

MISE EN PLACE D'UN SYSTEME DE GEOCODAGE POUR BREST METROPOLE OCEANE

Adresse	Code_postal	Commune
RTE DE QUIMPER	29590	PONT DE BUIS LES QUIMERCH
17 R DE BREST	29290	ST RENAN
11 R DE LA ROCHE	29870	LANNILIS
15 R DU BEL AIR	29860	BOURG BLANC
21 BD MICHEL BRIANT	29804	BREST
38 R DE QUIMPER	29590	PONT DE BUIS LES QUIMERCH
2 R DU SQUIRIOU	29590	PONT DE BUIS LES QUIMERCH
21 BD MICHEL BRIANT	29215	GUIPAVAS
21 BD MICHEL BRIANT	29490	GUIPAVAS
2 R ROSEMONDE GERARD	29850	GOUESNOU
18 R DE L EAU BLANCHE	29804	BREST
2 PL DE LA GARE	29870	LANNILIS
2 PL DE LA GARE	29870	LANNILIS
2 PL DE LA GARE	29870	LANNILIS
21 BD MICHEL BRIANT	29215	GUIPAVAS
38 R DE QUIMPER	29590	PONT DE BUIS LES QUIMERCH
260 R AUGUSTIN FRESNEL	29801	BREST
32 RTE DE LANNILIS	29260	LE FOLGOET
80 R GENERAL DE GAULLE	29260	LESNEVEN

G
é
o
c
o
d
a
g
e



Damien Couppé

Promotion : 2010/2011

UNIVERSITÉ

La Rochelle

Rapport de stage effectué dans le cadre de la licence professionnelle SIG de La Rochelle,
à Brest métropole océane

Maître de stage : Laurent DUPONT

Tuteur de stage : Alain LAYEC

Remerciements

Je tiens en premier lieu à remercier Laurent Dupont en qualité de maître de stage qui a su m'accompagner, m'encadrer et me guider tout au long de ces 5 mois de stage.

Je remercie également toutes les autres personnes du service SIG : Élisabeth Naja, Fanny Gallou, Adrien Hamel, François Vigouroux, Grégoire Vourc'h, Jean-Christophe Le Guen qui ont su m'accorder du temps et répondre à mes questions lorsque j'en avais besoin.

Je remercie aussi Julien Sichler et Laurent Méar à la DSIT (Direction des Systèmes d'Informations et Télécommunications) pour leur appui technique et les solutions qu'ils m'ont apportées, notamment en ce qui concerne l'application Web.

Enfin, je remercie mes professeurs Frédéric Pouget et Alain Layec pour leurs enseignements et leur savoir-faire qu'ils m'ont transmis tout au long de l'année.

Sommaire:

INTRODUCTION :	1
1 CONTEXTE DU STAGE :	3
1.1 LA COMMUNAUTE URBAINE DE BREST METROPOLE OCEANE (BMO) :	3
1.1.1 <i>Présentation générale :</i>	3
1.1.2 <i>L'administration :</i>	3
1.2 LE SERVICE SIG DE BMO :	5
1.2.1 <i>Présentation générale :</i>	5
1.2.2 <i>Ses missions :</i>	5
1.2.3 <i>Son portail Web :</i>	5
2 LES BESOINS DU SERVICE SIG ET LES OBJECTIFS DU STAGE :	7
3 LE GEOCODAGE DANS LA SOLUTION ARCGIS :	9
3.1 FONCTIONNEMENT GENERAL :	9
3.2 LA STANDARDISATION :	9
3.3 CONSTRUCTION D'UN GEOCODEURS ARCGIS :	10
3.3.1 <i>Les styles de localisateurs d'adresses :</i>	10
3.3.2 <i>Les référentiels utilisés par le géocodeur :</i>	11
3.3.3 <i>Les tables de noms secondaires :</i>	13
3.3.4 <i>Les tables d'alias :</i>	13
3.4 LES GEOCODEURS COMPOSITES :	15
3.5 LE RE-APPARIEMENT INTERACTIF :	15
4 MISE EN PLACE DU SYSTEME DE GEOCODAGE INTERNE DE BMO :	17
4.1 LES REFERENTIELS UTILISES :	17
4.2 L'UTILISATION DE TABLES DE NOMS SECONDAIRES, UNE EVIDENCE :	17
4.3 CREATION DES STYLES DE LOCALISATEURS :	18
4.4 CREATION DES GEOCODEURS :	19
4.4.1 <i>À l'adresse :</i>	20
4.4.2 <i>À l'adresse extrapolée :</i>	20
4.4.3 <i>À la voie :</i>	22
4.4.4 <i>Au lieu-dit :</i>	22
4.4.5 <i>À la commune :</i>	23
4.5 CREATION DES GEOCODEURS COMPOSITES :	23
4.6 UTILISATION DU SYSTEME DE GEOCODAGE :	24
4.6.1 <i>Emplacement sur le réseau :</i>	24
4.6.2 <i>Interface :</i>	24
4.6.3 <i>Accompagnement de l'utilisateur :</i>	24
4.6.4 <i>Outils complémentaires :</i>	25
4.7 CONCLUSION :	25
5 DEVELOPPEMENT D'UNE APPLICATION WEB :	27
5.1 ARCGIS SERVER :	27
5.1.1 <i>Description :</i>	27
5.1.2 <i>Création des services et publication des géocodeurs composites :</i>	27
5.2 LE CHOIX DU LANGAGE DE DEVELOPPEMENT :	29
5.3 L'APPLICATION WEB GEOCODAGE :	29
5.3.1 <i>Description générale :</i>	29
5.3.2 <i>Ses caractéristiques :</i>	29
5.3.3 <i>Fonctionnement :</i>	30
5.4 LE DEPLOIEMENT DE L'APPLICATION :	31
5.5 CONCLUSION :	31

6	LA MISE A JOUR DU SYSTEME DE GEOCODAGE :	33
6.1	AUTOMATISATION GRACE A PYTHON :	33
6.2	L'OUTIL DE MISE A JOUR :	33
6.2.1	<i>Les fichiers de configurations :</i>	33
6.2.2	<i>Outils.py :</i>	34
6.2.3	<i>Maj.py :</i>	34
6.3	EXECUTION DE L'OUTIL DE MISE A JOUR :	35
6.3.1	<i>Procédure automatique :</i>	35
6.4	CONCLUSION :	36
	CONCLUSION :	37

BIBLIOGRAPHIE/WEBOGRAPHIE

ANNEXES : Document joint

ANNEXES NUMERIQUES : CD-ROM joint (Voir 4^{ème} de couverture)

Introduction :

Ce rapport finalise 5 mois de stage au service SIG de Brest métropole océane (BMO). C'est Laurent Dupont, ingénieur SIG dans ce service qui a proposé de prendre un stagiaire pour réaliser un travail nécessaire dans le cadre de l'évolution des outils SIG de BMO : la mise en place d'un outil de géocodage performant, simple d'utilisation et accessible à tous.

En effet, de plus en plus de services et directions de BMO utilisent régulièrement les SIG, que ce soit directement avec ArcGIS ou simplement en consultation avec TatukGIS Viewer. Cependant, jusqu'à maintenant, ils étaient dépendants du service SIG pour tout ce qui concerne le géocodage de listes d'adresses.

Ce dernier recevant également des listes d'adresses d'organismes extérieurs (Ex. listes d'élèves des écoles, listes des inscrits de bibliothèques, listes des commerces de la CCI¹...). Le problème est qu'il dispose d'un outil relativement ancien développé en Avenue pour ArcGIS v3 (version en voie de disparition chez BMO). De plus cet outil nécessite beaucoup de travail manuel en prétraitement et en post-traitement. Il est donc relativement chronophage et compliqué à mettre en œuvre.

Aujourd'hui, la version d'ArcGIS la plus répandue à BMO est la v9.3.1. La v10 commence à être installée mais elle est loin d'être majoritaire pour le moment. Depuis la version 9, ArcGIS dispose d'outils de géocodage, l'idée est donc de mettre en place un nouvel outil de géocodage pour BMO en s'appuyant sur l'outil proposé par ArcGIS.

¹ CCI : Chambre de Commerce et de l'Industrie



Figure 1 : Carte de situation
Source : Fond Google Maps



Figure 2 : Carte des communes de Brest métropole océane

1 Contexte du Stage :

1.1 La communauté urbaine de Brest métropole océane (BMO) :

1.1.1 Présentation générale :

Créée en mai 1974, la Communauté Urbaine de Brest, la CUB, regroupe 8 communes du nord Finistère. Il s'agit de : Bohars, Brest, Gouesnou, Guilers, Guipavas, Le Relecq-Kerhuon, Plougastel-Daoulas et Plouzané (cf. Figure 2). C'est en novembre 2004, que la communauté urbaine s'est dotée de sa nouvelle et actuelle dénomination : "Brest métropole océane"(BMO), évocatrice de ses projets de développement pour l'agglomération, ses habitants et toute la Bretagne Occidentale.

Aujourd'hui, BMO s'étend sur 220 km² et compte environ 214 504 habitants², ce qui fait d'elle la 10^{ème} plus grosse communauté urbaine de France en termes de nombre d'habitants. Cette communauté est également très active, elle consacre une grande part de son budget à l'investissement pour se moderniser, notamment avec la construction de son tramway (fin des travaux prévus pour juin 2012).

Budget (CA 2 009) :

	Dépenses	Recettes
Fonctionnement	241,85 M€	277,53 M€
Investissement	161,25 M€	163,84 M€
Total	403,10 M€	441,34 M€

1.1.2 L'administration :

La partie administrative de BMO est organisée en 5 pôles :

- Pôle Développement culturel, Éducatif et Sportif
- Pôle Solidarités – Citoyenneté – Proximité
- Pôle Espace Public et Environnement
- Pôle Développement Économique et Urbain
 - Dont dépend le service SIG.
- Pôle Ressources

Ces différents pôles sont eux-mêmes subdivisés en différents ateliers, services, directions ou missions, qui eux aussi peuvent être redécouverts.

² INSEE Recensement de la population 2007

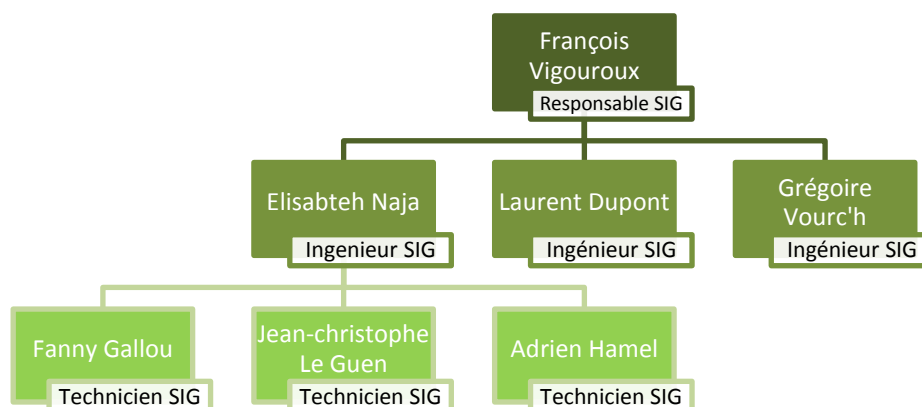


Figure 3 : Organigramme du service SIG

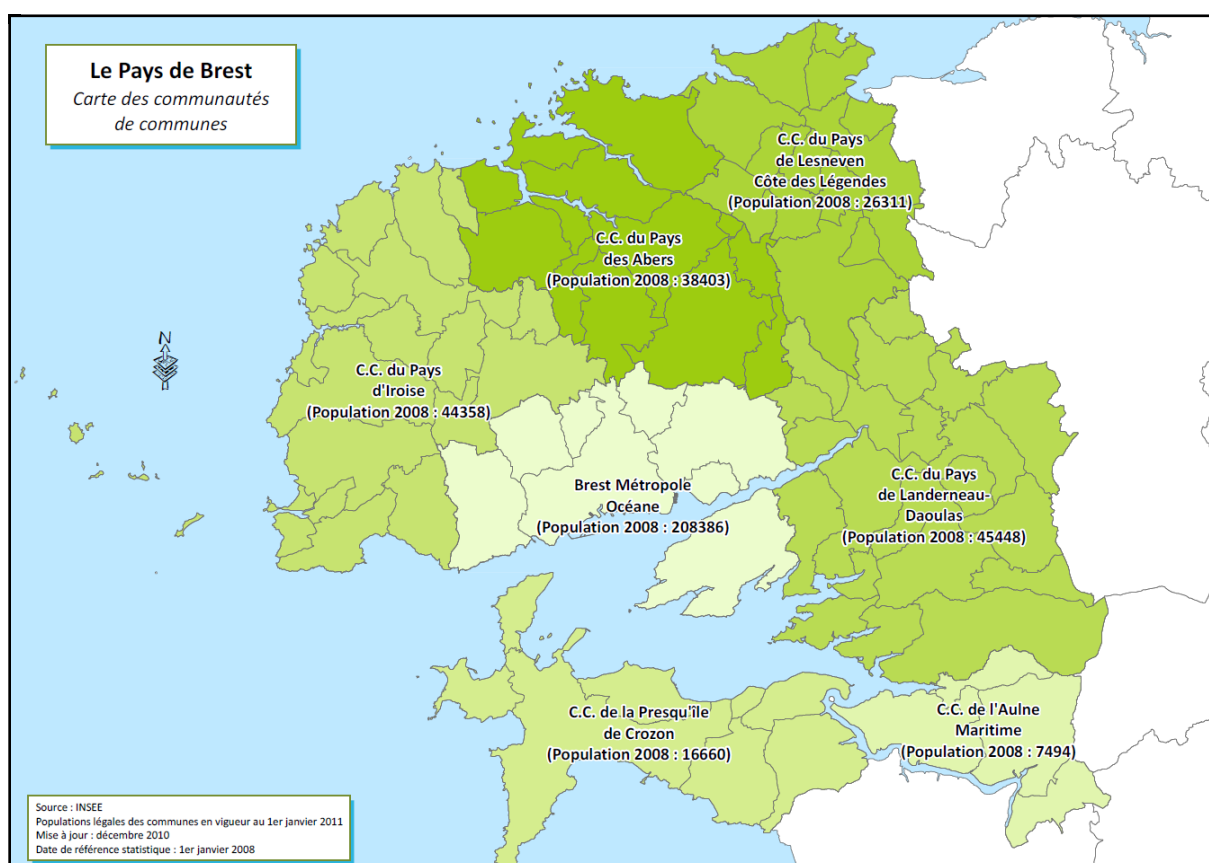


Figure 4 : Carte des communautés de communes du Pays de Brest
Source : http://sig.pays-de-brest.fr/Tout%20public/Pdb_comcom_pop.pdf

1.2 Le service SIG de BMO :

1.2.1 Présentation générale :

Le service SIG dépend de l'Atelier de Données et d'Études Urbaines (ADEUR) qui lui-même fait partie du pôle Développement Économique et Urbain (cf. ANNEXE 1: Organigramme du pôle développement économique et urbain).

Avant janvier 2011 :

Il comptait 5 personnes et gérât un territoire de 8 communes, soit environ 220 km².

Après janvier 2011 :

À cette date, le service SIG du Pays de Brest (anciennement rattaché à l'ADEUPa³) et celui de BMO ont fusionné pour n'en former plus qu'un. Cette fusion porte les effectifs du service à 7 personnes (cf. Figure 3) et a multiplié presque par 8 la superficie de territoire géré. En effet, le Pays de Brest compte 89 communes (dont les 8 de BMO), pour une superficie d'environ 1 690 km² (cf. Figure 4).

Aujourd'hui, les couches SIG de BMO et du Pays de Brest n'ont pas encore totalement fusionné et on se retrouve donc plus ou moins avec deux SIG, celui de BMO et celui du Pays de Brest, regroupés au sein d'un même service.

1.2.2 Ses missions :

Elles sont variées et peuvent évoluer au cours de l'année. Néanmoins elles peuvent être regroupées en quatre grandes catégories :

- Produire de l'information géographique 2D et 3D : pour enrichir le SIG et répondre à des besoins ponctuels des services et des partenaires.
- Administrer le SIG de BMO et le SIG du pays : gérer son contenu et le documenter.
- Développer les usages du SIG : en interne et en externe.
- Partager ses données publiques, notamment via son portail Web.

1.2.3 Son portail Web :

Le service SIG entretient un espace Web de téléchargement de données. Cet espace donne accès à trois types de données en différents formats :

Type	Format
Des cartes prêtes à l'emploi	PDF
Des fichiers de données géographiques	SHP ou KML
Des cartes personnalisables : infographie ou DAO	EPS ou DXF

Ce portail est également la vitrine du service SIG. Il permet de suivre son actualité et d'être un lien avec ses partenaires. Il se compose en réalité de deux portails, l'un pour les données de BMO, l'autre

³ L'ADEUPa est l'agence d'urbanisme du pays de Brest (<http://www.adeupa-brest.fr>).

pour les données concernant le Pays de Brest. Respectivement accessible aux adresses suivantes :
"<http://sig.brest.fr>" et "<http://sig.pays-de-brest.fr>".

2 Les besoins du service SIG et les objectifs du stage :

Chaque année le service SIG de BMO reçoit de nombreuses demandes de géocodage (au moins une centaine). Jusqu'à maintenant, il les réalisait grâce à un outil développé en Avenue sur ArcView 3. Cet outil fonctionne, mais il présente plusieurs inconvénients. D'abord, il demande beaucoup de travail manuel en pré-traitement et en post-traitement et par conséquent beaucoup de temps. Ensuite, son utilisation est relativement complexe ce qui le rend utilisable uniquement par le service SIG et enfin, la version 3 d'ArcView est en voie de disparition chez BMO au profit de la version 9.3.1 et bientôt la version 10.

Le service souhaite donc rendre le géocodage plus simple, plus efficace et utilisable directement par les personnes qui en ont besoin. C'est-à-dire dans un premier temps, les utilisateurs ArcView de BMO et ensuite les personnes non équipées d'ArcView, mais ayant accès à l'intranet.

Les objectifs sont donc de mettre d'abord en place un outil de géocodage simple, performant, nécessitant le moins de travail manuel possible et accessible en interne à toutes les personnes équipées d'ArcGIS 9.3.1 (ou 10) sur leur poste informatique et ensuite de déployer cet outil de géocodage sur l'intranet (voir sur internet) au travers d'une application Web simple, s'adressant à des personnes manipulant les SIG ou pas.

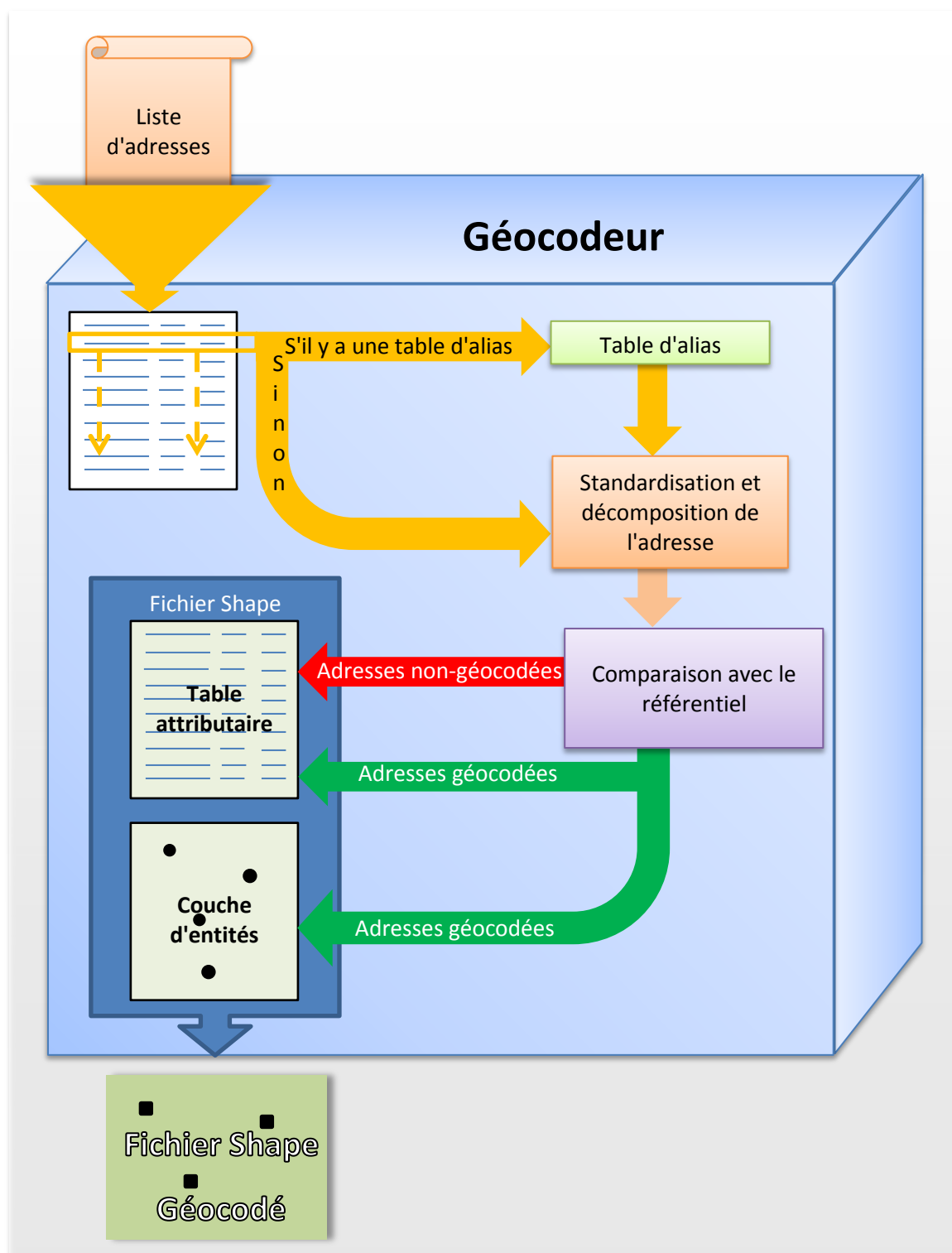


Figure 5: Schéma de fonctionnement d'un géocodeur ArcGIS

3 Le géocodage dans la solution ArcGIS :

3.1 Fonctionnement général (cf. Figure 5) :

Un géocodeur (que ce soit chez ArcGIS ou pas) s'appuie d'abord et avant tout sur un référentiel. Ce référentiel peut être une couche de points-adresses, un filaire de voies, une couche de parcelles cadastrales... À peu près n'importe quelle couche d'informations géographiques comportant des adresses.

Une particularité d'un géocodeur ArcGIS, c'est qu'il standardise les adresses qu'on lui demande de géocoder, avant de les comparer avec son référentiel. Nous détaillerons la standardisation un peu plus loin dans ce chapitre. Une fois que le géocodeur a trouvé l'adresse recherchée dans son référentiel, il lui ajoute les coordonnées correspondantes.

Lorsque le géocodage est terminé, le géocodeur ressort un fichier Shape contenant toutes les adresses présentes dans la liste d'adresse à géocoder, mais évidemment, seules les adresses géocodées sont associées à une entité géographique de type ponctuel, les autres ne sont présentes que dans la table attributaire.

Le ré-appariement interactif est également une autre particularité du géocodage chez ArcGIS. C'est un post-traitement semi-automatique permettant d'affiner le géocodage. Ceci sera abordé de façon plus approfondie dans ce rapport.

3.2 La standardisation⁴ :

La standardisation est un processus permettant en fonction de règles prédéfinies d'uniformiser les adresses du référentiel d'un géocodeur mais également les adresses qui y sont recherchées, afin d'obtenir un meilleur géocodage. Elle peut par exemple, retirer tous les mots non déterminants dans une adresse comme : le, la, les, de, du... ou encore abréger certains mots : Comme "Commandant" en "cmd", ou à l'inverse étendre certaines abréviations. Elle décompose également les adresses en leurs différents constituants : numéro, type de voie, nom de la voie...

Le référentiel ne se standardise pas lorsqu'il est ajouté au géocodeur. Il faut donc le standardiser préalablement. Évidemment, il est préférable d'utiliser le même système de standardisation pour le référentiel que celui utilisé par le géocodeur pour standardiser les adresses à géocoder.

Exemple de standardisation :

Avant standardisation :

Adresse	Code INSEE
11 bis rue Jean Jaures	29019
1 boulevard Léon Blum	29019

Après standardisation :

Numéro	Indice de rep.	Type	Nom	Code INSEE
11	B	R	JEAN JAURES	29019
1		BD	LEON BLUM	29019

⁴ Ecriture standardisée : Dans ce rapport, lorsque qu'une adresse sera écrite de la façon suivante : "11 | B | R | JEAN JAURES | 29019", "1 | | BD | LEON BLUM | 29019"... c'est qu'elle est écrite sous sa forme standardisée.

3.3 Construction d'un géocodeurs ArcGIS :

La construction d'un géocodeur ArcGIS peut aussi bien avoir lieu dans ArcView que dans ArcCatalog. Elle peut utiliser différents éléments obligatoires ou non :

- Obligatoire :
 - Un style de localisateur d'adresses⁵
 - Un référentiel
- Facultatif :
 - Une table d'alias

De plus, en fonction du style de localisateur d'adresses choisi, il peut être nécessaire d'utiliser une ou plusieurs tables de noms secondaires. Ces quatre éléments vont être abordés plus en détail ci-dessous.

3.3.1 Les styles de localisateurs d'adresses :

Ces styles sont les plans définissant comment le géocodeur doit être construit et comment il fonctionne. Ils s'appuient sur 6 fichiers. Par défaut en version 9.x et 10, ils se trouvent dans les dossiers suivants :

- (répertoire d'installation d'ArcGIS)\Geocode :
 - le "match file" (*.mat)
 - le "match key dictionary" (*.dct)
 - la "classification table" (*.cls)
 - le "command file" (*.stn)
 - le "pattern file" (*.pat)
- (répertoire d'installation d'ArcGIS)\Locators:
 - le "locator file" (*.lot)

3.3.1.1 Les Fichiers : *.dct, *.cls, *.stn, *.pat :

Ces quatre fichiers portant tous le même nom, définissent les règles de standardisation et de décomposition de l'adresse. Ils sont accompagnés si besoin est, de deux autres fichiers "Abbrev.tbl" et "expand.tbl". Ces deux fichiers contiennent des listes de mots étendus et leur version abrégée pour le premier et l'inverse pour le second. C'est dans ces fichiers que le processus de standardisation va puiser les abréviations et leurs contraires quand cela est nécessaire.

3.3.1.2 Le match file (*.mat) :

Il définit quels éléments de l'adresse seront pris en compte pour le géocodage et quel sera leur poids dans le score du géocodage. Ainsi, une erreur dans un élément de moindre importance (ex. indice de répétition) ne sera pas forcément éliminatoire. Cela permet par exemple de géocoder l'adresse "2 | B | R | JEAN JAURES | 29019" au "2 | | R | JEAN JAURES | 29019" si le 2 bis n'existe pas dans le référentiel.

NOTE : Pour plus d'information sur les fichiers *.mat, *.dct, *.cls, *.stn ou *.pat se référer à l'ANNEXE 7 : "Geocoding Rule Base Developer Guide".

⁵ Dans la suite ArcGIS, le terme "géocodeur" est remplacé par "localisateur d'adresses"

3.3.1.3 Le Locator file (*.lot) :

C'est le plan de construction du géocodeur. Il définit les champs que le géocodeur prendra en entrée, comment le référentiel doit être utilisé, quel processus de standardisation est à appliquer, le géocodeur utilise-t-il une table de nom secondaire ou pas... Il définit également les requêtes qui seront effectuées entre l'adresse recherchée et le référentiel.

3.3.2 Les référentiels utilisés par le géocodeur :

Le référentiel est la base du géocodeur. C'est en comparant les adresses de la liste à géocoder avec lui que le géocodeur peut les localiser. Étant dans une solution ArcGIS, ce référentiel est une couche Shape ou son équivalent en géodatabase.

Lors de sa création, le géocodeur intègre son référentiel, il le "duplique". L'avantage, c'est qu'il ne dépend plus de la couche d'origine, si celle-ci disparaît, le géocodeur fonctionne toujours. Elle peut également être modifiée et utilisée par n'importe qui sans problème de lecture seule (notamment pour les géodatabase personnelle). L'inconvénient est que si cette couche est modifiée, les mises à jour ne seront pas automatiquement répercutées dans le géocodeur.

Un référentiel adresses complet sur un territoire, peut suffire à lui seul à géocoder toutes les adresses de ce territoire qu'on lui soumet, à condition que toutes les personnes qui l'utilisent, emploient dans leurs listes d'adresses uniquement les noms officiels des voies et sans erreur. Or c'est rarement le cas. Une des erreurs la plus souvent commise sur le territoire de BMO est la "rue de Gouesnou" qui est appelé "route de Gouesnou". Le problème pour le géocodeur, c'est qu'il ne connaît que la version officielle qui est dans son référentiel.

La solution pour pallier ce problème, c'est l'utilisation de tables de noms secondaires.

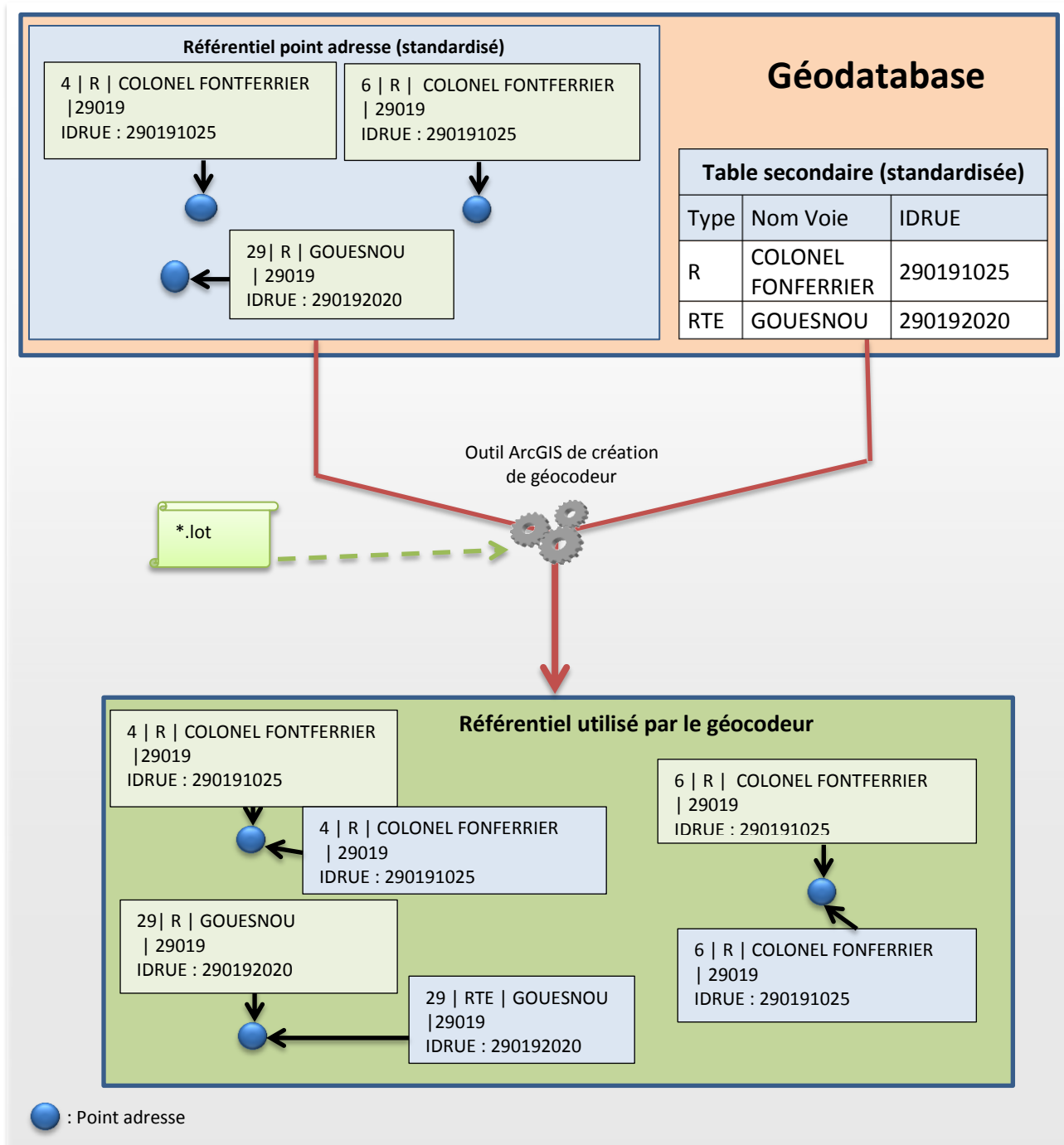


Figure 6: Fusion d'un référentiel et d'une table de noms secondaires, lors de la création d'un géocodeur

3.3.3 Les tables de noms secondaires :

Ces tables permettent d'associer plusieurs adresses à un même point au moment de l'inclusion du référentiel dans le géocodeur. Elles sont constituées de différents champs contenant des éléments d'adresses standardisées de la même manière que le référentiel (type de voie, nom de voie...) plus un identifiant qui correspond à l'identifiant de l'adresse dans le référentiel avec laquelle le géocodeur doit associer ces éléments.

Lors de sa création, le géocodeur va donc fusionner son référentiel et sa (ou ses) table(s) de noms secondaires pour ne former plus qu'un seul référentiel, selon les règles définies dans son Locator file (*.lot). C'est ce référentiel fusionné que le géocodeur utilisera lors du géocodage et c'est ce référentiel qui est modifié. La couche d'origine reste inchangée (cf. Figure 6).

Il y a toutefois une petite contrainte à respecter. Lorsque l'on utilise un géocodeur avec table(s) de noms secondaires, le référentiel du géocodeur et la (ou les) table(s) de noms secondaires doivent se trouver dans une même géodatabase.

3.3.4 Les tables d'alias :

Une table d'alias sert à apprendre au géocodeur l'adresse de lieux particuliers (Ex. La mairie de Brest). Ainsi, si le géocodeur trouve dans sa liste d'adresses à géocoder une adresse qui n'en est pas vraiment une mais bien un nom de lieu et que ce nom se trouve dans sa table d'alias, il la remplacera par l'adresse qui lui est indiquée dans cette table et pourra ainsi la trouver dans son référentiel.

Cette table d'alias contient un nom de lieu (l'alias) et une adresse associée. Ces alias doivent être écrits de toutes les façons possibles de les rencontrer, car le géocodeur passe par la table d'alias avant le processus de standardisation (cf. Figure 5), néanmoins, il n'est pas sensible à la casse.

Les adresses ne doivent pas nécessairement être écrites de façon standardisée, la standardisation intervenant juste après le passage par la table d'alias. Toutefois, il est préférable de le faire, ainsi on est sûr que les adresses seront correctement trouvées par le géocodeur.

Exemple de table d'alias :

Alias	Adresse	Code INSEE
Mairie Brest	2 R FREZIER	29019
Mrie Brest	2 R FREZIER	29019

Grâce à cette table, si le géocodeur trouve dans sa liste d'adresse à géocoder une adresse : "Mairie de Brest", il la remplacera par "2 R FREZIER 29019".

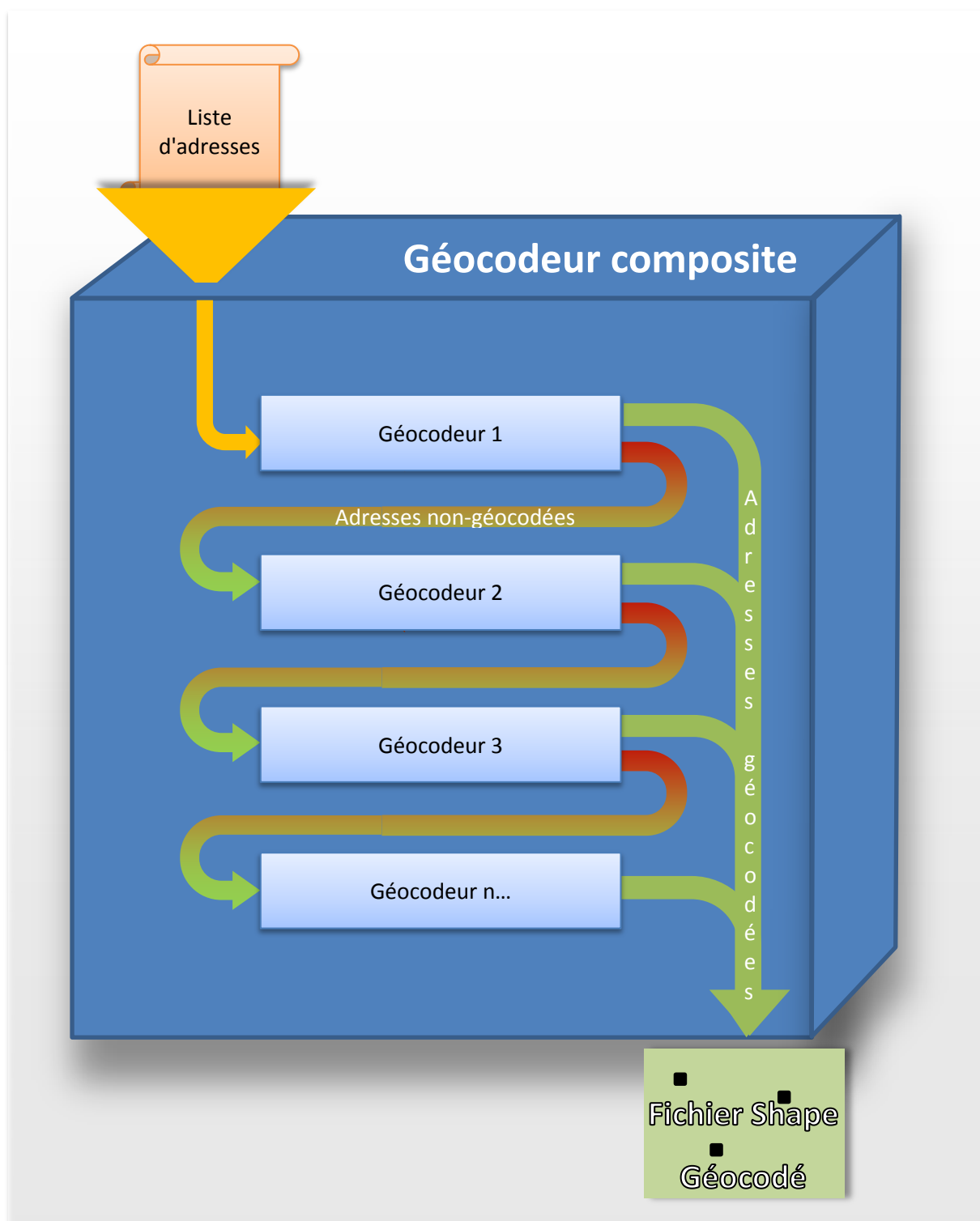


Figure 7 : Schéma de fonctionnement d'un géocodeur composite

3.4 Les géocodeurs composites :

ArcGIS propose par défaut, un type de localisateur appelé "Composite". Ce type de localisateur permet de créer un "gros" géocodeur intégrant plusieurs géocodeurs ArcGIS. L'énorme avantage est de pouvoir utiliser dans un seul géocodeur autant de référentiel qu'on le souhaite. Ainsi, on peut intégrer un géocodeur géocodant à l'adresse avec un référentiel points-adresses, un autre à la voie avec un référentiel voirie...

Le géocodeur composite utilisera ses "sous-géocodeurs" dans l'ordre qu'on lui impose. Cela permet de le forcer à utiliser les sous-géocodeurs du plus précis au moins précis.

Les adresses ne sont évidemment géocodées qu'une seule fois, par le premier sous géocodeur dans lequel elles ont été trouvées. Seules les adresses non géocodées passent dans le sous-géocodeur suivant (cf. Figure 7).

Exemple de géocodeur composite :

Géocodeur composite	
Géocodeur 1	Référentiel points-adresses
Géocodeur 2	Référentiel voies
Géocodeur 3	Référentiel communes

3.5 Le ré-appariement interactif :

Le ré-appariement interactif est un post-traitement applicable aux fichiers Shape géocodés permettant d'affiner le géocodage. Il est accessible uniquement dans d'ArcCatalog et se présente sous la forme de cette interface :

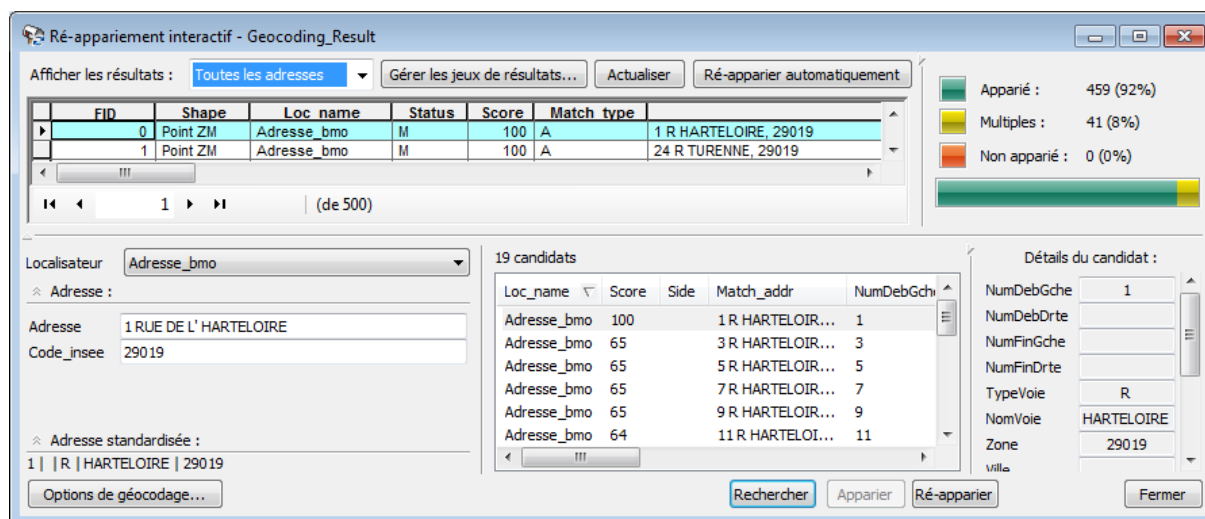


Figure 8 : Fenêtre de ré-appariement interactif

On peut y voir différents éléments comme la table attributaire du fichier géocodé, les statistiques du géocodage, mais le plus important est la partie inférieure. En effet c'est là que l'utilisateur va pouvoir modifier les adresses qui ont été mal géocodées et les ré-appariement de façon semi-automatique avec la bonne adresse. Parfois, il ne manque pas grand-chose pour qu'une adresse soit correctement géocodée, un espace manquant par exemple : "2 bis rueJean Jaures" ne sera pas géocodé. Pour plus d'information concernant le ré-appariement interactif, consulter l'ANNEXE 2 : "Manuel UTILISATEUR".

4 Mise en place du système de géocodage interne de BMO :

4.1 Les référentiels utilisés :

Huit référentiels ont été utilisés et regroupés dans des géocodeurs composites afin d'optimiser le géocodage. Ces huit référentiels sont de type et d'emprise différents. Ils ont également des précisions de géocodage différentes de par leur nature : point-adresse, voie, communes... En revanche, ils ont tous le même système de projection : Lambert 93.

Référentiel	Source	Type	Emprise
Adre_p	SIG de BMO	Points-adresses	BMO
Segm	SIG de BMO	Filaire de voie	BMO
Lxd_t_Bmo	SIG de BMO	Points (Correspond aux lieux-dits, zones commerciales, zones industrielles...)	BMO
Routes_Pays_de_Brest	SIG de BMO	Filaire de voie (<i>Note : ce référentiel est en cours de création</i>)	Pays de Brest (sans BMO)
Lxd_t_pays	SIG de BMO	Points (Correspond aux lieux-dits, zones commerciales, zones industrielles...)	Pays de Brest (sans BMO)
Bd adresse	IGN	Points-adresses	Pays de Brest
ROUTE_NOMMEE (issue de la Bd Topo)	IGN	Filaire de voie	Pays de Brest
ville	IGN	Points (correspond au chef-lieu de chaque commune)	France métropolitaine

Le référentiel "ville" provient du Répertoire Géographique des Communes (RGC) 2010 en téléchargement gratuit sur le site de l'IGN.

Dans chacun de ces référentiels hormis le référentiel "ville", les communes ne sont identifiées que par leur code INSEE. Une jointure sur le champ du code INSEE a donc été réalisée avec une table contenant le nom des communes, leur code postal et leur code INSEE (table issue du RGC), afin de pouvoir utiliser ces référentiels dans des géocodeurs utilisant le nom de commune par exemple.

4.2 L'utilisation de tables de noms secondaires, une évidence :

Les référentiels seuls apportent au géocodeur les noms officiels des adresses et des voies, mais bien souvent il existe des noms d'usages qui reviennent régulièrement. Soit parce que la rue a changé de nom ou bien c'est simplement une erreur qui est tellement répétée que l'on finit par se demander quel est le nom officiel. Une de ces erreurs la plus couramment rencontrée sur BMO est celle de la "rue de Gouesnou" qui est très souvent appelée "route de Gouesnou".

L'utilisation d'une table de noms secondaires permet de pallier ce problème et d' "apprendre" au géocodeur que le "15 | | RTE | GOUESNOU | 29200 | BREST", correspond en réalité au "15 | | R | GOUESNOU | 29200 | BREST".

4.3 Création des styles de localisateurs :

Par défaut dans ArcGIS 9.x et 10, on ne trouve pas de style de localisateur français utilisant des tables de noms secondaires. Seuls des styles de localisateurs US utilisent ce genre de tables. Le problème est que les adresses françaises et les adresses US ne sont pas structurées de la même manière. Il n'est donc pas possible d'utiliser un style de localisateur US pour géocoder des adresses françaises. Ces styles de localisateurs français utilisant des tables de noms secondaires ont donc été créés.

Les fichiers composant un localisateur d'adresses sont écrits en script lisible hormis le "*.pat" qui est encodé. Les scripts ne correspondent à aucun langage connu, il est propre à ArcGIS ou au moins aux styles de localisateurs.

Les styles de localisateur sont écrits dans leur langage propre et avec une logique qui leur est également propre. Les styles de localisateurs français utilisant une table de noms secondaires sont donc créés par comparaison entre un style de localisateur français et un style de localisateur US utilisant une table de noms secondaires. La consultation du document "Geocoding Rule Base Developer Guide" m'a également beaucoup aidé (cf. Annexe 7). L'encodage du "*. pat" n'est pas un problème, ce dernier n'ayant pas besoin d'être modifié.

Une fois la logique et la technique de création d'un style de localisateur avec table de noms secondaires acquise, il a été possible de créer tous les différents styles de localisateur suivant :

Type	Nom
Géocodeur à l'adresse nécessitant un numéro, un type de voie, un nom de voie, un identifiant de commune ⁶ et prenant en compte les indices de répétition.	<ul style="list-style-type: none"> Adresses fr avec code INSEE (ou postal) et nom secondaire Adresses fr avec code INSEE (ou postal)
	<ul style="list-style-type: none"> Adresses fr avec commune et nom secondaire Adresses fr avec commune
	<ul style="list-style-type: none"> Adresses fr avec cp, commune et nom secondaire Adresses fr avec cp, commune
Géocodeur à l'adresse extrapolée nécessitant un numéro de début et de fin de voie à gauche et à droite, un type de voie, un nom de voie, un identifiant de commune et ne prenant pas en compte les indices de répétition.	<ul style="list-style-type: none"> Adresses fr extrapolé avec code INSEE (ou postal) et nom secondaire Adresses fr extrapolé avec code INSEE (ou postal)
	<ul style="list-style-type: none"> Adresses fr extrapolé avec commune et nom secondaire Adresses fr extrapolé avec commune
	<ul style="list-style-type: none"> Adresses fr extrapolé avec cp, commune et nom secondaire Adresses fr extrapolé avec cp, commune
Géocodeur à la voie nécessitant un type de voie, un nom de voie et un identifiant de commune.	<ul style="list-style-type: none"> Voies fr avec code INSEE (ou postal) et nom secondaire Voies fr avec code INSEE (ou postal)
	<ul style="list-style-type: none"> Voies fr avec commune et nom secondaire Voies fr avec commune
	<ul style="list-style-type: none"> Voies fr avec cp, commune et nom secondaire Voies fr avec cp, commune
Géocodeur à la commune nécessitant uniquement un identifiant de commune	<ul style="list-style-type: none"> Commune fr avec code INSEE (ou postal)
	<ul style="list-style-type: none"> Commune fr avec cp et nom de commune
	<ul style="list-style-type: none"> Commune fr avec nom de commune

⁶ Cela peut être le code INSEE, le code postal, le nom de la commune ou le code postal et le nom de commune.

On a donc ici 21 styles de localisateurs différents permettant d'exploiter pleinement tous les référentiels cités précédemment.

(Tous les styles sont dans : Annexes_numeriques\styles_localisateurs)

Adresse extrapolée ?

Une adresse extrapolée est une adresse obtenue par extrapolation des numéros sur un segment de voie. Ce n'est possible que si l'on connaît le numéro de début et le numéro de fin de la voie, à gauche et à droite.

Exemple d'adresse extrapolée :

Dans le référentiel	<div> <div>2</div> <div>1</div> <div>R JEAN JAURES</div> <div>8</div> <div>7</div> </div>
Extrapolée dans le géocodeur	<div> <div>2</div> <div>1</div> <div>4</div> <div>3</div> <div>R JEAN JAURES</div> <div>6</div> <div>5</div> <div>8</div> <div>7</div> </div>

Deux particularités de ces styles de localisateur d'adresses :

La première, visible ici, est que pour chaque style de localisateur utilisant une table de noms secondaires, on trouve un style identique mais n'utilisant pas de table de noms secondaires. Nous avons vu précédemment qu'il est possible de réunir au sein d'un géocodeur composite, différents géocodeurs utilisant des référentiels eux aussi différents. Si l'on souhaite y regrouper des géocodeurs utilisant des tables de noms secondaires et d'autres n'en utilisant pas, il est préférable d'utiliser ces styles identiques pour obtenir un géocodage semblable et homogène, quel que soit le géocodeur dans lequel l'adresse recherchée a été trouvée.

Deuxième particularité, elle concerne les géocodeurs utilisant le nom de commune ou le nom de commune et le code postal, ils intègrent tous une table supplémentaire contenant les noms secondaires pour les communes. En effet, les communes ayant des noms relativement longs, ou composés, sont parfois abrégées. Le géocodeur ne pouvant le deviner, il faut lui apprendre ces "variantes" de noms.

Ex : La commune de Camaret-sur-mer est très souvent appelée Camaret.

4.4 Création des géocodeurs :

Tous les référentiels et tous les styles de localisateur cités précédemment sont utilisés. Seuls les géocodeurs utilisant les référentiels de BMO bénéficient d'une table de noms secondaires, car théoriquement toutes les adresses doivent s'y trouver. En revanche, tous les géocodeurs utilisant le nom de commune, ou le nom de commune et le code postal, utilisent une table de noms secondaires pour les noms de communes (nom_ville_secondaire).

4.4.1 À l'adresse :

Référentiel	Style de localisateur	Nom du géocodeur
Adre_p + nom_voie_secondaire + Table d'alias	Adresses fr avec code INSEE (ou postal) et nom secondaire	Adresse_insee_bmo
		Adresse_cp_bmo
	Adresses fr avec commune et nom secondaire	Adresse_commune_bmo
	Adresses fr avec cp, commune et nom secondaire	Adresse_com_cp_bmo
Bd Adresse + Table d'alias	Adresses fr avec code INSEE (ou postal)	Adresse_insee_bd_adresse
		Adresse_cp_bd_adresse
	Adresses fr avec commune	Adresse_commune_bd_adresse
	Adresses fr avec cp, commune	Adresse_com_cp_bd_adresse

Seuls les géocodeurs géocodant à l'adresse possèdent une table d'alias. En effet, il n'est pas intéressant de géocoder à un autre niveau que l'adresse, un lieu que l'on connaît précisément.

4.4.2 À l'adresse extrapolée :

Référentiel	Style de localisateur	Géocodeur en sortie
Segm + nom_voie_secondaire	Adresses fr extrapolé avec code INSEE (ou postal) et nom secondaire	Adresse_extra_insee_bmo
		Adresse_extra_cp_bmo
	Adresses fr extrapolé avec commune et nom secondaire	Adresse_extra_commune_bmo
	Adresses fr extrapolé avec cp, commune et nom secondaire	Adresse_extra_com_cp_bmo
Routes_Pays_de_Brest + nom_voie_secondaire	Adresses fr extrapolé avec code INSEE (ou postal) et nom secondaire	Adresse_extra_insee_pays
		Adresse_extra_cp_pays
	Adresses fr extrapolé avec commune et nom secondaire	Adresse_extra_commune_pays
	Adresses fr extrapolé avec cp, commune et nom secondaire	Adresse_extra_com_cp_pays
ROUTE_NOMMEE (issue de la Bd Topo)	Adresses fr extrapolé avec code INSEE (ou postal)	Adresse_extra_insee_topo
		Adresse_extra_cp_topo
	Adresses fr extrapolé avec commune	Adresse_extra_commune_topo
	Adresses fr extrapolé avec cp, commune	Adresse_extra_com_cp_topo

Il y a deux particularités concernant les champs contenant le numéro de début et de fin de la voie à gauche et à droite pour chacun de ces trois référentiels. En effet dans certain cas, il se peut que les champs ne soient pas remplis, par exemple s'il n'y a pas d'adresse à gauche de la rue, comme ceci :

Type	Nom	Num_deb_g	Num_fin_g	Num_deb_d	Num_fin_d
R	JEAN JAURES	2	8	1	7
RTE	LAVALLOT			1	25

L'inconvénient de ces champs vides, c'est que lorsque le géocodeur va trouver une adresse à géocoder du type "route de Lavallot" sans numéro, il va la standardiser de la manière suivante :

" | RTE | LAVALLOT" et il va la géocoder sur la gauche de la route de Lavallot car pour lui, le champ vide "Num_deb_g" correspond.

Pour pallier ce problème, avant de construire le géocodeur, on vérifie les 4 champs contenant les numéros et si l'on trouve un enregistrement vide on y inscrit "NR", pour "Non Renseigné".

Type	Nom	Num_deb_g	Num_fin_g	Num_deb_d	Num_fin_d
R	JEAN JAURES	2	8	1	7
RTE	LAVALLOT	NR	NR	1	25

Ainsi, " | RTE | LAVALLOT" ne correspond pas à " NR | RTE | LAVALLOT ".

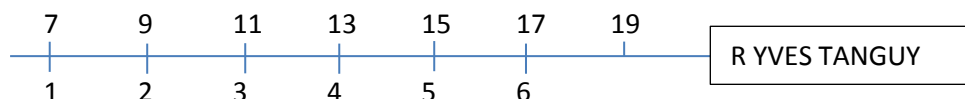
Deuxième particularité, la numérotation séquentielle, c'est-à-dire que les numéros sur un même côté de la voie ne sont plus 2, 4, 6, 8 mais 1, 2, 3, 4, 5, 6, 7, 8. Le problème pour les géocodeurs, c'est d'extrapoler correctement tous les numéros. Il est capable de comprendre de lui-même que les adresses sont séquentielles à condition que les numéros de début et de fin de la voie soient de parité différente.

Ex :

Ici nous avons une adresse avec une numérotation de type séquentiel (TYNUM = s). On observe que les numéros de début et de fin à gauche sont de parité différente (1 = impair, 6 = pair), alors qu'à droite, ils sont de même parité (7 = impair, 19 = impair)

LIBRU	Num_deb_g	Num_fin_g	Num_deb_d	Num_fin_d	TYNUM
RUE YVES TANGUY	1	6	7	19	s

Voici les adresses extrapolées par le géocodeur :



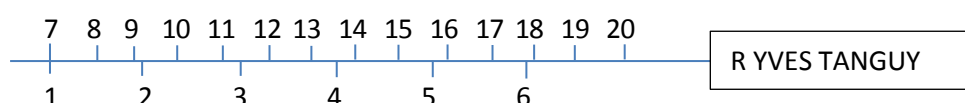
On constate bien ici que sur le côté gauche, tous les numéros ont été extrapolés de façon séquentielle, alors que sur le côté droit les numéros ont été extrapolés de façon "classique".

La solution, on vérifie le champ "TYNUM", lorsque celui-ci est égal à s (séquentiel). On vérifie la parité des numéros du début et de fin côté droit, s'ils sont de même parité, on ajoute au numéro de fin. Ensuite, on refait la même opération côté gauche.

Ainsi, après traitement (on a ajouté 1 à Num_fin_d) on obtient ceci :

LIBRU	Num_deb_g	Num_fin_g	Num_deb_d	Num_fin_d	TYNUM
RUE YVES TANGUY	1	6	7	20	s

Adresses extrapolées par le géocodeur :



4.4.3 À la voie :

Référentiel	Style de localisateur	Géocodeur en sortie
Segm + nom_voie_secondaire	Voies fr avec code INSEE (ou postal) et nom secondaire	Voie_insee_bmo
		Voie_cp_bmo
	Voies fr avec commune et nom secondaire	Voie_commune_bmo
Routes_Pays_de_Brest + nom_voie_secondaire	Voies fr avec code INSEE (ou postal) et nom secondaire	Voie_insee_pays
		Voie_cp_pays
	Voies fr avec commune et nom secondaire	Voie_commune_pays
ROUTE_NOMMEE (issue de la Bd Topo)	Voies fr avec code INSEE (ou postal)	Voie_insee_topo
		Voie_cp_topo
	Voies fr avec commune	Voie_commune_topo
	Voies fr avec cp, commune	Voie_com_cp_topo

4.4.4 Au lieu-dit :

Référentiel	Style de localisateur	Géocodeur en sortie
Lxd_t_bmo + nom_lxd_t_bmo_secondaire	Voies fr avec code INSEE (ou postal) et nom secondaire	Lxd_t_bmo_insee
		Lxd_t_bmo_cp
	Voies fr avec commune et nom secondaire	Lxd_t_bmo_commune
Lxd_t_pays + nom_lxd_t_pays_secondaire	Voies fr avec code INSEE (ou postal) et nom secondaire	Lxd_t_pays_insee
		Lxd_t_pays_cp
	Voies fr avec commune et nom secondaire	Lxd_t_pays_commune
	Voies fr avec cp, commune et nom secondaire	Lxd_t_pays_com_cp

Premier point important ici : ce sont des styles de localisateurs d'adresses de type "Voies" qui ont été utilisés pour créer des géocodeurs avec des référentiels ponctuels. Cela s'explique par la structure des référentiels lieux-dits. En effet, un lieu-dit est caractérisé par :

- Un type (LD, ZA, ZI...)
- Un nom
- Un identifiant de commune (Code INSEE, code postal...)

Ces trois éléments sont les mêmes que ceux qui caractérisent une voie, on peut donc apparenter un lieu-dit à une voie et utiliser un style de localisateur de type "Voies" pour créer un géocodeur avec un référentiel de lieux-dits.

Deuxième point, les tables secondaires pour les lieux-dits, ne sont pas les mêmes que pour les autres géocodeurs. En effet, celles-ci sont générées automatiquement lors de la création des géocodeurs pour pallier le problème suivant : un lieu-dit peut être nommé "Kergoat" ou "lieu-dit de Kergoat". Or, dans les référentiels lieux-dits de BMO, c'est la première dénomination qui est utilisée :

Type	Nom	Code INSEE
	LEUR GOZ	29019
CITE	SUFFREN	29019
	QUIZAC	29019
	KERGOAT	29019
ZAC	PORT COMMERCE	29019

Dans ce cas de figure, une adresse "LD | KERGOAT | 29019" ne peut pas être trouvée. Les tables secondaires des lieux-dits sont donc générées à partir du DBF de chacun des deux référentiels. Un calcul est effectué sur le champ "Type", lorsque celui-ci est vide, on y ajoute "LD".

Type	Nom	Code INSEE
LD	LEUR GOZ	29019
CITE	SUFFREN	29019
LD	QUIZAC	29019
LD	KERGOAT	29019
ZAC	PORT COMMERCE	29019

Ainsi, "LD | KERGOAT | 29019" peut être trouvé, tout comme " | KERGOAT | 29019".

4.4.5 À la commune :

Référentiel	Style de localisateur	Géocodeur en sortie
Communes	Commune fr avec code INSEE (ou postal)	Commune_insee
		Commune_cp
	Commune fr avec nom de commune	Commune_commune
	Commune fr avec cp et nom de commune	Commune_com_cp

4.5 Création des géocodeurs composites :

Les 44 géocodeurs peuvent être regroupés en 4 catégories en fonction de l'identifiant de commune qu'ils utilisent :

- Géocodeurs utilisant le **Code INSEE**
- Géocodeurs utilisant le **Code postal**
- Géocodeurs utilisant le **Nom de commune**
- Géocodeurs utilisant le **Code postal et le Nom de commune**

Quatre géocodeurs composites ont donc été créés en fonction de ses quatre catégories. Ils sont tous les quatre composés de 11 géocodeurs ordonnés comme ceci :

Géocodeur composite	
Nom du géocodeur	Précision de géocodage
Adresse_bmo	Adresse
Adresse_bd_adresse	Adresse
Adresse_extra_bmo	Adresse extrapolée
Adresse_extra_pays	Adresse extrapolée
Adresse_extra_topo	Adresse extrapolée
Voie_bmo	Voie
Voie_pays	Voie
Lxdt_bmo	Lieu-dit
Lxdt_pays	Lieu-dit
Voie_topo	Voie
Commune	Commune

On observe que Voie_topo qui est un géocodeur à la voie se trouve après les géocodeurs au lieu-dit. C'est parce que dans la couche ROUTE_NOMMEE de la BD topo, il y a des lieux-dits, mais qui sont placés à la voie. Si ce géocodeur avait été placé avant Lxdt_bmo et Lxdt_pays, certains lieux-dits auraient été géocodés à la voie avec Voie_topo.

4.6 Utilisation du système de géocodage :

4.6.1 Emplacement sur le réseau :

L'outil de géocodage (référentiels, géocodeurs et géocodeurs composites) est placé sur un lecteur (J:\ en l'occurrence) qui est visible en lecture par tous les utilisateurs d'ArcGIS, mais disponible en écriture seulement pour le service SIG et les administrateurs réseaux. Ceci permet aux utilisateurs d'avoir accès à l'outil de géocodage, sans pouvoir le modifier.

4.6.2 Interface :

L'utilisateur lors d'un géocodage, n'est confronté qu'à quatre géocodeurs. Ce sont les quatre géocodeurs composites :

- Geocodage_avec_code_insee
- Geocodage_avec_code_postal
- Geocodage_avec_nom_commune
- Geocodage_avec_nom_commune_et_code_postal

Ils portent des noms clairs qui permettent à l'utilisateur de rapidement trouver quel géocodeur il doit utiliser en fonction de la définition des communes dans sa liste d'adresses. Il est conseillé d'utiliser les géocodeurs via ArcCatalog afin d'accéder directement à l'interface de ré-appariement interactif.

4.6.3 Accompagnement de l'utilisateur :

L'utilisateur dispose d'un manuel utilisateur détaillé (cf. ANNEXE 2 : "Manuel UTILISATEUR") ou d'un mémo de géocodage (cf. ANNEXE 3 : "Mémo de géocodage") pour le guider à travers l'outil de géocodage.

De plus, 4 formations sur cet outil ont déjà été dispensées et d'autres seront sans doute proposées à l'avenir par le service SIG.

4.6.4 Outils complémentaires :

Un outil complémentaire a également été développé en langage Python. Il permet à l'utilisateur de retirer facilement les caractères spéciaux de sa liste d'adresses (Excel ou dBASE) avant de la géocoder. En effet les adresses dans les référentiels sont écrites sans caractères spéciaux et les géocodeurs ne sont pas capables de retirer eux-mêmes ces caractères.

4.7 Conclusion :

Grâce à ces styles de localisateurs français utilisant des tables de noms secondaires, il a été possible de créer 44 géocodeurs différents, exploitant pleinement tous les référentiels disponibles et intéressants pour le géocodage sur BMO et le Pays de Brest. Ces 44 géocodeurs regroupés en 4 géocodeurs composites permettent une utilisation simple, rapide et performante de ce système de géocodage. De plus, le ré-appariement interactif proposé par ArcGIS dans ArcCatalog permet d'affiner le géocodage.

Le premier objectif est atteint : Cet outil répond au premier besoin exprimé par le service SIG de BMO et à celui de ce stage, à savoir : "Mettre en place un outil de géocodage, simple, performant et utilisable par toutes les personnes équipées d'ArcGIS 9.3.1 ou 10".

L'objectif suivant est rendre cet outil disponible pour toutes les personnes non-équipées d'ArcGIS 9.3.1 ou 10.

Accueil

Services

Gérer des services

Publier une ressource SIG

Ajouter un nouveau service

Liens réseau KML

Services OGC

Paramètres

Serveur SIG

Sécurité

Gérer des services

Publier une ressource SIG

Ajouter un nouveau service

Services dans : HCU51 (racine)

Démarrer

Arrêter

Pause

Redémarrer

Supprimer

Gérer les dossiers

Nom

Type

Etat

Instances (Utilisées/En cours d'exécution)

Autorisations

Modifier

EL_decoupage3

Map Service

Démarré

0/1

Geocodage_avec_code_insee

Geocode Service

Démarré

0/1

Geocodage_avec_code_postal

Geocode Service

Démarré

0/1

Geocodage_avec_nom_commune

Geocode Service

Démarré

0/1

Geocodage_avec_nom_commune_et_code_postal

Geocode Service

Démarré

0/1

PlanVille

Map Service

Arrêté

0/0

Sites_patrimoniaux

Map Service

Arrêté

0/0

Description :

Source : \\dit.cb\dfs\v116\OUTILS\geocodage\geocodeur\Localisateurs_Server\localisateur_composite\Geocodage_avec_code_insee

Fonctions : Géocodage

Groupe : Groupé

Démarrage : Automatique

Chemin où se trouve le géocodeur sur le volume réseau

Les quatre géocodeurs

Figure 9 : Les géocodeurs dans l'interface ArcGIS SERVER Manager

26

Page : 26

5 Développement d'une application Web :

Pourquoi ?

Dans un premier temps, cette application Web permettra aux personnes internes à BMO, ne disposant pas d'ArcGIS v9.3.1 ou 10, de géocoder eux-mêmes leurs listes d'adresses. Dans un deuxième temps, elle pourrait être ouverte sur le Web et permettre à toutes personnes souhaitant géocoder une liste d'adresses sur le territoire du Pays de Brest, de disposer d'un outil performant utilisant des référentiels mis à jour régulièrement et d'une grande précision.

Les ressources disponibles :

BMO dispose de serveurs ArcGIS SERVER (v10), de serveurs Windows server IIS ainsi que du logiciel Visual Studio 2010.

5.1 ArcGIS SERVER :

5.1.1 Description :

ArcGIS SERVER est, comme son nom l'indique, le système de serveur de la solution ArcGIS. Il prend en charge les tâches des plateformes du système et peut créer, gérer et distribuer des services SIG sur le Web, pour des clients bureautiques, mobiles et Web.

Les fonctionnalités d'ArcGIS Server sont nombreuses, elles permettent de gérer et répliquer des données, de faire de la cartographie, de l'analyse spatiale... mais elles permettent aussi, et c'est qui nous intéresse, de mettre en place des services Web SIG.

5.1.2 Création des services et publication des géocodeurs composites :

Pour mettre en place un service Web SIG utilisant nos géocodeurs, il faut d'abord les publier sur ArcGIS SERVER. La publication des géocodeurs se fait via ArcGIS SERVER Manager⁷. Elle consiste à créer des services Web SIG sur le serveur pour chaque géocodeur. Ces services sont de type "Geocode Service" et s'appuient chacun sur un géocodeur existant sur un volume réseau (ou un disque local). En publiant les quatre géocodeurs composites, on crée donc quatre services (cf. Figure 9).

La publication des géocodeurs composites sur ArcGIS SERVER entraîne la prise d'un "lock" sur les 4 géocodeurs composites ainsi que sur les 44 géocodeurs qui les composent, les rendant inutilisables via ArcCatalog. De plus, les géocodeurs composites utilisés via ArcGIS SERVER doivent employer les chemins relatifs vers leurs "sous-géocodeurs" alors que ceux utilisés via ArcCatalog doivent au contraire employer les chemins absolus.

Pour éviter ces problèmes, les 44 géocodeurs et les 4 géocodeurs composites ont été dupliqués avant d'être publiés. Ainsi, on a un système de géocodeurs pour ArcCatalog et un autre pour ArcGIS SERVER.

⁷ ArcGIS SERVER Manager : interface web permettant de gérer les services et autres paramètres du serveur

5.1.2.1 Caractéristiques des services : "Geocode Service" :

Une fois les 4 géocodeurs composites publiés sur ArcGIS SERVER, nous disposons de 4 services Web de géocodage. Le fonctionnement, les différentes fonctions, les objets utilisés par ces services sont décrits dans des fichiers WSDL⁸ (cf. Annexes_numeriques\exemple_de_fichiers_WSDL).

Les services de types "Geocode Service" d'ArcGIS SERVER disposent des fonctions suivantes :

Fonction	Description
FindAddressCandidates	Geocodes a single address and returns multiple results (candidates).
GeocodeAddress	Geocodes a single address and returns a single result.
GeocodeAddresses	Geocodes multiple addresses and returns a single match result for each address.
GetAddressFields	Fields used to define address information submitted to a geocode service.
GetCandidateFields	Fields contained in a list of candidates returned from geocoding an address using the FindAddressCandidates method.
GetDefaultInputFieldMapping	Acceptable field name mappings (inputs) to address fields in a geocode service.
GetIntersectionCandidateFields	Fields contained in a list of candidates returned from geocoding an intersection with the FindAddressCandidates method.
GetLocatorProperties	The default properties of a geocode service (locator). Often used to modify characteristics of a geocode operation.
GetResultFields	Fields contained in the results of a geocode operation using the GeocodeAddress and GeocodeAddresses methods.
GetSingleLineAddressField	Field used to define the single line address information to be submitted to a geocode service (not supported for pre-ArcGIS 10 locators).
GetStandardizedFields	Fields contained in an standardized address returned from the StandardizeAddress method (only supported for pre-ArcGIS 10 locators).
GetStandardizedIntersectionFields	Fields contained in a standardized intersection returned from the StandardizeAddress method (only supported for pre-ArcGIS 10 locators).
ReverseGeocode	Returns the address closest to the given point, and within a specified search distance.
StandardizeAddress	Standardizes an address using rules defined within the geocode service (only supported for pre-ArcGIS 10 locators).

Source : "http://tasks.arcgisonline.com/ArcGIS/SDK/SOAP/index.htm#SOAP_Geocode_Overview.htm"

Pour ce qui concerne l'application Web de géocodage que l'on souhaite mettre en œuvre, seule la fonction GeocodeAddresses nous intéresse.

⁸ Le WSDL (Web Services Description Language) décrit une Interface publique d'accès à un Service Web, notamment dans le cadre d'architectures de type SOA (*Service Oriented Architecture*). C'est une description fondée sur le XML qui indique « comment communiquer pour utiliser le service » ;
Source : http://fr.wikipedia.org/wiki/Web_Services_Description_Language

5.2 Le choix du langage de développement :

BMO dispose de serveurs équipés d'**Internet Information Services**, communément appelé IIS. C'est le logiciel de serveur services Web (ou FTP, SMTP, HTTP etc.) de la plateforme Windows NT. Il permet entre autre de publier sur le Web et d'exécuter des applications Web développées en ASP.NET.

BMO disposant également du logiciel Visual Studio 2010.

"Visual Studio est un ensemble complet d'outils de développement permettant de générer des applications Web ASP.NET, des Services Web XML, des applications bureautiques et des applications mobiles. Visual Basic, Visual C++, Visual C# et Visual J# utilisent tous le même environnement de développement intégré (IDE, Integrated Development Environment), qui leur permet de partager des outils et facilite la création de solutions faisant appel à plusieurs langages. Par ailleurs, ces langages permettent de mieux tirer parti des fonctionnalités du Framework .NET, qui fournit un accès à des technologies clés simplifiant le développement d'applications Web ASP et de Services Web XML grâce à Visual Web Developer."

Source : http://fr.wikipedia.org/wiki/Visual_Studio

Ce logiciel permet également d'intégrer facilement les fonctions et les objets utilisés par un service Web au travers de son fichier de description WSDL.

Enfin, la personne qui va suivre l'application Web et réaliser sa maintenance, développe généralement en C#.NET. C'est donc ce langage retenu pour le développement de l'application Web de géocodage.

5.3 L'application Web géocodage :

5.3.1 Description générale :

Dans son fonctionnement, l'application Web prend en entrée une liste d'adresses au format CSV séparateur tab ou point-virgule. En retour, l'utilisateur a la possibilité de télécharger la liste géocodée au format CSV séparateur point-virgule ou au format KMZ.

Visuellement, l'application se présente en 3 pages principales et 2 pages d'aides. Une page d'accueil qui présente rapidement l'application, la page de saisie de la liste d'adresse et la page de résultats. De plus, les pages de saisie et de résultats sont accompagnées chacune d'une page d'aide détaillant les étapes à suivre.

5.3.2 Ses caractéristiques :

Caractéristique	Valeur
Nombre maximal d'adresses en entrée	20 000
Poids maximal de la liste d'adresses en entrée	1 000 000 de caractères
Vitesse moyenne de géocodage	450 000 adresses/heure
Format en entrée	CSV séparateur tab ou point-virgule
Formats en sortie	- CSV (;) avec coordonnées en Lambert93 - Format GoogleEarth : KMZ

Les limites du nombre maximal d'adresses en entrée et du poids maximal en nombre de caractères de la liste ont été définies par rapport aux limites du serveur et des navigateurs Web à envoyer, recevoir et traiter le volume d'informations (testées sur internet explorer 6 et Firefox 4).

5.3.3 Fonctionnement :

Il est conseillé, pour mieux appréhender le fonctionnement de l'application de lire ce tableau en ayant en vis-à-vis l'ANNEXE 4 : "Vues de l'application Web".

Côté client	Côté serveur
<u>Étape1 :</u> L'utilisateur copie/colle sa liste d'adresses au format CSV dans la zone de texte prévue à cet effet et clique sur "Validation de la saisie"	
En Java Script, la liste d'adresses est récupérée depuis la zone de texte. On vérifie qu'elle n'est pas trop lourde. Si c'est le cas on la tronque. Puis on la réinscrit dans la zone de texte.	
	On récupère la liste depuis la zone de texte. On vérifie si la liste est correcte. Si oui, on identifie les champs que l'on stocke dans un tableau et on identifie chaque adresse, que l'on stocke également dans un tableau.
	On renvoie vers le navigateur du client le nom des champs pour enrichir les listes déroulantes de l'étape 3.
<u>Étape2 :</u> L'utilisateur sélectionne le type de géocodeur qu'il va utiliser (en fonction des identifiants de communes qu'il a dans sa liste d'adresses)	
<u>Étape3 :</u> Il renseigne au travers des listes déroulantes les champs qui contiennent les informations nécessaires au géocodage.	
<u>Étape4 :</u> Il clique sur le bouton géocodage pour lancer le géocodage	
	On boucle sur les tableaux de champs et d'adresses précédemment remplies, pour mettre en forme les adresses comme ArcGIS SERVER souhaite les recevoir.
	Envoi vers ArcGIS SERVER et récupération du résultat.
	Redirection des résultats vers la page de résultats.
	Restructuration des adresses en liste au format CSV séparateur point-virgule, enrichie avec les champs du géocodage et notamment X et Y en Lambert 93.
	Envoi vers la zone de texte de la page résultat
Affichage de la page résultat	
Téléchargement au format CSV	
	Envoi du contenu de la zone de texte dans un fichier au format CSV
Téléchargement au format KMZ	
	Génération du fichier KML, compression en KMZ et envoi du fichier au format KMZ

5.4 Le déploiement de l'application :

Le déploiement de cette application sur le Web soulève les points suivants :

-La légitimité de BMO, en tant que collectivité locale, à proposer gratuitement un service de géocodage alors que certaines entreprises privées le proposent aussi, mais payant.

-La part de responsabilité de BMO, si un usage illégal est fait des résultats obtenus par ce service (Ex : géocodage de toutes les personnes de confession musulmane sur le territoire de BMO par un groupuscule extrémiste)

-Une liste d'adresses géocodées est considérée par la CNIL (Commission nationale de l'informatique et des libertés) comme de l'information à caractère personnel, il est donc nécessaire de faire une demande d'autorisation auprès de cette dernière.

Cependant, en attendant que tous ces "problèmes" soient résolus par les personnes et organismes compétents, l'application sera publiée sur l'Intranet de BMO, pour permettre aux personnes n'ayant pas ArcGIS sur leur poste de pouvoir utiliser l'outil de géocodage.

5.5 Conclusion :

Cette application répond au second objectif de ce stage : rendre utilisable l'outil de géocodage pour les personnes internes à BMO non-équipées d'ArcGIS 9.3.1 ou 10. En effet, elles peuvent consulter leurs résultats dans GoogleEarth grâce au KMZ ou dans TatukGIS VIEWER grâce au CSV.

Cette application ne permet pas de faire du ré-appariement interactif. Néanmoins, les services de type "Geocode Service" d'ArcGIS SERVER possèdent les fonctions nécessaires pour en réaliser. Cela pourrait-il être une future amélioration de cette application Web.

Si cette application vient à être ouverte sur Internet, une fois encore, grâce au CSV le résultat d'un géocodage pourra être intégré dans de nombreux logiciels SIG (Ex. QGIS, ArcGIS...)

Il nous reste à expliquer comment mettre à jour ce système de géocodage.

6 La mise à jour du système de géocodage :

6.1 Automatisation grâce à Python :

Le système de géocodage de BMO est composé de 8 référentiels, 44 géocodeurs et 4 géocodeurs composites pour ArcCatalog et autant pour ArcGIS SERVER. Nous l'avons vu précédemment, les géocodeurs lors de leur construction incluent leurs référentiels, ils ne sont donc plus mis à jour. De plus, ces référentiels doivent être standardisés et placés dans la même géodatabase que les tables de noms secondaires.

Mettre à jour manuellement l'ensemble du système de géocodage prendrait des heures pour une personne qui connaîtrait parfaitement la procédure et qui ne ferait aucune erreur.

Il est donc impensable de mettre à jour ce système manuellement.

La solution ArcGIS propose d'exécuter des tâches et d'utiliser différents outils au travers de scripts Python. La mise à jour du système de géocodage se fait donc via un script Python automatisant toutes les tâches à effectuer.

6.2 L'outil de mise à jour :

Cet outil se compose de 6 scripts :

- config_localisateur.py
- config_referentiel.py
- outils.py
- maj.py (Mise à jour complète)
 - maj_geocodeurs.py (Mise à jour uniquement des géocodeurs)
 - maj_referentiel.py (Mise à jour uniquement des référentiels)

Le code complet et les scripts se trouvent dans : Annexes_numeriques\Python.

6.2.1 Les fichiers de configurations :

config_localisateur.py :

Il contient toutes les caractéristiques propres à chaque géocodeur : les référentiels qu'il utilise, le style de localisateur, son emplacement, son nom, son appariement de champs... Chacune de ces informations est contenue dans une fonction qui lui est propre. Ces fonctions, lorsqu'elles sont appelées renvoient l'information demandée.

Ex. Cette fonction renvoie le style de localisateur à utiliser par le géocodeur à l'adresse utilisant le référentiel de BMO (Adre_p) et le code INSEE comme identifiant de commune.

Code : Python

```
#Style de localisateur
def Adresse_insee_bmo_locator_style():
    return "Adresses fr avec code INSEE (ou postal) et nom secondaire"
```

config_referentiel.py :

Il contient toutes les caractéristiques et informations nécessaires à chaque référentiel lors du processus de mise à jour des référentiels du système de géocodage : l'emplacement d'origine du référentiel, le nom du champ contenant le code INSEE, la structure du référentiel... Il contient également les chemins de différents éléments du système de géocodage, comme par exemple celui de la géodatabase referentiels.gdb ou celui des tables de noms secondaires.

Il est structuré de la même manière que config_localisateur.py, c'est-à-dire avec des fonctions qui renvoient les informations demandées.

6.2.2 Outils.py :

Ce script comme son nom l'indique contient des outils. Ce sont des fonctions qui regroupent plusieurs actions permettant la mise à jour des éléments du système de géocodage. Les trois principales fonctions de ce script sont :

Code : Python

```
def standardiseur ([12 paramètres]):
    """Fonction permettant la standardisation des référentiels, la
    jointure avec la table nom_insee_cp la conservation des champs utiles au
    géocodage et le transfert dans la géodatabase servant aux géocodeurs"""
    [...]

def maj_localisateur ([7 paramètres]):
    """Fonction servant à la mise à jour des géocodeurs simples"""
    [...]

def maj_localisateur_composite ([8 paramètres]):
    """Fonction servant à la mise à jour des géocodeurs composites"""
    [...]
```

Ces trois fonctions prennent en entrée de nombreux paramètres (masqués ici), chacun de ces paramètres correspond à une information contenu dans le fichier config_localisateur.py ou config_referentiel.py

6.2.3 Maj.py :

Ce script est le "moteur" de la mise à jour. C'est lui qui va aller chercher les informations contenues dans les fichiers de configurations (config_localisateur.py et config_referentiel.py) et qui va les utiliser avec les outils contenus dans outils.py.

Il crée également un rapport de mise à jour dans lequel est indiqué pour chaque étape son bon déroulement. Dans le cas contraire, est indiqué le message d'erreur du problème rencontré. (Exemple de rapport de mise à jour ANNEXE 5 : "Exemples rapport de mise à jour").

En fin d'exécution, il renvoie une valeur : 0 (aucune erreur rencontrée) 1 une ou plusieurs erreur(s) rencontrée(s)

maj_geocodeurs.py et maj_referentiel.py

Ce sont des copies partielles de maj.py sans la partie référentiels pour le premier et sans la partie géocodeurs pour le second. Ces deux scripts permettent d'effectuer une mise à jour d'une partie du système de géocodage, référentiels ou géocodeurs, pour gagner du temps lorsque l'on souhaite faire une mise à jour partielle.

6.3 Exécution de l'outil de mise à jour :

Les trois scripts maj.py, maj_localisateurs.py, maj_referentiel.py peuvent s'exécuter directement en double-cliquant dessus ou bien depuis des fichiers de commandes (*.bat). Ces trois fichiers de commandes regroupés dans un même dossier " outils administrateurs " facilite le lancement des trois scripts Python pour un utilisateur. (cf. Annexes_numeriques\outils_administrateur).

Conditions pour l'exécution des scripts :

Pour pouvoir exécuter un des scripts précités, le poste doit être équipé d'ArcGIS v10 ou à défaut v9.3.1 en sachant que sur cette dernière, les géocodeurs composites ne seront pas mis à jour ainsi que le système de géocodage pour ArcGIS SERVER.

Temps d'exécution moyen des scripts de mise à jour en fonction de la version d'ArcGIS :

Script	V10	V9.3.1
maj.py	16 minutes	13 minutes
maj_referentiel.py	5 minutes	5 minutes
maj_localisateurs.py	11 minutes	16 minutes

On note ici que la mise à jour complète sur la version 10 est un peu plus longue que sur la v9.3.1, en effet, cette dernière ne met pas à jour les géocodeurs composites, ni le système de géocodage pour le serveur.

On observe également qu'il y a une incohérence entre le temps de la mise à jour complète et le temps de la mise à jour des géocodeurs seuls dans la version 9.3.1. Je n'ai trouvé aucune explication à ce phénomène.(?)

6.3.1 Procédure automatique :

Une exécution automatique du script de mise à jour complète est effectuée tous les dimanches dans la nuit, sur un poste équipé de la version 10 d'ArcGIS. Elle s'effectue par le lancement d'un fichier de commande (*.bat), ce fichier lance l'exécution du script maj.py, récupère la valeur en fin d'exécution et envoie un mail à la personne référente du service SIG pour le géocodage. Ce mail, renvoie un lien vers le rapport de mise à jour et indique le bon ou mauvais déroulement de la mise à jour.

6.4 Conclusion :

Une mise à jour manuelle serait une tâche trop difficile à réaliser, peu fiable et très longue à mettre en œuvre.

Grâce au langage Python, il est possible de mettre à jour automatiquement, de façon fiable et rapide l'ensemble de ce système de géocodage, garantissant ainsi aux utilisateurs un système à jour et en état de marche.

Conclusion :

Un système de géocodage de A à Z...

Cette petite phrase pourrait être le plus court résumé de mon stage.

En effet, lorsque je suis arrivé, il n'existait rien sur quoi me baser, excepté l'outil développé en v3 mais je n'ai rien pu en tirer pour l'appliquer en v9.3.1 ou 10. J'ai donc fait "naître" ce projet à partir de zéro, en collaboration avec mon maître de stage Laurent Dupont et le reste de l'équipe SIG, pour aboutir au système de géocodage complet, performant et accessible à tous, présenté dans ce rapport.

Cette absence de base a été à la fois un avantage et une difficulté. Un avantage dans le sens où je n'étais pas influencé par l'existant, ni contraint de respecter des normes ou habitudes. Une difficulté car effectivement il fallait tout trouver à partir de zéro.

Là, où j'ai rencontré le plus de difficulté, c'est à la création des styles de localisateurs, car je n'ai trouvé que peu de documentation sur le sujet, hormis le Geocoding Rule Base Developer Guide. Néanmoins, cette difficulté a pu être surmontée.

Outre la conduite d'un projet du début à la fin, ce stage m'a permis d'apprendre de nouveaux langages de programmation : Python et C#.net. J'ai pu également me mettre à la place de mes différents professeurs, devenant moi-même formateur durant quelques heures auprès des agents de BMO pour les initier à la pratique de ce nouvel outil de géocodage.

Enfin l'ensemble de ce stage, et plus particulièrement la partie importante de programmation occasionnée par ce projet de système de géocodage, a conforté mon choix de suivre cette licence professionnelle SIG.

Cette expérience enrichissante renforce mon choix de me diriger le plus possible vers des postes de technicien SIG ayant une composante programmation.

Bibliographie/Webographie :

Bibliographie :

ESRI (2003) : "*ArcGIS 9 – Geocoding Rule base Developer Guide*"

Webographie :

ArcGIS :

<http://Webhelp.esri.com/arcgisDEsktop/9.3/index.cfm?TopicName=welcome>

ArcGIS SERVER :

http://tasks.arcgisonline.com/ArcGIS/SDK/SOAP/index.htm#SOAP_Geocode_Overview.htm

Python :

<http://www.siteduzero.com/tutoriel-3-223267-apprendre-python.html>

<http://docs.python.org>

<http://python.developpez.com>

Divers :

<http://www.forumsig.org>