

Production et mutualisation de données dans la ZABR

Questions et réflexions à propos du SIG historique du Y lyonnais

Thierry JOLIVEAU, UMR CNRS 5600, Université de Saint-Etienne

Le site Atelier Y lyonnais, objectifs

Le site Atelier "Y Lyonnais", dont nous assurons la responsabilité en collaboration avec J-P. Bravard, est un des sites supports de la ZABR (Zone Atelier Bassin du Rhône). Dans le dispositif de la ZABR, un site atelier est un site d'observation ou d'expérimentation pouvant servir de support à des recherches portant sur des thèmes pluridisciplinaires et croisant les regards de scientifiques, de gestionnaires et d'acteurs sociaux. Le site du Y lyonnais, appelé ainsi de par la forme et l'orientation du Rhône et de la Saône qui confluent à Lyon, a vocation à porter des recherches à dimension géo-historique sur l'aménagement du fleuve et de l'espace alluvial à Lyon.

Le site combine deux aires d'observation, qui correspondent à des approches thématiques et des échelles d'observation différentes. Un premier niveau d'analyse concerne le corridor fluvial qui correspond globalement à l'espace alluvial de la Saône et du Rhône et qui forme le Y lyonnais proprement dit, de la confluence de l'Ain et du Rhône, Trévoux au Nord sur la Saône et Givors au Sud sur le Rhône. Le deuxième niveau d'analyse est à plus petite échelle et concerne l'aire métropolitaine. Il correspond grosso modo à la Communauté Urbaine du Grand Lyon actuelle, augmentée des communes proches de Lyon riveraines de la Saône et du Rhône :

Les thématiques de recherche potentielles sont multiples :

- extension urbaine et corridor fluvial,
- artificialisation et dépérissement des milieux naturels liés au fleuve (endiguement, chenalisation, atterrissement, ...)
- vulnérabilisation des activités humaines dans les plaines alluviales,
- dynamique de patrimonialisation des cours d'eau,
- dynamiques paysagères.

Par ailleurs le site doit aussi servir d'espace de référence pour des recherches sur des thématiques transversales : généalogie des politiques de gestion du fleuve, veille sociale rhodanienne ou l'organisation de réflexions avec des acteurs tel que l'Atelier "Prospective du Rhône".

Le site Atelier Y lyonnais, recherches

Le Site Atelier vise à rassembler, mettre en relation et fédérer des recherches.

Exemple de travaux réalisés ou en cours sur le Site Atelier:

Travaux réalisés :

- DEA C. Combe C. (2002). Pour une approche historique et spatiale du risque en milieu urbain et périurbain, DEA Interface Nature-Société, 87 p.
- DEA M Maarouf (2003). La dynamique de l'aléa inondation, approche historique et quantitative. Le cas de la plaine de Miribel-Jonage. DEA Interface Nature-Société, 52 p.

Travaux en cours :

- Mémoires de 1ère année de Master Territoire, patrimoine, Environnement de l'Université de Saint-Etienne (Isabelle Laffay et Audray David)
- DEA de M. Fischer

- Thèse de Claire Combe sur l'approche historique du risque inondation dans l'agglomération lyonnaise

Travaux partenaires

- Thèse d'Emmanuelle Delahaye sur les villes situées à l'aval de Lyon a une problématique comparable sur d'autres villes du Rhône.

Le SIG du YLyon, les objectifs

Un des outils de fédération est la construction d'un Système d'Information Géographique historique permettant de spatialiser une analyse sur la longue durée des relations complexes entre une société urbaine et un environnement fluvial.

Trois échelles sont prévues :

1. L'aire métropolitaine. Les données doivent couvrir si possible toute la zone avec une échelle d'utilisation comprise entre le 1/50 000 et le 1/10000.
2. Le corridor fluvial Y lyonnais. L'échelle d'utilisation doit varier entre le 1/15 000 et le 1/50 000. Il y a une volonté d'exhaustivité de la donnée à terme sur tout l'espace, mais la réalisation peut être échelonnée.
3. Des zones d'intérêt particulier, sur lesquelles des données plus précises peuvent être constituées en fonction des recherches (exemple : reconstitution de la crue de 1856 dans la partie centrale de Lyon).

La constitution d'un SIG historique est un travail de longue haleine, prévu pour s'étaler sur plusieurs années. Les deux premières années ont été financées par la région Rhône-Alpes dans le cadre d'un programme Emergence.

Les fonctions du SIG historique varient en fonction des publics.

- Les chercheurs : le SIG est pensé pour supporter un travail scientifique pluridisciplinaire : il se nourrit et contribue en retour aux travaux et réflexions des chercheurs qui peuvent mobiliser les informations et des données du SIG YLyon et proposer d'autres données.
- Les gestionnaires de milieu, de territoire ou de système technique : l'objectif est de coopérer en fournissant un recul ou une dimension historiques sur les phénomènes liés au fleuve sous forme de la production conjointe de données ou d'analyse historiques (exemple des données sur l'histoire des digues pour le service de la Navigation).
- Les formateurs : qu'ils s'agisse des établissements de l'éducation nationale ou d'autres structures, de formation à l'environnement par exemple, une approche spatio-temporelle des phénomènes est en effet une manière efficace de réfléchir à un phénomène. Une des valorisations du SIG passe par la construction d'activités pédagogiques autour de ses données.
- Le grand public : les cartes ou illustrations anciennes ont la faveur du grand public et peuvent être des vecteurs efficaces pour faciliter une réflexion sur les thèmes d'environnement ou d'aménagement locaux.

Le SIG YLyon, les réalisations

La base de données Aire métropolitaine

Une base de d'images cartographiques géoréférencées

Nous avons fait scanner par l'IGN trois séries de cartes topographiques anciennes que nous avons recalées nous-même à partir des scan 25 de l'IGN

- 1835 : Les quatre feuilles des minutes de terrain dessinées au lavis par quart, de la feuille Lyon de la Carte d'Etat-Major, couleur, échelle : 1/40 000.
- 1902: Une feuille : Lyon XXX-31 Type 1900, couleur, échelle : 1/50 000

- 1949 (1942-1952): Trois feuilles : 3031(Lyon), 3131(Montluel) et 3032 (Givors), deux factures différentes, couleur, échelle : 1/50 000.

Le travail de recalage n'est pas aisé, étant donné les caractéristiques variées et pas toujours connues dans le détail des types de levés. On peut estimer que la précision de recalage permet une superposition des différentes cartes au 1/30 000 approximativement avec de fortes variations locales. Pour les visualiser, un système de double-visualiseur a été développé au CRENAM. Ce système qui permet de lier la navigation entre deux cartes visualisables simultanément à l'écran au même niveau de zoom est un moyen aisé de comparer visuellement deux cartes anciennes. Il est moins sensible à des décalages liés aux imprécisions de géoréférencement que la visualisation classique multicouche des SIG. Cet outil sera mis à la disposition des chercheurs pour la constitution des bases de données aux différentes dates et pourrait être diffusé avec les cartes à destination du public scolaire. Nous testons aussi un visualiseur Internet.

Une base de données historiques à l'échelle de l'agglomération

Elle concernera à terme la population, l'occupation du sol et les communications.

L'occupation du sol sur l'ensemble de l'agglomération est reconstituée à 5 dates (environ 1835, environ 1900, environ 1950, 1975, 1990, 2000). Pour 1835, 1900, 1949, 1990 et 2000 la couverture d'occupation du sol est reconstituée par interprétation visuelle de la carte topographique et attribution des caractéristiques repérables sur une grille de points espacés de 50 m. Pour 1950, 1975 et 1990 une typologie simple (construit/non construit) a été produite par rasterisation et filtrage d'une carte papier de l'Agence d'Urbanisme du Grand-Lyon. Ce jeu de données est quasiment terminé. Les données de population par commune depuis 1800 ont été saisies.

Une base de données d'images cartographiques non géoréférencées

Nous avons photocopié les cartes topographiques anciennes au 1 :25 000 et au 1 : 20 000 de manière à les rendre disponibles pour les chercheurs. La réflexion sur la manière de gérer ces documents est en cours. Une solution envisagée est de les scanner pour les conserver dans un état correct et de les repérer dans le SIG

La base de données Corridor fluvial :

La base de données géohistoriques

L'objectif est de spatialiser les interactions société-environnement liées à la dynamique fluviale. Les objets saisis dans la base dépendent des objectifs des recherches mises en œuvre, des besoins des partenaires et des documents disponibles. Le principal thème en cours lié à la thèse de C. Combe porte sur la vulnérabilisation des activités humaines dans la plaine alluviale. Les autres thèmes comme l'analyse du dépérissement des milieux naturels liés au fleuve (endiguement, chenalisation, atterrissement, ...), la patrimonialisation ou l'évolution paysagère des cours d'eau pourront s'appuyer sur le jeu de données constitué et sur d'autres données collectées dans d'autres projets de recherche ou d'autres partenariats.

Les objets saisis dans le cadre du thème vulnérabilisation l'ont été pour certains objets à l'échelle du corridor. D'autres font l'objet d'une maquette sur la zone amont du Rhône (Confluent de l'Ain/Parc de la Tête d'Or).

- le système hydrographique : réseau hydrographique en lit majeur, îles, ruisseaux, plans d'eau, nappes
- les obstacles à l'écoulement en lit majeur : digues de chenalisation affectant la géométrie du chenal, digues de protection contre les crues, remblais
- l'extension des crues inondantes et l'enveloppe de mobilité du fleuve,
- les enjeux du corridor : occupation du sol, usines et stockages, voies de communication
- les documents de réglementation (PSS, R111-3, PPR, POS et PLU).

L'exemples du Projet Digue

Suite à la crue de 1856, un double système de quais et de digues assure la protection « définitive » de Lyon. Le « rempart de protection » a été en majeure partie achevé avant 1870. Il constitue depuis l'essentiel de la défense de Lyon intra muros contre les inondations. Un risque résiduel demeure dû à la perte de mémoire progressive de la présence de digues de protection contre les crues et de leur fonction première. Or certains remblais ont accueilli de nouveaux ouvrages (routes notamment). Quel est l'état des digues aujourd'hui ? La protection est-elle toujours efficace contre une crue de type 1840 ou 1856 ? C'est pour répondre à ces questions que le Service Navigation Rhône-Saône (SNRS) a commandé à C. Combe la réalisation d'une reconstitution historique du système de digues, organisé dans un SIG. Il s'agit d'un outil d'aide à la gestion pour l'expertise technique du risque de rupture de digues. Mais il permettra aussi d'un point de vue scientifique de réfléchir sur la notion de risque résiduel lié à la présence de protections structurelles et de penser la protection physique comme un élément complexe qui a une histoire et peut être vulnérable.

Une base de données géoarchéologiques

La ZABR a récemment choisi d'inclure le temps long des sociétés fluviales dans ses centres d'intérêt, bénéficiant de plus de 20 ans de recherche, notamment à Lyon. Il est prévu d'intégrer les héritages de la paléo-dynamique fluviale dans le SIG YLyon (niveaux alluviaux depuis le début de l'Holocène, paléo-chenaux depuis le début de l'occupation gallo-romaine du site, travaux de protection à l'époque historique, etc...). Une collaboration sera développée avec le Service Archéologique de la Ville de Lyon et le Service régional de l'Archéologie.

La base de données iconographiques

L'objectif est de stocker dans le SIG l'iconographie du corridor fluvial. Un test est en cours à partir du fond de cartes postales de la Maison du Rhône.

Le repérage des sources

Les données intégrées dans le SIG Historique de l'YLyon proviennent des données d'archives. Pour être utile, il faut au minimum stocker dans le système (métadonnées) les cotes de source et d'archives qui permettront à d'autres chercheurs d'aller plus loin. Vu les nouvelles méthodes de collecte des données en archives (photo numérique des documents), la question se pose de l'intégration dans la base de données d'une copie des documents d'archive eux-mêmes. Problèmes qui peuvent se poser : dérive vers une recherche de l'exhaustivité, difficultés liées aux droits des données, nécessité d'outils techniques de gestion de données documentaires.

Partenariat et mutualisation

La constitution du SIG YLyon met en œuvre différents modes de relation entre fournisseurs de données originelles, créateurs de nouvelles données et utilisateurs des données produites. Nous abordons les questions de partenariat et de mutualisation susceptibles de se poser dans le cadre du SIG historique en distinguant deux aspects : les contraintes techniques et les règles juridiques.

Les contraintes techniques de l'échange de données géographiques

Pour les données géographiques trois paramètres sont primordiaux : le géoréférencement, les métadonnées et l'existence d'un support technique de partage des données.

Le géoréférencement

Pour être partagées, les données doivent être repérées dans un référentiel explicite, de manière à ce qu'elles puissent être transformées éventuellement dans un autre référentiel. Nos données sont en Lambert 2 étendu dans le système NTF.

Le système officiel français est dorénavant le système RGF93 avec des coordonnées exprimées soit sous la forme tridimensionnelle géographique (longitude, latitude, hauteur ellipsoïdale) soit sous forme bidimensionnelle, selon la projection Lambert 93, unique pour l'ensemble du territoire. Ce système assure une précision centimétrique. Le passage de l'ancien système au nouveau ne pose pas de problème compte tenu de la précision de nos données qui est toujours supramétrique. En revanche, il

semble que le référentiel du Grand Lyon ne corresponde pas exactement à celui du scan 25 de l'IGN. L'axe des digues digitalisées sur les fonds de plan fournis par le Grand Lyon au Service de la Navigation peut ainsi s'éloigner de 5 à 30 m de l'axe des routes sur le scan 25. Cela ne pose pas de problème grave pour le SIG historique de l'aire métropolitaine dont la précision est de l'ordre de 50 mètres mais cela peut être gênant pour le SIG historique du corridor fluvial dont la précision attendue est plutôt de l'ordre de 10 mètres.

Les métadonnées

L'échange suppose que les données soient décrites de manière précise et professionnelle par des métadonnées permettant à un utilisateur différent du producteur de les utiliser sans contresens. Dans le cas du SIG Historique du Y Lyon, l'interprétation des cartes anciennes commande un certain nombre de précautions. En effet, il est souvent difficile d'estimer la validité planimétrique et thématique des cartes topographiques anciennes. Certaines sont des minutes de terrain et entachées d'une erreur de localisation parfois considérable. Les cartes éditées ont souvent fait l'objet de corrections successives sur les plaques originales, qui empêchent de dater exactement l'état de l'occupation du sol. La description de leurs limites de précision doit être très clairement expliquée dans les métadonnées.

Un système support

Il est nécessaire de disposer d'un outil technique d'information et de diffusion du système. Je ne le mentionne que pour mémoire, puisque c'est l'objet du projet de Géorépertoire de la ZABR

Les règles juridiques du partage

Les règles à établir concernent essentiellement les droits de propriété intellectuelle et de réutilisation d'éléments empruntés à un autre partenaire. Elles dépendent à la fois du type de partenaire, de la nature et de leur intérêt à la participation au projet et de l'utilisation envisagée des données. La mutualisation et le partage peuvent concerner différents éléments du travail: les données bien sûr, mais aussi les méthodes ou modèles, les outils et les résultats. Un projet comme le SIG YLyon nécessite de mobiliser un fort volume de données et d'outils d'origine externe ou extrascientifique, soit chez des producteurs de données commerciaux ou assimilés soit auprès de collectivités ou de bureaux d'études.

Les fournisseurs commerciaux

Avec les fournisseurs commerciaux de données (y compris l'IGN qui refuse d'envisager un partenariat, même sur les données historiques), le seul échange possible est une transaction commerciale sur le droit d'utilisation des données. Celle-ci prend souvent la forme d'une licence, variable selon le type d'utilisation. L'IGN propose ainsi 4 types de licence : licence standard, licence serveur, licence de représentation graphique, licence de représentation électronique, selon que l'on utilise les données sur un seul poste de travail, sur un serveur, qu'on les diffuse sous forme de document papier ou au moyen d'un support numériques (CD ou Internet)...

Dans le cas des cartes anciennes, l'IGN a un copyright sur l'ensemble des cartes qui remonte à 70 ans plus les années de guerre, ce qui conduit à payer les droits d'utilisation pour tous les documents postérieurs à 1926, c'est-à-dire dans notre cas les cartes topographiques de 1949 et de la période contemporaine. Si, comme on l'envisage dans le SIG YLyon, on veut pouvoir diffuser les cartes géoréférencées à l'intérieur de la ZABR, il faut payer le coût de diffusion des cartes récentes, en fonction du type de média envisagé. La dégradation de la qualité du document peut être dans certains cas un élément de négociation (mais pas avec l'IGN ?) Les données digitalisées à partir des cartes topographiques ne relèvent pas a priori du droit d'auteur car elles ne reconstituent pas tout ou partie des objets figurant sur la carte.

Dans le cadre d'un projet comme la ZABR, la question est : qui prend en charge le coût de rediffusion des données à l'intérieur et à l'extérieur de la ZABR ?

Les partenaires non scientifiques

Les fournisseurs non commerciaux de données : collectivités locales, services de l'Etat sont des partenaires potentiels du projet. Alors qu'ils demandent souvent une redevance pour l'utilisation de

leurs données, ils peuvent être conduits à les fournir gratuitement dans le cadre d'un projet de recherche en échange par exemple de la fourniture des nouvelles données produites. Il est bien sûr interdit de rediffuser les données, et la diffusion des résultats doit porter mention de la source utilisée. L'intérêt pour le fournisseur de données est que les données qu'il récupère sont compatibles avec les siennes. Par exemple, si le Grand Lyon fournit le réseau routier actuel en référence, les données sur l'évolution historique de la voirie seront compatibles avec les données contemporaines. Dans les faits, c'est souvent plus compliqué. Par exemple, le Grand Lyon ne fournit ses données que sur un espace qui mesure moins de 10 % de la superficie communale. Il est donc impossible de disposer hors convention spécifique de l'ensemble du réseau routier. La solution est donc la signature d'une telle convention, mais avec quelle entité : équipe, regroupement d'équipes (UMR), Zone Atelier ?

Les données des archives sont accessibles 30 ans après l'élaboration du document, sauf si celui-ci contient des données personnelles ou si les documents ne sont qu'en dépôt. Il faut alors demander l'autorisation aux ayant droits. Les centres des archives ou de documentation font habituellement payer un coût de mise à disposition de données numériques, mais peuvent aussi laisser l'utilisateur procéder par ses propres moyens à la reproduction. La diffusion de la reproduction dans un cadre non commercial est normalement possible après accord et fourniture d'un exemplaire du document diffusé.

Les bureaux d'étude fournissent moins souvent des données que des méthodes ou des modèles (dans le cas du SIG YLyon : CNR, Hydrotec). La question est alors de pouvoir disposer de ces modèles pour des recherches spécifiques. Dans le cas du travail de M. Maarouf, la CNR a ainsi intégré les données fournies par le Site Atelier du Y lyonnais, pour faire tourner en interne son modèle de simulation de crue dans les conditions correspondant à une année antérieure. Est-il possible d'envisager une mise à disposition des partenaires de leurs modèles pour tester différents scénarios ?

Les partenaires scientifiques

Dans le domaine de la recherche scientifique, les résultats, les modèles et les méthodes sont habituellement partagés au travers des publications. Il n'en va pas forcément de même des outils et des données, pour lesquelles le droit d'auteur est souvent appliqué. La concurrence scientifique peut parfois conduire à refuser, différer ou ralentir la diffusion des données auprès d'autres équipes ou à ne pas diffuser des outils spécifiquement développés pour une recherche.

Il faut donc réfléchir à une politique de mutualisation des données de la ZABR à deux niveaux :

- en interne entre les équipes scientifiques participantes
- en externe : avec les partenaires non scientifiques et avec d'autres équipes scientifiques

En interne : le minimum est une mise à disposition systématique des partenaires des données produites dans le cadre des activités de la ZABR. On peut dans ce cadre garder une approche classique de protection du droit d'auteur de chaque équipe ou partenaire. On peut aussi réfléchir à une optique de type "open source" en adoptant un principe inspiré des logiciels libres (GNU)¹, qui donnerait le droit à l'ensemble des partenaires de la ZABR d'utiliser les données produites, de les adapter et de les améliorer collectivement. Cela pourrait porter aussi sur les outils logiciels éventuellement développés. Il faut cependant tenir compte des droits liés au caractère composite des œuvres, c'est à dire des droits de propriété incorporés dans l'œuvre produite, qu'il s'agisse de données ou de logiciels).

En externe : on doit décider d'une politique de mise à disposition à d'autres acteurs des données constituées dans le cadre de la ZABR, qu'il s'agisse d'équipes scientifiques, de bureaux d'étude ou de collectivités. Les solutions peuvent s'échelonner d'une non communication systématique à une diffusion libre de type open-source.

Une approche de type Open Research ?

Il existe une proposition d'organisation de recherche sur un modèle « open-source », visant à diffuser « librement » les données, les modèles et les logiciels d'un projet de recherche (voir le site Open Research System <http://www.open-research.org/> et la brochure <http://www.open-research.org/ORS/ORSbrochurev3.pdf>). L'article de référence du projet (Schweik et Grove 2001) passe en revue l'intérêt et les difficultés d'un tel dispositif, en particulier les réticences légitimes des thésards à une diffusion trop large et trop rapide des données qu'ils auraient collectés. L'article propose certaines réponses adaptées. Même si nous sommes du point de vue du droit des données

dans un système beaucoup plus contraignant que le système américain, la ZABR pourrait réfléchir dans cette direction. Il n'est pas indifférent que le groupe qui est à la source de cette initiative appartienne à la LTER de Baltimore et travaille dans une problématique d'étude d'un système environnemental complexe très proche de celle de la ZABR, avec des données portant sur des territoires qui peuvent être assimilés aux sites de la zone atelier (bassin de la Chesapeake Bay et Adiron Parck).

Schweik C. M. et Grove J. M. (2001). *Fostering Open-Source Research Via a World-Wide Web System*. 2004. http://www.pamij.com/5_4/5_4_2_opensource.html

¹ Il existe un projet de licence de données géographiques libres, la licence PGL ("Public GeoData License") ou "DGL", pour "Donnée Géographique Libre"

cf : <http://sig.cwriter.org/index.php/PGL>:

« La licence PGL doit faire partie des métadonnées, lorsque la norme de métadonnées utilisées le permet, ou accompagner les données sinon. Une donnée géographique libre doit respecter les quatre libertés suivantes:

- La liberté d'utiliser les données, pour tous les usages (liberté 0).
- La liberté d'étudier les données, de les adapter à vos besoins (liberté 1).

Pour ceci l'accès aux données et aux métadonnées est une condition requise.

- La liberté de redistribuer des copies, donc d'aider votre voisin, (liberté 2).
- La liberté d'améliorer les données, de les adapter, et de publier vos améliorations, pour en faire profiter toute la communauté (liberté 3).

Pour ceci l'accès aux données et aux métadonnées est une condition requise. »