

# « Vrais trous » et « fausses erreurs » : le contrôle de la numérisation de contours des formations à silex en Dordogne à l'aide d'un SIG.

"True holes" and "false errors": the control of the digitization of shapes of flint formations in the Dordogne department by using a GIS.

Christophe Tuffery<sup>1</sup>, David Talec<sup>2</sup>, Jean-Baptiste Caverne<sup>3</sup>, Pascal Tallet<sup>3</sup>, Geneviève Pinçon<sup>2</sup>, Stéphane Konik<sup>2</sup>, Jean-Pierre Platel, Alain Turq<sup>4</sup>, André Morala<sup>4</sup>, Paul Fernandes<sup>3</sup>

<sup>1</sup> Inrap - Direction Scientifique et Technique, christophe.tuffery@inrap.fr

<sup>2</sup> CNP, talec.david@free.fr, genevieve.pincon@culture.gouv.fr, stephane.konik@culture.gouv.fr

<sup>3</sup> Paléotime, jean-baptiste.caverne@laposte.net, pascal.tallet@paleotime.fr, paul.fernandes@paleotime.fr

<sup>4</sup> MNP, alain.turq@orange.fr, andre.morala@orange.fr

**RÉSUMÉ.** Dans le cadre de plusieurs Projets Collectifs de Recherches<sup>1</sup>, un partenariat avec le Centre National de Préhistoire à Périgueux a permis de réaliser la numérisation au 1/10 000 de tous les polygones identifiés comme correspondant à des formations à silex dans le département de la Dordogne. Ce travail, réalisé avec les logiciels ArcGIS et QGis en utilisant les cartes géologiques au 1/50 000 numérisées et géoréférencées, disponibles dans le service web Infoterre du BRGM, a permis la production d'une base de données unique sur les géoressources. Une fois numérisées et caractérisées de façon homogène, les formations à silex servent à visualiser les géoressources actuelles et donc à émettre des hypothèses sur les stratégies d'approvisionnement et de déplacement des populations préhistoriques. Les données sur les formations à silex ont été transférées sur ArcGISOnLine (AGOL), une plateforme cartographique d'ESRI<sup>2</sup> qui permet de diffuser et de partager des données géoréférencées de façon sécurisée. Lors de cette étape, des différences de géométrie sont apparues entre les données d'origine et les données après leur importation sur la plateforme AGOL. Un contrôle systématique de la qualité de la numérisation et des transformations des fichiers après leur transfert sur la plateforme AGOL, est apparu indispensable. Il s'agit d'une garantie de la qualité des données pour les partenaires du projet. Si ce contrôle n'avait pas été fait, des erreurs d'interprétation auraient pu apparaître, notamment lors de l'utilisation de fonctionnalités d'analyse spatiale avec ces données, pour lesquelles la qualité de la géométrie est essentielle. Cette communication vise à souligner l'importance des étapes du contrôle qualité des données géoréférencées, non seulement lors de leur numérisation initiale mais aussi lors de leurs traitements et de leurs transferts vers diverses plateformes de diffusion et de partage de données.

**ABSTRACT.** In the framework of several collective research projects, a partnership with the Centre National de Préhistoire in Périgueux has achieved digitization at 1:10 000 scale of all flint formations in Dordogne department. This work, carried out with ArcGIS and QGis software, using 1:50 000 scale geological maps scanned and georeferenced, available on BRGM Infoterre web service, enabled the production of a unique database on georesources. Once digitized and uniformly characterized, flint formations are used to view current georesources and therefore to make hypothesis on supply strategies and movements of prehistoric populations. Data on flint formations were uploaded on ArcGISOnLine (AGOL), a mapping web platform of ESRI for sharing securely geo-referenced data. During this step, differences in geometry arose between original data and data after uploading on AGOL platform. A systematic quality control of digitization and transformation of files after their transfer to AGOL platform, has appeared necessary to ensure the data quality for project partners. If this control had not been made, errors of interpretation could have occurred, especially when using functionalities of spatial analysis with these data, for which the quality of geometry is essential. This communication aims to highlight importance of quality control of georeferenced data, not only during their initial digitalization but also throughout their treatments and transfers to various diffusion and sharing platforms.

**MOTS-CLÉS.** Silex, cartographie, SIG, numérisation, contrôle qualité.

**KEYWORDS.** Flint, mapping, GIS, digitalization, quality control.

<sup>1</sup> « Réseau de lithothèques en Rhône-Alpes », « Réseau de lithothèques en Auvergne », « Réseau de lithothèques en Grande Aquitaine » et « Réseau de lithothèques en Centre Val de Loire »

<sup>2</sup> www.arcgis.com

## 1. Cadre du projet et sa chronologie

Les résultats présentés ici sont le fruit de travaux conduits depuis une dizaine d'années entre un nombre croissant de chercheurs (préhistoriens, géologues, minéralogistes, géographes, etc.), engagés dans la connaissance des zones de formations à silex en France et dans la compréhension des processus physico-chimiques intervenant dans la formation des silex et dans l'évolution leur divers micro-faciès (Fernandes P., 2012, Fernandes P. et al., 2016).

Parmi les travaux réalisés pendant cette décennie de collaboration scientifique, on peut mentionner les étapes suivantes qui intéressent plus particulièrement notre propos :

- 2006 : premières cartes de localisation de gîtes à silex (Massif central, Alpes, vallée du Rhône)
- 2006-2010 : lancement du Projet Collectif de Recherche (PCR) « *Réseau de lithothèques en Rhône-Alpes* » dont les objectifs étaient et restent d'harmoniser les pratiques, les modèles, les méthodes descriptives
- 2011-2012 : harmonisation de plusieurs bases de données de plusieurs laboratoires
- 2013-2016 : élargissement des zones géographiques concernées par les travaux des divers PCR (régions PACA, Aquitaine, Centre Val de Loire, Midi-Pyrénées, etc.) qui ont fonction de mettre en commun et d'harmoniser les informations recueillies à l'échelle régionale
- 2016 : lancement de l'enquête nationale sur les lithothèques sous l'égide du Ministère de la Culture et de la Communication, Sous-Direction de l'Archéologie (SDA), du Centre National de la Préhistoire (CNP), du CNRS (instituts InEE et InSHS), de l'Inrap et de Paléotime dont l'une des retombées envisagées est la constitution d'un réseau national.

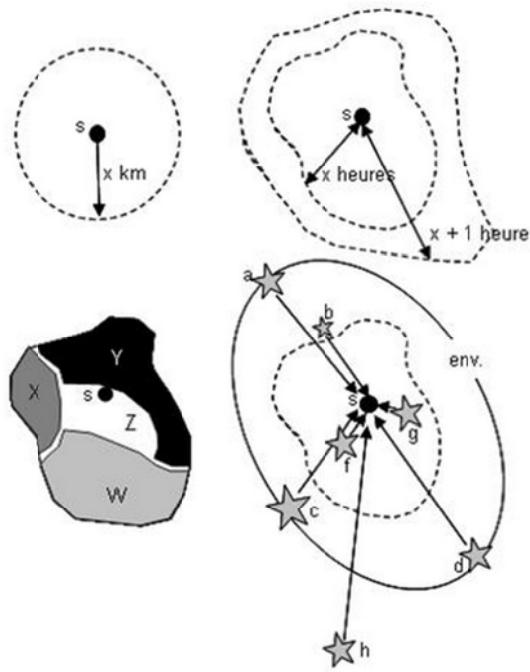
Ces différents travaux et les résultats auxquels ils ont permis d'aboutir, visent à répondre à une thématique de recherche impliquant des problèmes méthodologiques.

## 2. Thématique et problèmes méthodologiques

L'identification des sources d'approvisionnement en matières premières est une question déjà ancienne, qu'ont peut dater de l'apparition de la préhistoire comme discipline scientifique dans la deuxième moitié du XIXème siècle.

L'étude des espaces parcourus par les hommes durant la préhistoire s'appuie depuis longtemps et essentiellement sur l'identification des sources d'approvisionnement en matières premières utilisées pour la fabrication des séries lithiques trouvés dans les sites archéologiques ou en prospection.

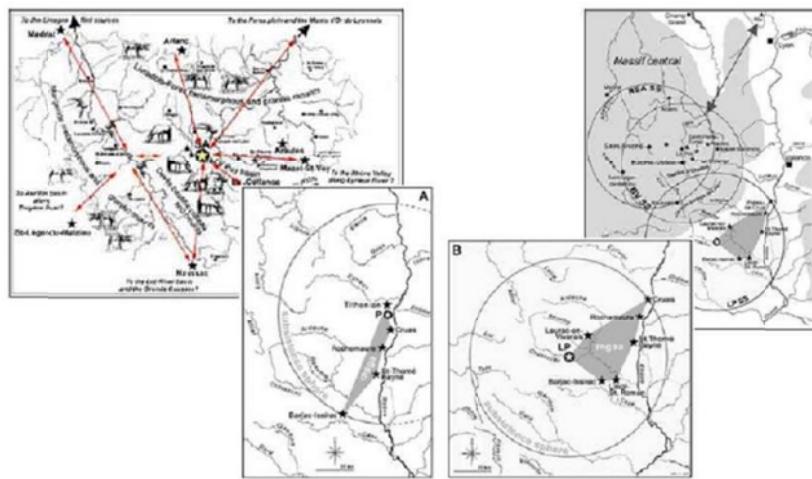
Ph. Boissinot (2015) a proposé de schématiser la partition des espaces accessibles (en distance et en temps) pour les populations du passé (figure 1). Ces espaces sont découpés selon des potentialités économiques supposés des différents secteurs d'approvisionnement en ressources (W, X, Y, Z) et des ressources exploitées dans plusieurs gisements archéologiques proches (a, b, c, d, e, f, g, h) du gisement principal S situé au centre.



**Figure 1.** Schéma de partition des espaces accessibles (en distance et en temps) pour les populations du passé (in . Boissinot Ph., 2015)

On peut retrouver dans cette figure, les principes du "site catchment analysis" proposés par des auteurs anglo-saxons, parmi lesquels de P. Binford (1980) qui a proposé de distinguer les chasseurs de type "foragers" et ceux de type "collectors", les premiers désignant des populations à forte mobilité mais avec des stratégies opportunistes d'improvisation dans leurs activités d'approvisionnement en ressources, les seconds se déplaçant moins souvent mais de façon davantage programmée et avec une stratégie de prévoyance.

Appliqués à des travaux réalisés sur les territoires d'approvisionnement des populations préhistoriques en matières premières dans le Massif-central, ces principes ont donné lieu à la production de plusieurs cartes (figure 2).



**Figure 2.** Cartes des zones théoriques d'approvisionnement en matières premières dans le Massif central pour le paléolithique moyen (in Raynal J-P. et al., 2010)

Sur ces cartes, les cercles tracés à partir des sites archéologiques délimitent les territoires compris dans une zone de 20 à 50 km à l'intérieur desquels les populations étaient censées s'approvisionner en matières premières.

Habituellement, la détermination des sources d'approvisionnement s'appuie sur la connaissance des géoressources : les lithothèques, la cartographie, les prospections de terrain, la consultation de la littérature spécialisée, les bases de données de référence.

Nous menons, depuis plusieurs années, un travail méthodologique pour aboutir à la mise en place d'un protocole adapté à une meilleure reconnaissance de la provenance du mobilier lithique. Dans ce but, nous avons formé le concept de « chaîne évolutive » qui élargit à la préhistoire les raisonnements sur le rapport entre silicifications et environnements (Fernandes 2012). Notre protocole est fondé sur l'analyse comparative des transformations rencontrées dans les échantillons géologiques et dans les objets archéologiques. L'enregistrement des caractères acquis et mémorisés nécessite des mesures systématiques. Ce protocole renouvelé implique à la fois démarche qualitative et quantitative, qui intègre les caractères évolutifs les plus discriminants. Il s'appuie sur la caractérisation des phases minéralogiques, l'analyse des éléments figurés et l'examen des états de surface. La prise en compte des évolutions prédépositionnelles et post-dépositionnelles participent donc à la reconnaissance des géoressources exploitées ainsi qu'à l'évaluation de l'intégrité du site dans une perspective taphonomique.

L'analyse des silex nécessite l'utilisation de modèles dynamiques afin de décrypter les étapes de la phase diagénétique et l'historique des processus post-génétiques. En fait, chaque environnement du silex a imposé une série de traits communs caractéristiques dans une unité paléogéographique déterminée. C'est sur le postulat de décryptage de la variabilité, qu'elle soit génétique et/ou post-génétique, qu'est fondée notre conception de la pétroarchéologie du silex.

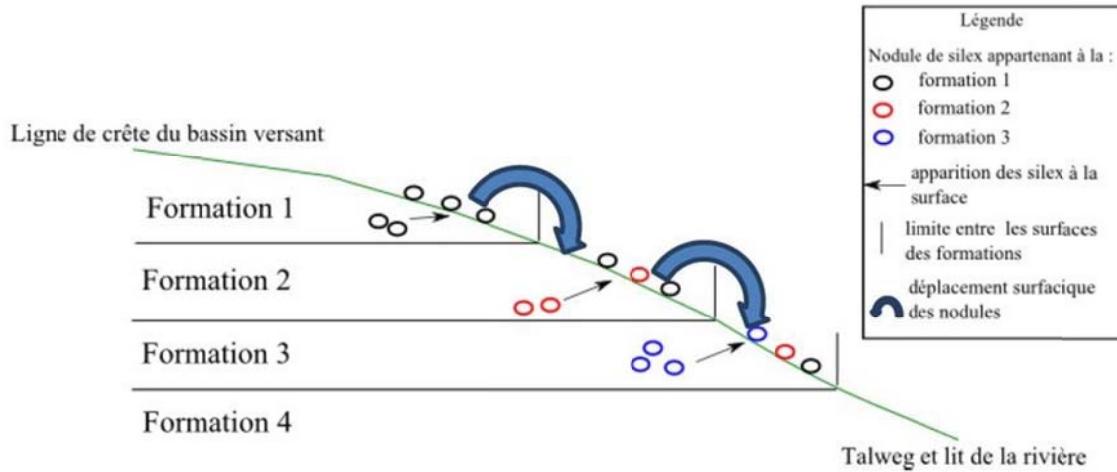
L'étude et la description des matériaux géologiques s'effectuent à trois échelles complémentaires :

- échelle macroscopique : vestige archéologique ou échantillon géologique,
- échelle mésoscopique : microscope optique,
- échelle microscopiques : microscope électronique à balayage.

Ces nouveaux outils de lecture, qui prennent en compte les processus génétiques et l'histoire post-génétique des silex, sont couplés aux méthodes de caractérisation structurale (diffractométrie X, Raman) et géochimique (analyses élémentaires), ce qui permet une approche considérablement affinée des sources. Leur détermination s'appuie désormais aussi sur une caractérisation pétrologique, tenant compte de l'évolution physico-chimique des matériaux géologiques et archéologiques.

L'étape suivante consiste en la confrontation entre assemblages archéologiques et référentiels géologiques. Cette confrontation permet d'émettre ensuite des hypothèses sur des logiques d'approvisionnement qui nécessitent néanmoins toujours une grande prudence dans leur interprétation.

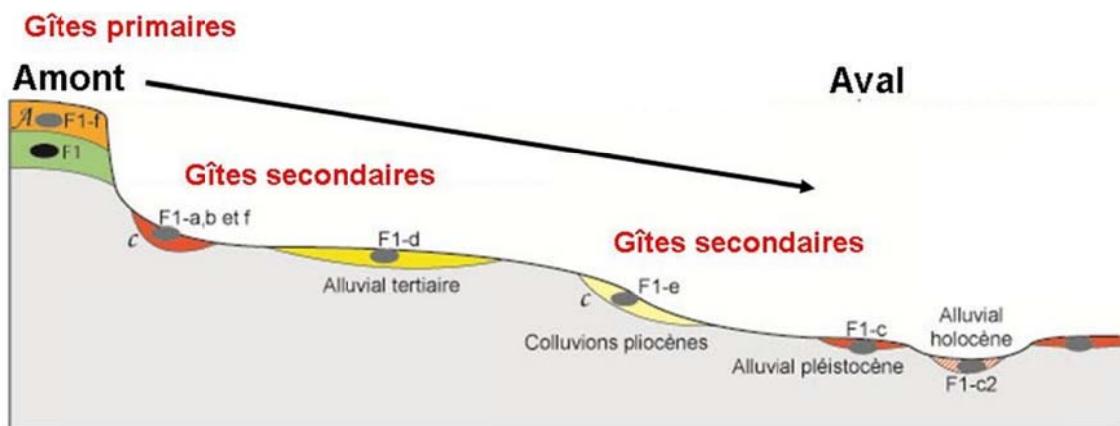
Concernant la caractérisation spatiale, les processus d'évolution des matières siliceuses (ou concept de Chaine évolutive) interviennent dans trois types de sites (gîtes primaires, gîtes secondaires et sites archéologiques). Une cartographie basée sur la position des formations actuelles doit prendre en compte l'évolution des reliefs, ce qui implique une étude géomorphologique pour chaque gîte secondaire étudié. Notre expérience montre que la prise en compte de l'évolution du réseau hydrographique et des bassins versants est essentielle. Les matériaux présents dans ce type de formation fluviatile témoignent des polarités dynamiques (types d'itinéraires parcourus par les silex) dans la distribution entre les familles génétiques (silex en position primaire) et les variantes gîtologiques (silex issus de l'évolution des types génétiques en position secondaire). Le but final est de constituer des cartes sur lesquelles chaque gîte primaire est en relation avec l'ensemble des formations qu'il alimente. Chaque gîte secondaire est considéré comme le maillon d'une chaîne évolutive, de l'encaissant vers les gîtes les plus éloignés.



**Figure 3.** Déplacement schématisé des silex le long d'une pente

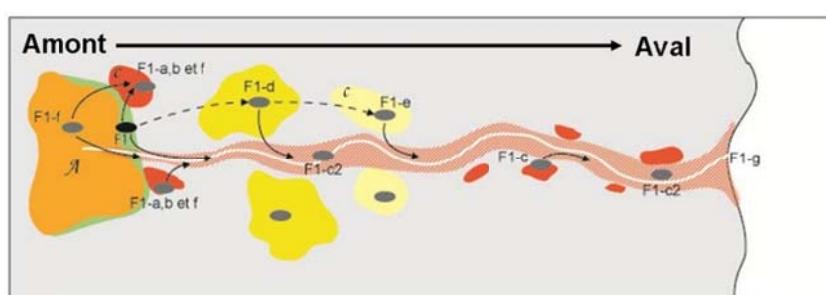
Les déplacements le long d'une pente des silex provenant de différentes formations qui en comportent peuvent être schématisés (figure 3). On constate une accumulation de l'amont à l'aval des nodules de silex provenant de plusieurs formations successives. Dans la formation la plus en aval, on peut retrouver des nodules provenant des différentes formations se trouvant en amont, ce qui n'est pas le cas des formations les plus en amont.

Dans un secteur (autour de Fumel) de la vallée du Lot, ces principes théoriques ont pu être identifiés sur le terrain (figure 4), en s'appuyant sur la distinction entre gîtes primaires et gîtes secondaires.



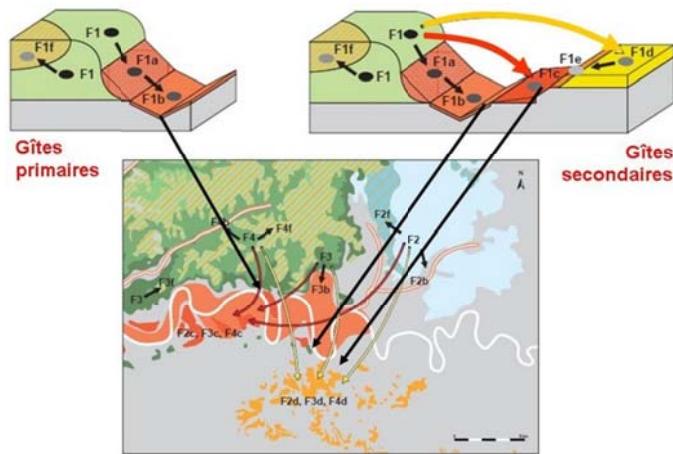
**Figure 4.** Déplacement schématisé des silex le long d'une pente

Cette distinction entre les différentes étapes de la chaîne évolutive des silex de l'amont vers l'aval se retrouve en planimétrie, dans les différents types de formations correspondant aux divers types génétiques et à leurs évolutions respectives (figure 5).



**Figure 5.** Positionnement en plan des divers types génétiques et de leurs types évolutifs respectifs

Dans une même vallée (figure 6), on peut observer des transports naturels relativement simples, alors que dans d'autres secteurs on peut voir des transports successifs entre gîtes primaires et gîtes secondaires qui sont plus difficiles à interpréter, notamment par rapport à la topographie actuelle qui a été largement modifiée du fait de l'érosion fluviale.



**Figure 6.** Les déplacements des matières siliceuses de part et d'autre du thalweg actuel d'une vallée

Pour pouvoir émettre des hypothèses recevables quant aux scénarios possibles de mise en place des divers types gîtologiques c'est-à-dire ayant évolué au cours de leur itinéraire post-génétique, il convenait de renforcer la qualité des observations faites sur ces matières aux différentes échelles indiquées : macro, méso, micro.

Les informations disponibles dans les lithothèques devaient permettre d'alimenter ces travaux. Or force fut de constater que les lithothèques constituées en France, l'avaient été de façon non cordonnée pour des raisons d'émergence et de développement hétéroclite de ce type de référentiel.

Lors de la confrontation de plusieurs lithothèques, divers problèmes méthodologiques sont apparus, parmi lesquels :

- une diversité des pratiques et des modalités de description des formations à silex *in situ*, des échantillons géologiques, des pièces lithiques,
- des données disparates et souvent inadéquates pour répondre aux questionnements archéologiques,
- un cloisonnement inutile des initiatives dû aux limites administratives,
- l'absence de prise en compte de la notion de chaîne évolutive.

C'est pourquoi il est apparu indispensable de viser deux objectifs majeurs :

- améliorer le niveau de qualité de la précision spatiale de la connaissance des formations à silex,
- améliorer l'interopérabilité entre les corpus de données, de dictionnaires de données et parfois de métadonnées,
- concevoir des lithothèques intégrant des données dynamiques (l'altération post-génétique) jusqu'à lors sous exploitées.

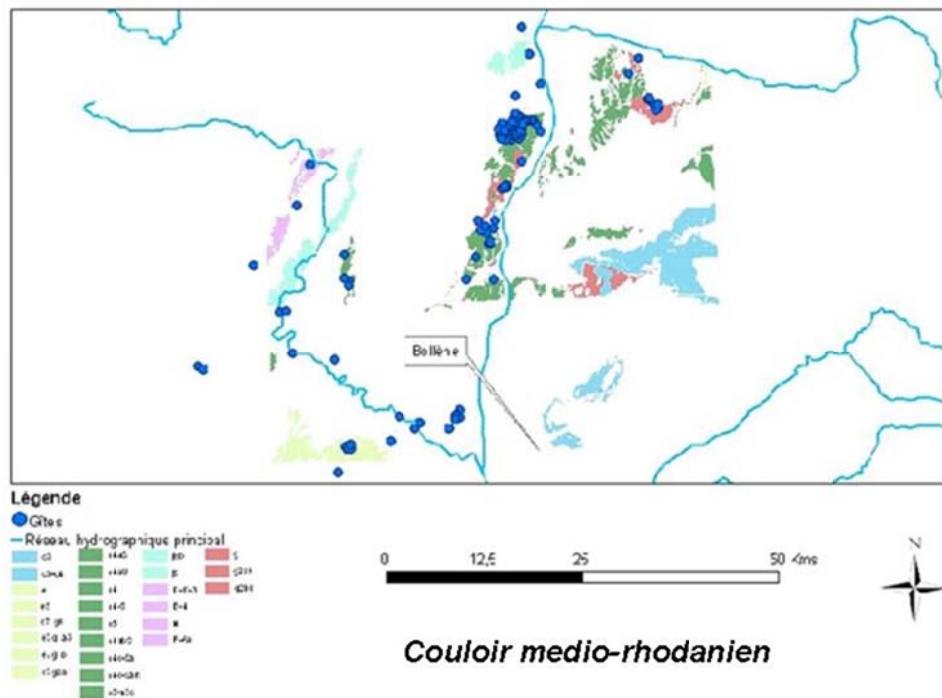
### 3. Les travaux réalisés dans ce cadre

Les travaux des PCR réseaux de lithothèques ont précisément permis d'aboutir à plusieurs résultats :

- des cartes de localisation des lieux de collecte de silex,

– une carte des limites des principales formations à silex dans 6 régions du sud de la France (Aquitaine, Auvergne, Languedoc-Roussillon, Midi-Pyrénées, Provence-Alpes-Côte-d'Azur, Rhône-Alpes).

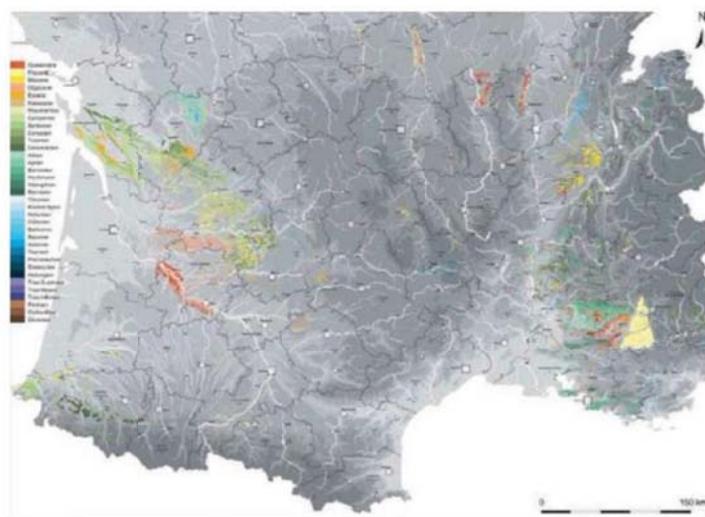
Les premiers travaux cartographiques conduits avec le PCR en région Rhône-Alpes, datent de 2006. Réalisés avec ArcGIS, ils ont consisté à reporté la localisation des gîtes à silex en liaison avec les formations géologiques ayant pu les alimenter (figure 7).



**Figure 7.** Premières cartes de localisation de gîtes à silex dans le couloir médio-rhodanien (Fernandes P. et Tufféry C., 2005)

A ces premiers travaux ont succédé ceux qui ont permis d'aboutir à la production d'une carte réalisée avec Adobe Illustrator (figure 8) à partir d'une lecture interprétative des cartes géologiques du BRGM scannées, géoréférencées et mosaïquées dans ArcGIS (Fernandes P. et al. 2013).

Mais les conditions de production de ces cartes limitaient leur utilisation pour des travaux systématiques à grande échelle.

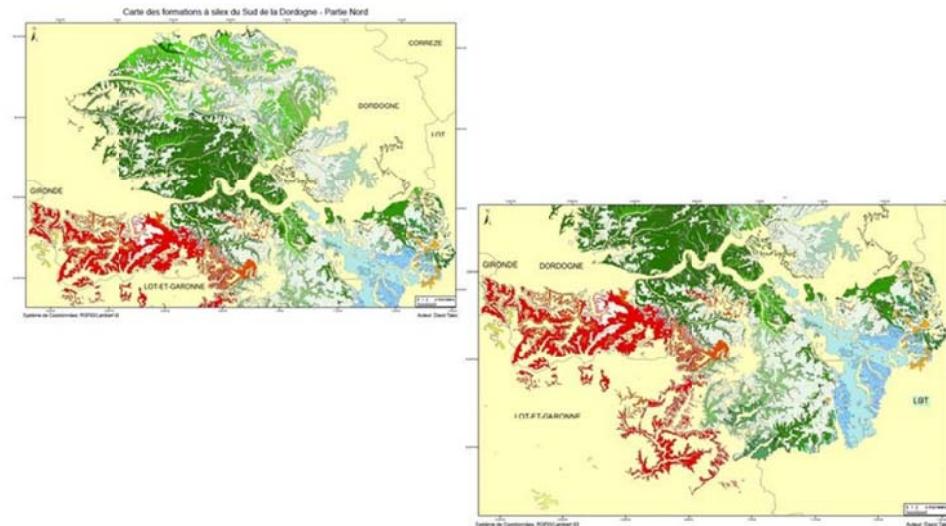


**Figure 8.** Premières cartes de localisation de gîtes à silex pour la France méridionale (Fernandes P., et al., 2012)

Après ces premiers travaux qui ont produit des cartes statiques produites par un seul chercheur et consultable que par un faible nombre d'autres, il a été décidé de réaliser des cartes interactives et évolutives, produites par un collectif de chercheurs et rendues disponibles en ligne.

Les contours des formations à silex ont alors été numérisées, avec les logiciels QGis et ArcGIS, en utilisant les cartes géologiques au 1/50 000 du BRGM disponibles sur le webservice WMS Infoterre du BRGM. La numérisation a été effectuée selon un protocole homogène consistant à les afficher à l'écran au 1/10 000. Puis des attributs descriptifs sur la géologie et sur la géographie ont été associés aux polygones des formations numérisées.

Après avoir été mise au point pour la région Rhône-Alpes puis pour la région PACA, ce même protocole a été utilisé en 2015 pour le département de la Dordogne (figure 9), dans le cadre d'un travail co-financé et encadré par le CNP (Talec D., 2015).



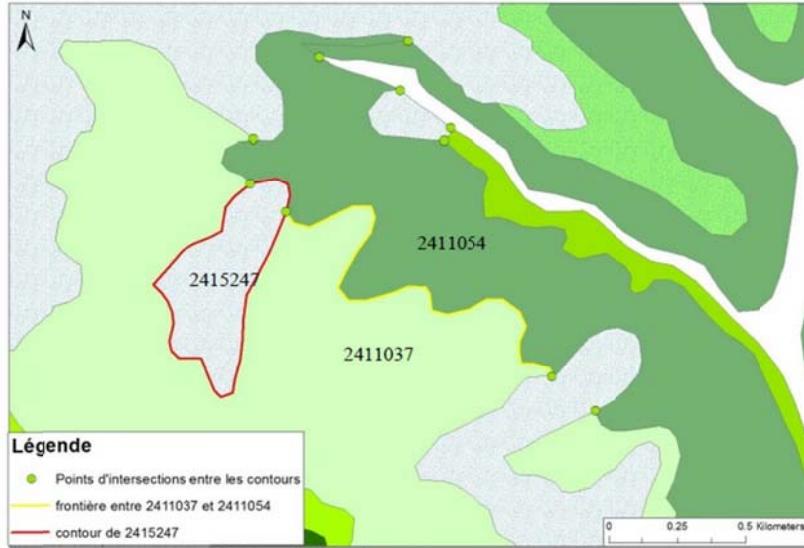
**Figure 9.** Carte des formations à silex de Dordogne (parties nord et sud) (Talec D., 2015)

Une fois numérisées, les données ont été stockées, traitées et mises à jour dans le système de gestion de base de données PostgreSQL et dans son module de gestion de données spatiales PostGIS. Dans le but de modéliser la notion de chaîne évolutive sous SIG des scripts ont été développés notamment pour le parcours de graphe sur les réseaux hydrographiques ou encore dans la recherche des formations contigües (Fernandes P. et Dalphinet D., 2014). Ces scripts font intervenir les fonctions de récursivité pour constituer la dynamique et les liens qui unissent les formations.

Lors de ces développements, le problème de la qualité des données est apparu crucial, que ce soit dans la qualité de numérisation des contours mais aussi dans la pertinence de l'attribution des données descriptives.

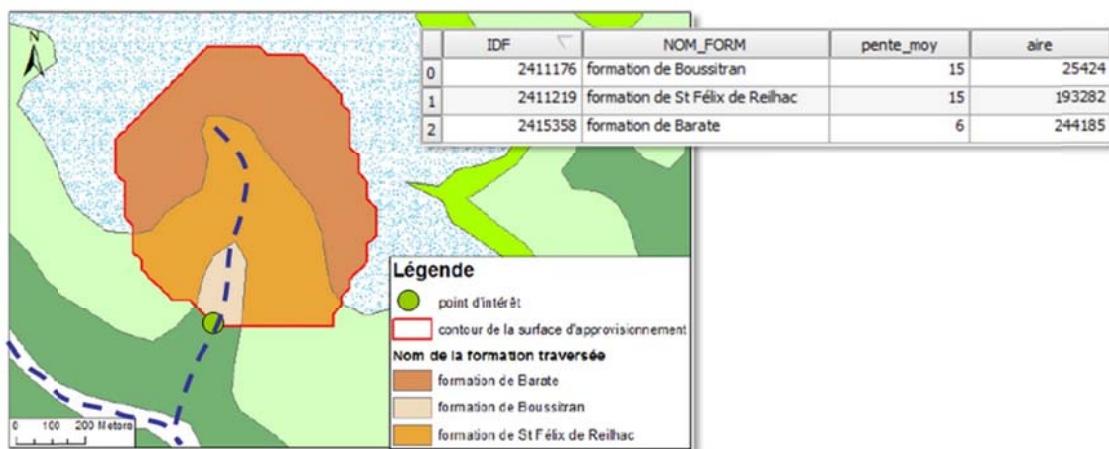
#### 4. La qualité des données et diffusion

Une topologie de qualité s'est avérée indispensable pour pouvoir aborder l'étude des relations spatiales entre les polygones de formations à silex. C'est le cas, notamment, lorsqu'il s'agit d'identifier les frontières entre formations contigües (figure 10).



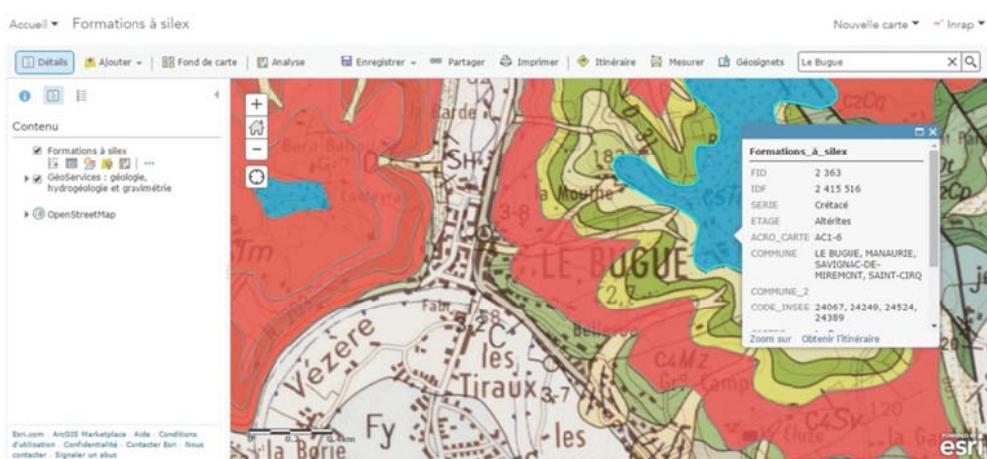
**Figure 10.** Identification des frontières entre formations à silex contigües

Il en est de même pour le calcul de la surface théorique d'approvisionnement en silex à partir d'un point d'intérêt dans un bassin-versant dans lequel se trouvent des formations siliceuses (figure 11).



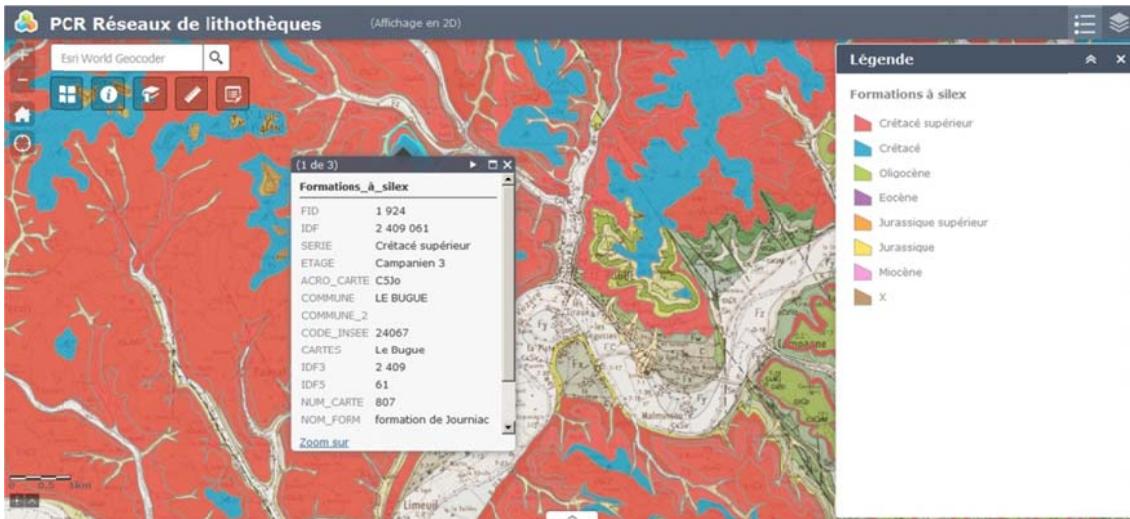
**Figure 11.** Identification de la surface théorique d'approvisionnement en silex à partir d'un point d'intérêt dans un bassin-versant

Afin de permettre la diffusion et le partage des données, des tests ont été réalisés avec la plateforme ArcGisOnLine (AGOL) d'ESRI (figure 12).



**Figure 12.** Affichage dans AGOL des polygones des formations à silex et des données attributaires associées

Une fois importées dans AGOL, il est possible de développer une application de webmapping grâce à des modèles disponibles et l'usage de boutons et de fonctionnalités associées (widgets), très simples à insérer dans l'interface du site de consultation (figure 13).



**Figure 13.** Affichage dans AGOL des polygones des formations à silex et des données attributaires associées

A l'issue de l'importation des polygones des formations à silex dans AGOL, nous nous sommes aperçus que les données avaient subi des transformations de leur géométrie. En effet, les contours des polygones des formations avaient été "généralisés", c'est-à-dire simplifié. Dans ce cas, le nombre de vertex des contours des polygones est diminué, mais d'une façon qui n'est pas systématique. Cela dépend de la forme et du nombre de vertex par polygone.

Cette généralisation est l'option cochée par défaut dans AGOL lors de l'importation. Pour éviter la généralisation, il faut veiller à cocher l'option « Conserver les entités d'origine » lors de l'importation dans AGOL. De plus, il faut modifier la limite par défaut des 1000 entités au maximum qu'il est possible d'importer.

L'identification de la modification des contours des polygones a été réalisée grâce à la fonction de vérification de topologie du logiciel QGis. Cette extension, utilisée pour ce projet dans la version 2.8 de QGis, s'installe dans le menu Vecteur de QGis. Cette extension a été complétée depuis la version 2.12 par l'extension "Vérificateur de géométrie" qui enrichit grandement les possibilités de vérification et qui s'installe dans les outils de géométrie du logiciel.

Le vérificateur de topologie de QGis propose 5 options de vérification :

- Doublons (2 polygones de géométries identiques).
- Polygones multi-parties (2 polygones distincts mais ne formant qu'une seule entité dans la table attributaire).
- Polygones à géométrie invalide.
- Superpositions entre polygones.
- Trous.

Avec l'option "Trous", le vérificateur a détecté 408 "erreurs".

Une vérification systématique visuelle de toutes ces "erreurs" a été faite. Elle s'est imposée. Certes le développement d'un script de nettoyage aurait pu être développé. Mais celui-ci aurait pu identifier par erreur des trous là où ils existent effectivement sur les cartes géologiques du BRGM. Sur le terrain

ce sont de véritables polygones inclus dans les polygones des formations où le silex n'apparaît plus en surface.

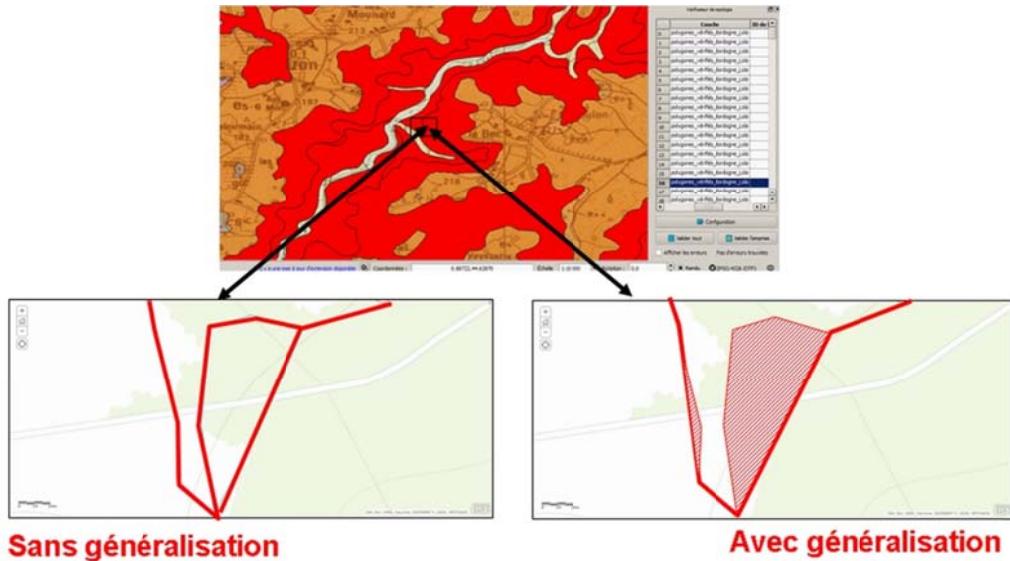
Dans ce cas, il s'agit de "vrais trous", justifiés qui sont donc de "fausses erreurs" et qu'il faut conserver.

A l'inverse, il peut y avoir aussi des polygones supprimés qui résultent de la limitation du nombre de nœuds par polygones (ce qui est le principe de la généralisation). Dans ce cas, il faut retrouver les polygones injustement supprimés.

C'est en zoomant sur les "trous" identifiés qu'on peut s'apercevoir que les contours de certains des polygones ont été modifiés lors de leur importation dans AGOL.

Ces "trous" sont des petits polygones créés en limites de polygones plus grands. Il faut les supprimer et revenir à la topologie d'origine.

Dans ce cas, les "faux polygones" sont de "vraies erreurs" du vérificateur de topologie (figure 14).



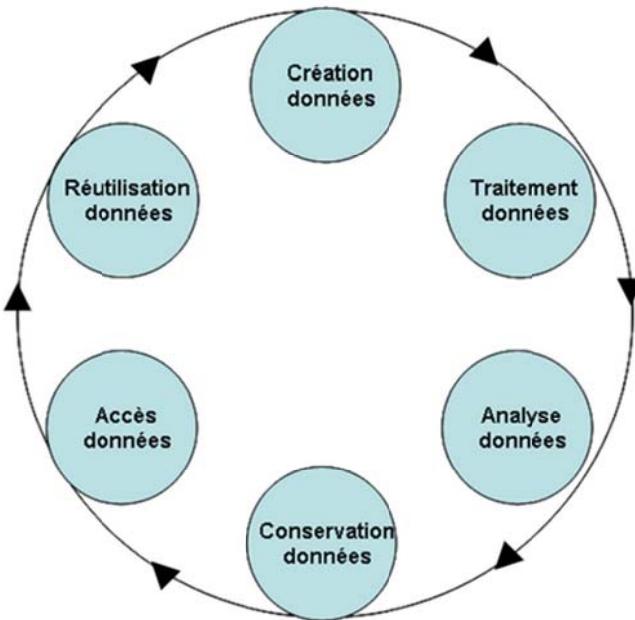
**Figure 14.** Résultat de la généralisation des contours de polygones contigus de formations à silex

Les conséquences des erreurs de topologie sur la géométrie et sur les possibilités d'utiliser des fonctionnalités d'analyse spatiale sont multiples :

- des données au niveau de qualité hétérogène,
- des relations spatiales artificielles, tronquées, disparues entre les polygones des formations à silex,
- des erreurs dans la recherche des contiguités entre polygones,
- des erreurs dans l'utilisation des scripts de parcours de graphe du réseau hydrographique par rapport aux formations à silex.

De ces exemples, on peut distinguer quelques préconisations générales sur la qualité des données.

En archéologie, comme tous les domaines de la recherche, la connaissance de la qualité des données s'inscrit dans la gestion des données. Cette gestion doit s'appuyer sur un Plan de Gestion des Données (PGD) ou *Data management plan* (DMP). Un PGD doit prendre en compte toutes les étapes du cycle de vie des données qui comporte 6 grands domaines d'activité : Création, Traitement, Analyse, Conservation, Accès, Réutilisation (figure 15)



**Figure 15.** Cycle de vie des données (d'après Corti L. et al., 2014)

Le PGD définit ce que les chercheurs feront de leurs données pendant et après le projet, explicitant notamment la mise à disposition des données (cf. l'initiative pilote pour le libre accès aux données de la recherche qui oblige les projets financés sur le programme Horizon 2020 de disposer d'un PGD).

Il existe deux types de qualité des données sur les données géoréférencées (CERTU, 2010).

- une qualité interne : liée aux spécifications de contenus qui doivent être définies en amont qui doit conduire à s'interroger sur "Que doit-on numériser et comment ?";
- une qualité externe : liée aux spécifications des besoins d'utilisation en aval et qui doit faire se demander "Quels usages sont possibles des données numérisées ?".

Le contrôle de la qualité de la numérisation des données archéologiques géoréférencées devrait répondre à trois exigences principales :

- il devrait s'agir d'une activité indispensable tout au long du processus d'acquisition et de traitement des données ("bonne pratique"),
- ce processus devrait être décrit dans les métadonnées,
- il devrait s'appuyer sur les normes existantes dans le domaine de l'information géographique.

De façon plus générale, la qualité des données se trouve à l'interface entre quatre domaines :

- l'exactitude,
- la complétude,
- la fraîcheur,
- la cohérence.

Les problèmes rencontrés dans l'usage des normes traitant de la qualité des données archéologiques géoréférencées, sont communs à toutes les données géoréférencées. Ces normes sont nombreuses et souvent difficiles et fastidieuses à appliquer. Elles sont très peu ou très mal connues, notamment en raison du manque de formations sur le sujet.

Pour faire face à ces difficultés, il conviendrait :

- d'ouvrir le dialogue entre chercheurs, producteurs, utilisateurs...mais aussi juristes, techniciens, grand public,
- de mettre en place des actions d'accompagnement (formation, sensibilisation, séminaires, ateliers, et... des communications et des publications sur le sujet etc.), notamment sur les métadonnées permettant de décrire la qualité des données.

A côté des actions nécessaires à la sensibilisation au problème de la qualité des données, il faut aussi que des actions soient conduites dans le domaine des métadonnées qui comprend, notamment, des informations sur la qualité des données elles-mêmes.

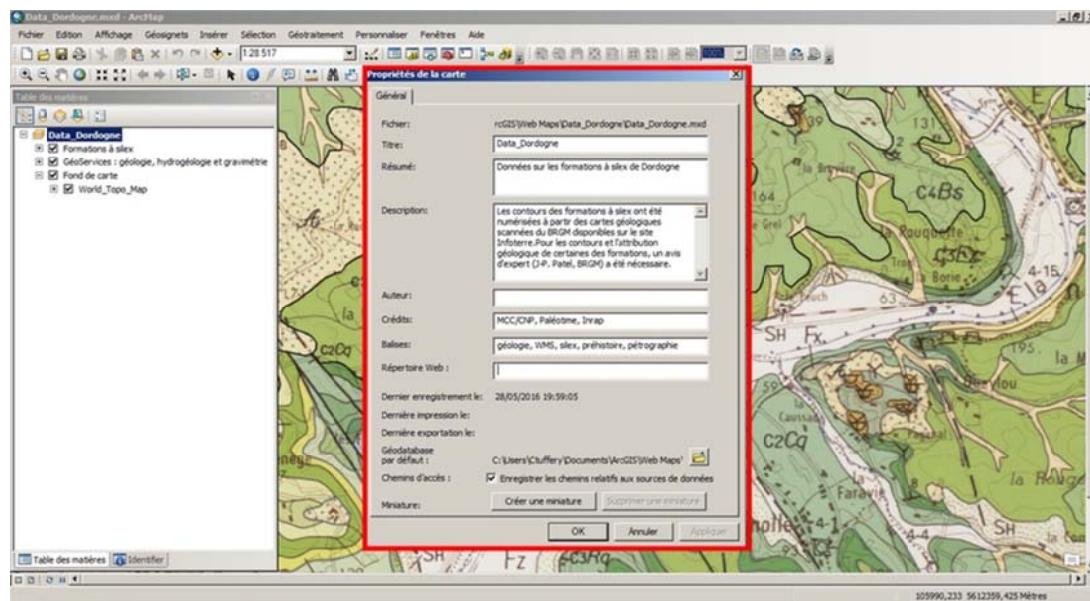
## 5. L'importance des métadonnées

Les métadonnées sont indispensables pour permettre aux données d'être correctement diffusées entre les partenaires d'un projet collectif de recherche présentés ici, en transportant avec la carte toutes les descriptions définies dès le PGD.

Les métadonnées permettent aux données d'être publiées et moissonnées sur les webservices dédiés aux données archéologiques qui se multiplient ces derniers temps, à l'échelle française ou européenne (cf. projet ARIADNE).

L'intérêt final de ces webservices est de permettre aux données scientifiques ainsi publiées et moissonnées, d'alimenter le web sémantique.

La constitution des métadonnées doit s'effectuer dès les phases amont de la production des couches de données géoréférencées. Pour cela, les logiciels ArcGIS et QGis proposent tous les deux des outils facilitant la saisie des métadonnées (figure 16).



**Figure 16.** Panneau dans ArcGis permettant la saisie des métadonnées

Dans AGOL, il existe un premier niveau de description de la ressource ; ce sont les métadonnées minimales (figure 17) :

- Description.
- Contraintes d'accès et d'utilisation.
- Contenu des couches de la carte.

## – Propriétés.

The screenshot shows the AGOL metadata panel for a feature layer named 'Formations à silex'. The top navigation bar includes links for Accueil, Bibliothèque, Carte, Scène, Groupes, Contenus, and Organisation. A search bar and a user dropdown are also present. The main content area has a blue header 'Formations à silex'. Below it, there's a map thumbnail with a legend indicating 'Feature Layer (Hébergé) par INRAPESRI', 'Source : Feature Service', 'Dernière modification : 2 juin 2016', and '0 évaluations, 0 vues'. A series of buttons for actions like OUVRIRE, PARTAGER, MODIFIER, EFFACER, PUBLIER, DÉPLACER, CHANGER DE PROPRIÉTAIRE, UTILISATION, and REMPLACER are shown. Below these are sections for 'Description', 'Contraintes d'accès et d'utilisation', 'Contenu de la carte', and 'Propriétés'. The 'Propriétés' section contains detailed information about the resource, including its sharing status, credits to MCC/CNP, Paléodime, and Inrap, and its dimensions (9 MB). At the bottom, there are buttons for Calcul d'itinéraire, Outil de mesure, Sélecteur de fonds de carte, and Recherche par adresse.

**Figure 17.** Panneau dans AGOL permettant l'affichage des métadonnées

Mais AGOL propose aussi de décrire les données selon plusieurs normes de métadonnées dont la norme ISO 19139 qui comprend de nombreux descripteurs, organisés de façon hiérarchique en rubriques et sous-rubriques (figure 18). Par exemple, dans les rubriques décrivant la ressource, apparaissent les sous-rubriques sur les titres et dates, les URL pour accéder aux données, les identifiants, la forme, les contacts, etc.

The screenshot shows the AGOL data entry tool for ISO 19139 metadata. The window title is 'Formations à silex'. The top menu bar includes Valider, Enregistrer, Enregistrer et fermer, Enregistrer une copie locale, Ouvrir, Effacer, and Fermer. Below the menu is a toolbar with buttons for Ressource, Qualité, Répartition, Représentation, contenu, Champs, Référence, and Métadonnées. The 'Qualité' tab is selected. A sub-section titled 'Qualité des données' is open, showing a 'Domaine d'applicabilité' section with 'Rapport' selected. Under 'Type de rapport', 'Cohérence topologique' is chosen. The 'Mesurer' tab is active, showing a 'Dimension' section with 'Horizontale' selected. A 'Description de la mesure' text area contains the following text: 'Vérification de la topologie du fichier des polygones pour distinguer les vrais polygones extrudés (absence de la formation en surface) et les faux polygones (résultant de la généralisation lors de l'importation du fichier shape dans ArcGisOnLine.)'. At the bottom, there's a 'Nom de la mesure' field with icons for file, folder, and cloud.

**Figure 18.** Outil de saisie dans AGOL des métadonnées selon la norme ISO 19139

Dans ces rubriques sur la qualité des données, il faut fournir les informations décrivant les mesures effectuées sur la qualité, l'évaluation de la qualité qui a été faite et les résultats des mesures sur cette qualité. Il peut y avoir plusieurs types de rapport (ex. Cohérence topologique).

## 6. Conclusion

La connaissance et la caractérisation des formations à silex dans le sud de la France présenté ici, s'appuient sur le travail de nombreux chercheurs d'horizons très variés (archéologues, préhistoriens, géologues, géographes, etc.) dont les apports et les compétences sont complémentaires depuis une dizaine d'années.

La constitution, selon un protocole harmonisé, de bases de données sur les formations à silex ouvre la voie à la comparaison des corpus de données descriptives sur ces formations, longtemps restés difficiles à confronter du fait d'une hétérogénéité dans leurs modes de constitution et des descripteurs utilisés.

Le partage des données et des métadonnées devient dorénavant possible, même si un travail important reste à faire pour consolider l'approche proposée. Celle-ci présente néanmoins l'intérêt d'offrir un cadre renouvelé pour une approche véritablement pluridisciplinaire sur le silex comme ressource essentielle et pérenne (puisque parvenue jusqu'à nous) des populations du passé. Il s'agit de mieux comprendre leurs logiques d'approvisionnement et leurs déplacements sur leurs territoires, en ayant pu mettre en connexion toutes les formations à silex.

La constitution des bases de données, si elle devient de plus en plus aisée à mettre en œuvre, ne doit pas faire l'impasse sur plusieurs aspects. Notamment elle doit intégrer la connaissance et la gestion de la qualité des données en s'appuyant sur la production et le maintien des métadonnées associées. Il en va de la crédibilité de la recherche effectuée et de ses possibilités de partage et de diffusion.

## 7. Bibliographie

- Binford P., 1980 – Willow Smoke and Dogs' Tails: Hunter-Gatherer Settlement Systems and Archaeological Site Formation, *American Antiquity* 45:420.  
[https://anthropology.boisestate.edu/wp-content/uploads/2010/06/article10\\_willow-smoke-and-dogs-tails.pdf](https://anthropology.boisestate.edu/wp-content/uploads/2010/06/article10_willow-smoke-and-dogs-tails.pdf)
- Boissinot Ph., 2015 – Qu'est-ce qu'un fait archéologique ? Ed. EHESS, volume : 57, 365 pages (Collection : En temps & lieux)
- CERTU, 2010 – La qualité des données géographiques. État des lieux pour un débat. Rapport d'étude du CERTU, 32 pages. <http://www.certu-catalogue.fr/la-qualite-des-donnees-geographiques.html>
- Corti L., Van den Eynden V., Libby Bishop L., Woppard M., 2014 – Managing and sharing research data : a guide to good practice. SAGE, 240 pages.  
<https://www.ukdataservice.ac.uk/manage-data/handbook>
- Fernandes P., 2012 – Itinéraires et transformations du silex : une pétroarchéologie refondée, application au Paléolithique moyen. Thèse de l'université de Bordeaux 1 sous la direction de Jean-Paul Raynal soutenue le 12 juin 2012. 2 vol., 623 pages  
[http://ori-oai.u-bordeaux1.fr/pdf/2012/FERNANDES\\_PAUL\\_2012.pdf](http://ori-oai.u-bordeaux1.fr/pdf/2012/FERNANDES_PAUL_2012.pdf)  
[http://www.prehistoire.org/offres/file\\_inline\\_src/515/515\\_P\\_40036\\_578298164ec72\\_6.pdf](http://www.prehistoire.org/offres/file_inline_src/515/515_P_40036_578298164ec72_6.pdf)
- Fernandes P. et Tufféry C., 2005 – Utilisation d'ArcGIS 9 et d'un SGBD pour la recherche de corrélations spatiales entre sites archéologiques et gîtes de matériaux siliceux au Paléolithique moyen dans le Massif central. Conférence annuelle francophone d'ESRI France, 5-6 octobre 2005 <http://www.esrifrance.fr/sig2005/communications2005/eheess/eheess.htm>
- Fernandes P., Tufféry C., Raynal J.-P., Piboule M., Séronie-Vivien M., Séronie-Vivien M.-R., Turq A., Morala A., Tallet P., Affolter J., Millet D., Millet F., Bazille F., Schmidt P., Foucher P., Delvigne V., Liagre J., Gaillot S., Morain A., Moncel M.-H., 2013 –Une carte et une base de données pour les formations à silex du sud de la France : un outil pour la pétroarchéologie, PALEO [En ligne], 24 | 2013, mis en ligne le 17 avril 2014, consulté le 01 novembre 2016. URL : <http://paleo.revues.org/2633>
- Fernandes P. et Dalphinet D., 2014 – Vers une cartographie dynamique des géoressources en silex dans le sud de la France, contribution à l'étude des comportements humains au cours de la Préhistoire, Communication à la Conférence SIG 2014, Versailles, 1-2 octobre 2014

Fernandes P., Tufféry C., Binder D., Leandri-Bressy C., Bracco J-P., Tallet P., Morala A., Turq A., Davtian G., Caverne J-B., Dalphinet D., Delvigne V., Liagre J., Gaillot S., Millet D., Millet F., Piboule M., Picavet R., Schmidt P., Tomasso A., Affolter J., Bazile F., Garnier J-F., Bintz P., Pinçon G. et Raynal J-P., 2016 – Les formations à silex dans le Sud de la France : élaboration en multipartenariat d'une base de données géoréférencées, premiers résultats, in Ressources lithiques, productions et transferts entre Alpes et Méditerranée. Actes de la journée de la Société Préhistorique Française. Nice, 28-29 mars 2013, Textes publiés sous la direction de Antonin Tomasso, Didier Binder, Gabriele Martino, Guillaume Porraz, Patrick Simon et Nicolas Naudinot, Paris, Société préhistorique française, 2016 (Séances de la Société préhistorique française, 5), p. 137-150

Raynal J-P. et Moncel M-H., 2010 – Matières et cultures au Paléolithique moyen : un retour sur expérience. Les journées silex de Lyon. Vers une harmonisation méthodologique. Octobre 2010, Lyon, France. <hal-00660526>

Talec D., 2015 – Intégration des phénomènes de diffusion dans la cartographie. Rapport de Stage CNAM-ESGT, 40 pages.