

POTENTIALITES DU GEOWEB
(Internet sémantique géographique)

Robert Laurini, Sylvie Servigne
LIRIS – CNRS 5205
INSA de Lyon
Université de Lyon

Depuis quelques années arrivent de nouveaux concepts et outils sous le vocable générique de web sémantique¹ géographique, appelé plus communément GeoWeb. Le web sémantique se veut l'Internet du futur en proposant aux utilisateurs des moteurs de recherches beaucoup plus sophistiqués que ceux existant actuellement (de type Google ou Yahoo). Parallèlement, suite aux travaux de Ben Shneidermann (1997), l'approche dominante est basée sur son mantra « *Overview, zoom and fliter, details on demand* » ; ce qui veut dire que dans un premier temps, un panorama général est visualisé, puis le champ de recherche est graduellement restreint pour aboutir enfin aux détails demandés. Cette approche progressive s'est ainsi imposée dans des outils comme Bing ou Google Earth.

Devant ce contexte, l'utilisateur dispose de nouveaux outils de recherche et de visualisation qui remettent en question les approches traditionnelles. L'objet de cet article sera d'analyser ces nouvelles tendances et de montrer en quoi certaines pratiques doivent être renouvelées.

1 – Nouveaux concepts et outils

Sous le vocable de GeoWeb (Fu-Sun, 2010), de nouveaux concepts et outils ont fait leur apparition depuis quelques années. Analysons-les rapidement.

1.1 – Geographic Information Retrieval

Nous sommes en présence d'un changement de paradigme : le problème typique n'est plus de rechercher ce qu'il y a en un point dans une base de données géographiques, mais sur Internet tout entier. Ce problème peut s'étendre facilement à une zone, ou une recherche de chemin optimal multimodal. La résolution de ces problèmes nécessite le geocoding, le geonaming, le geoparsing, des répertoires de toponymes, le geotagging et l'indexation spatiale.

Sous le terme de géocoding, il s'agit de conférer les coordonnées (au sens longitude/latitude) à un lieu en partant de son nom, par exemple la Tour Eiffel, ou tout Paris.

Le geonaming est l'opération inverse, c'est-à-dire de donner un nom à un lieu : si l'on clique sur un point, ce peut-être un bâtiment, la commune où il se trouve, le canton, le département etc.

Le rôle du geoparsing est tout autre. Il s'agit d'analyser un texte (par exemple une page web) pour en extraire les références de lieu. Dans un premier, devant un mot, il faut savoir s'il s'agit réellement d'un lieu, et si oui lequel. L'exemple classique est « Washington ». S'agit-il de George Washington, la capitale des USA (Washington D.C.), ou l'état de Washington, ou bien l'une des multiples villes appelées ainsi aux Etats-Unis, ou bien d'autres personnes de ce nom ?

Un autre exemple de geoparsing est un texte décrivant les prévisions météorologiques du genre « Au nord d'une ligne Bordeaux-Strasbourg, il pleuvra. Sur la façade est, il neigera et sur les régions allant de l'Aquitaine à Nice, il fera beau, sauf une poche sur le Pays Basque avec des nuages ». A partir d'un tel texte, il s'agira de produire une carte.

¹ Voir http://fr.wikipedia.org/wiki/Web_s%C3%A9mantique

Dans ce but, un élément important est l'existence d'un répertoire de toponymes (gazetteer² en anglais) lié à une ontologie géographique. De fait, il s'agit d'une base de données de termes géographiques avec leur catégorie (villes, fleuves, etc.), leurs variantes d'écriture (Saint Etienne et St Etienne), les noms dans différentes langues (Venise, Venezia, Venice, Venedig, etc.), les anciens noms (du type Byzance, Constantinople, Istanbul) sans oublier la forme géométrique (le Paris du XIXème siècle est plus petit que celui d'aujourd'hui). C'est grâce à un tel outil que l'on pourra distinguer le Rhône comme fleuve ou comme département.

Un autre exemple difficile est par exemple l'analyse d'un portail d'une agence de voyage : distinguer les offres de différents lieux de séjour ; la localisation de l'agence est délicate du point de vue sémantique.

Par geotagging on dénomme les annotations géographiques. Sous ce vocable, existent plusieurs catégories d'annotations. La plus simple est l'annotation localisée : à un endroit précis, une annotation est mise, par exemple une explication sur un lieu touristique. Il existe aussi des annotations géographiques d'images et de vidéos qui sont plus compliquées techniquement. L'annotation peut être réalisée sous forme de texte, d'images, de sons, de vidéos, etc.

Dès lors, chaque document se voit attribuer une empreinte spatiale, soit généralement un polygone, soit le contour exact ou approximatif, par exemple pour un document sur la Corse, ou bien encore le rectangle englobant minimum. Répondre à une requête, revient à déterminer sa pertinence thématique, selon les procédés classiques des moteurs de recherche textuelle et sa pertinence spatiale généralement liée à l'intersection entre l'empreinte du document et la zone de requête (voir Figure 1a).

Ainsi, à partir d'une question du type « Je veux les châteaux à Zurich », il s'agira :

- grâce à l'indexation classique de type textuel (mots-clés) retrouver les châteaux dans différentes langues,
- grâce à l'indexation spatiale, délimiter les documents à Zurich, éventuellement complétés par le voisinage,
- déterminer pour chaque document leur pertinence à la fois thématique et spatiale,
- et finalement choisir les documents les plus pertinents sémantiquement et géographiquement parlant (voir Figure 1b).

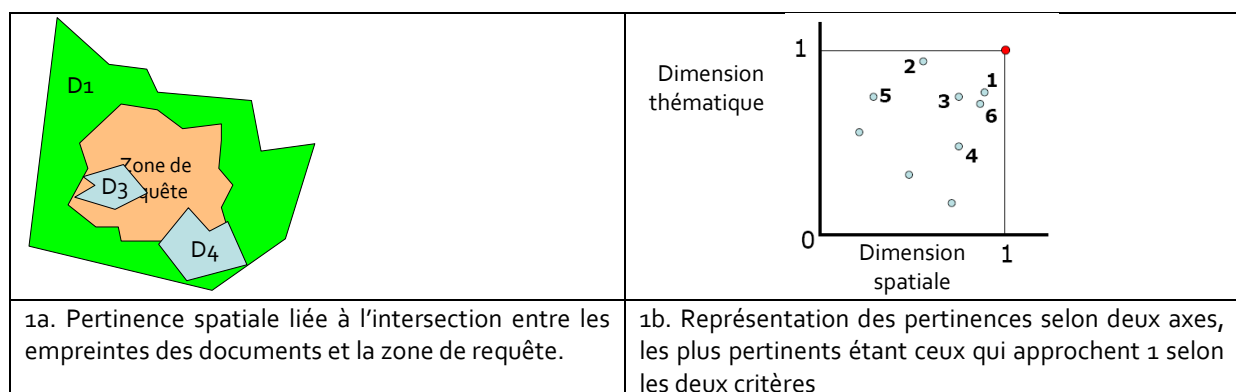


Figure 1. Classification des documents géographique selon deux critères, thématique et géographique.

Le langage SpatialML³, développé par la Mitre Corporation⁴ est un langage de balises pour annoter les lieux et les relations spatiales. Dès lors, à côté du texte original, une version annotée avec ce langage est générée, ce qui facilitera la recherche (Figure 2).

² Voir <http://en.wikipedia.org/wiki/Gazetteer>

³ Voir <http://sourceforge.net/projects/spatialml/>

⁴ Voir <http://www.mitre.org/>

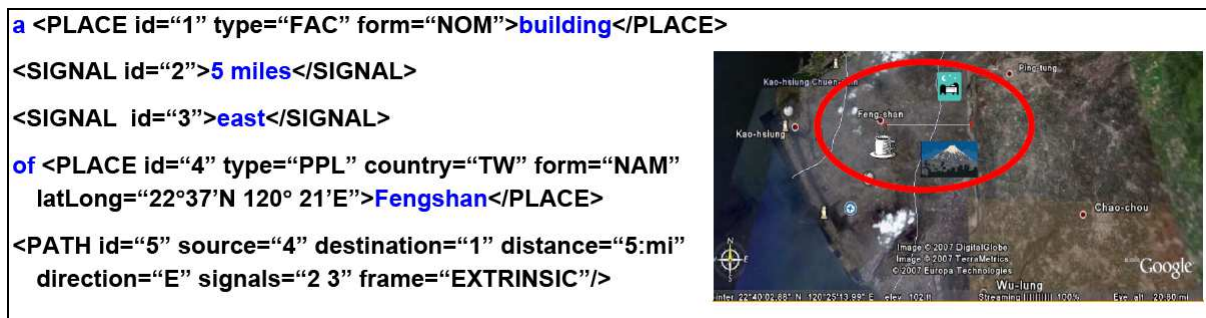


Figure 2. Exemple de codage SpatialML

1.2 – Villes à trois dimensions

La représentation des villes à trois dimensions est un sujet à part entière aujourd’hui. A partir de photos aériennes multiples à haute résolution (pixel de l’ordre de 5 à 15 cm au sol), il est possible de générer un modèle d’un bâtiment.

L’initiative allemande CityGML⁵ s’est donné pour objectif de représenter les villes à trois dimensions. Afin de permettre différents niveaux de modélisation, plusieurs niveaux de détails (LOD = Levels of Details) ont été définis et ceci pour deux raisons totalement différentes. La première dépend de la résolution avec laquelle les données ont été prises et stockées tandis que la seconde permet de dégrader la représentation des objets urbains et des bâtiments les plus éloignés, car cette dégradation permet d’envoyer moins d’information sur le réseau et ainsi d’obtenir des temps d’affichage plus rapides. Ces niveaux sont les suivants (Figure 3) :

- LOD 0 (Modèle régional) : il s’agit d’un modèle numérique de terrain (2,5D) permettant de monter l’ensemble d’un paysage et dont les données proviennent essentiellement de laser aéroporté ;
- LOD 1 (Modèle urbain) : « modèle bloc » dans lequel les bâtiments sont schématisés sous forme de blocs sans structures de toit, donnant ainsi une idée de la répartition de la hauteur des bâtiments ; les données de base proviennent du cadastre et de laser aéroporté ; dans ce cas sont intégrées des données provenant de photogrammétrie verticale ;
- LOD 2 (Modèle urbain) : même modèle que précédemment auquel sont ajoutées des textures pour les façades et les toitures (photogrammétrie oblique) ;
- LOD 3 (Modèle urbain) : même modèle que précédemment mais avec un niveau plus détaillé du point de vue architectural ;
- LOD 4 (Modèle intérieur) : modèle véritablement architectural « parcourable », c’est-à-dire avec un modèle de l’intérieur des bâtiments.

La Figure 4 donne un aperçu de l’usage concomitant de plusieurs niveaux de détails dans la même image.

⁵ Voir <http://www.citygml.org/>

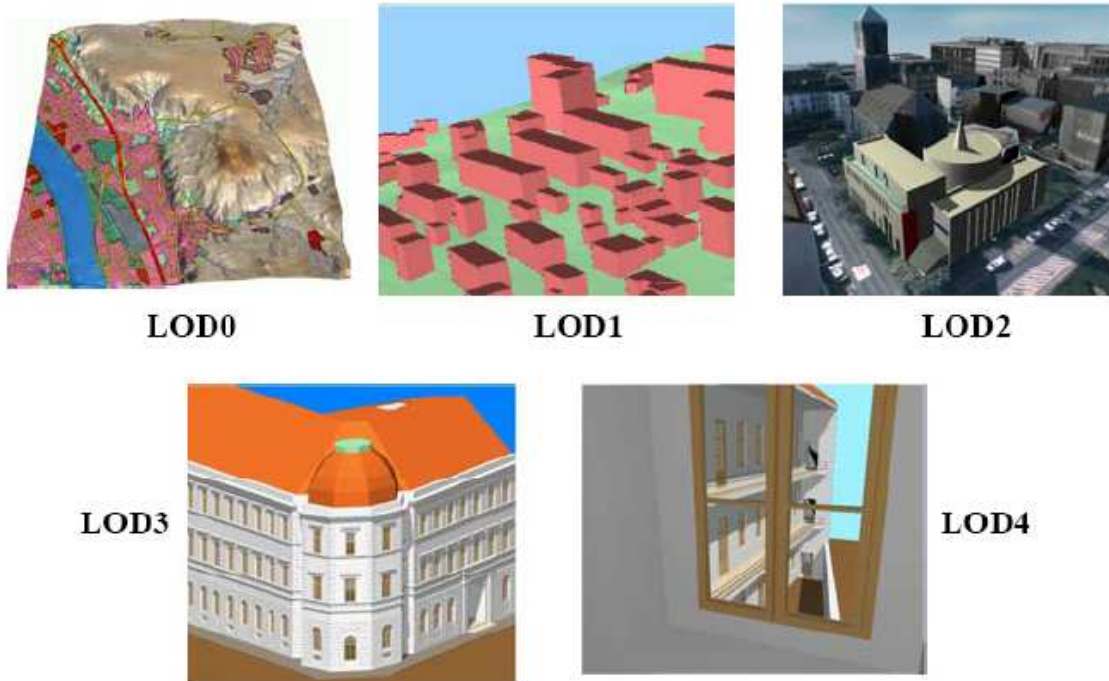


Figure 3. Exemple de niveaux de détails (LOD) proposés par la norme CityGML (Source <http://www.ikg.uni-bonn.de/sig3d>)



Figure 4. Intégration de différents niveaux de détails selon la distance. Source http://www.citygml.org/docs/Kolbe_CityGML_OGC_NYC_CAD+GIS_WG.pdf

1.3 – Les geobrowsers

Si les moteurs de recherche classique sont basés sur du texte, les geobrowsers se présentent comme des moteurs de recherche ou des navigateurs basés sur la géographie : à partir d’une carte, l’utilisateur peut retrouver ce qu’il recherche, par exemple un hôtel.

A l’heure actuelle, les dominants sont Google Earth (de la société Google) et BING (initialement appelé Virtual World de la société Microsoft).

En 2004, la société Google a acheté la société Keyhole afin de lancer le produit Google Earth qui a été téléchargé des millions de fois. Les utilisateurs de Google Earth ont apporté leur contribution (images, modèles 3D, ...) au logiciel. Rappelons que l'objectif de la société Google est d'organiser les savoirs du monde entier et de les rendre disponibles et accessibles universellement. Devant cet objectif, la carte du monde de Google est la base sur laquelle seront disposés tous les renseignements. En d'autres termes, cette carte est un chemin d'accès à tout type d'information localisée.

Cependant, de son côté la société Microsoft a acheté la société Vexcel en mai 2006 qui possédait un grand savoir-faire en photogrammétrie tridimensionnelle, notamment grâce à sa caméra UltraCam-X⁶ qui peut prendre des images de 216 megapixels au rythme de 3 gigabits par seconde avec des précisions de l'ordre de 4 cm par pixels. C'est ainsi qu'est né le projet Virtual Earth de Microsoft qui couvre à l'heure actuelle une quinzaine de villes américaines (San Francisco, San José, Seattle, Boston, Philadelphia, Los Angeles, Las Vegas, Detroit, Phoenix, Houston, Baltimore, Atlanta, Denver, Dallas-Fort Worth et New York), alors que plusieurs centaines sont en cours de saisie. On pourra constater (Figure 5) le caractère saisissant du réalisme ainsi obtenu en animation.

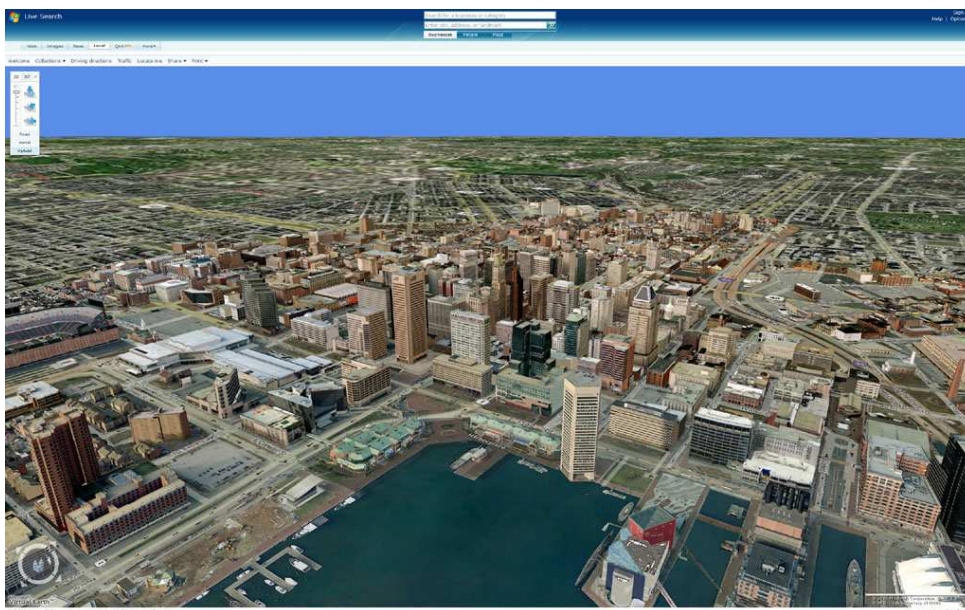


Figure 5. Exemple de modèle de ville virtuelle (Baltimore) avec Virtual Earth (Source : <http://www.3pointd.com/index.php?tag=mapping&paged=2>)

Pour une municipalité, disposer d'un modèle générique de l'ensemble de sa ville s'avère un atout indispensable pour des applications potentielles comme :

- la simulation des niveaux de bruit à trois dimensions, c'est-à-dire non seulement à la hauteur des oreilles d'un piéton, mais aussi aux divers étages ; cette modélisation peut être étendue à d'autres pollutions comme celle de l'air ;
- la modélisation des inondations et des risques naturels ; dans ce cas, il s'agit de simuler la hauteur du niveau d'eau partout dans la ville y compris les constructions souterraines du type parking souterrain, métro, tunnel, etc.
- la modélisation des autres risques environnementaux comme les tremblements de terre, et la visualisation de leurs conséquences,
- la comparaison des hauteurs réelles des bâtiments et les hauteurs maximales autorisées par la réglementation ;
- l'impact visuel d'un futur lotissement ou d'un grand projet d'équipement dans son environnement ;

⁶ Voir <http://www.microsoft.com/ultracam/en-us/UltraCamXp.aspx>

- la vérification rapide des déclarations d'impôts locaux ;
- l'aide à la définition de la stratégie d'approche pour la sécurité civile ou la lutte contre les incendies dans les villes ;
- etc.

Mais également d'autres acteurs peuvent bénéficier de telles représentations. Sans être exhaustifs, signalons quelques pistes :

- en géomarketing, disposer d'une maquette permettant de tester l'impact visuel d'un panneau publicitaire,
- pour les agents immobiliers, permettre aux acquéreurs potentiels d'avoir une idée claire sur les bâtiments aux alentours du bien qu'ils désirent acheter ou louer,
- pour des touristes, avoir une vue plus précise des monuments et des endroits à visiter,
- pour les opérateurs de téléphonie mobile, trouver les lieux idéaux pour installer les antennes (carte d'intervisibilité),
- pour les installateurs de panneaux solaires, repérer les toitures les plus adaptées,
- pour les pilotes d'hélicoptères, repérer les toitures où il est possible de se poser,
- pour des historiens et des archéologues, visualiser ou mieux modéliser une ville aux siècles/millénaires précédents,
- etc.

Le projet StreetView⁷ de Google est différent : son objectif est de naviguer, non plus comme un oiseau, mais comme un piéton dans les rues. La plupart des villes d'Amérique du Nord et d'Europe sont couvertes. Sous le nom de géo-panoramas StreetView, les applications possibles⁸ sont les suivantes :

- Opération d'analyse pour l'industrie pétrolière / gazière,
- Analyse des installations sur sites sensibles,
- Préparation de chantiers et travaux publics,
- Immobilier, tourisme, etc.

1.4 – Les Mash-ups

Le terme « mash-up » est originellement une expression désignant un mélange de musiques. Il a été repris en informatique pour exprimer le fait de mélanger des données provenant de sources différentes, par exemple de plusieurs bases de données.

La société Google a popularisé le langage KML⁹, depuis normalisé, afin de faciliter la création de cartes sur lesquelles l'utilisateur peut adjoindre ses propres données. En d'autres termes, Google fournit des fonds de cartes du monde entier sur lequel on peut ajouter des objets géographiques ponctuels, linéaires ou surfaciques et des annotations. Ci-dessous on trouvera un exemple de programme KML permettant d'obtenir la Figure 6.

```
<?xml version="1.0" encoding="UTF-8"?>
<kml xmlns="http://earth.google.com/kml/2.0">
<Placemark>
  <name>INSA Lyon, Dept IF</name>
  <description><![CDATA[<a
    href="http://liris.insa-lyon.fr/robert.laurini/cours4IF.htm"
    target=_blank>"Cours 4IF"</a><br>
    Pr. R. Laurini<br>
    ]]></description>
  <Point id="khPoint600">
```

⁷ Voir <http://maps.google.com/intl/fr/help/maps/streetview/>

⁸ Voir <http://www.globalvision.ch/fr/360/panorama-geo-reference.html>

⁹ Voir <http://code.google.com/intl/fr/apis/kml/documentation/>

```
<coordinates>4.8720471, 45.782474</coordinates>
</Point>
</Placemark>
</kml>
```

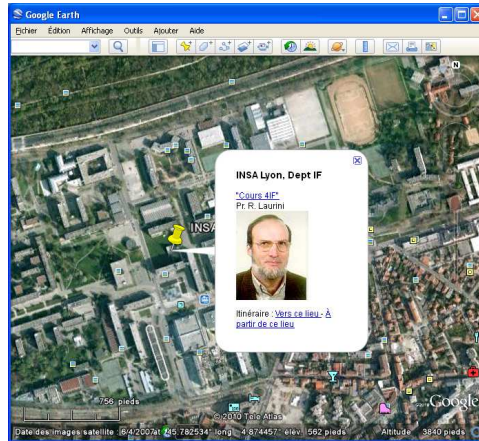


Figure 6. Exemple de résultat du programme KML précédent.

De plus, il existe une possibilité de créer des mash-ups sans connaître la programmation KML, simplement par une série de clics. Dès lors, une fois terminée, sont générées au choix, une URL que l'on pourra insérer dans un mail, ou bien une séquence HTML qui pourra être insérée dans une page web ; c'est cette deuxième solution qui est généralement adoptée par un nombre immense de d'entreprises, d'universités, d'hôtels, de restaurant comme base pour leur plan d'accès.

Parmi les mash-ups¹⁰, on trouvera des applications comme des circuits de dégustation de vin, les routes jugées dangereuses, des hôtels bon marché, le inondations, la pollution, des lieux mentionnés dans la bible, etc.

1.5 – Les SIG basés sur le volontariat

Par la technique des annotations, tout le monde peut insérer des compléments d'information. Par exemple, le projet d'OpenStreetMap¹¹ permet à chacun d'ajouter des informations localisées. C'est donc une nouvelle vision dans l'esprit du Web 2.0 où non seulement des experts donnent à lire, mais où chacun peut contribuer à sa guise. Bien évidemment se posent des problèmes de contrôle de qualité, d'éthique et de confidentialité.

1.6 – Les capteurs

Une autre direction est celles des capteurs en temps réel connectés à Internet, capteurs mesurant par exemple la température, la qualité de l'air, le niveau de bruit, la vitesse et la direction du vent. Dans le projet CitySense¹², de même chacun peut raccorder son propre capteur, et ainsi contribuer à donner des visions différentes de celles fournies par des experts. Des problèmes de contrôle de qualité se posent également.

1.7 – Les services localisés

L'informatique ne se base plus sur les machines de bureau. Au-delà des portables, de nouveaux appareils se font jour comme les téléphones intelligents, les tablettes, regroupés sous le vocable anglais de PDA (Personal Digital Assistant) qui non seulement sont connectés à Internet, mais aussi possèdent des dispositifs de localisation de type GPS. Dès lors tout un chacun peut poser la question prototypique « où se trouve le restaurant le plus proche ? »

¹⁰ Voir par exemple <http://mashable.com/2009/01/08/google-maps-mashups-tools>.

¹¹ Voir <http://www.openstreetmap.org/>

¹² Voir <http://www.citysense.net/>

Mais l'utilisateur laisse une trace de son passage. Des recherches actuelles utilisent ces traces (notamment à partir de l'analyse de fichier de logs) par exemple pour déterminer les embouteillages à partir des vitesses des véhicules, où connaître les déplacements de masse.

Devant ces potentialités, se pose la grave question des libertés individuelles et de préservation de la vie privée.

2 – Conséquences sur la géographie

Un des points forts du Web Sémantique est que chacun peut contribuer. Alors que dans le passé, des « clercs » possédant la connaissance rédigeaient le contenu des pages, les « ouailles » peuvent aujourd'hui contribuer. Si cette démocratisation est une excellente évolution, différentes questions se font jour comme le contrôle de qualité des contributions et la nécessité d'une éthique. Dans cet article, nous n'aborderons rapidement que quelques questions sur la géographie parfois appelée néo-géographie.

2.1 – Sur la définition de la géographie

En France vers la fin des années 1970, lorsque les GIS sont apparus, les discussions étaient les suivantes : nous hésitions sur la traduction de « geographic » en nous disant que l'on était loin de la géographie traditionnelle. Des traductions comme « géolocalisées », « géospaciales » ont vu le jour. Mais devant la déferlante ESRI, la paresse aidant, la traduction « géographique » bien que mal adaptée s'est imposée. Lorsque l'on parle de tuyaux d'eau, ou de gestion des rues, a-t-on vraiment affaire à la géographie ? La question reste posée, mais revient plus fortement actuellement : on ne parle pas de services géographiques, mais de services localisés !

Les géographes classiques ont été « secoués » par l'apparition des systèmes d'information géographiques (SIG), non pas tellement par la cartographie automatisée, mais par la constitution de bases de données. La carte support millénaire de l'information géographique devenait un simple état de visualisation des informations contenues dans la base de données.

Maintenant, avec l'apparition des SIG basés sur le volontariat, le statut social des géographes nécessiterait d'être réexaminé.

2.2 – Sur la cartographie

Avec l'apparition des mash-ups, les fonds de cartes sont gratuits et de bonne qualité, et l'on peut même s'appuyer sur les photos aériennes pour visualiser les objets géographiques, donnant plus d'importance à la représentation des phénomènes socio-économiques et aux nouvelles sémiologies graphiques. A titre de conséquence, certains cours de cartographie pourraient être repensés.

A titre d'exemple, un office du tourisme d'une petite commune peut en quelques heures refaire l'intégralité de son site web.

3 – Conclusions

L'avènement de l'Internet sémantique géographique remet en cause un certain nombre d'habitudes et de pratiques voire même de paradigmes.

- les GIS (Geographic Information Systems) se transforment en GIR (Geographic Information Retrieval); l'objectif devient la recherche géographique dans Internet tout entier sous réserve d'affecter à chaque document son empreinte géographique ;
- les navigateurs géographiques permettent d'accéder aux informations en partant des cartes du monde entier ;
- des sociétés proposent des fonds de cartes gratuits, à la fois vecteur et photos aériennes ; alors que les collectivités locales ont investi des sommes énormes pour la constitution de ces fonds de plans, se pose la question cruciale du positionnement des services cartographiques vis-à-vis de cette offre ;

- l'écriture de mash-ups permet de produire des cartes très facilement et de les coupler à d'autres sources de données géographiques ; dès lors les bases de données géographiques doivent être conçues dans cet esprit ;
- les SIG basés sur le volontariat permettent à tout un chacun de pouvoir contribuer ;
- la prolifération des capteurs connectés à Internet permet d'avoir d'autres visions pour suivre en temps réel de nombreux phénomènes environnementaux ;
- les modèles de villes à trois dimensions basés sur la photogrammétrie deviennent courants ;
- les services localisés se popularisent.

De plus, cet avènement soulève nombre de questions sur les usages et au niveau sociétal notamment :

- celle du copyright de telles cartes obtenues à partir de mash-up,
- celle de la pérennité de ces solutions, car en cas de la disparition d'une société comme Google, les cartes disparaîtront complètement,
- celle de la migration d'un geobrowser à un autre laquelle semble actuellement impossible,
- celle de la confidentialité et peut-être d'un nouveau visage de Big Brother qui proviendrait non plus d'un état totalitaire, mais de sociétés privées.

Bibliographie

Fu P., Sun J. (2010) « Web GIS: Principles and Applications ». EuroSpan, December 2010, 312 p.

Kolbe T., Bacharach S. (2006) « CityGML: An Open Standard for 3D City Models ». <http://www.directionsmag.com/articles/citygml-an-open-standard-for-3d-city-models/123103>

Shneiderman B. (1997) "Designing the User Interface". Third edition. Addison-Wesley, Publishing Company. 1997. pp. 600.