A Web more Geospatial Insights into the Location inside

Susanne Boll
Media Informatics and Multimedia Systems
University of Oldenburg
Oldenburg, Germany

Dirk Ahlers
OFFIS Institute for Information Technology
Oldenburg, Germany

Understanding Web Evolution:
A Prerequisite for Web Science
April 22, 2008
Beijing, China
The Web & the role of location

The Web
- an unlimited resource of interlinked information
- subject to keyword-based search

Location
- has always been part of Web content
  - companies, travel diaries, ...
- information more relevant than ever to (mobile) Web search
- will be a driving force behind the next generation Web

So, if the location information is there, why is more insight needed?
Let’s look at location in the Web today

We find
- Directories: Yellow Pages, Qype
- Maps and Route Planning: Google Maps, Map-based Mashups
- Localized Content: Locr

The location we see is not necessarily the location inside the Web
- Location content is “manually” created, edited
- From companies, user-generated
- But if you want to place something on a map, you must know its location

But
- There is much location hidden inside
- > 25% of Web pages / Web sites refer to a location
Let's look at location in the Web today.

We find:
- Directories: Yellow Pages, Qype
- Maps and Route Planning: Google Maps, Map-based Mashups
- Localized Content: Locr

The location we see is not necessarily the location inside the Web. Location content is "manually" created and edited by companies, user-generated, but if you want to place something on a map, you must know its location.

But there is much location hidden inside:

> 25% of Web pages / Web sites refer to a location.
Let's look at location in the Web today. We find:

- Directories: Yellow Pages, Qype
- Maps and Route Planning: Google Maps, Map-based Mashups
- Localized Content: Locr

The location we see is not necessarily the location inside the Web. Location content is "manually" created, edited by companies, or user-generated. But if you want to place something on a map, you must know its location.

But there is much location hidden inside:

> 25% of Web pages / Web sites refer to a location.
Let's look at location in the Web today.

We find:
- Directories: Yellow Pages, Qype
- Maps and Route Planning: Google Maps, Map-based Mashups
- Localized Content: Locr

The location we see is not necessarily the location inside the Web. Location content is "manually" created, edited from companies, user-generated. But if you want to place something on a map, you must know its location.

But there is much location hidden inside:

> 25% of Web pages / Web sites refer to a location.
Let's look at location in the Web today.

- Directories: Yellow Pages, Qype
- Maps and Route Planning: Google Maps, Map-based Mashups
- Localized Content: Locr

The location we see is not necessarily the location inside the Web. Location content is "manually" created, edited by companies, or user-generated. But if you want to place something on a map, you must know its location.

But there is much location hidden inside:

> 25% of Web pages/Web sites refer to a location.
Let's look at location in the Web today.

- Directories: Yellow Pages, Qype
- Maps and Route Planning: Google Maps, Map-based Mashups
- Localized Content: Locr

The location we see is not necessarily the location inside the Web. Location content is “manually” created, edited by companies, and user-generated. But if you want to place something on a map, you must know its location.

But there is much location hidden inside. More than 25% of Web pages/web sites refer to a location.
The Web and its undiscovered locations

The location we see is not the location inside
But, it is there
The Web and its undiscovered locations
The Web and its undiscovered locations
The Web and its undiscovered locations
What do we want to do?

Understand the spatial character of the Web

We want to

- search content by location, area, corridor, ...
- retrieve in-depth geospatial information about the pages we visit
  - its location, area and spread
  - its community rating
  - spatially neighboring pages
  - its connectivity
  - its geographic audience distribution
  - the density of similar content close by
  - ...

Your site: medien.informatik.uni-oldenburg.de

Number of sites referencing your site:

- 5000 – 4000
- 4000 – 3000
- 3000 – 2000
- 2000 – 1000
- 1000 – 0

Total number of referencing sites: 220,878
You searched for sites about: diving

Number of relevant sites:

- 5000 – 4000
- 4000 – 3000
- 3000 – 2000
- 2000 – 1000
- 1000 – 0

Total number of relevant sites: 220,878
Dolore magna aliquyam
(Area: around Dortmund)
Erat, sed diam voluptua
(City: Frankfurt)
Getting insights into the location inside

- identification
- geoparsing
- link analysis
- crawling
- feedback
- annotation
- interaction
- mining
- explore
- visualization
- integration
- extension
- analysis
- semantic enrichment
- scope

**discover**

**understand**

**augment**
Conclusion

**Location**

- is hot
- is already there
- is more than placing points on maps
- is not (only) a question of markup and ontologies
- needs a multidisciplinary approach from discovery to exploration

**Location in Web Science means**

- engineering models and methods to describe, understand, augment, search, find and (visually) explore location-related information on and in the Web
Contact

Susanne Boll
Media Informatics and Multimedia Systems
University of Oldenburg, Oldenburg, Germany

medien.informatik.uni-oldenburg.de
susanne.boll@informatik.uni-oldenburg.de