

LE GAZETIER COMME OUTIL D'ACCÈS A L'INFORMATION GEOGRAPHIQUE

BEDEL O.(1), QUESSEVEUR E.(2), LEPRINCE F.(1)

(1) Alkante SAS, CESSON SÉVIGNÉ, FRANCE ; (2) Université Rennes 2/ Laboratoire Reso, RENNES, FRANCE

Dans sa forme commune, un gazetier est un dictionnaire d'éléments remarquables du territoire, un index associant des noms de lieux à une description de ces lieux. Les éléments décrits peuvent être hétérogènes (continent, pays, ville, montagne, rivière, site archéologique...) et leur description variées (position géographique, profondeur, superficie, population, photographie du lieu...). Un certain nombre de gazetiers sont disponibles sous forme numérique sur le web (Geonames, World Gazetteer, gazetier de l'Alexandria Digital Library...). La plupart s'appuient sur une structuration et des protocoles ad'hoc et sont de fait difficilement exploitables en dehors de leur contexte premier d'utilisation.

L'Open Geospatial Consortium (OGC) a proposé une déclinaison de sa norme Web Feature Service (WFS) pour fournir un service de type gazetier. Cette déclinaison baptisée WFS-G (ref. *OGC Best practice document : Gazetteer service - application profile of the web feature service implementation specification, 2006, 05-035r2*) est encore peu exploitée actuellement, mais fournit un cadre générique concernant, d'une part l'organisation des données d'un gazetier dans une base de données, et d'autre part, les modalités d'interrogation d'un gazetier prenant la forme d'un service web géographique.

Sur la base d'un service web WFS-G, nous proposons un cadre d'utilisation d'un gazetier pour interroger la donnée géographique disponible sur un lieu ou une zone donnée. Pour cela, nous décrivons un dispositif construit autour d'un gazetier et présentant l'information relative aux lieux dans trois vues complémentaires : un index de lieux, une projection cartographique et une vue descriptive. Dans la suite de cette communication, nous décrivons le principe fonctionnel du dispositif ainsi que les caractéristiques techniques d'une implémentation que nous avons réalisée. Nous comparons ensuite sur le plan fonctionnel notre dispositif aux gazetiers et aux catalogues de données disponibles actuellement. Enfin nous présentons un contexte d'application de ce dispositif à des données réelles : le projet CHARM 3.

Les données à la base du dispositif sont un ensemble de lieux. Un lieu est un objet possédant un ancrage géographique (ex. : un village, un lieu d'échouage d'une épave, un banc de sable, un département administratif...). Un lieu est décrit par un ensemble de descripteurs, certains assimilables à des mots clés (nom, type, thématique associée au lieu), d'autres prenant une forme plus complexe (ex. description spatiale associée au lieu). Enfin des ressources peuvent être liées à un lieu, ces ressources sont des éléments permettant de décrire le lieu comme un texte ou une image représentant le lieu (dessin, photographie, carte postale...). Deux descripteurs sont nécessaires pour le fonctionnement du système : le nom du lieu (son toponyme) et sa description spatiale. La description spatiale peut être une position (latitude, longitude) ou une géométrie (ligne, polygone) caractérisant les limites spatiales du lieu. La description spatiale est utilisée pour calculer une relation d'inclusion entre les lieux. Par exemple, à partir des descriptions spatiales des zones « Europe » et « France », le système est en mesure de déterminer que « l'Europe » est un lieu englobant le lieu « France ».

Le système fournit une représentation de l'ensemble de lieux sous la forme de trois vues complémentaires : une vue index, une vue cartographique et une vue description. La vue index est construite comme un arbre de toponymes. Cet arbre traduit l'inclusion spatiale de lieux (par exemple, « Bretagne » est un nœud apparaissant sous « France »). Chaque nœud de l'arbre est matérialisé par un toponyme accompagné d'un nombre. Le nombre recense les ressources disponibles sur le serveur de données du système et associées au lieu identifié par le toponyme. Les nœuds sont des éléments sélectionnables et la sélection d'un nœud parent entraîne la sélection de l'ensemble de ses nœuds fils. La vue cartographique est une carte dynamique où sont projetées les descriptions spatiales des lieux sur un fond de plan cartographique (images satellites, limites administratives, réseaux routiers...). La vue cartographique n'est pas statique, il est possible de s'y déplacer (déplacement, zoom) à la manière des applications Google Maps ou OpenStreetMap. Enfin la vue description permet de présenter les ressources liées à un lieu sous la forme d'une liste avec une présentation adaptée au format de chaque ressource (ex. : texte, image...).

Les différentes vues sont interactives et liées entre elles de façon dynamique comme le présente le schéma de la figure 1. Il est possible de sélectionner plusieurs éléments dans la vue index, ce qui entraîne la mise en surbrillance sur la vue cartographique de la position ou des limites du lieu concerné et la mise à jour de la liste des éléments apparaissant dans la vue description. A l'inverse, le dessin d'une zone de sélection sur

la carte entraîne à la fois la mise en surbrillance des lieux inclus dans cette zone sur la carte et des toponymes associés dans la vue index, et la mise à jour de la liste des ressources disponibles dans la vue description pour que celle-ci soit cohérente avec la sélection courante de lieux. Le dispositif permet à l'utilisateur de placer le toponyme comme point d'entrée dans sa recherche d'information. Le parcours de l'arbre de toponymes lui donne directement un aperçu des ressources disponibles sur une zone donnée.



Figure 1

Sur le plan technique, le dispositif se présente comme une application web classique reposant sur une architecture client-serveur. Les informations sont structurées sur le serveur qui les délivre au client sous la forme de services web. Sur le poste client, les différentes vues sont présentées à l'utilisateur dans une page web. Les actions menées sur les vues se traduisent par des requêtes envoyées au serveur et dont les réponses entraînent la mise à jour des vues.

L'ensemble de lieux est structuré dans une base de données dont la structure suit les recommandations du profilé WFS-G. Les données sont accessibles via un service web WFS et WFS-G. Un service web spécifique a été mis en place pour la construction de l'arbre de toponymes et la gestion de la vue index. Les données apparaissant dans la vue cartographique sont délivrées suivant le protocole WFS. Les sélections spatiales effectuées sur la vue cartographique sont traduites en requêtes spatiales au format Filter Encoding. Ce format est utilisé dans les différents protocoles OGC (WFS, CSW). Cela permet à l'intérieur du dispositif de générer des requêtes d'inclusion spatiale dans les zones associées aux toponymes. Il est donc possible à partir de ce dispositif d'interroger des serveurs de données et de méta-données géographiques externes et qui n'ont pas été à l'origine structurés pour répondre à des requêtes construites avec des toponymes, pour peu que ces serveurs supportent les services web OGC.

Une version du dispositif a été implémenté en s'appuyant sur un ensemble de composants géomatiques open-source relativement répandus. Le système de gestion de bases de données PostgreSQL et sa cartouche spatiale PostGIS ont été utilisés pour la base de données. Les différents services web (WFS, WFS-G et service de gestion de la vue index) s'appuient sur le serveur Deegree. La vue cartographique a été réalisée au moyen de la librairie OpenLayers et les autres vues (index et description) ont été construites en utilisant la librairie JQuery et quelques uns de ses plugins.

Comme la plupart des gazetiers disponibles aujourd'hui sur le web, notre dispositif présente pour une zone géographique donnée un inventaire de noms de lieu et leur associe une description (coordonnées géographiques, typologie du lieu, thématique...). Dans une certaine mesure, il informe sur les relations spatiales entre ces lieux (inclusion spatiale de « Bretagne » dans « France »). Sur le fond, les informations présentées sont identiques. Sur la forme cependant, le couplage dynamique de la vue index et de la vue cartographique offre une plus grande flexibilité dans l'accès à l'information. De plus, de part sa construction autour des normes et des protocoles OGC, l'outil n'est pas limité à l'exploitation de la base de

toponymes, mais permet d'utiliser les toponymes comme critères de recherche spatiaux d'un entrepôt de données géographiques supportant les services OGC (WFS, WMS, CSW...).

Le dispositif se démarque également des outils de catalogage de données géographiques (geonetwork par exemple) qui permettent de référencer les données géographiques au travers de fiches de métadonnées. Ces outils proposent des fonctionnalités de recherche de métadonnées (et donc, in fine, de données géographiques) au travers d'une interface permettant à l'utilisateur de préciser plusieurs critères thématiques (mots clés libres, thème sélectionné dans une liste...), critères spatiaux (sélection graphique d'une zone d'intérêt sur une carte) ou autres (date de publication, producteur de la donnée...). Un intitulé de zone caractérisant la zone de validité spatiale de la donnée et constituant donc un toponyme est souvent renseigné dans la métadonnée en complément de l'emprise géographique de la donnée référencée. Cependant les interfaces de ces outils ne proposent ni une présentation hiérarchique arborescente de toponymes, ni la sélection d'une zone prédéfinie ou la saisie graphique d'une zone arbitraire sur une carte comme critère de recherche. De ce point de vue, le dispositif de gazetier offre un service complémentaire aux interfaces de recherche des catalogues de métadonnées existants en plaçant le toponyme et l'organisation spatiale des lieux au cœur de la recherche de données. Comme précisé précédemment le dispositif de gazetier peut être couplé facilement à un outil de catalogage en s'appuyant sur le protocole CSW pour définir des filtres spatiaux.

Le dispositif de gazetier que nous proposons est en cours d'expérimentation dans le cadre du projet européen CHARM 3 (Interreg IV.A). Le projet CHARM 3 vise à accroître la connaissance de l'espace maritime de la Manche dans le but de contribuer à la gestion intégrée de cet espace frontalier. Il rassemble 17 partenaires scientifiques anglais et français aux compétences très complémentaires. Le projet comporte ainsi de nombreux volets thématiques portant notamment sur :

- l'inventaire des données physiques, environnementales, biologiques et d'activité humaine, disponibles en Manche,
- l'inventaire des taxons planctoniques,
- la classification des habitats marins,
- l'étude de l'exploitation commerciale des espèces halieutiques,
- l'analyse de la législation en vigueur,
- la cartographie et la modélisation de l'habitat,
- la modélisation des réseaux trophiques,
- ...

Un volet destiné à la diffusion des connaissances rassemblées doit permettre de les porter à la connaissance des services de l'État, des collectivités territoriales et des acteurs socio-économiques.

Dans cette logique un atlas interactif accessible par un portail web sera mis en œuvre. Le gazetier est un des éléments de cet atlas, complémentaire d'outils plus traditionnels de présentation dynamique de cartes (sur le modèle du Géoportail de l'IGN). Dans ce contexte, la vocation de l'outil est de proposer et d'expérimenter un accès à l'information géographique guidé par les toponymes auprès des utilisateurs du portail : d'une part, les scientifiques produisant de l'information sur la zone Manche, et d'autre part, le grand public.

L'outil gazetier est utilisé pour accéder à trois jeux de données :

- une base de données toponymiques des éléments remarquables des fonds marins (épaves, bancs de sables...) sur l'ensemble de la Manche,
- un recueil de lieux culturellement liés à la mer (magasins de pêche, mouillages de bateaux, pubs...) à proximité du port d'Hastings (Angleterre),
- une liste de lieux du bord de mer sur la commune de Wimereux (France).

Le premier jeu offre une couverture complète de la zone Manche mais avec des descriptions relativement sommaires (nom d'une épave, date de naufrage...). A l'inverse, les deux derniers jeux concernent des zones géographiques plus réduites mais contiennent des descriptions plus riches et plus hétérogènes (mots clés, photographies, cartes postales, descriptions textuelles, bandes sonores...). Ces descriptions riches peuvent être facilement intégrées au gazetier sous la forme de documents liés aux lieux. Ces documents sont ensuite présentés à l'utilisateur sous une forme adaptée (fiche textuelle, image ou encore enregistrement audio avec un lecteur incorporé) au travers de la vue description de l'application. Du point de vue de la présentation de description au format hétérogène, la vue description du gazetier offre un espace dédié et flexible qui n'existe pas dans les autres composants de l'atlas web (applications web-mapping traditionnelles, fiches synthétiques ou cartes sous la forme de pages web statiques) du projet CHARM 3 et

qui, à notre connaissance, n'est pas encore intégré aux solutions de cartographie en ligne disponibles actuellement.

Nous avons décrit un cadre d'utilisation d'un gazetier pour faciliter l'accès à l'information associée à des noms de lieux, voire l'accès à une information géographique distante. De notre point de vue, cette approche est innovante par plusieurs points. Cette approche est de notre point de vue innovante sur plusieurs points. Premièrement, le système propose une représentation arborescente dynamique des noms de lieux offre . Le couplage dynamique entre cette vue index et la projection spatiales des positions et des limites spatiales de lieux une vue cartographique offre une flexibilité au niveau de l'interrogation. Enfin, l'exploitation des protocoles OGC dans le dispositif permet d'interroger des serveurs de données externes en utilisant un critère toponyme. De ce point de vue, ce système se veut complémentaire des systèmes de gestion de méta-données actuels. Nos travaux futurs visent à enrichir l'arbre de toponymes pour y intégrer la typologie des lieux, leur thématique ou des informations temporelles qui pourraient ainsi servir de critère de sélection pour les différentes vues. Nous nous intéressons également à la caractérisation des lieux dont la description spatiale peut être floue (ex. : un banc de sable dont la position évolue avec les mouvements marins).