

Chapitre 2

Introduction au GeoWeb

2 – Introduction au GeoWeb

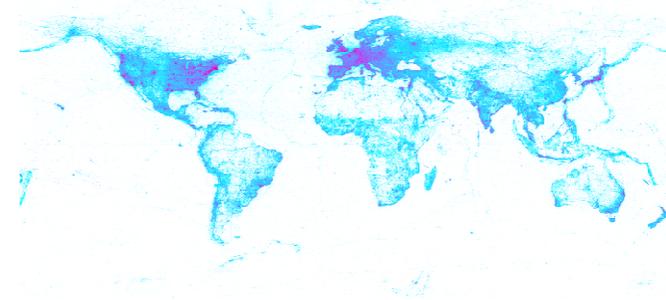
- 2.1 – Web sémantique/géographique
- 2.2 – Introduction aux mashups
- 2.3 – Geocoding, Geonaming, Geoparsing et Geotagging
- 2.4 – Conclusions

GeoWeb, il y a 2 ans



Copyright 2009 Google Inc.

GeoWeb, maintenant



Copyright 2009 Google Inc.



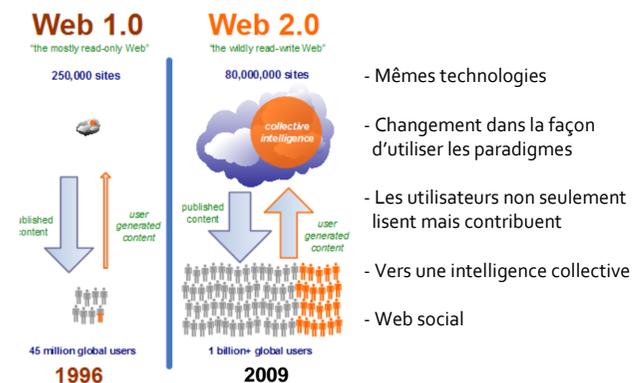
2.1 – Web sémantique/géographique

- Evolution des SIG
- Nouvelles fonctionnalités
- Nouvelles mentalités
- Tous les citoyens peuvent contribuer

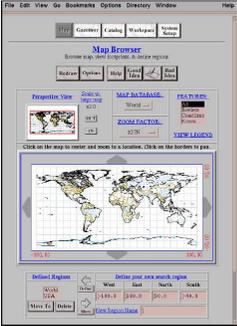
Les mots-clés du Web d'aujourd'hui

Communication, collaboration, partage, participation, tagging, video online, sharing, widget, filtering, online documents editing, wiki, blog, corporate blogging, peer production, viral marketing, proximity marketing, internet of things, bookmarklet, social media, social networks, links, social news, citizen journalism, messaging, pooling, Structures émergentes, ordonner, prioriser, mashup, feeds...

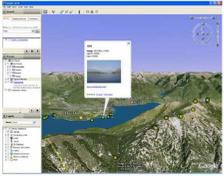
De hier à aujourd'hui



Evolution



Après 1990 →



Depuis 2005 →

Evolution



Après 1990 →
Données spatiales structurées

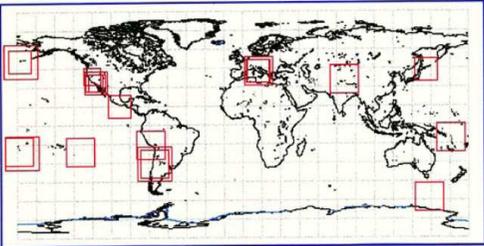


Depuis 2005 →
Données spatiales non-structurées

Global Earthquake Map

The earthquakes are shown on this world map as red squares, with the size of the square proportional to the magnitude of the earthquake.

You may click on any location on this map to zoom in much closer. Once you zoom, you will have access to the standard Xerox PARC Map Viewer with its full set of functions.



[RETURN TO INTRO PAGE](#)

Cartographie sur le Web

- Génération d'images statiques
- Interactivité → carte raster
 - Clic et quelque chose survient
 - Filtrage et zoom
 - Quelques changements possibles (couleurs, etc.)
 - Données provenant d'un seul serveur

Web sémantique

- Transformation du WWW en un environnement de publication de documents (pages HTML, fichiers, images, etc.)
- Association de données et de métadonnées
- Format adapté à l'interrogation, l'interprétation et, en général, à l'élaboration automatique.

- Interprétation → autres types de recherche beaucoup plus évoluées que celles basées sur les mots-clés
- construction d'opérations spécialisées comme la construction de réseau de relations
- entre les documents selon des logiques plus élaborées que le lien hypertexte classique

But du Web Sémantique

- Développement d'applications pour
 - Extraction d'information à partir de collection dynamique de documents
 - Contrôle de validité des contenus
 - Identification de style
 - Reconnaissance de liaisons hypertextuelles plus sophistiquées
 - Agents intelligents

GeoWeb

- GeoWeb est un néologisme qui implique la fusion d'informations basées sur la localisation, avec celles plus abstraites qui dominent Internet
- Question prototypique « **qui y a-t-il ici ?** »
- → trouver sur Internet toutes les informations relatives à un lieu

GeoWeb

- Localisation intelligente
- Usage d'Internet
- Toponymes → localisation sur le globe
- Adresse postale → localisation sur le globe
- Relations de localisation
- Gazetteer = répertoire de noms de lieux

Trois ingrédients du web 2.0

- Technologique
 - Web interactif
 - non plus logiciels mais des services
- Sociologique
 - Se mesurer à une communauté, interagir
- Economique
 - Celui qui fournit un service dépense peu
 - mais peut gagner beaucoup

Partage

	Web 1.0	Web 2.0
Information	Experts	Blogs
Connaissance	De celui qui écrit à celui qui lit	Wikipedia
Images	Base de données	Communautés (Flickr)
Video	Base de données	Communautés (YouTube)
Marque-page	Privé (utilisateur)	Partagés Del.icio.us
Classification	Taxonomie	Folksonomie

Taxonomie / Folksonomie

- Catégories
- Classes
- Règles
- Planifiée
- Précise
- Déterministe
- Laborieuse
- Web sémantique
- Utilisée - catalogues
- Individus
- Attributs
- Suggestions
- Implicite
- Floue
- Probabilistique
- Immédiate
- Web 2.0
- Utilisée par une communauté

Mashup ?

- Terme qui provient de la musique pop
- Fusionner plusieurs services existantes pour générer de nouvelles applications
 - SOAP
 - REST
 - JavaScript
 - RSS/Atom

1 browser, 1000 applications

- Courrier électronique (Gmail)
- Cartographie (Gmaps, Yahoo!maps, VirtualEarth)
- Word processor (Writely, officelive)
- Feuille de calcul (Google spreadsheet)
- Encyclopédie (Wikipedia)
- Agenda (30 boxes, Gcalendar)
- Bookmarks (del.icio.us)
- News, podcast, etc. (Bloglines)
- Programmation (zimki)

Avantages pour l'utilisateur

- Ubiquité des applications
- Aucune installation
- Mises à jour continues (chaque heure)
- Données « dans l'air »
- L'ordinateur devient un terminal intelligent vers Internet

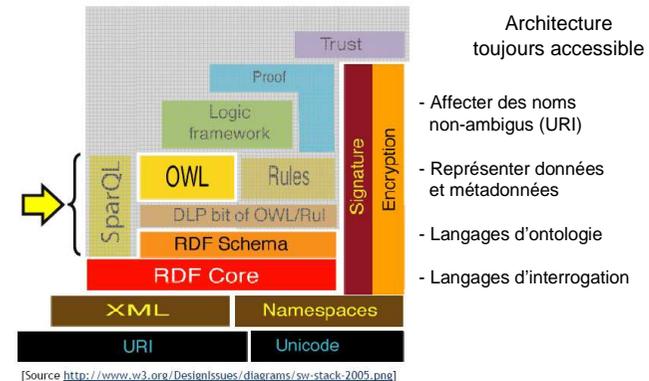
Problèmes immenses pour les utilisateurs

- Que font-ils de nos données ?
- Comment ne plus se satisfaire d'un service et passer à un service concurrent ?
- Comment faire migrer données et métadonnées ?

Monopole de nouveaux géants

- Chaque développeur finira par écrire programmes qui tournent dans leur datacenter ?
- Le desktop n'est plus le champ de bataille
- Ils ont investi des milliard de dollars dans l'acquisition et dans le développement des services Web 2.0, comme YouTube,

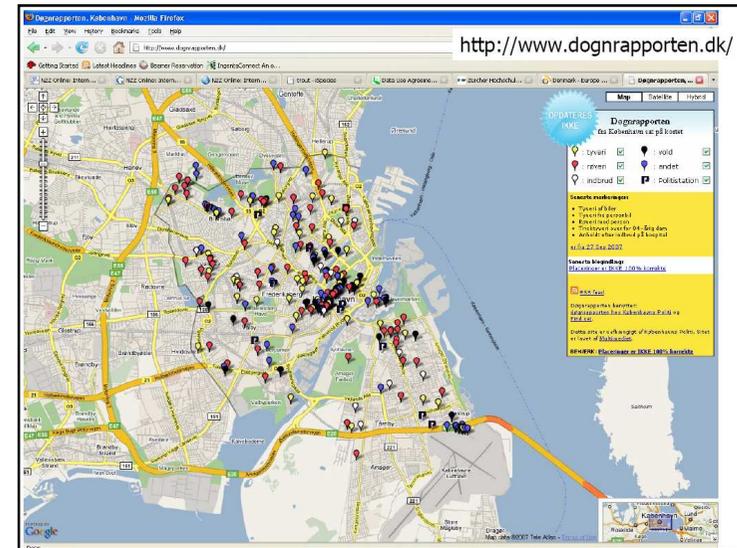
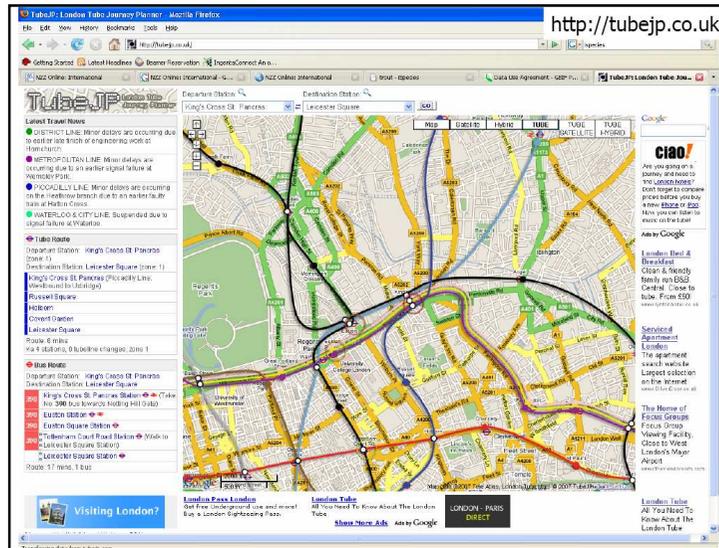
Gâteau du web sémantique



2.2 – Introduction aux mashups

- Les mashups s'appuient sur des données provenant de sources multiples pour réaliser de nouveaux produits et applications.
- Pour permettre la création des mashup avec des données et services, les fournisseurs doivent fournir les interfaces à leurs données de façon à créer une chaîne de services sur le web
- Les mashups permettent la cartographie interactive des données en temps réel. Car n'était pas possible avec les cartes papier, par exemple où est MAINTENANT mon train ?





Exemples de Mashup

- Interroger les espèces, demander des informations sur leur distribution, sur les images, et sur les articles scientifiques ou livres à la demande
- Cartographier un itinéraire avec les stations du métro de Londres
- Cartographier les avions et les informations les concernant en temps réel au voisinage de Zurich
- Cartographier les crimes récents à Copenhague

Aspects des mashups

- De nombreux mashups sont basés seulement encore sur une géographie de type point ; par exemple GeoNames rend un centroïde pour toute la Suisse,
- Les mashups peuvent fusionner des données de sources différentes, avec un objectif différent de celui avec lequel ils avaient été conçus originalement
- Les aspects de qualité de données sont ignorés en général
- Les aspects de confidentialité sont rarement traités
- Certains services peuvent devenir populaires très rapidement

Recherches des informations géographiques (GIR)

- Une grande partie des données que nous voyons quotidiennement n'est pas structurée ou partiellement structurée,
- Une recherche sur les documents textuels montre que
 - 85% de 20 000 documents britanniques incluent des noms de lieux,
 - et 13% de 4 millions de requêtes sur le web contiennent un lieu géographique

Éléments-clefs du GIR

- Identification de la trace : identifier les endroits mentionnés dans les textes non- structurés.
- Expansion d'une requête : ajouter des noms non présents dans la question initiale.
- Indexation spatiale et indexation textuelle.
- Classements : selon le sujet et la position.
- Formulation de requête et visualisation des résultats : une interface qui permet aux utilisateurs de formuler et d'explorer les requêtes spatiales.

Traces spatiales dans les documents



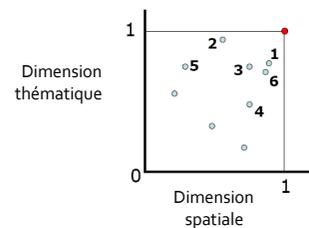
Scotland
Cairngorm National Park
Norway
Alps
Nepal
Greenland
01479 861256
Glenmore Lodge
Glenmore
Aviemore
Inverness-Shire
PH22 1PL
Tel: 01479 861256

Expansion de requêtes et indexation spatiale

- Expansion de requêtes signifie que si un utilisateur recherche des documents sur Lyon, il faut ajouter Villeurbanne, Caluire, etc. à la question.
- Ainsi besoin de connaître la topologie, les lieux voisins et leurs relations spatiales.
- On doit typiquement utiliser des dictionnaires géographiques. Si un utilisateur demande des documents à propos des "châteaux à Zurich", un index spatial réduit le numéro de documents seulement à cette ville et à son voisinage.

Classification

- Un système de recherche des informations ramène des documents avec un score
- En GIR les scores ont typiquement, deux dimensions (thématique et spatiale)



2.3 – GeoCoding, GeoNaming GeoParsing et GeoTagging

- GeoCoding
 - Trouver les coordonnées d'un lieu
- GeoNaming
 - Affecter un nom à un lieu
- GeoParsing
 - A partir d'un texte, retrouver le ou les lieux correspondants
 - Parfois des ambiguïtés
- GeoTagging
 - Annoter un lieu sur une carte

GeoCoding

- Affecter les coordonnées à un lieu généralement par longitude et latitude
- Deux représentations
 - degrés, minutes, secondes ($gg^{\circ}mm'ss''$)
 - degrés décimaux (gg,ddd)

$$ddd = mm/60 + ss/3600$$

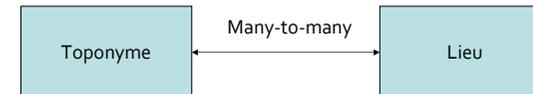
- Dans certains cas : interpolation des adresses postales
 - Interpolation linéaire à partir des numéros dans les rues

GeoNaming

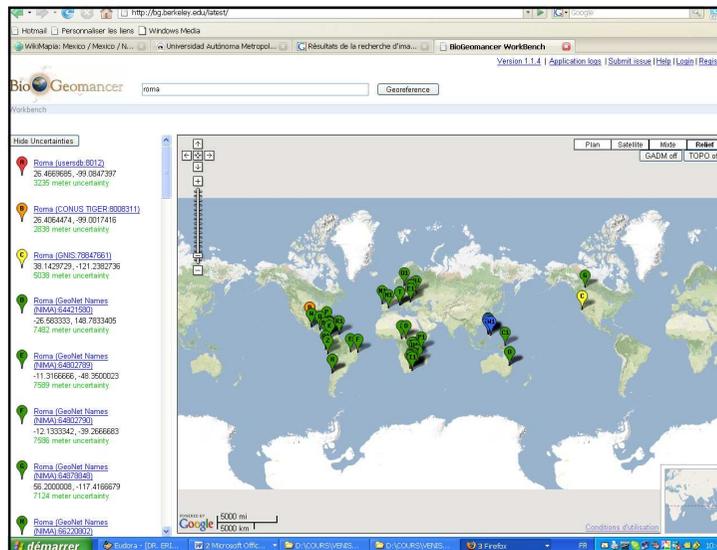
- Des coordonnées au nom de lieu
- Définition du lieu
 - Point, ligne, zone
 - Problèmes de linguistique
 - Problèmes multilingues
- Quel nom ?
 - Celui dans la langue officielle du pays
 - Celui dans la langue de l'utilisateur
 - Celui dans la langue du système d'exploitation

GeoParsing

- Analyser un texte pour localiser



- Exemple : Rhône (fleuve ou département ?)
- Exemple : Roma



GeoParsing : 3 définitions

- Toponyme → localisation
 - Où se trouve le Campus de la Doua ?
- Relation avec un toponyme → localisation
 - A 15 Km au sud-est de Lyon
- Analyse de texte → localisation
 - Bible, textes historiques, etc.

Instruments

- Annuaire de toponymes = Gazetteers
- Dictionnaires multilingues
 - Venezia, Venice, Venise, Venecia, Venedig, Benetke, Benátky... etc.

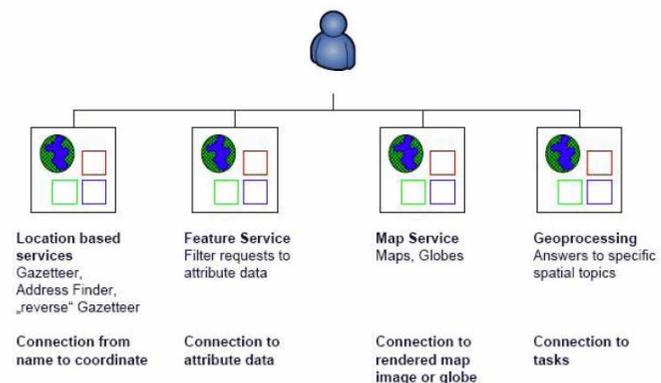
Problématique du facteur

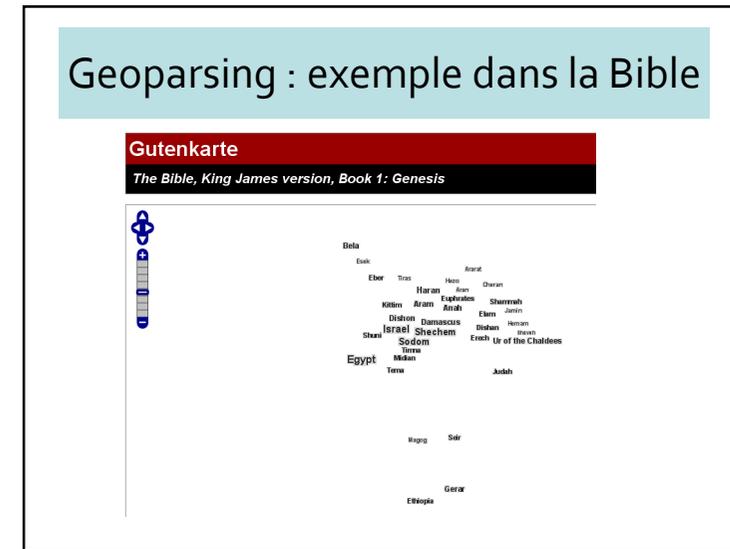
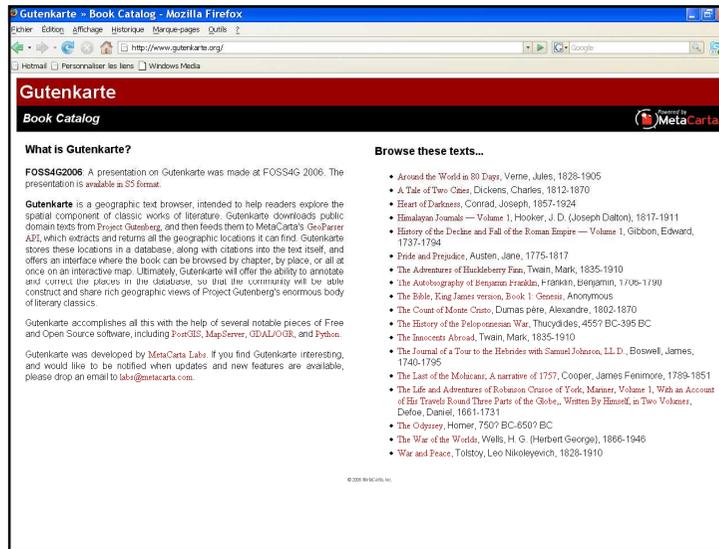
- Mrs Florence Manchester
- 2345 New York Avenue
- 97347 Aberdeen, WA
- =====
- Sgra Emilia Romagna
- 145 via Enesto Viale
- Nicosia

Sites web de GeoParser

- NGA GEOnet Names Server (GNS)
 - <http://earth-info.nga.mil/gns/html/>
- BioGeoMancer
 - <http://bg.berkeley.edu/latest/>
- Edina GeoParser
 - <http://edina.ac.uk/projects/geoxwalk/geoparser.html>
- Ecc.

Gazetteer part of a service architecture



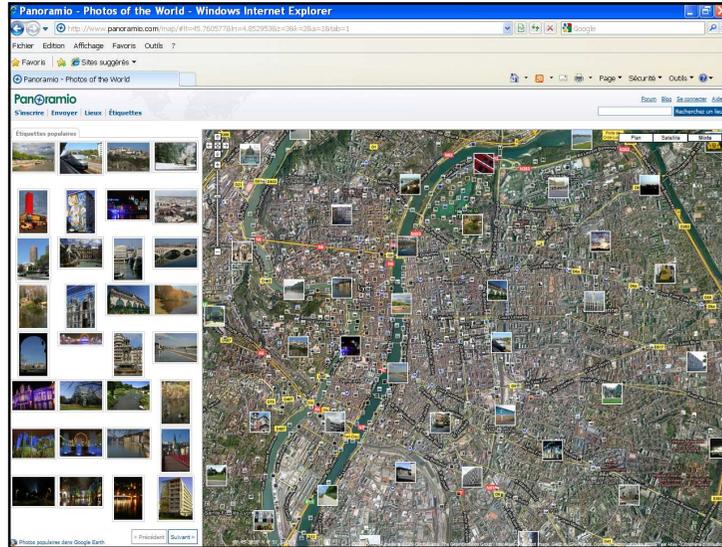


Transformer un texte en carte

- Autres applications
 - Reconstruction des itinéraires d'explorateur
 - Reconstruction de voyage dans l'ancien temps
 - Météorologie (du texte à la carte)
 - Etc.

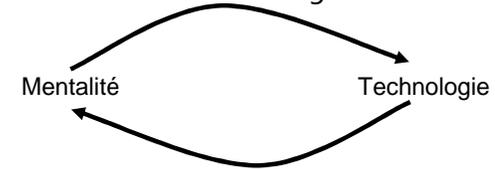
Geotagging

- Annotations géographiques généralement multimédias
 - Photo
 - Textes
 - Films
 - Voix, musique
- Systèmes existants
 - Panoramio
 - Flickr
 - Everyscape
 - Wikimapia
 - Trippermap



2.4 – Conclusions

- Evolution des mentalités
- Evolution de la technologie



- Homo Faber → Homo Connectus
 - Du dogmatisme des informations
 - Au partage des informations