis juin 2022) et site web www.cartosilex.fr (ouvert depuis novembre 2022)

Données

Les données sont publiées progressivement sur l'entrepôt Nakala au sein d'une collection commune : <https://nakala.fr/collection/10.34847/nkl.dd1f3l6v> (sous réserve d'être autorisé à accéder à cette collection et à y déposer des données).

Date de publication du référentiel : 01/07/2022

Licence

Etalab 2.0
<https://www.etalab.gouv.fr/wp-content/uploads/2017/04/ETALAB-Licence-Ouverte-v2.0.pdf>

Type de données

Données accessibles sous la forme de services web de type WFS (ArcGIS Feature Services).

Il est possible de se connecter aux URL depuis WFS depuis ArcGIS, QGIS, etc. à partir des URL³.

Couverture spatiale

France entière (métropole et Dom).

Coordonnées géographiques

France entière (métropole et Dom).

³ Pour l'instant, ces URL sont disponibles sur simple demande, elles pourront être publiées avec un DOI si besoin. Mais le DOI n'est pas un format reconnu comme WFS pour un logiciel de SIG.

2. MÉTHODES

Étapes

Les données de la Bd Charm-50 du BRGM ont été téléchargées sur le site Infoterre et les données attributaires ont été adaptées pour les besoins selon une liste de critères définis. Ainsi, lors du téléchargement, pour chacun des départements, les données sont regroupées en 5 thèmes⁴ :

- L_DIVERS (objets linéaires « Divers »)
- L_FGEOL (objets linéaires « Formations géologiques »)
- L_STRUCT (objets linéaires « Structuraux »)
- S_SURCH (objets surfaciques)
- S_FGEOL (objets surfaciques « Formations géologiques »)

Dans le cadre de nos travaux, le thème que nous souhaitons récupérer est le thème S_FGEOL, qui décrit les formations géologiques de surface. Les données attributaires contenues dans les fichiers comprennent 16 champs descriptifs, dont certains sont indispensables pour nos recherches et d'autres, considérés comme inutiles, ont été supprimés. Des champs descriptifs supplémentaires ont été ajoutés. Le nombre final des champs est de 21 (cf. étape 5).

L'intérêt majeur des données du BRGM est d'une part de disposer de données harmonisées pour tous les départements, et d'autre part, la qualité de ces données est bien adaptée pour répondre aux besoins de cartographie des formations à silicites aux échelles régionale et nationale. Cette qualité d'ensemble n'exclut toutefois pas l'étape de vérification des contours des formations, soit par le biais d'un retour sur le terrain, soit par le biais de la comparaison avec les données provenant des lithothèques (voir enquête 123Survey et *infra*). Enfin, la licence Etalab 2.0 autorise des traitements et des adaptations comme celles réalisées pour les besoins des projets.

Pour exploiter les données ici décrites, il faut disposer d'un logiciel de SIG capable de lire les formats .shp, .gpkg et les webservices de type WFS (comme par exemple le logiciel libre QGIS⁵). Nous conseillons l'installation de la version LTR (*Long Terme Release*) la plus récente, plus stable que la « dernière version » également disponible en téléchargement sur le site de QGIS. À la date de rédaction de cet article, la version LTR est la 3.22.

⁴ Voir le document du BRGM fourni avec les données de la BD-Charm50 (Janjou 2004).

⁵ QGIS est un logiciel open source distribué sous licence GNU GPL 3.0 (<https://www.qgis.org/fr/site/>).

Étape 1 : téléchargement des données Bd Charm 50 sur le site du BRGM

L'adresse est la suivante :

<http://infoterre.brgm.fr/formulaire/telechargement-cartes-geologiques-departementales-150-000-bd-charm-50> (URL raccourcie: <https://page.hn/2ie938>)

Les données pour la métropole sont téléchargeables par départements. Les produits fournis sont des archives au format .zip qui comprennent une série de fichiers au format shapefile (voir description du format plus loin). Avant de pouvoir télécharger ces données, un formulaire doit être renseigné.

Une fois les fichiers téléchargés, il faut les décompresser puis ouvrir le répertoire comprenant tous les fichiers. Parmi ceux-ci, il faut choisir les six fichiers intitulés GEO050K_HARM_[n°_dep]_S_FGEOL_2154, dont les extensions diffèrent et sont les suivantes : .shp, .dbf, .shx, .prj, .lyr et .qml.

Les données au format shapefile sont constituées de fichiers au nom identique, mais aux extensions différentes. Les fichiers obligatoires sont ceux dont les extensions sont .shp (fichier principal décrivant la géométrie), .shx (fichier d'indexation permettant la fonctionnalité de recherche), .dbf (fichier de base de données contenant les données attributaires associées aux géométries). Les fichiers optionnels sont .prj (métadonnées relatives aux coordonnées et au système de projection) et .lyr (fichier contenant une représentation persistante de la couche). Le dernier fichier (extension .qml), optionnel également, est un fichier de description du style de la couche spécifiquement destiné à QGIS. Excepté les fichiers .lyr et .qml dont nous n'avons pas besoin ici, l'ensemble de ces fichiers devront être systématiquement accessibles et stockés au même endroit pour que la donnée SIG soit utilisable.

Étape 2 : affichage des données dans QGIS

Les données du BRGM doivent ensuite être affichées dans QGIS. Il faut choisir le ou les fichiers au format shape pour le ou les départements souhaités.

Étape 3 : fusion des données pour les départements concernés

S'il y a besoin de fusionner les fichiers de plusieurs départements, il faut d'abord les afficher dans QGIS puis

choisir l'outil de fusion disponible dans le menu Vecteur > Outils de gestion de données > Fusionner des couches vecteurs.

Une fois les données fusionnées, il faut enregistrer le fichier résultant au format .gpkg et lui donner le nom de son choix. Le format .gpkg présente plusieurs avantages par rapport au format .shp :

- C'est un format de données ouvert et libre ;
- Un seul fichier (et non plus 4 à 6 fichiers comme pour le format shapefile) regroupe toutes les données de la couche ;
- Il n'y a pas de limite de caractères pour le nom des champs ;
- Il est admis dans AGOL pour une publication en WFS, ainsi que dans la plupart des logiciels de SIG.

Étape 4 : découpage des données par carte géologique

Afin de disposer des données selon le découpage des cartes géologiques, il faut d'abord afficher le fichier de découpage des cartes géologiques sur la zone concernée. Ce fichier est disponible au format shapefile sur le site Infoterre du BRGM. Une fois affiché dans QGIS, il faut sélectionner une des cartes avec l'outil de sélection.

Il faut ensuite utiliser l'outil de découpage (Couper) qui se trouve dans le menu Vecteur > Outils de géotraitement > Couper.

Dans le panneau qui apparaît, il faut choisir le fichier à découper (les données de la BD Charm-50 pour la zone concernée) puis le fichier de découpage à savoir dans le fichier des cartes géologiques la carte préalablement sélectionnée. Pour cela, il faut cocher la case Entité(s) sélectionnée(s) uniquement.

Une fois les données découpées, il faut enregistrer le fichier résultant au format .gpkg. Il faut choisir un nom de fichier du résultat. Il est proposé d'intégrer le nom et/ou le numéro de la carte géologique dans le nom du fichier résultant.

Dans le panneau qui s'affiche, il faut supprimer la mention *Output* et cliquer sur OK puis sur le bouton Exécuter. Le résultat du traitement apparaît dans la fenêtre puis dans QGIS.

Étape 5 : modifier la table attributaire du fichier résultant (fig. 1)

Pour le fichier issu du découpage, il faut afficher la table attributaire puis activer le mode édition pour pouvoir modifier cette table attributaire. Il faut ajouter les champs suivants :

- "AUTEURS_CREATION" (texte, 80)
- "DATE_CREATION" (date)
- "AUTEURS_MODIF" (texte, 80)
- "DATE_MODIF" (date)
- "ERE" (texte, 20)
- "SERIE" (texte, 30)
- "ETAGE" (texte, 30)
- "TYPE_GITE" (texte, 50)
- "NOM_NUM_CARTE" (texte, 20)
- "LISTE_TYPER_PRESENTS" (texte, 80)
- "VERIFIE_SUR_TERRAIN" (texte, 80)
- "COMMENTAIRES" (texte, 80)
- "ETAGE_FINAL" (entier, 80)
- "CODE_HEX" (texte, 80)
- "CODE_DEPT" (entier, 80)
- "NOM_DEPT" (texte, 80)
- "CODE_DEPT" (texte, 80)

Les champs suivants du fichier d'origine du BRGM doivent être conservés :

- "NOTATION" (texte, 80)
- "DESCR" (texte, 80)
- "C_FOND" (texte, 80)
- "M_FOND" (texte, 80)
- "J_FOND" (texte, 80)
- "N_FOND" (texte, 80)

Il faut ensuite supprimer les champs suivants :

- MI_PRINX
- CARTE
- CODE
- CODE_LEG
- NOM_SYMB
- C_SYMB
- M_SYMB
- J_SYMB
- N_SYMB
- ROT_SYMB
- Layer
- path

Il faut ensuite charger l'ordre des champs selon celui d'une autre couche .gpkg déjà ordonnée. Ce traitement s'appelle refactoriser les données. Pour cela il faut ouvrir la

boîte à outils⁶ de traitements de QGIS puis trouver l'outil nécessaire en tapant dans la recherche « refactoriser ». L'outil apparaît ; il faut le sélectionner et l'ouvrir.

Il faut choisir le fichier à réorganiser (source) et le fichier de réorganisation (cible). Il faut choisir d'enregistrer le résultat dans un fichier au format GeoPackage (e.g. GEO50K_HARM_Amboise_458_version_finale.gpkg). Le résultat du traitement apparaît dans la fenêtre, puis dans QGIS dans lequel il faut supprimer le fichier précédent (cf. fichier cible).

Étape 6 : autre méthode à partir d'un fichier .xlsx

Il existe une autre méthode qui s'appuie sur un fichier au format .xlsx qui peut être préalablement établi pour faciliter le travail et éviter l'étape de refactorisation. Ce fichier .xls ou .xlsx ne comprend que le champ d'identifiant (fid).

Dans QGIS, il faut afficher ce fichier tableur et le fichier des données géologiques du BRGM pour une carte. Il faut ensuite afficher le panneau des propriétés de cette couche des données géologiques et choisir l'onglet Jointure puis cliquer sur le symbole +.

Dans le panneau de la jointure, il faut choisir le fichier tableur comme fichier source puis choisir le champ de nom et numéro de la carte géologique comme champ source puis le champ NOM_NUM_CARTE dans le fichier des données du BRGM qui est le fichier cible.

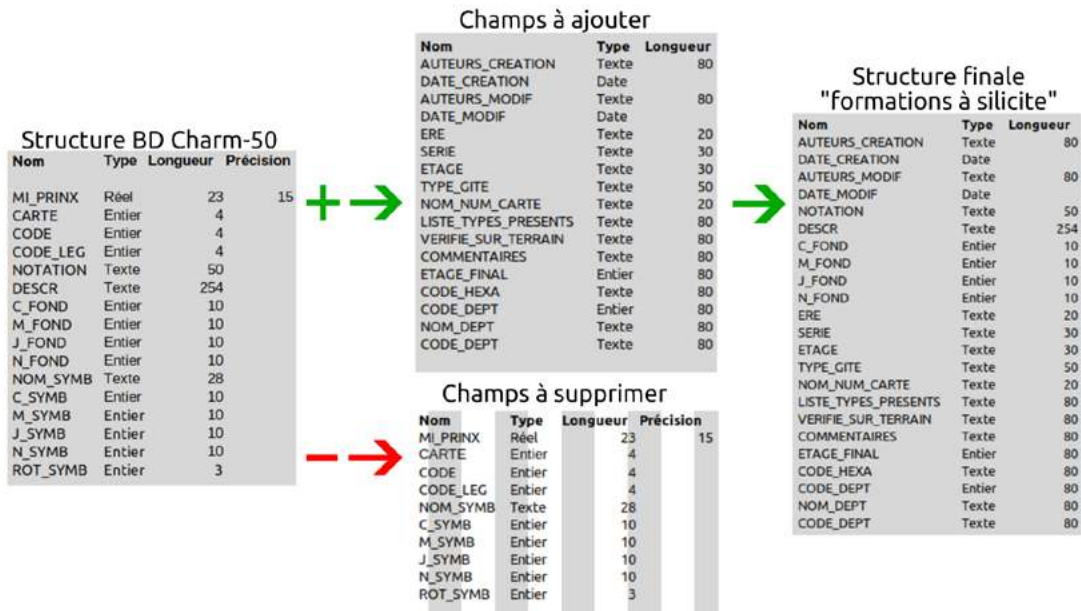
Une fois cette jointure effectuée, le résultat apparaît dans la table attributaire.

Puis dans la table attributaire, on peut sélectionner la valeur de la carte souhaitée dans le champ NOM_NUM_CARTE avec l'outil de sélection par expression.

On peut par exemple choisir les polygones dont le champ NOM_NUM_CARTE est le numéro_458. L'expression est la suivante : "NOM_NUM_CARTE" = 'numero_458'

Le résultat montre que toutes les formations d'une carte n'ont pas le numéro de carte. Il faut donc homogénéiser ces valeurs pour tous les polygones compris à l'intérieur d'une même carte géologique.

⁶ cf. documentation de QGIS : https://docs.qgis.org/3.22/fr/docs/user_manual/processing/toolbox.html?highlight=bo%C3%A0Ete%20outils



1. Processus d'adaptation des données attributaires issues de la BD Charm-50. Crédit : GDR SILEX (J. Garniaux).

Étape 7 : homogénéiser certains champs

Pour cela, il faut mettre à jour le champ "NOM_NUM_CARTE" avec le numéro de la carte géol.

Il faut ouvrir la calculatrice de champs puis saisir la formule suivante : `if("NOM_NUM_CARTE" IS "", "Amboise_458')`

Le résultat est l'homogénéisation du champ NOM_NUM_CARTE avec comme seule valeur 'Amboise_458'.

Étape 8 : séparer les formations selon les valeurs du champ NOTATION

Pour séparer les formations d'une carte selon les valeurs du champ NOTATION, il faut créer au préalable un dossier avec le gestionnaire de fichiers du poste de travail (p. ex: `Data_extraites_GEO50K_HARM_Notation_[Nom_num_carte]`)

Dans QGIS, il faut ouvrir l'outil de séparation d'une couche vecteur qui se trouve dans le menu Vecteur > Outils de géotraitement > Couper

Il faut choisir la couche à séparer (les données des formations d'une carte géologique) puis le champ de séparation ("NOTATION"), puis le répertoire de destination préalablement créé pour y enregistrer le

résultat de la séparation puis cliquer sur Sélectionner un dossier, puis sur Exécuter⁷.

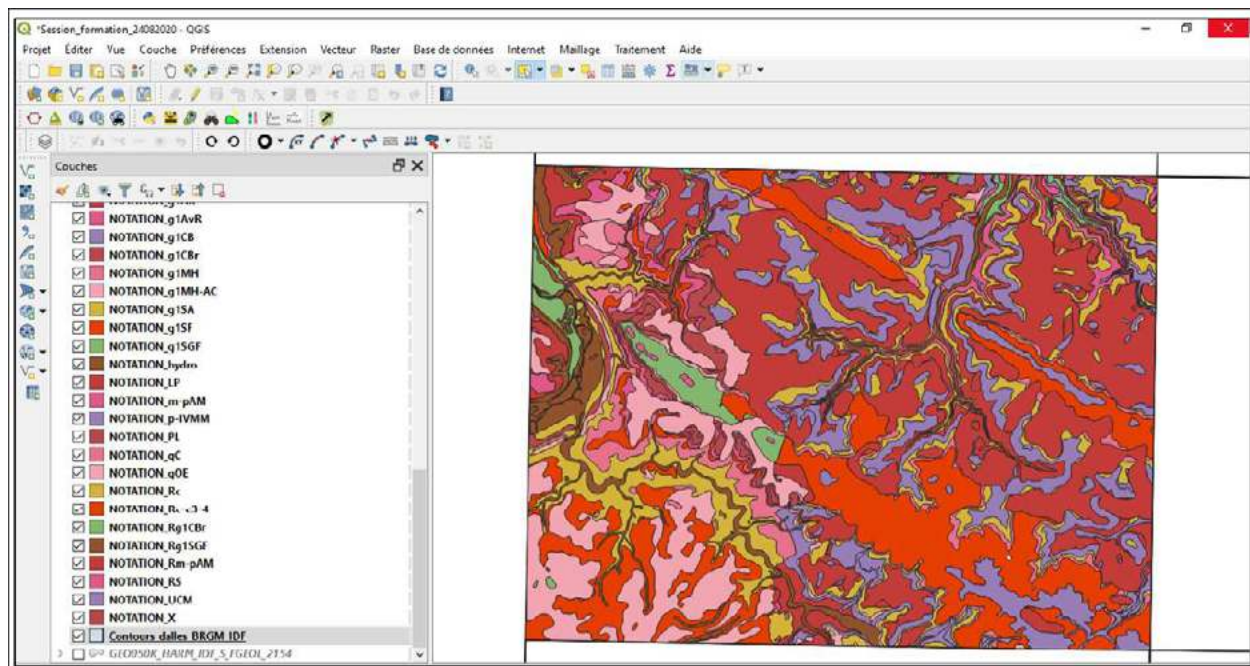
Une fois le traitement effectué, le résultat peut être affiché dans QGIS en ajoutant la ou les couches vecteurs issues de la séparation comme le montre la figure 2 ci-dessous.

À partir de cette étape, les données peuvent être sélectionnées et certaines supprimées si on ne souhaite pas conserver l'ensemble de celles-ci.

Étape 9 : établissement de la liste des couches et des valeurs à renseigner pour les polygones

Cette étape consiste à choisir les couches à retenir selon la valeur du champ "NOTATION", puis les valeurs à renseigner par lots pour les polygones concernés. Cette étape nécessite, la lecture des cartes géologiques, des données de sondages et de tous documents relatifs aux silicites dans la zone concernée. L'avis d'un ou de plusieurs experts est un plus non négligeable lors de cette étape.

⁷ Si des messages d'erreur apparaissent et empêchent d'exécuter entièrement cette étape, cela est souvent dû à des caractères spéciaux qui ne sont pas admis lors de l'étape de découpage selon un champ. Il faut donc d'abord corriger ces caractères spéciaux partout où ils apparaissent dans le champ qui sert au découpage. Une fois cette correction préalable effectuée, l'opération de découpage selon le champ corrigé, peut être exécutée.



2. Exemple de découpage selon le champ "NOTATION" d'une couche de données. Crédit : GDR SILEX (C. Tufféry)

Étape 10 : fusionner des couches

Une fois la table attributaire renseignée pour les polygones concernés, il peut être nécessaire de fusionner des couches distinguées lors du découpage par le BRGM, mais similaires dans la « problématique silicite » (c'est typiquement le cas des argiles à silex, noté de différentes manières en fonction des cartes géologiques harmonisées) en reprenant les outils de QGIS décrits pour l'étape 3. Le résultat doit correspondre aux données des seules formations à silicites retenues, renseignées pour un département.

Étape 11 : vérifier le .gpkg

Cette étape peut se faire en consultant les données attributaires et, le cas échéant, en ajoutant ou supprimant certains polygones. Ce travail est effectué à partir de l'avis du ou des experts.

Étape 12 : publier la donnée

Une fois le fichier au format .gpkg vérifié et validé, il doit être exporté depuis QGIS au format .geojson⁸.

Pour ce faire, sélectionnez la couche à exporter, puis cliquez sur Exporter et choisissez le format GeoJSON.

Une fois le fichier exporté au format GeoJSON, celui-ci et le fichier .gpkg peuvent être importés tous les deux sur les plateformes AGOL et QGIS Server dédiées (et doivent avoir strictement le même nom). Lors de l'importation il faut choisir l'option de publication de ces données en WFS. L'étape suivante consiste à rendre les données homogènes avec les données des autres départements. La symbologie des couches doit être commune pour tous les départements. Pour cela, nous avons décidé d'utiliser le champ ETAGE_FINAL. Pour chacune des valeurs du champ "ETAGE_FINAL", nous avons décidé d'appliquer une couleur à partir de la valeur de son Code Hexa issu de la charte stratigraphique internationale. Enfin, un taux de transparence est appliqué selon les valeurs du champ "ETAGE_FINAL":

- valeurs sans le signe _ (correspondant aux valeurs avec _P) : 0%
- _AA : 25% (alluvions anciennes)
- _RS : 50% (formations résiduelles à silicites)

géographique. Il utilise la norme JSON (JavaScript Object Notation), créée pour l'échange de données sur le web. Il permet de décrire des données de type point, ligne, chaîne de caractères, polygone, ainsi que des ensembles et sous-ensembles de ces types de données et d'y ajouter des attributs d'information qui ne sont pas spatiales (source : Wikipedia).

⁸ Geojson est un format ouvert d'encodage de données à caractère

- _CRS : 75% (formations colluviales à silicites)

Cette solution permet de pallier l'absence de symboles de trames transparentes dans l'outil de construction des légendes d'AGOL et de QGISServer.

Stratégie d'échantillonnage

Les données sur les formations à silicites font l'objet d'échantillonnages réguliers pour en vérifier la qualité dans le cadre des besoins et des possibilités des programmes de recherche (e.g. PCR « Réseau de lithothèques » et GDR « SILEX »).

Contrôle de qualité

Les données sur les formations à silicites issues des données du BRGM sont considérées globalement d'une qualité adaptée aux besoins des programmes de recherche qui les ont utilisées et les utilisent encore.

Néanmoins, selon les résultats des échantillonnages effectués, une information sur la fiabilité des données attributaires et/ou la géométrie des polygones des formations pourra être fournie dans la table attributaire. Cette information se trouvera dans le champ « Fiabilité ».

Pour chacun des polygones, le champ de type texte comportera une valeur comprise entre 0 et 3 selon les significations suivantes pour chacune des valeurs possibles :

- Valeur 0 : aucune vérification n'a été faite sur la formation
- Valeur 1 : des vérifications ponctuelles ont révélé que les contours du polygone et/ou des attributs descriptifs sont imprécis ou faux
- Valeur 2 : des modifications ponctuelles ont été apportées aux contours du polygone et/ou aux attributs descriptifs
- Valeur 3 : des modifications systématiques ont été apportées aux contours du polygone et/ou aux attributs descriptifs

Contraintes

Les données du BRGM réutilisées et adaptées correspondent à un état des connaissances à la date de leur publication. Elles peuvent faire l'objet de corrections ou de mises à jour en fonction de nouvelles observations ou de nouvelles analyses ou des résultats du contrôle qualité décrit au paragraphe précédent.

3. DESCRIPTION DU JEU DE DONNÉES

Nom de l'objet

« Données sur les formations à silicites » ou « Les silicites en France »

Type de données

Données traitées

Format

Les données sont accessibles sous la forme de WFS de type webservices (ArcGIS Feature Services).

À l'aide d'un logiciel de SIG pouvant accéder à ce type de WFS, divers formats d'export sont disponibles selon le logiciel de SIG utilisé.

Dates de création

01/01/2006 - 30/03/2022

Créateurs du jeu de données

Le BRGM reste l'auteur des données d'origine de la Bd Charm-50.

Les noms des contributeurs ayant apporté des modifications sur ces données dans le cadre des travaux des PCR et du GDR sont disponibles dans deux des champs descriptifs des données : auteurs_création, auteurs_modif

Langue

Français

4. REMERCIEMENTS/PARTENARIATS

Remerciements

Ces données ont été produites grâce à la collaboration de nombreuses personnes. La liste en est trop longue pour pouvoir les citer nominativement toutes, nous souhaitons toutefois particulièrement remercier Erwan Vaissié, Pierrick Tigréat et Alix Gibaud pour leur investissement dans le test et la production de données primaires.

Partenariat/ Financement

Ces données ont été adaptées et financées dans le cadre des travaux des PCR : « Réseau de lithothèques en Centre-Val-de-Loire », « Réseau de lithothèques en Nouvelle Aquitaine », « Réseau de lithothèques en Auvergne Rhône-Alpes », « Les Silicites Cénozoïques d'Ile de France » et du GDR « SILEX ». Ces programmes sont cofinancés par le ministère de la Culture via les DRAC concernées, par le ministère de l'Enseignement Supérieur, de la Recherche et de l'Innovation via les agents impliqués dans les projets, par le CNRS, par l'Inrap, par la S.A.R.L. Paléotime et par les Conseils départementaux de l'Ain, de l'Allier, de l'Ardèche, de la Drôme, de l'Isère, du Puy-de-Dôme et de la Savoie.

5. RAPPORT D'EXPERTISE

Sylvain Rassat
Aix Marseille Univ, CNRS, Minist Culture, LAMPEA,
UMR 7269, 5 rue du Château de l'Horloge, 13097,
Aix-en-Provence, France – ORCID: 0000-0002-6451-
687X – sylvain.rassat@univ-amu.fr

La cohérence et l'aspect didactique de l'article sont appréciables. Cet article, qui est la description d'une méthode commune de production et de validation pour une description harmonisée des formations à silicites par départements à partir des données de la Bd Charm-50 du BRGM, permet la compréhension globale du process. Le modèle de métadonnées est conforme aux recommandations. Les données sont organisées de manière logique et cohérente et sont facilement lisibles et utilisables : les données hébergées par le BRGM répondent à ces critères, de même que pour les données liées au référentiel qui sont accessibles à travers la plateforme AGOL. L'indication de la version du logiciel de cartographie utilisé (QGIS) et de son cadre d'utilisation (licence, accessibilité au code, etc.) permet la reproductibilité complète du processus. On notera tout particulièrement l'utilisation du thesaurus multilingue Pactols Frantiq permettant l'alignement ontologique ultérieur et l'interopérabilité des thesaurus. La rédaction par les auteurs d'un lexique pétrographique (publication en cours) permettra d'explicitier et de proposer la définition (ou le renvoi par URL aux définitions de l'entité source) des différents acronymes et abréviations usitées et développés par des tiers extérieurs aux projets (BRGM en particulier).

L'accessibilité à la plateforme AGOL est possible aux membres des PCR et du GDR, aux agents de l'Inrap ainsi qu'à toute personne intéressée par le domaine, permettant l'ouverture des données aux publics et aux débats contradictoires. Dans l'optique d'une « FAIRisation » des données, une collection Nakala (TGIR HumNum) allouée à chacun des éléments constitutifs du PCR « Réseau de lithothèques » et du GDR « SILEX » a été créée. Les données traitées et le process de traitements peuvent donc être utilisées pour d'autres problématiques et données. Les cas d'utilisation décrits dans le protocole sont cohérents avec les données BRGM, les données brutes au sein du BRGM sont suffisamment documentées. Le processus de traitements mérite d'être publié.

Citer comme

Rassat S., 2022, « Rapport d'expertise » in Tufféry C., Delvigne V., Fernandes P., Garniaux J. (2022), Protocole de préparation et de republication des données du BRGM sur les formations à silicites - article de protocole, *Préhistoires méditerranéennes*, 2022, 10.1 – [report] Nakala, <https://doi.org/10.34847/nkl.0b45tgqx>

BIBLIOGRAPHIE

- Fernandes et al. 2013** Fernandes P., Raynal J.-P., Tallet P., Tufféry C., Piboule M., Seronie-Vivien M., Seronie-Vivien M.- R., Turq A., Morala A., Affolter J., Millet D., Millet F., Bazile F., Schmidt P., Foucher P., Delvigne V., Liagre J., Gaillot S., Morin A., Moncel M.-H., Garnier J.-F., Léandry-Bressy C., Une carte et une base de données pour les formations à silex du sud de la France : un outil pour la pétroarchéologie, *Paléo*, n° 24, p. 219-228. <http://doi.org/10.4000/paleo.2633>
- Janjou 2004** Janjou D., *Descriptif des cartes géologiques à 1/50 000 format "vecteurs"*, rapport final, BRGM/RP-53473-FR, 21 p., 6 fig., 9 tabl.
- Tufféry et al. 2021** Tufféry C., Delvigne V., Fernandes P., Bressy-Leandri C., À propos de quelques outils de collecte de données : réflexions sur les pratiques numériques en archéologie, *Humanités numériques*, 3 | 2021. <http://doi.org/10.4000/revuehn.1603>
- Tufféry et al. 2019** Tufféry C., Fernandes P., Delvigne V., Bressy-Leandri C., Collecte, gestion et valorisation des données sur les formations et les gîtes à silex en France : pour des plateformes cartographiques et de publications partagées. Proceedings of the session III-3 (CA) of the XVIII^e UISPP congress, Paris, June 2018 Construire des référentiels partagés : Webmapping et archéologie. *Archéologies numériques*, 1, vol. 3. <http://doi.org/10.21494/ISTE.OP.2019.0351>