
L'Atlas Historique du Limousin. Questions et méthodes vers un système de données spatio-temporelles contributif

Juliette Morel

*Criham, Université de Limoges
Faculté des Lettres et Sciences Humaines
Campus Vanteau, 39^E rue Camille Gérin, 87036 Limoges, France
Juliette.morel@unilim.fr*

Nous proposons de présenter à l'atelier SAGEO HumaNS'2018 l'Atlas Historique du Limousin (AHL), et plus précisément les enjeux méthodologiques impliqués par le développement d'un outil de contribution directe pour l'alimentation de la base de données spatio-temporelle de cet Atlas historique.

L'Atlas Historique du Limousin est un projet mené au Laboratoire Criham de l'Université de Limoges et a bénéficié depuis 2014 des financements de la Région Limousin et de l'Institut de Recherche des Sciences de l'Homme et de la Société de l'Université de Limoges. Il a pour objectif de rassembler des données spatio-temporelles et des sources historiques sur l'histoire du Limousin, et de proposer un outil d'interrogation spatiale, temporelle et/ou thématique sur le web pour les représenter et les comparer. L'Atlas a été mis en ligne en 2015 et agrémenté d'une interface de *webmapping* sur l'histoire de Limoges en 2017. Ses perspectives d'évolution concernent actuellement le mode d'alimentation de la base de données géo-historique. L'objectif est de promouvoir la contribution directe de différents acteurs grâce aux outils du web 2.0 pour permettre une croissance importante de la quantité de données et donc un potentiel de croisements élargi et inédit. Nous entendons solliciter des contributions de deux types : d'une part des contributions spécialisées liées aux champs d'expertise des collaborateurs plus ou moins directs de l'AHL, et d'autre part des contributions du grand public sur des aspects précis et encadrés. Cette ouverture à la contribution engendre des questionnements méthodologiques, techniques, voire déontologiques. Ces questionnements et les statistiques présentés dans la suite de l'article se fondent notamment sur une enquête menée en 2018 auprès des collaborateurs de l'AHL, contributeurs spécialisés

potentiels (contributions de premier type), pour évaluer les besoins et les problèmes liés au développement d'outils contributifs¹.

TABLE 1 : Récapitulatif du contexte de développement des outils contributifs de l'AHL

Contributeurs potentiels	Historiens du Criham (23% du total des contributeurs spécialisés) ou d'autres laboratoires de recherche (11%) ; Archéologues opérant en Limousin (8%) ; Sociétés savantes (15%) ; Dépôts d'archives et collectivités territoriales (archives municipales et départementales, ville de Limoges, etc.) (27%) ; le grand public.
Données	Des données et sources de données hétérogènes localisée (90%) et datées (100%) : littérature grise, plans anciens, photographies, descriptions littéraires et documentaires, données archéologiques, données issues de recherches historiques, données statistiques, démographiques, etc.
Implémentations réalisées ou prévues	1 ^{er} étape : Une base de données relationnelle spatio-temporelle et documentaire et une plateforme d'interrogation web (réalisée et en ligne) ; 2 ^e étape : des outils de contribution/versement automatique de données spatio-temporelles dans la base de données (en cours). 3 ^e étape : intégration des données à un entrepôt de données intégré au web des données (à développer).
Média de visualisation et consultation	Un moteur de recherche spatial, temporel et thématique. Webmapping et datavisualisation dynamique pour la visualisation des sélections (réalisé et en ligne. À rendre automatique / à adapter à l'intégration automatique de données).

1. Hétérogénéité géographique et qualité disparate des données historiques collectées

D'après l'enquête menée auprès des contributeurs potentiels de l'AHL, les jeux de données historiques dont l'emprise correspond au Limousin ne concernent que 40% des contributions potentielles et la granularité de la localisation est variable : 46% des données sont localisées à l'échelle infra-communale, 36% à l'échelle communale, et 9% respectivement à l'échelle départementale et à l'échelle régionale. La base de données spatio-temporelle contributive de l'AHL ne pourra donc offrir une couverture géographique continue ni homogène. Les méthodes de géolocalisation (notamment en termes d'implantation cartographique et de précision de la géométrie) sont différentes d'un chercheur ou d'une étude à l'autre ; et ce d'autant plus dans un contexte interdisciplinaire. Les géométries issues de travaux archéologiques, souvent très précises mais correspondant à un état partiel des objets historiques (puisque

¹ Cette enquête est un questionnaire web (créé avec le site *Drag'n'Survey*) envoyé à 130 personnes ayant des liens avec l'Atlas Historique du Limousin (voir la liste des contributeurs potentiels dans la Table 1) ou l'Atlas Historique de la Nouvelle-Aquitaine (qui a vocation à fusionner dans les années à venir avec l'AHL). 33 personnes interrogées ont déclaré avoir des données géo-historiques sur le Limousin et répondu à la totalité du questionnaire. Les pourcentages indiqués dans le reste de l'article concernent ces 33 personnes.

correspondant à l'état des découvertes matérielles), sera par exemple difficile à comparer à la localisation des camps de résistants en Corrèze pendant la seconde guerre mondiale, agrégés à la commune faute d'archives plus précises.

À cette hétérogénéité s'ajoutent la difficulté intrinsèque de la géolocalisation des données historiques – celles-ci se limitant aux informations dont témoignent les traces conservées (écrites ou matérielles), forcément partielles –, ainsi que les cas de contributions incomplètes. L'enjeu sera donc de renseigner la cause de l'incomplétude des données : si les données n'existent pas du fait d'une absence de sources, si les données n'ont pas encore été intégrées à la base de données mais qu'elles pourront l'être plus tard, ou si le contributeur ne possède pas les données sur ce territoire et/ou à cette époque parce que ceux-ci sortent de son cadrage spatio-temporel. Nous entendons répondre à ces diverses sources d'imperfection par un sourçage précis, systématique et obligatoire des données, de leur(s) source(s) et de leur intégration dans la BDD : qui a produit les données ? dans quel cadre institutionnel ou scientifique ? à quel moment ? qui les a entrées dans la base ? y-a-t'il eut des mises à jour ?

De la consultation de nos différents collaborateurs est également ressortie une difficulté concernant la formalisation numérique de la datation, problème récurrent de la modélisation en histoire. Pour répondre à cette difficulté et pour que chaque contributeur puisse formaliser l'aspect temporel de ses données de manière autonome mais standardisée, nous travaillons à la construction d'un outil de traduction d'une datation complexe en une modélisation en quatre dates (voir figure 1).

Figure 1 : Prototype d'interface pour l'outil de modélisation automatique des datations

Paramètres de la datation

Format de la datation
☐ Date ☒ Période

Précision des dates
☐ jour ☐ mois ☒ année ☐ siècle

Précision de la date de début de la période	Précision de la date de fin de la période
☐ date exacte connue : en ☐ approximative : vers 1878 Précisez l'intervalle d'imprécision : (Par défaut ce nombre est de 10 ans +/-) 5 ☐ intervalle connu : entre et ☐ avant ☐ après	☐ date exacte connue : en 1909 ☐ approximative : vers Précisez l'intervalle d'imprécision : (Par défaut ce nombre est de 50 ans +/-) ☐ intervalle connu : entre et ☐ avant ☐ après

VALIDER

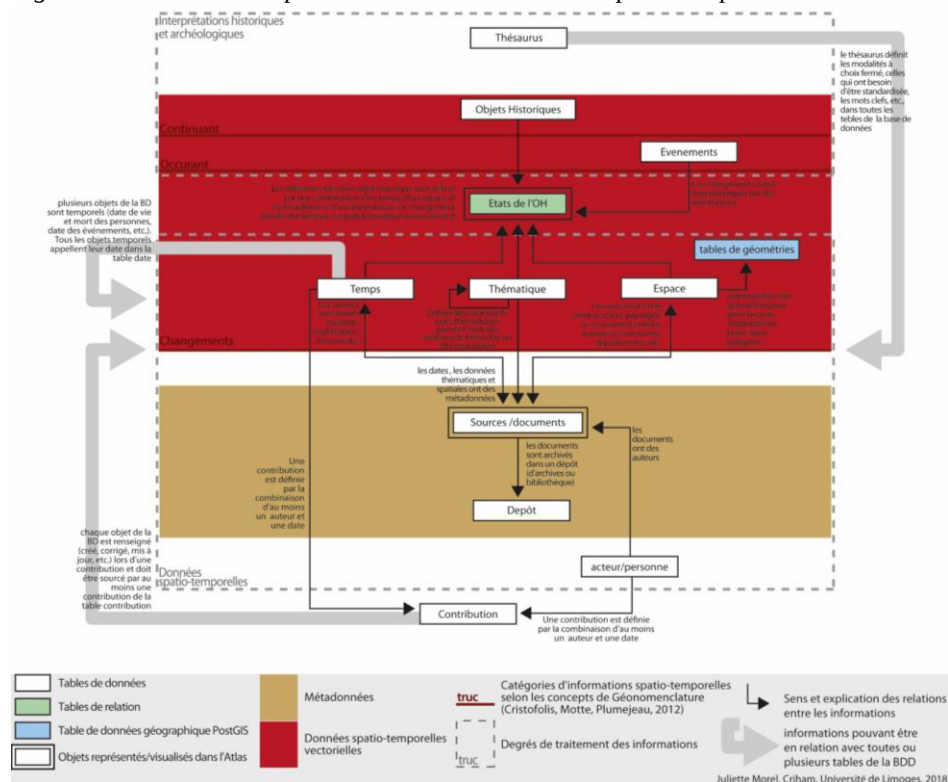
Résultat de la modélisation en quatre dates

date de début minimum (date_begininf) 1873-01-01	date de début maximum (date_beginsup) 1883-12-31	date de fin minimum (date_endinf) 1909-01-01	date de fin maximum (date_endsup) 1909-12-31
--	--	--	--

2. Un modèle pour prendre en charge l'hétérogénéité des données et des sources

En conséquence de ces diverses contraintes, le modèle de données sur lequel repose l'Atlas doit être équilibré entre flexibilité, pour accueillir l'hétérogénéité des données, et genericité, pour pouvoir les comparer. Nous avons choisi de construire un modèle de données spatio-temporel classique autour de la triade de Peuquet (Peuquet, 1994), c'est-à-dire que nous définissons les entités spatio-temporelles comme une combinaison d'attributs de trois types : spatial, temporel et thématique. Les dimensions temporelle et spatiale sont obligatoires – ce sont les deux points communs, donc de comparaison, minimaux entre toutes les données – et la dimension thématique est libre et capable d'absorber la diversité des travaux de recherche sur l'histoire du Limousin. Le changement historique est modélisé par la relation entre des éléments « continuant » (identité des objets historiques malgré leur modification lorsque l'analyse historique estime qu'il y a permanence), des états changeant (définis par la combinaison : temporel + spatial + thématique) et des événements « occurring », expliquant ponctuellement le passage d'un état à l'autre des objets historiques (Plumejeaud, Cristofoli, Motte, 2012).

Figure 2 : Modèle conceptuel de la base de données spatio-temporelles de l'ATLAS



La fixation d'un tel modèle contraint les contributeurs à adapter leurs informations à une structure de données prédéfinie et générique, qui n'est pas spécifique à leurs données. L'enquête que nous avons menée révèle qu'entre 20 et 60% de nos contributeurs potentiels – selon qu'ils sont historiens universitaires ou professionnels de l'information spatiale et/ou historique – possèdent des informations géo-historiques sous forme numérique (tableur ou fichier SIG) et que 30 à 60% d'entre eux (toujours selon la même distinction de corps professionnels) sont prêts à adapter le format de leurs informations au modèle conceptuel imposé.

3. L'enjeu de la participation

Comme dans tout projet de sciences participative, l'enjeu tient précisément dans l'ampleur de la participation qu'il rencontrera. L'enquête et les chiffres précédemment cités indiquent que le nombre de contributeurs spécialisés est limité, mais que leur motivation et la pertinence de leur contribution est grande. C'est exactement ce que nous attendons des contributions spécialisées dans le cadre de l'Atlas Historique du Limousin.

Comme nous le disions en introduction, un deuxième vivier de contributeurs est visé : le grand public. Il s'agit cette fois de solliciter une connaissance locale, liée à la mémoire collective, à la pratique du territoire et à l'investissement citoyen dans l'histoire locale. Cette sollicitation se fera concernant des aspects très précis de l'histoire et de la représentation du territoire Limousin, pour nous aider à préciser l'estimation spatiale et temporelle et à numériser un grand nombre d'informations : photographies anciennes et paysages à localiser et dater, cartes anciennes à vectoriser, etc. Comme dans le cas des grands programmes de production collective de connaissance (*Wikipédia*, *OpenStreetMap*), seule une participation significative (variant selon l'échelle et l'ambition du projet) garantit la qualité des informations collectées. Ici comme ailleurs, nous serons sans aucun doute soumis à la loi du « 90-9-1 » (Meriskay, Roche, 2011). Il faut donc tout mettre en œuvre pour rendre simple et attractif les outils permettant de telles contributions. Voilà pourquoi nous attachons un soin particulier à leur graphisme, leur ergonomie et leur pédagogie, en nous inspirant de l'outil *Footprint checker* créé par la *New-York Public Library* pour inviter le grand public à vectoriser ou à valider la vectorisation des bâtiments présents sur les anciennes cartes de New-York².

4. Conditions d'intégration et de diffusion des données

L'ouverture à la contribution – qu'elle soit spécialisée ou grand public – implique d'explicitier des conditions d'intégration et de diffusion des données collectées et de

² Outil *Footprint checker* de la *New-York Public Library*, en ligne, consulté le 25/10/2018, <https://buildinginspector.nypl.org/footprint>

s'assurer de leur acceptation par les contributeurs. Or tous les contributeurs ne sont pas d'accord sur le sujet, et il semble nécessaire d'adapter au cas par cas le degré de diffusion des données pour ne pas risquer d'en « perdre ». Celui-ci peut aller du contributeur qui permet le téléchargement de ces données sans conditions de réutilisation (60% des contributeurs spécialisés qui se sont prononcés sur cette question), jusqu'à l'utilisateur qui souhaite seulement visualiser ses données de manière temporaire dans le *webmapping* sans les intégrer à la base de données, en passant par le contributeur qui refuse le téléchargement (33.3%) ou qui n'accepte qu'une consultation restreinte par ses seuls collègues (6.7%).

Pour résumer, notre intention est de multiplier les possibilités de contributions en multipliant les conditions d'intégration et de diffusion et les interfaces de saisie, pour les adapter aux divers données, usages et publics. Les contributeurs pourront :

- Intégrer des données à la BDD en dessinant les géométries une par une sur une carte et en renseignant les modalités temporelles et thématiques concernant chaque objet. Le contributeur sera considéré comme l'auteur des données et définira leurs conditions de réutilisation selon les combinaisons de la licence *Creative Commons* (cette solution intéresse 46,7% des contributeurs spécialisés qui se sont prononcés);
- Intégrer des données à la BDD en uploadant des tableurs ou fichiers SIG complexes correspondant à une structure uniformisée préalablement imposée et documentée. Le contributeur sera considéré comme l'auteur des données et définira leurs conditions de réutilisation selon les combinaisons de la licence *Creative Commons* (26,7%) ;
- Ajouter temporairement des données géographiques au *webmapping* de l'AHL pour les visualiser avec/superposées aux données de l'AHL³;
- Proposer des campagnes de numérisation participative précises au grand public en utilisant des outils ad-hoc simples et ergonomiques (la première concernera la localisation des prises de vue des photographies anciennes à Limoges).

Conclusion : perspectives et limites de la contribution dans le projet AHL

Les choix méthodologiques effectués pour répondre aux problématiques posées par l'ouverture de l'Atlas Historique du Limousin à la contribution prennent le parti de la flexibilité face à la diversité des contributeurs, des données et des usages ; celle-ci nous paraissant nécessaire à l'attractivité et donc à la réussite du projet contributif. Néanmoins, ces choix pourront devoir être reconsidérés selon les directions que

³ Comme le propose par exemple *Atlas Historique des territoires politiques*, qui permet de « Glissez-déposez un fichier texte [csv] dans [une zone] pour projeter vos points sur la carte », GEO-LARHRA, LARHRA – UMR5190, CNRS, ISH Lyon, mis en ligne entre 2012 et 2015, consulté le 25/10/2018, <http://geo-larhra.ish-lyon.cnrs.fr/?q=atlas-historique/regroupement-de-territoires/evolution-des-territoires-en-europe>.

prendra l'Atlas dans le futur, comme l'intégration de ses données au web sémantique. Une telle intégration pose la question suivante : une fois que toutes les données de l'AHL seront moissonnables, comment maîtriser les hybridations et les éventuelles déformations résultant de la circulation des données ? La question des conditions juridiques de la diffusion des données collectées devra alors être reposée et uniformisée.

Plus globalement, c'est la pérennité institutionnelle de l'Atlas Historique du Limousin qui pourrait être remise en doute – et qui redéfinirait les choix méthodologiques. Le projet contributif de l'AHL vise à accumuler des données Géo-historiques sur le long terme. La pérennité nécessaire à cet objectif est garantie par l'écosystème du web qui est promis à durer (à condition d'actualisations techniques régulières) ; mais elle est confrontée aux modes de financement à court terme de la recherche universitaire. Le projet AHL est en effet financé jusque début 2019. Il doit fusionner en 2019 avec l'Atlas Historique de la Nouvelle-Aquitaine, dont le financement (Région Nouvelle-Aquitaine) court lui jusqu'en 2021. Mais après cela, comment garantir l'actualisation technique et la modération scientifique minimales nécessaires au fonctionnement de l'aspect contributif de l'Atlas ?

Bibliographie

Butez C. (2013). Conception d'un atlas historique numérique et d'une plateforme de travail collaborative à partir de la méthode SyMoGIH. Partie 1 - Naissance et conception d'un système d'information géo-historique collaboratif. *Géomatique Expert*, n°91, pp 30-35.

Joliveau T., Noucher M., Roche S. (2013). La cartographie 2.0, vers une approche critique d'un nouveau régime cartographique. *L'Information géographique*, vol. 77, n° 4, pp. 29-46.

Mericskay B., Roche S. (2011). Cartographie 2.0 : le grand public, producteur de contenus et de savoirs géographiques avec le web 2.0. *Cybergeog : European Journal of Geography* [En ligne], Science et Toile, document 552, mis en ligne le 20 octobre 2011, consulté le 08 avril 2018. URL : <http://journals.openedition.org/cybergeog/24710> ; DOI : 10.4000/cybergeog.24710

Peuquet D. (1994). « It's about Time: A Conceptual Framework for the Representation of Temporal Dynamics. Geographic Information Systems », *Annals of the Association of the American Geographers*, n°84 (3), pp. 441-461.

Plumejeaud C., Cristofoli P., Motte C. (2012). « De l'étude des nomenclatures territoriales à la modélisation des dynamiques des territoires administratifs en France », *Revue Internationale de Géomatique* – n° 1/2012, pp. 1-5

Atlas Historique du Limousin (2017), Criham, <http://www.unilim.fr/atlas-historique-limousin/>