

ALPAGE : Compte-rendu de la réunion tenue à La Rochelle le 19 octobre 2006

Présents : Pierre Heroux, Sébastien Adam, Romain Raveaux, Brigitte Boissavit-Camus, Jean-Marc Ogier, Jean-Christophe Burie, Frédéric Rousseaux, Guillaume Champlong, Laure Paradis, Frédéric Pouget, Karell Bertet, Alexandre Hamon, Hélène Noizet.

Excusé : Laurent Costa.

Après un tour de table pour faire connaissance avec les nouvelles personnes arrivées dans le projet, et notamment des informaticiens de Rouen, anciens collègues de JM Ogier intéressés par le projet, un point est fait sur la gestion administrative et financière. Un contrat établi par l'ANR doit être bientôt envoyé, sans doute vers les 23-25 octobre, à chaque laboratoire partenaire bénéficiant d'une dotation financière, à savoir le Lamop, L3i et Otelo. Les besoins financiers d'ArScAn ont finalement été intégrés dans ceux du Lamop en raison des règles financières instaurées par l'ANR. L'argent devrait être concrètement versé peu après, sans doute au début du mois de novembre : le versement des deux premières années sera à chaque fois de 30 % de la somme totale, les 40 % restants étant versés lors de la 3^{ème} et dernière année du projet. Les contrats indiquant la somme allouée à chaque partenaire doivent permettre à l'université de La Rochelle d'engager les premières dépenses, notamment pour payer les salaires des deux personnes recrutées pour le projet durant 3 ans. Il s'agit d'une part de Romain Raveaux, ingénieur recruté par le L3i, et d'autre part de Guillaume Champlong, assistant-ingénieur recruté par Otelo. R. Raveaux, qui fera de son travail un sujet de thèse, aura en charge la vectorisation des images et la reconnaissance des symboles, tandis que G. Champlong s'occupera du géoréférencement des données, de la mise en place du SIG à partir du modèle conceptuel de données (qui sera élaboré par les différentes personnes compétentes en la matière), et de l'intégration dans le SIG des données historiques fournies par le groupe des historiens-archéologues.

Les images-sources de plans parcellaires parisiens données par les Archives de Paris ont été récupérées par H. Noizet le 17 octobre. Elles sont copiées durant cette réunion sur un disque dur externe de J.M. Ogier, puisque leur poids total est d'environ 65 Go. Il est également transmis aux Rochelais la base de données, faite sous Access, par Archimaine (l'entreprise privée qui a procédé à la numérisation) et/ou les Archives de Paris, pour repérer les fichiers qui nous intéressent, puisqu'il n'y a pas que les images de l'Atlas Vasserot dans la collection fournie par les Archives de Paris.

Une convention a été élaborée durant l'été 2006 entre le Lamop et les Archives de Paris, propriétaire de ces images raster, afin d'indiquer précisément l'usage qu'il peut être fait de ces fichiers, notamment à des fins de publication. Cette convention, qui sera prochainement envoyée pour signature aux coordinateurs du L3i, d'Otelo et d'ArScAn, permet la publication de ces images dans la mesure où sont respectées les mentions institutionnelles. La convention distingue 2 types de publications : d'une part les publications collectives (par exemple les actes d'un colloque qui pourrait avoir lieu à la fin du projet) et les expositions, qui devront reproduire les logos de toutes les institutions ayant financé la numérisation de ces images (ministère de la Culture, Région IdF, département de Paris, préfecture de la région IdF...); d'autre part, les publications d'articles qui doivent comporter la mention suivante sous l'image reproduite : CHAN, F 31 (côte à indiquer)/Cliché Département de Paris-Archives de Paris

Il n'y aura aucun droit à payer.

Le protocole de travail de la société Archimaine a été transmis à J. M. Ogier. Si les Rochelais souhaitent avoir plus de renseignements sur leur méthode de numérisation, ils peuvent contacter directement cette société, en rappelant notamment l'entrevue récente (21/09/2006) entre S. Aufray (des Archives de Paris) et Jacques de Baglion (contact : Martine Cosson : contact@archimaine.fr). La technologie utilisée correspond à une scannérisation en ligne, et non pas à des prises de vue photographiques, ce qui permet de n'avoir aucune déformation liée à la numérisation. Les images comportent en revanche une déformation inhérente aux documents originaux, qui est la pliure centrale, puisqu'il s'agit d'atlas reliés. Ces images ont une résolution de 200 dpi, ce qui est peu.

Déjeuner au restaurant Le Mistral

1) Parcellaire

La première question qui se pose est la gestion du géoréférencement et de la vectorisation des parcelles : faut-il procéder d'abord à l'un puis à l'autre, et dans quel sens ? Une discussion s'engage à ce sujet de laquelle émerge le besoin, à la fois pour les géographes s'occupant du géoréférencement, et pour les informaticiens s'occupant de la vectorisation, de travailler à partir des images-sources, et non pas à partir des images transformées par l'autre partie. En effet, le géoréférencement des images induit des déformations importantes, que l'on peut concrètement observer grâce à un premier test de géoréférencement effectué par Laure Paradis sur une des images récupérées. L. Paradis remarque qu'il y a beaucoup de décalage entre le parcellaire début 19^e et le parcellaire actuel. Pour cette image, dont 4 points homologues ont été définis avec le parcellaire actuel, la torsion apparaît nettement. Cela poserait des problèmes aux informaticiens notamment pour reconnaître les caractères, comme les numéros des parcelles. Cela peut aussi affecter la couleur des pixels puisque le géoréférencement induit une redéfinition des caractéristiques des pixels.

Inversement, il est préférable, pour les géographes, de géoréférencer les images raster, sur lesquelles on a plus de prise, que des fichiers déjà vectorisés. Surtout, ils ont absolument besoin d'avoir les noms des rues sur les images pour les situer et recalibrer les images : s'ils travaillaient à partir des données déjà vectorisées, et non des images raster-source, cela impliquerait que les données vectorisées contiennent non seulement les parcelles mais aussi les noms de rues, ce qui, pour le coup, nécessite un très gros travail pour les informaticiens.

A ces contraintes techniques, il faut ajouter la contrainte de temps : si l'on devait procéder à l'une des opérations, puis seulement après à l'autre, cela prendrait beaucoup de temps et ne permettrait sans doute pas de fournir aux historiens des plans assemblés suffisamment tôt. P. Heroux demande s'il est possible que les historiens reportent leurs données sur les images raster de l'Atlas Vasserot, ce qui permettrait d'évacuer la contrainte de temps pour la vectorisation et le géoréférencement. Après discussion, il apparaît difficile de procéder ainsi pour des raisons pratiques de gestion des documents : cela impliquerait de gérer plusieurs milliers de documents, à multiplier par le nombre de couches historiques fournies. Il faut fournir aux historiens au moins des cartes d'arrondissement, avec des îlots déjà assemblés.

Il n'est en revanche pas nécessaire de fournir toutes les données parcellaires en bloc aux historiens : les Rochelais peuvent fournir au fur et à mesure de l'avancée de leur travail les parcellaires, par exemple en suivant l'ordre des arrondissements.

Pour toutes ces raisons, il est décidé de travailler parallèlement au géoréférencement et à la

vectorisation : ainsi les géographes géoréférenceront et les informaticiens vectoriseront quelques images en parallèle; puis les géographes recalcront ensuite la version vectorisée de ces images, cette opération devant être facilitée par le fait que la version raster de ces images aura déjà été recalée.

Un premier test pour tester la validité de la méthode va être fait : les géographes doivent proposer à R. Raveaux une petite sélection d'images représentatives des difficultés qu'ils peuvent rencontrer ; R. Raveaux complètera éventuellement cet échantillon pour intégrer un panel suffisamment large des difficultés liées à cette fois-ci à la vectorisation. Une fois défini un échantillon commun, les informaticiens développeront les algorithmes permettant la vectorisation des parcelles, pendant que les géographes géoréférenceront ces images. Puis, les géographes récupéreront les données vectorisées pour les recalcr sur les images raster elles-mêmes déjà géoréférencées. La validité de la méthode sera ainsi discutée lors de la prochaine réunion, afin de voir si elle peut ensuite appliquée à toutes les images.

JM Ogier pose la question de la précision exigée : est-on à un pixel près ? Pour les historiens, plus c'est précis, mieux c'est, mais il est difficile de dire si l'exigence se situe au pixel près. Les géographes ont en revanche des exigences fortes, pour obtenir un géoréférencement optimum : à première vue, il semblerait que l'on soit au pixel près. Mais cela reste à déterminer précisément. En revanche, ce qui est fondamental pour les historiens, c'est de connaître l'échelle d'imprécision pour chaque image, afin de savoir si on est à 1 ou à 5 mètres près, car cela conditionne par la suite les questions historiques qui seront ensuite posées à cette documentation. Cette échelle d'imprécision devrait pouvoir être connue à partir du modèle de déformation de l'image, qui se concrétiserait par un indice traduisible en une échelle métrique. Il est impératif de connaître l'échelle d'imprécision pour ne pas surinterpréter la couche parcellaire, qui va devenir une source pour l'historien (type calculs métrologiques au centimètre près alors que l'échelle d'imprécision serait par exemple à 2 mètres). Dans la mesure où cette imprécision est une donnée inhérente à la source elle-même, il appartiendra à chaque utilisateur du SIG de vérifier cette échelle d'imprécision avant de se lancer dans des raisonnements par trop hasardeux.

Les informaticiens de Rouen proposent de faire faire comme petits travaux à leurs étudiants des opérations de vectorisation manuelle sur quelques images, afin d'avoir une évaluation de la vectorisation automatique.

F. Pouget demande s'il est possible de proposer aux historiens de rendre leurs données sous forme numérique. Il apparaît vraisemblable que certains d'entre eux puissent dessiner directement en mode vecteur. Pour ceux qui ne le pourraient pas, les informaticiens du L3i se proposent de leur laisser dessiner leurs cartes sur du papier puis de reconnaître automatiquement leurs dessins, en créant des procédures de reconnaissance de symboles ad hoc. Pour ces historiens qui souhaiteront dessiner sur papier leurs données, les informaticiens devront définir des consignes de dessin et une charte graphique que les historiens devront respecter.

F. Pouget demande s'il sera nécessaire d'acheter un traceur format AO, qui coûte à lui seul au moins 7000 €, et qui aujourd'hui n'est pas couvert par le budget alloué par l'ANR à Otelo. Dans la mesure où il avait été décidé que les images du parcellaire seront données aux historiens sous forme numérique, et non pas en format papier pour éviter les problèmes matériels de transfert, il n'apparaît pas vital de faire cet achat pour ce projet.

2) Filiale des rues début 19^e s

Cette couche sera créée, non pas à partir des documents sources, mais de la couche parcellaire

géoréférencée et vectorisée. Puisque les îlots y seront dessinés comme des objets surfaciques et les rues apparaîtront comme les vides entre les îlots, une procédure devra être définie pour créer des objets linéaires dans les vides, c'est-à-dire créer en quelque sorte un négatif linéaire des surfaces vectorisées.

Après discussion, il n'apparaît pas nécessaire que les informaticiens reconnaissent les noms des rues écrits à la main. Comme cela leur paraît très lourd en terme de quantité de travail, qu'il ne s'agit que de quelques unités par îlot, il est décidé que les géographes reporteront les noms de rues manuellement dans les données attributaires. Il y aura en revanche une coopération nécessaire avec les historiens pour déterminer précisément où commence et où s'arrête chaque rue.

3) Filaire des rues début 14^e s

Dans la mesure où un archéologue du Lamop, Cédric Roms, a déjà vectorisé les îlots du plan de Paris au 14^e s (plan du CNRS), et que ce dernier a eu la gentillesse de nous donner son travail, il ne sera pas nécessaire de repartir de ce document pour établir le filaire des rues du 14^e s. On pourra appliquer la même procédure définie pour le filaire précédent, en créant des lignes à la place des vides entre îlots. Un tel filaire permettra de recaler des informations historiques médiévales : si la précision de la localisation restera relative, ce filaire pourra néanmoins autoriser une spatialisation de ces données, par exemple les métiers connus à partir des rôles de taille (travail de C. Bourlet). Cela impliquera en revanche de géoréférencer le document fourni par les archéologues.

H. Noizet demande s'il est envisageable de créer une seule couche filaire, commune pour les 2 périodes (19^e et 14^e), en indiquant la date de chaque segment de rue comme une donnée attributaire : le filaire du 14^e s. pourrait ainsi être obtenu par une requête attributaire. Mais l'opération, et le bénéfice induit, paraissent a priori faibles par rapport à la complexité de l'opération : autant créer directement 2 états du filaire.

Concernant les opérations de vectorisation, un ordre des priorités est défini :

1. Les parcelles en reprenant l'unité de couleur et en se fondant sur l'axe central des traits qui séparent 2 plages de couleur
2. La présence d'un numéro de rue (et éventuellement la reconnaissance du numéro)
3. Le filaire des rues 19^e
4. Le filaire des rues 14^e
5. Le bâti sur un secteur à définir par B. Boissavit
6. L'extension du parcellaire au-delà de la limite des Fermiers Généraux, à partir d'autres plans cadastraux

La prochaine réunion de ce groupe de travail est fixée à **La Rochelle le jeudi 25 janvier 2007**, aux mêmes heures permises par les liaisons ferroviaires entre Paris et La Rochelle (10h30-17h).