

INTÉGRATION SÉMANTIQUE ET CARTOGRAPHIQUE DES SERVICES LOCALISÉS MULTI-FOURNISSEURS

par Roula Karam, Robert Laurini, Frank Favetta

LIRIS CNRS UMR 5205, INSA de Lyon, Université de Lyon, F-69621 Villeurbanne, France
Naval Academy Research Institute, Lanveoc-Poulmic, BP 600, 29240, Brest Naval, France
Roula.karam@insa-lyon.fr ; Roula.karam@ecole-navale.fr ; Franck.favetta@insa-lyon.fr ;
Robert.laurini@insa-lyon.fr

Rima Kilany

ESIB, Faculté d'Ingénierie, Université Saint Joseph, B.P:11-514, Beirut, Lebanon
Rima.kilany@usj.edu.lb

L'utilisation de systèmes portables étant aujourd'hui très répandue, de nombreuses applications comportant des services localisés ont été conçues pour délivrer des informations utiles aux utilisateurs et ce, partout, à tout moment, en se basant sur leurs profils et leurs positions géographiques. De plus, avec le nombre croissant de données et d'applications géographiques distribuées, issues de bases de données hétérogènes, de nombreux problèmes peuvent se présenter concernant : 1) l'appariement des bases de données géographiques, 2) l'intégration des données/métadonnées géographiques de plusieurs fournisseurs, et 3) le développement de portails conviviaux unifiés sur des systèmes portables de type PDA. Alors que de nombreux standards tel qu'OGC et des applications Web tels que Google Maps et Bing démontrent la faisabilité de portails issus d'un seul fournisseur, l'objectif de notre travail est d'améliorer l'approche qui génère des portails visuels autorisant plusieurs fournisseurs à commercialiser leurs services superposés sur une carte unifiée. En outre, plusieurs algorithmes de fusion, des raisonnements par intelligence artificielle, dont le concept d'ontologie, seront utilisés et modifiés avec d'autres propositions. Tout ceci sera testé par des prototypes afin de résoudre la problématique d'intégration des objets homologues et d'éviter leur duplication sur écran. Ces objets « multi-fournisseurs », se référant au même point d'intérêt mais ayant des données spatiales/non spatiales légèrement différentes, seront appariés par leur noms de lieu, leurs positions géographiques, leurs détails sémantiques et leurs symboles cartographiques sur un seul fond de carte.

1 Introduction

Les services géolocalisés (LBS) sont destinés à délivrer de l'information adéquate aux utilisateurs quels que soient le temps et l'endroit, et ceci en se basant sur leur profil, contexte et position géographique (Dulizia, 2007). À travers n'importe quelle application dans ce domaine, par exemple *m*-tourisme, les utilisateurs souhaitent toujours recevoir une réponse rapide et précise en se déplaçant. Cependant, la qualité de service proposée par les fournisseurs cartographiques actuels (i.e. Google Maps, Bing, Yahoo Maps, Mappy ou Via Michelin) dépend de leurs données géographiques. En général, ces données sont stockées dans plusieurs bases de données géographiques (BDG) au niveau mondial. D'autre part, le nombre croissant des différentes BDG couvrant la même zone géographique et la récupération des données/métadonnées non erronées pour un service quelconque impliquent de nombreux raisonnements et des contrôles d'accès aux BDG afin de résoudre les ambiguïtés dues à la présence des objets homologues dupliqués sur l'écran mobile. Un exemple type nous a inspiré pour diriger notre

recherche vers un problème plus important nécessitant l'intégration cartographique pour le même service géolocalisé provenant de plusieurs fournisseurs. Considérons une requête utilisateur pour trouver les restaurants les plus proches. Cet utilisateur pourra récupérer la réponse d'un restaurant américain, visualisé par deux fournisseurs différents, donc deux icônes différentes placées à 50 mètres l'une de l'autre et avec des noms différents. La première icône porte le nom « KFC » et la seconde porte le nom « Kentucky Fried Chicken ». De même, les fonds de cartes sont différents, provenant de fournisseurs différents, et les détails sémantiques diffèrent légèrement entre les deux icônes pour le même restaurant (fig.1).

Une interprétation classique de ce phénomène est que les fournisseurs de services devaient se servir jusqu'à présent d'une seule base de données géographique comme les pages jaunes, les pages blanches, le guide des restaurants dans le monde, etc. Le refus d'accéder à plusieurs BDG ajoute de l'incohérence et de la rigidité aux résultats

fournis pour les applications géolocalisées. De plus, certaines entités d'un point d'intérêt peuvent être modifiées au fil du temps (changement de nom de lieu ou du numéro de téléphone) et ne sont pas mises à jour systématiquement dans la BDG concernée.

Une solution envisagée propose de récupérer l'information depuis plusieurs BDG. Mais on doit savoir comment résoudre les incohérences surtout entre des objets homologues, candidats à l'intégration. Mon travail consiste à permettre cette intégration cartographique pour les applications m-tourisme et ceci en récupérant les informations spatiales/non-spatiales (noms des lieux, positions géographiques, catégorie du service, détails sémantiques et symboles cartographiques) de plusieurs fournisseurs. Cependant, ceci peut conduire à visualiser des objets dupliqués pour le même point d'intérêt et imposer des difficultés au niveau de la gestion des données. En outre, l'utilisateur sera dérouté par la présence de résultats multiples pour un même point. Donc, mon but ultime sera de générer automatiquement une carte unique intégrant plusieurs interfaces des fournisseurs, sur laquelle les objets homologues seront intégrés avant de les visualiser sur l'écran mobile.

Dans cet article, nous présentons MPLoM (Multi-Providers LBS on Mobile), un prototype que nous avons développé pour tester l'intégration des données/métadonnées au niveau de leur positions géographiques, noms des lieux, catégorie et détails sémantiques. Au niveau cartographique, et pour générer une carte unifiée, nous montrons la possibilité de construire un nouveau type d'ontologie dont les concepts sont décrits sémantiquement et visuellement (icône, texture, couleur, etc.). Pour atteindre cette approche, on propose une extension du standard Web Ontology Language, nommé CartOWL.

De plus, nous avons aussi étudié les avantages de l'utilisation des services Web géographiques au lieu d'un accès direct non sécurisé aux BDG. De même, on a proposé de créer un prototype pour simuler une orchestration intelligente des Web Services nécessaires entre OGC et W3C afin de répondre à n'importe quelle requête utilisateur, tout en contactant plusieurs fournisseurs. Nos nouveaux concepts, basés sur certains algorithmes de fusion, sur l'ontologie pour assurer l'intégration au niveau sémantique et cartographique, sur l'orchestration des géo-Web Services, sont implémentés dans des prototypes modulaires et évalués.

Ce papier est divisé en plusieurs parties : après

l'introduction au contexte, l'état de l'art est ensuite présenté, puis notre contribution pour chaque concept tout en tenant compte de la conception et l'implémentation du prototype concerné ; finalement notre travail se conclut sur quelques perspectives.

2 État de l'art

En dépit d'initiatives fructueuses et d'une utilisation répandue de normes, les solutions d'aujourd'hui ne concernent pas l'intégration de services issus de plusieurs fournisseurs tels que le portail visuel présenté précédemment en figure 1, ni leurs problématiques de visualisation intégrée. L'intégration est le procédé utilisé pour réaliser la correspondance entre des objets géographiques issus de différentes bases de données, modélisant le même phénomène dans la réalité.

La connexion à plusieurs bases de données spatiales des fournisseurs LBS, hétérogènes dans leurs conceptions et leurs données, c'est-à-dire l'interopérabilité, n'est pas une tâche aisée (Gordillo, 2008).

L'interopérabilité des données géographiques représente une solution au problème de partage et d'intégration de ces données. Puisque notre travail consiste à étudier l'interopérabilité de plusieurs fournisseurs de services localisés sur appareil mobile, on va donc, dans cet article, se contenter de citer les technologies et standards décrits dans la littérature, qui assurent la base de cette interopérabilité, avant de proposer, dans notre contribution, une nouvelle avancée dans l'état de l'art.

L'Open Geospatial Consortium (OGC) est une organisation pour la standardisation des Systèmes de l'information géographiques (SIG), qui a pour objectif d'améliorer l'interopérabilité entre les applications en créant des langages communs d'échange par des standards communs ; on peut citer : 1) des fonctions de cartographie utilisant les services Web Mapping Service (WMS), 2) une recherche de données géo-localisées utilisant les services Web Features Service (WFS), 3) les services de catalogue des métadonnées avec Catalog Service et surtout le langage d'interopérabilité, basé sur XML, nommé Geography Mark Up Language (GML). En plus des standards de l'OGC incluant Open Location Services (OpenLS), on ne peut pas négliger le rôle des ontologies et des Web Services pour réussir l'interopérabilité des LBS provenant de plusieurs fournisseurs avec d'autres algorithmes de fusion aux niveaux lexical, géométrique et sémantique.

2.1 Les standardisations OGC

2.1.1 Geography Markup Language (GML) et Compact GML

GML¹ est un codage de XML utilisé pour les informations géographiques. Ce langage permet le stockage, l'échange, et la modélisation d'informations géographiques contenant des attributs spatiaux et non spatiaux. Étant basé sur XML, les langages de requêtes disponibles pour XML peuvent également être employés pour GML. Idéalement, ils devraient permettre le traitement des données spatiales et temporelles qui rendent GML puissant.

Les coordonnées dans GML représentent les coordonnées des objets de la géométrie (ex. points, lignes) et peuvent être indiquées par l'un des éléments suivants de GML :

```
<gml:coordinates>
<gml:pos>
<gml:posList>
```

ou de cette manière :

```
<gml:Point gml:id="p21"
srsName="urn:ogc:def:crs:EPSG:6.6:4326">
<gml:coordinates>45.67,
88.56</gml:coordinates>
</gml:Point>
```

Puisque GML est un langage textuel, il fournit également des moyens efficaces pour archiver des données géo-spatiales, assurant ainsi sa compatibilité avec de futurs logiciels. L'utilisation de GML sur les téléphones portables présente deux inconvénients majeurs. Tout d'abord, il consomme beaucoup de mémoire et de bande passante lors de la sauvegarde et du transfert des données. De plus, les cartes décrites doivent être projetées et leurs échelles doivent être recalculées avant d'être affichées sur l'écran du mobile. Pour cela, dans notre cas, on utilisera Compact GML (De-Vita, 2003) (cGML). C'est une spécification de XML qui vise à fournir les cartes géographiques digitales aux appareils sans fils. cGML a été testé sur les appareils J2ME (ex. Nokia 7650).

Dans cGML, les balises provenant de GML sont devenues plus courtes. Le résultat final peut toujours être lu par les humains, mais sa taille est considérablement réduite. L'occupation dans la mémoire est diminuée de l'ordre de 60%. cGML n'est pas conçu pour être mis à l'intérieur d'un fichier XML. Il est indépendant et contient un entête qui contient toutes les

informations sur l'échelle choisie, l'appareil utilisé ainsi que d'autres informations utiles (fig.2).

2.1.2 Web Features/Mapping Services

Les services WMS² servent à afficher des collections de couches sous forme de cartes de type image et les services WFS³ pour afficher des données sous forme d'entités vectorielles SVG⁴ en collectant les informations des fournisseurs sous format GML.

2.1.3 Styled Layer Description/ Symbology Encoding (SLD/SE)

Défini par l'OGC, le Styled Layer Descriptor⁵ (SLD) est un format de description en XML, permettant la mise en forme de données géographiques provenant d'un flux WMS. Pour simplifier, il joue le même rôle qu'un fichier CSS pour une page HTML, le but étant de séparer complètement le style de la donnée. Concrètement, lors de la réception d'un flux WMS, un style défini lui est rattaché alors que ce fichier n'est pas physiquement lié au moteur cartographique. En effet, il est tout à fait possible d'interroger un serveur cartographique distant, de recevoir le flux WMS et de lui appliquer un style que vous aurez vous-même défini auparavant.

Le fonctionnement en lui-même n'a rien de bien compliqué, il suffit simplement de bien comprendre les différentes interactions ci-dessous :

- Un client effectue une requête WMS vers un serveur cartographique en spécifiant un SLD.
- Le moteur cartographique analyse la validité de la requête et, le cas échéant, envoie le flux correspondant.
- Le client récupère ce flux et peut ensuite l'afficher.

Le Symbology Encoding⁶ (SE) est construit sur SLD pour définir de la même façon la symbologie par défaut de la carte à visualiser (image, font, etc.), et n'implique donc aucune créativité de la part du « map-designer » tout en tenant compte des préférences des utilisateurs (Domingues, 2009).

2.1.4 Les ontologies (géographiques)

Plusieurs domaines ont recours aux ontologies comme solutions pour présenter, partager et enrichir les connaissances, à savoir le Web sémantique, le

1 <http://www.opengeospatial.org/standards/gml>

2 <http://www.opengeospatial.org/standards/wms>

3 <http://www.opengeospatial.org/standards/wfs>

4 www.w3.org/Graphics/SVG/

5 <http://www.opengeospatial.org/standards/wms>

6 <http://www.opengeospatial.org/standards/se>

commerce électronique, l'intelligence artificielle et les systèmes d'information géographiques (Fonseca, 2002 ; Cullot, 2003 ; Baglioni, 2007 ; Casati, 1998). Cependant, les concepts et les relations sont identifiés textuellement alors que certains concepts ont un aspect visuel surtout dans le domaine cartographique. Par exemple, le symbole d'un point d'intérêt dans une légende (fig.3) est identifié comme un concept par une icône et /ou abréviation, une couleur, une texture, une police de caractère, une orientation ou un certain numéro. Notre approche concerne le développement et l'intégration de ces ontologies visuelles, représentant des légendes, vers une ontologie référence du domaine touristique (fond de carte et symboles provenant de plusieurs fournisseurs).

OWL⁷ (Web Ontology Language) est conçu comme une extension de Resource Description Framework⁸ (RDF) et RDF Schema⁹ (RDFS). Il fournit les moyens pour définir des ontologies Web structurées. Il permet de décrire des ontologies, c'est-à-dire de définir des terminologies pour décrire des domaines concrets. Les composants de ce langage sont : les individus (instances de classes), les propriétés (relations entre les individus) et les classes. Protégé¹⁰ est un système auteur pour la création d'ontologies. Quant aux ontologies géographiques, leur construction et leur alignement se fera de la même manière que pour les autres ontologies avec Protégé ou G-Match, etc. Mais selon la littérature disponible, ces logiciels traitent seulement l'alignement des aspects lexicaux, sémantiques et topologiques sans tenir compte de l'aspect cartographique ou visuel.

2.1.5 La technologie des Web Services

Les touristes ont besoin de s'enquérir de la météo, des sites et des directions qui les entourent sans avoir recours à plusieurs sources. De plus, ils ne savent pas toujours comment rechercher l'information qu'ils veulent. Pour la plupart des services demandés, leurs positions géographiques servent de paramètres pour la recherche. L'utilisation des Web Services¹¹ permet de créer des applications sans avoir besoin de garder les outils SIG du côté client, et on peut même assembler plusieurs Web Services pour créer une solution complète. Les Web Services peuvent communiquer avec n'importe quelle application Web et mobiles (Alameh, 2003).

Étant standardisés avec W3C, plusieurs fournis-

seurs préfèrent développer et commercialiser leurs propres Web Services pour des raisons de sécurité en limitant l'accès direct à leurs bases de données. OGC a de même développé ses propres services Web géographiques en respectant les bases de W3C pour permettre une certaine compatibilité et bien sûr envisager l'interopérabilité. Ces géo-Web Services comme WMS, WFS, WCS, WPS¹², etc. sont utilisés par de nombreux logiciels cartographiques à savoir ESRI, MapInfo, GeoConcept, etc. (Schaeffer, 2008).

2.2 Intégration des BDG au niveau applicatif

2.2.1 Identification des données géographiques homologues

L'appariement est le processus consistant à établir les correspondances entre les objets géographiques des différentes bases de données, représentant le même phénomène du monde réel.

Dans le cadre des bases de données classiques, pour identifier ces objets, les valeurs des identifiants ou attributs clés sont utilisées. En revanche, les informations géographiques sont recueillies dans une logique d'autarcie. Les identifiants définis sont, pour la plupart, particuliers à chaque service. Par conséquent, ils peuvent rarement être utilisés pour identifier des données homologues (Beerli, 2004). Il faut donc intégrer les données géographiques à l'aide d'autres techniques nommées techniques d'appariement.

On distingue trois types d'outils d'appariement :

- l'appariement géométrique qui consiste à appairier les données géographiques par leur localisation et leur forme ;

- l'appariement sémantique qui est similaire au mécanisme d'intégration de données classiques (les objets sont alors appariés par la valeur de leur sens commun) ;

- l'appariement lexical qui compare les noms des lieux, caractère par caractère, pour assurer une certaine confiance si les noms sont similaires mais non mis à jour dans les BDG.

7 www.w3.org/TR/owl-features/

8 www.w3.org/RDF/

9 www.w3.org/RDFS/

10 protege.stanford.edu

11 <http://www.w3.org/standards/Webofservices/>

12 <http://www.opengeospatial.org/standards/wps>

• Appariement géométrique

Les données à appairer possèdent des géométries de précision différentes (et donc des localisations différentes). Par conséquent, l'appariement géométrique doit prendre en compte cette imprécision. Trois types d'appariements géométriques peuvent être adoptés : 1) en définissant une zone d'appariement connue sous le nom de la bande Epsilon ou zone tampon, 2) en se dotant d'une mesure de distance entre objets, 3) en utilisant d'autres caractéristiques géométriques de la forme de l'objet. Ces trois types peuvent être utilisés isolément ou conjointement.

On a décidé de choisir le critère de la distance entre objets. Pour définir la ressemblance entre deux objets, les distances entre ces objets semblent être des mesures pertinentes. Deux objets de classe correspondante sont appariés si la distance sélectionnée est inférieure à un seuil. Pour appairer deux objets ponctuels, la distance euclidienne¹³ s'impose. Pour appairer deux objets linéaires, trois distances vont être décrites (distance moyenne, distance de Hausdorff, distance de Fréchet). Pour comparer deux objets ponctuels homologues, nous calculons la distance euclidienne entre ces deux objets et nous les comparons à un seuil. Le choix de ce seuil est très important. Un seuil trop grand entraîne la sélection d'un grand nombre d'appariements litigieux (seuil de 9 m), un seuil trop bas provoque des solutions minimales. Pour l'exemple de la figure 4, un seuil de 9 m puis de 3 m permet de retenir l'ensemble des appariements à 9 m et de résoudre les appariements litigieux par un deuxième appariement à 3 m. Ainsi, la ligne B qui est litigieuse à 9 m est appariée à 2 ou 3 m. L'appariement géométrique à partir du calcul de la composante de Hausdorff¹⁴ avec deux seuillages donnera donc un résultat plus robuste. Dans notre application, un seuil de 5 m est suffisant pour appairer à partir de deux points litigieux les hôtels et les restaurants.

• Appariement lexical

Pour les noms des lieux, il y a plusieurs algorithmes de fusion dans la littérature comme la distance de Hamming¹⁵ (si le nombre de caractères est identique) ou la distance de Levenshtein¹⁶ pour un nombre différent. On a fait l'hypothèse que le nombre des caractères est différent et on a décidé d'utiliser celle de Levenshtein.

• Appariement sémantique

Pour voir si les détails sémantiques des deux objets se réfèrent à un seul point d'intérêt, donc à une nomenclature commune (ou ontologie du domaine), il sera nécessaire d'avoir un dictionnaire commun de référence (i.e. WordNet ou ImageNet). L'intégration sémantique par ontologie sera ajustée par des algorithmes de similarité sémantique (Word by Word similarity, Normalized Google Distance, Levenshtein Distance, etc.). Un certain accord pour visualiser les détails intégrés sera finalisé par code ainsi que par la nomenclature de l'ontologie du domaine.

3 Contribution

Chaque fournisseur de service possède sa propre carte et les icônes géographiques, différentes des autres fournisseurs. Dès lors, lorsqu'un utilisateur cherche un même emplacement, le résultat représenté par une icône diffère d'un fournisseur à un autre selon des critères de représentation. On a donc intérêt à intégrer ces différentes icônes désignant le même emplacement dans une carte globale. Une simple interopérabilité n'est pas suffisante car elle assure l'intégration au niveau lexical mais néglige l'aspect sémantique et les icônes d'une carte ont un aspect visuel et non pas textuel. Pour cette raison, on a recours aux ontologies qui offrent une représentation des informations et des connaissances d'une manière plus riche que l'intégration par schéma des bases des données.

Dans ce travail, on cherche à intégrer les différentes icônes des cartes désignant un même emplacement afin de construire une carte unifiée tout en la personnalisant suivant les caractéristiques et le profil de l'utilisateur. Cependant la construction de cette nouvelle carte globale présente de nombreux défis, surtout pour gérer les hétérogénéités de la représentation des données. Le même restaurant pourra être visualisé par des noms de lieux, des positions géographiques et des détails sémantiques légèrement différents dus à la présence de différentes données de différentes BDG couvrant la même zone géographique.

Donc la récupération des données/métadonnées non erronées pour un service quelconque, implique de nombreux raisonnements et des contrôles d'accès aux BDG afin de résoudre les ambiguïtés engendrées par la présence des objets homologues dupli-

13 en.wikipedia.org/wiki/Euclidean_distance

14 en.wikipedia.org/wiki/Hausdorff_distance

15 en.wikipedia.org/wiki/Hamming_distance

16 en.wikipedia.org/wiki/Levenshtein_distance

qués sur l'écran mobile. Ma contribution consiste alors à permettre cette intégration cartographique pour les applications de type *m*-tourisme.

3.1 Processus d'intégration

3.1.1 Intégration géographique

Dans l'implémentation de la plateforme que nous développons (MPLoM), pour définir si deux objets sont les mêmes, nous choisissons un seuil de cinq mètres entre les candidats et utilisons la distance euclidienne pour l'intégration de deux objets ponctuels. Ainsi, tant que la distance entre les deux objets est inférieure au seuil de cinq mètres, un certain degré de confiance est assigné mais une hypothèse supplémentaire sera requise pour conclure s'ils sont homologues ou non.

3.1.2 Intégration des noms de lieux

La technique de fusion utilise la distance de Levenshtein dL pour comparer les noms de lieux (en tant que chaînes de caractères) de deux objets issus de deux fournisseurs différents. Cette distance augmente si le nombre de différences entre les caractères augmente également.

3.1.3 Intégration sémantique

Le troisième type d'intégration de deux objets concerne leurs données/métadonnées ou détails sémantiques. Afin d'éviter la duplication des informations pour les services issus de deux fournisseurs différents, une méthode de correspondance a été utilisée dans l'application et une approche sémantique par ontologie avec Protégé et l'API Jena. Par exemple, si l'utilisateur d'un navigateur souhaite connaître les restaurants proposant des « Hamburgers », la plateforme doit lister tous les restaurants de type américain ou de restauration rapide fast food, etc. De nombreuses solutions ont été proposées dans le code du prototype et dans l'alignement vers une ontologie référence dans Protégé. Finalement, la théorie des croyances a été appliquée pour confirmer la plus forte pondération entre deux objets homologues en se basant sur leurs pondérations séparées de positionnement géographique, de noms de lieux et de détails sémantiques. Le résultat final sera basé sur toutes les pondérations aux niveaux géométrique, lexical (noms des lieux) et sémantique par la théorie de Dempster-Shafer¹⁷.

3.1.4 Intégration de symboles cartographiques

Après notre discussion concernant l'intégration des points objets, une intégration de symboles cartographiques par une approche de correspondance d'ontologies spatiales est proposée dans cette partie. Étant

donné que l'utilisation de plusieurs fournisseurs implique l'utilisation de plusieurs cartes, une par fournisseur, notre proposition est de visualiser les symboles d'un service localisé commun et superposé à une carte conviviale, prenant en considération certaines règles de sélection, les préférences de l'utilisateur et les règles de sémiologie graphique. Un portail visuel bien intégré sur appareil portable est notre objectif (Cruz, 2004).

Pour atteindre ce but, nous proposons de construire une ontologie de domaine touristique qui apparie les ontologies locales issues des fournisseurs de services (symboles des services et fond de la carte utilisé). En se basant sur cette ontologie de domaine, d'autres critères de sélection sont à prendre en compte : 1) Les préférences de l'utilisateur en terme d'icônes et de cartes, 2) les limitations du matériel, 3) la zone géographique à visualiser (zone de couverture), 4) les contraintes de sémiologie graphique.

Pour réaliser notre objectif, les étapes sont les suivantes :

1) récupération de palettes d'icônes cartographiques à partir des légendes de chaque fournisseur ;

2) sauvegarde de l'image de chaque carte pour une visualisation 2D/3D ;

3) réalisation d'un test psycho-cognitif à travers une application Web pour tester la meilleure connaissance des icônes sans légende en prenant en compte l'influence des cultures potentiellement différentes des utilisateurs ; les icônes les mieux connues seront retenues en priorité ;

4) création d'une bibliothèque commune d'icônes conviviaux pour tous les services localisés utilisées par notre application et leurs attributs (couleur, taille, nombre, forme, abréviation, texture, police, etc.) ; ces icônes peuvent être utilisés comme une solution par défaut pour résoudre tout autre conflit ;

5) développement d'une application «Building /Matching prototype» pour construire l'ontologie locale de symboles cartographiques de chaque fournisseur ; définition des relations entre les concepts des différentes ontologies des différents fournisseurs afin d'obtenir une ontologie de domaine applicatif ; cette application a la capacité d'exporter et d'importer une ontologie issue d'un nouveau fournisseur dans un langage OWL étendu nommé CartOWL (Cartographic Web Ontology Language) gérant les classes, sous-classes, relations et attributs des symboles cartographiques ;

17 en.wikipedia.org/wiki/Dempster-Shafer_theory

6) l'ontologie de domaine peut être vérifiée dans l'application pour générer et visualiser les symboles cartographiques correspondants pour tout service intégré et superposé sur une carte unique appropriée ;

7) les règles de sémiologie graphique (Chesneau, 2005) et (Christophe, 2009) doivent être respectées pour le choix des couleurs des icônes par rapport à celles de la carte ainsi que d'autres techniques d'adaptation/généralisation afin d'obtenir un portail visuel clair et compréhensible. Il faut conserver les couleurs des logos de marques, utiliser les icônes propres aux fournisseurs pour les objets hétérogènes, un macro-icône pour les objets homologues avec agrégation, une bibliothèque d'icônes en cas de conflit, etc. ;

8) rendre le système mieux apte à accepter plusieurs fournisseurs et donc développer dans Protégé une application qui permette d'extraire automatiquement des icônes de différentes légendes, construire à partir de ces icônes les ontologies cartographiques de leurs fournisseurs, puis aligner ces ontologies vers une ontologie référence du domaine touristique contenant les icônes les plus connues en fonction des préférences des utilisateurs (tests psycho-cognitifs) et des règles de sémiologie (ontologie et intelligence artificielle).

3.2 Architecture du prototype MPLoM

La plateforme MPLoM est développée pour tester la faisabilité de l'intégration et la localisation des symboles cartographiques dans un portail visuel unique sur un appareil portable. La première phase couvre l'intégration sémantique globale à partir de deux fournisseurs offrant des services de type pull (services de recherche d'hôtels et de restaurants) et de type push (services météo envoyés). La deuxième phase couvre l'intégration cartographique, notamment avec des solutions supplémentaires concernant les normes d'applications Web et de services Web localisés pour objectif d'interopérabilité.

Les services pull (restaurant et hôtels) implémentés sont visualisés sur une carte 2D Google Map télé-chargée par API et les composants sont superposés en tant que marqueurs Google (« R » pour restaurants et « H » pour hôtels). L'interface utilisateur est affichée sur un émulateur Nokia S60 équipé d'un middleware ou base de données intermédiaire de services localisés. Le langage utilisé pour la plateforme est Java. XQuery sert à interpréter des fichiers cGML afin d'intégrer les informations concernant les objets dans un fichier cGML unifié avant de visualiser les résultats sur l'écran. Une base de données

médiatrice, maintenue par l'administrateur, est utilisée pour stocker, d'une part les préférences de l'utilisateur, et d'autre part le fichier de sortie au format cGML qui peut contenir toutes les informations des deux fournisseurs relatives aux objets répondant à la requête. La conception de MPLoM se décompose en trois phases : la pré-intégration, l'identification des correspondances et l'intégration.

Notons que l'architecture employée est l'architecture médiatrice ; ainsi, chaque fournisseur possède une base de données spécifique et un adaptateur transformera ses données en un fichier CGML. Le schéma local consiste en un fichier CGML que nous avons nommé Nomenclature commune qui transformera les données des fournisseurs en un format unique que nous exploiterons dans nos requêtes. Après avoir effectué l'intégration complète, il faut construire une réponse unique sous un format standard. Donc nous devons générer un fichier cGML unique pour le rendre au client et l'afficher sur la carte. Le format du fichier final de la requête «hôtel» est donné en figure 2. Les données vont être envoyées au client dans ce format standard. Du côté client, nous allons récupérer tous les détails à l'aide d'un parser, et nous allons tout afficher sous forme de liste et sur la carte géographique (Karam1, 2010) et (Karam2, 2010).

3.3 Fonctionnalité et implémentation de l'application m-tourism

Deux scénarios décrivant les services de recherche d'hôtels et restaurants et les services météo ont été implémentés en deux phases. L'utilisateur peut démarrer l'application en décrivant ses préférences et en se connectant par une interface utilisateur textuelle. Les figures suivantes (fig.5, 6, 7, 8, 9) illustrent ces implémentations.

3.4 Conception des prototypes des ontologies cartographiques

Le concept est d'intégrer plusieurs icônes représentant un même emplacement géographique (fig.3) afin de construire une carte géographique globale. Au début, on doit reconnaître la catégorie de chaque icône afin de l'identifier. Pour cela on utilise les ontologies qui répondront aux requêtes par le type de l'icône : un hôpital, une école... On a choisi l'approche mono-ontologique qui consiste à construire une seule ontologie globale et lui envoyer des requêtes pour la classification. Cette approche a été faite en deux phases : la première par le développement d'un prototype intitulé « Building and Matching prototype » (fig.10, 11). L'expert du domaine construit manuellement les ontologies des légendes des

fournisseurs en intégrant leurs concepts textuellement et avec leurs aspects visuels ; ensuite l'alignement sémantique sera fait manuellement au début avec deux ontologies puis chaque nouvelle ontologie sera ajoutée à l'ontologie dite de référence. Les aspects visuels seront transmis dans le même fichier du type OWL que l'on nomme CartOWL. La deuxième phase consiste à tester la première dans Protégé pour des raisons de scalabilité et afin d'assurer une extraction automatique des icônes, puis viennent la construction des ontologies de chaque légende et leur alignement d'une façon quasiment automatique sans intervention humaine.

Dans cette phase, on construit les classes désignant une icône et ses caractéristiques. On réduit ces caractéristiques à ceux qui nous intéressent le plus et nous sont les plus utiles : texture, font, objet (image), URL, couleur (couleur de fond), abréviation, nombre. Ensuite, pour chaque caractéristique, on liste les instances, soit les options qui peuvent se présenter dans la réalité. Ainsi, la couleur du fond d'une icône peut être du vert, du bleu, du violet... Enfin, on doit lier ces instances à leur classe à l'aide des relations. Tout ceci est réalisé dans Protégé (fig.12, 13) (Karam3, 2010 ; Karam4, 2011 ; Karam5, 2010).

3.5 Conception de l'orchestration des services Web

Finalement, l'accès par des Web Services de W3C et d'OGC nécessite une orchestration intelligente et un « processing » à ne pas négliger. WPS-T (Transactionnel) d'OGC peut jouer ce rôle comme l'illustre le diagramme de séquence (fig.14). L'implémentation de ce concept est en cours : la littérature a traité l'orchestration des services Web W3C et l'orchestration des services géo-Web OGC, mais pas encore l'orchestration des deux standardisations W3C et OGC (Karam3, 2010).

4 Conclusions et perspectives

En conclusion, notre contribution peut être identifiée en deux parties :

- L'appariement en une ontologie de domaine spatial, utilisée pour intégrer les symboles cartographiques issus de différents fournisseurs ; des extensions de standards tels que CartOWL et cGML basés sur XML incluant les attributs des symboles cartographiques avec les données spatiales et non-spatiales.

- La plateforme proposée MPLoM dans laquelle des services localisés de type pull et push peuvent être intégrés en un portail unifié ; ceci est réalisé en définissant et utilisant des ontologies et d'autres raisonnements de fusion pour assurer l'interopérabilité au niveau applicatif parmi de nombreux fournisseurs, et en tenant compte des hétérogénéités lexicales (nom des lieux), géographiques (emplacements), sémantiques (détails) et cartographiques (attributs visuels) .

Nous prévoyons à l'avenir : 1) d'améliorer la plateforme MPLoM comme décrit dans cet article pour en assurer l'extensibilité ; 2) de proposer des extensions de formats XML pour décrire les métadonnées de services localisés (attributs de symboles cartographiques, prix des services, SPAM, etc.) ; 3) de préconiser la création de nouveaux services Web localisés « composites » et plus étendus afin de gérer des fournisseurs multiples (valeurs ajoutées aux services d'OGC) ; 4) de créer des APIs ou Web Services dédiés pour assurer la recherche des informations nécessaires, dans le cas où l'on ne dispose pas d'un accès original aux bases de données géographiques des fournisseurs, afin d'automatiser l'extraction automatique des icônes vers la construction des ontologies cartographiques et la collection des requêtes/réponses d'une façon plus adéquate aux « mobinautes ».

Bibliographie

Alameh N., 2003, "Chaining Geographic Information Web Services", *IEEE Internet Computing* 7, p. 22—29.

(Baglioni, 2007) **Baglioni M., Masserotti M. V., Renso C. et al. (2007)**, "Building Geospatial Ontologies from Geographical Databases", in *Proceedings of GeoSpatial Semantics, Second International Conference, GeoS 2007*, Mexico City, Mexico, November 29-30, 2007, Springer, p. 195-209.

(Beeri, 2004) **Beeri C., Kanza Y., Safra E. et al.**, 2004, "Object fusion in geographic information systems", in *Proceedings of the Thirtieth International Conference on Very Large Data Bases - Volume 30, VLDB Endowment*, p. 816-827.

(Casati, 1998) **Casati R., Smith B., Varzi A. C.**, 1998, "Ontological Tools for Geographic Representation", in *Formal Ontology in Information Systems*, p. 77-85.

(Chesneau, 2005) **Chesneau E., Ruas A., Bondan O., 2005**, “Colour Contrasts Analysis For A Better Legibility Of Graphic Signs For Risk Maps”, in *Proc. of the Dansternational Cartographic Conference: Mappdansg Approaches Dansto A Changdansg World, A Coruna, Spadans, 9-16 July 2005*, p. 10.

(Christophe, 2009) **Christophe S., Ruas, A., 2009**, “A process to design creative legend on-demand”, in *24th ICA conference - 14-19 November 2009, Santiago de Chile, Chile*.

(Cruz, 2004) **Cruz I. F., Sunna W., Chaudhry A., 2004**, “Semi-automatic Ontology Alignment for Geospatial Data Danstegration”, in Max J. Egenhofer; Christian Freksa & Harvey J. Miller, ed., *Proceedings of Geographic Dansformation Science, Third Dansternational Conference, GIScience 2004, Adelphi, MD, USA, October 20-23, 2004, Sprdansger*, p. 51-66.

(Cullot, 2003) **Cullot N., Parent C., Spaccapietra S., Vangenot C., 2003**, « Des SIG aux ontologies géographiques », *Dansternational journal of géomatique* 13, p. 285-306.

(De-Vita, 2003) **De-Vita E., Piras A., Sanna S., 2003**, “Using Compact GML to Deploy Interactive Maps on Mobile”, dans *Proceedings of the Twelfth International World Wide Web Conference, WWW2003, Budapest, Hungary, 20-24 May 2003 - (Posters)*.

(Domingues, 2009) **Domingues C., Christophe S., Jolivet L., 2009**, « Connaissances opérationnelles pour la conception automatique de légendes de carte », dans Fabien Gandon, ed., *Actes des 20e Journées francophones d'ingénierie des connaissances, IC2009, Plate-forme AFIA / Hammamet, Tunisie, 25-29 Mai 2009*, pp. 253—264.

(Dulizia, 2007) **D’Ulizia A., Ferri F., Grifoni P., 2007**, “A semantic and structural similarity approach to personalize location-based services”, in *Proceedings of the 5th dansternational conference on Databases in networked in formation systems*, Sprdansger-Verlag, Berldans, Heidelberg, p. 33—47.

(Fonseca, 2002) **Fonseca F. T., Egenhofer M. J., Agouris P. et al., 2002**, “Usdansg Ontologies for Danstegrated Geographic Dansformation Systems”, *Transactions GIS* 6(3), p. 231—257.

(Gordillo, 2008) **Gordillo S., Laurdansi R., Mostaccio C. et al. ,2008**, “Towards multi-provider LBS visual portals”, dans *The 14th Dansternational Conf. on Distributed Multimedia Systems*, Published by Knowledge Systems Dansstitute, p. 208—213.

(Karam1, 2010) **Karam R., Kilany R., Laurini R. et al., 2010**, “Uncertain GeoInformation Representation and Reasoning: a Use Case in LBS Integration”, dans A Min Tjoa & Roland Wagner, ed., *Proceedings of Database and Expert Systems Applications, DEXA, International Workshops, 5th International Workshop FlexDBIST, Bilbao, Spain, August 30 - September 3, 2010*, IEEE Computer Society,.

(Karam2, 2010) **Karam R., Kilany R., Laurini R. et al., 2010**, “Integration of Similar Location Based Services Proposed by Several Providers”, dans Filip Zavoral; Jakub Yaghob; Pit Pichappan & Eyas El-Qawasmeh, ed., *Proceedings of 2nd International Conference NDT 2010: Network and Digital Technologies, Springer*, p. 136-144.

(Karam3, 2010) **Karam R., Kilany R., Laurini R. et al., 2010**, “Cartographic Integration on mobile devices from several providers LBS by means of map symbol ontology”, in *Proceedings of WebMGS 2010, 1st International Workshop on Pervasive Web Mapping, Geoprocessing and Services*.

(Karam4, 2011) **Karam R., Chamoun R. K., Laurini, R. et al., 2011**, “Location and Cartographic Integration for Multi-Providers Location Based Services”, dans *Advances in Cartography in GIScience vol.1, LNG&C, Selection from 25th International Cartographic Conference (ICC’11), 3-8 July, Paris (France)*.

(Karam5, 2010) **Karam R., Chamoun R. K., Laurini, R. et al., 2010**, “Integrating many legends through ontology for cartographic symbols”, in *Proceedings of 6th International Conference SAGEO2010*.

Schaeffer B., 2008, “Towards a transactional Web Processing Service (WPS-T)”, in Bishr M. & T. Bartoschek Pebesma, E., ed., *GI-Days 2008 - Proceedings of the 6th Geographic Information Days*, IfGI prints, Institut für Geoinformatik, Münster’.



Figure 1 : Exemple du même objet décrit par deux fournisseurs

```

<cGML version="1.0">
<Head>
<View zoom=""></View>
</Head>
<FtCl>
<Ft id="" name="">
<Point>
<cds> </cds>
</Point>
<Info>
<Name></Name>
<etoile></etoile>
<Region></Region>
<Tel1></Tel1>
<Tel2></Tel2>
<Tel3></Tel3>
<Capacity></Capacity>
<email></email>
<website></website>
<Address></Address>
<roomrate></roomrate>
<Singrate></Singrate>
<Doubrate></Doubrate>
<suite></suite>
<fax></fax>
</Info>
</Ft>
</FtCl>
</cGML>

```

Figure 2 : Fichier CGML



Cartographic Symbols	Ordnance Survey	Rand McNally	IGN
Museum	icon	icon	icon
Park	icon	icon color	icon
Tourist Info. Center	icon	icon	icon
Picnic Site	icon	icon	icon
Parking, Park ride	icon	icon	icon
Selected places of Tourist Interest	texture	icon/color	icon
Main Road	A 30 Abb./Number texture/color	icon	N 170 Abb./Number color

Figure 3 : Extraits des aspects visuels des légendes

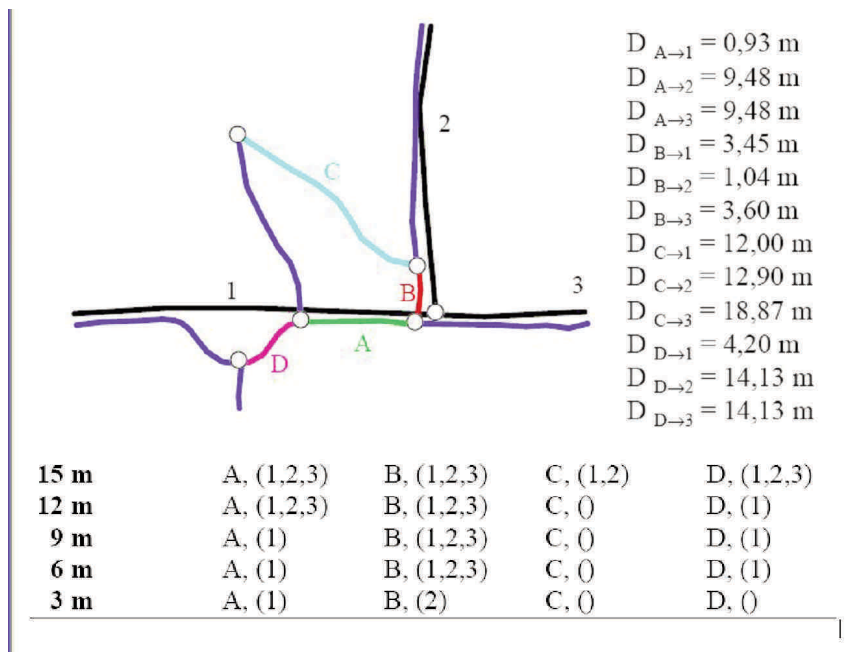


Figure 4 : Technique de Stricher et choix de seuil

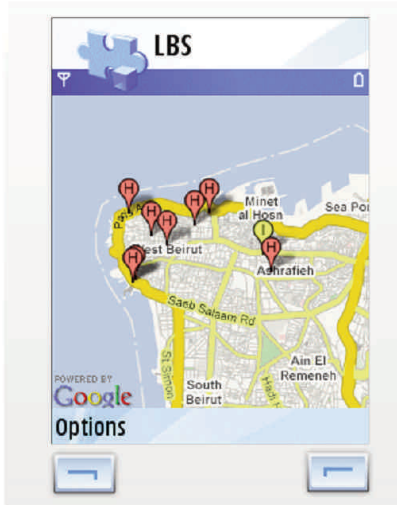


Figure 5 : Hôtels

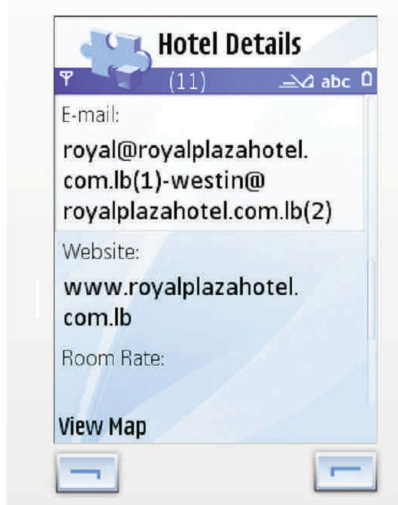


Figure 6 : Détails des hôtels

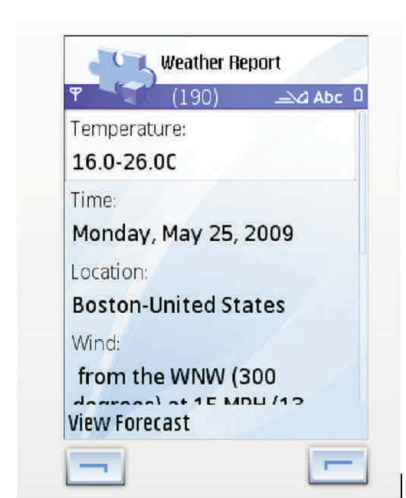


Figure 7 : Prévisions courante météo

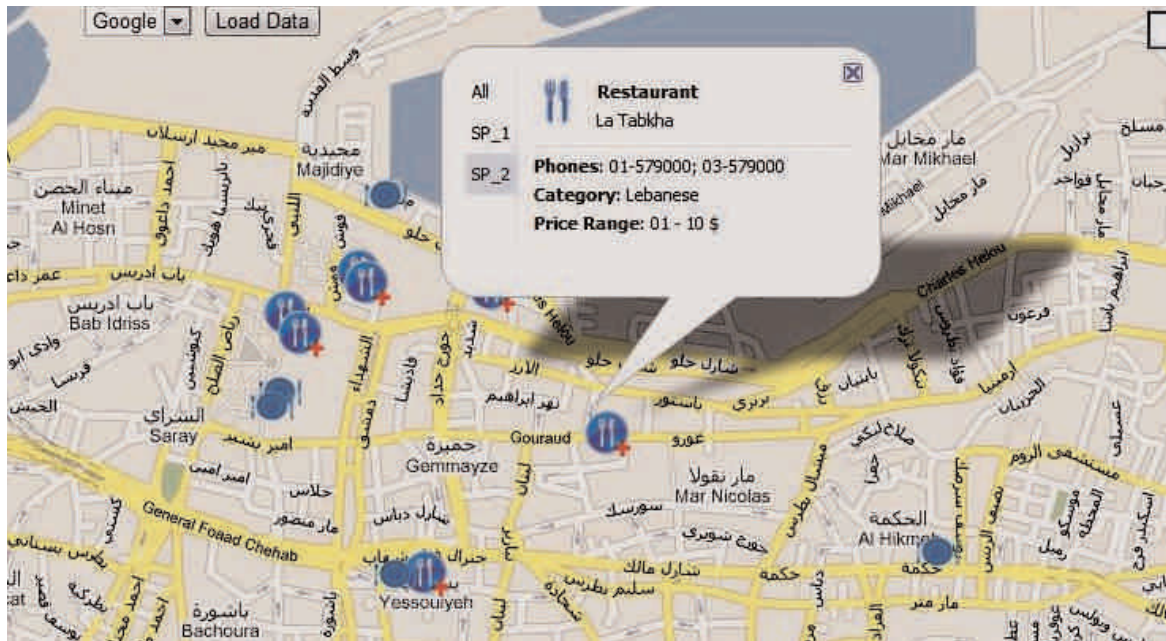


Figure 8 : Représentation d'un icône du fournisseur 2



Figure 9 : Représentation d'un icône intégrant les deux fournisseurs

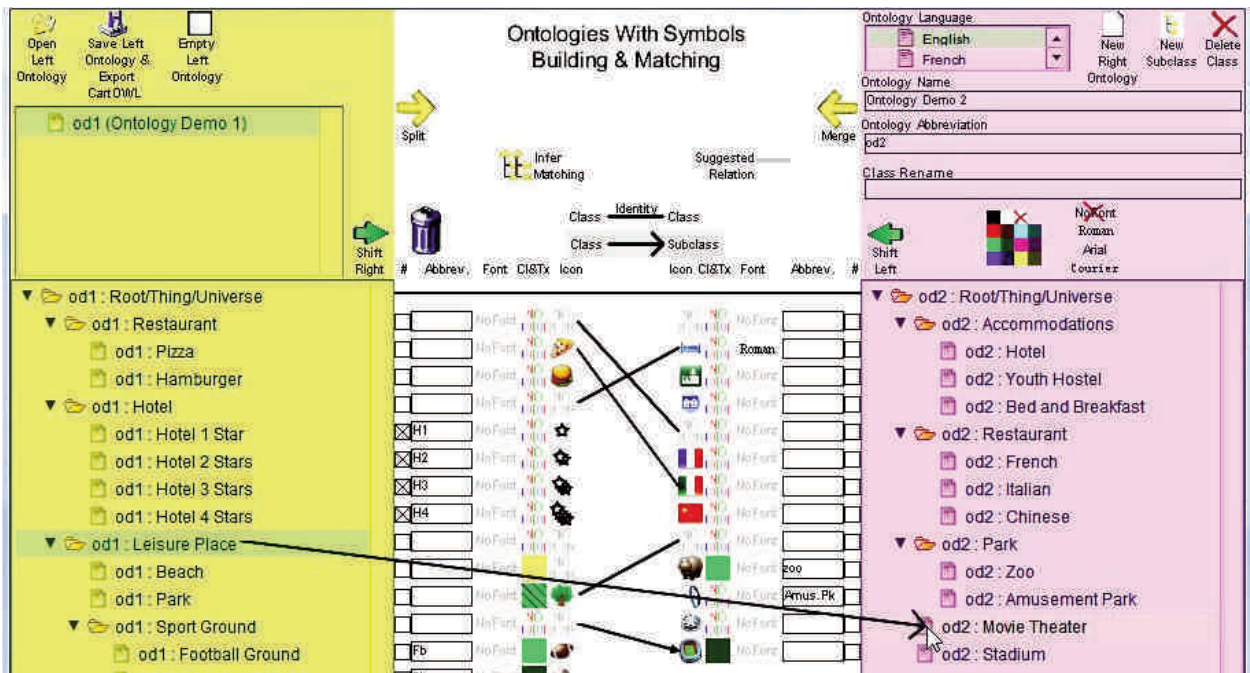


Figure 10 : Construction des ontologies cartographiques

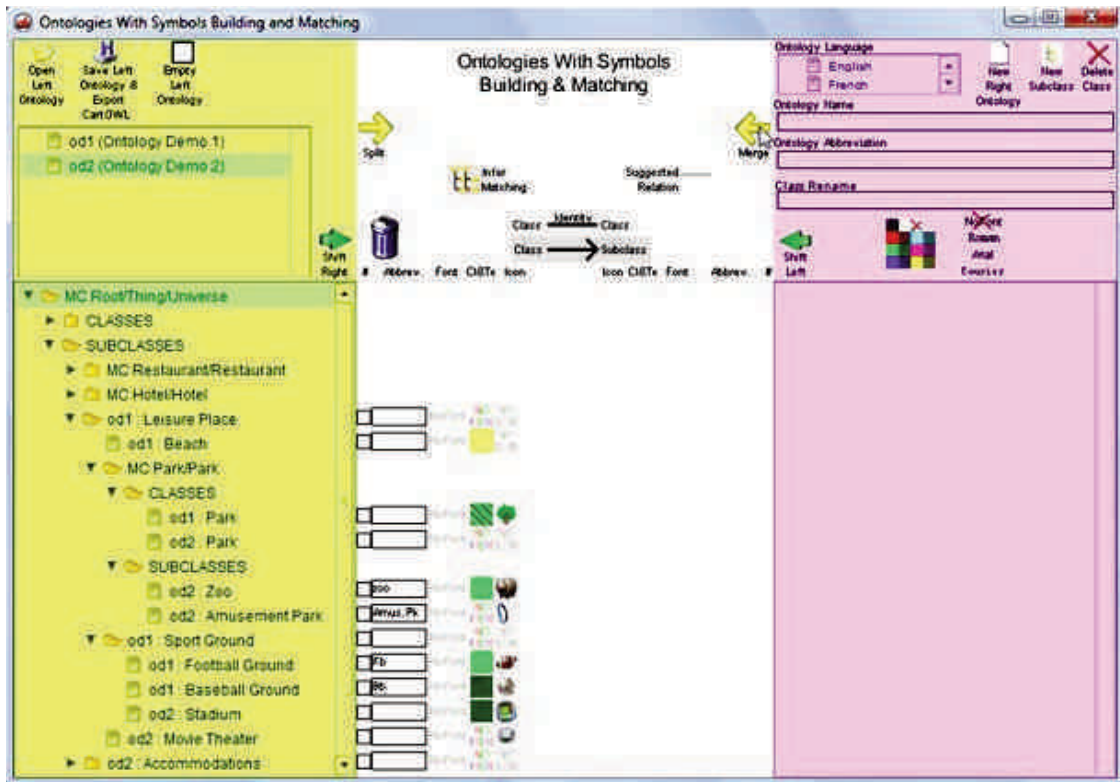


Figure 11 : Alignement des ontologies cartographiques

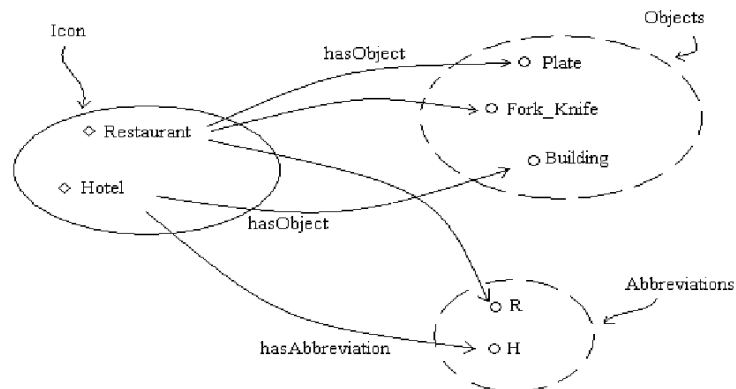


Figure 12 : Modèle des relations dans notre ontologie



Figure 13 : Classes et propriétés des symboles dans Protégé

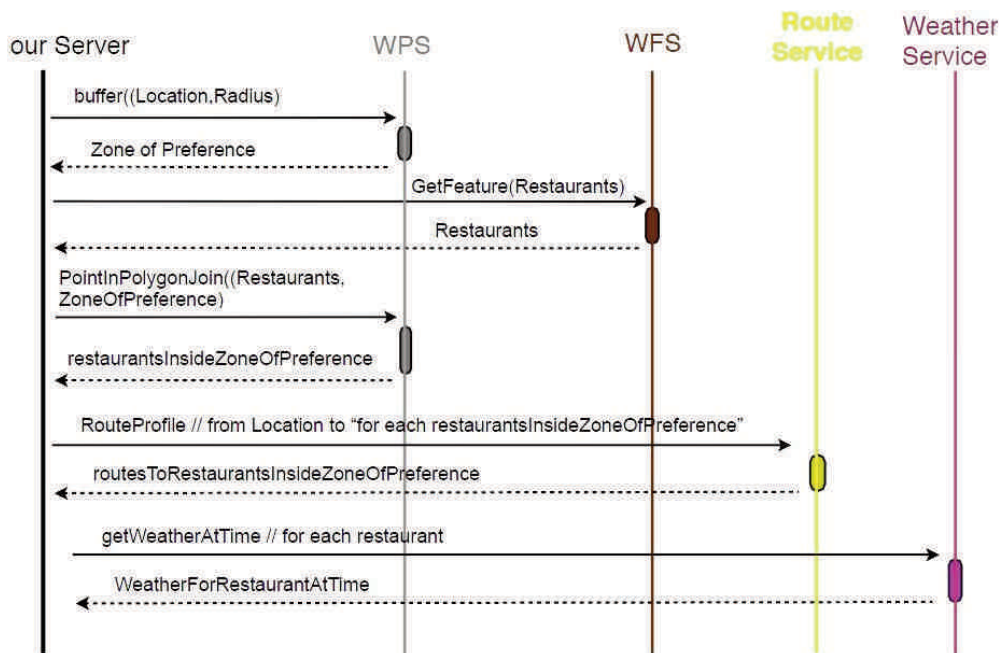


Figure 14 : Diagramme de séquences pour les différents services web

