

Workshop on Geographic Information Retrieval held at SIGIR'06

Ross Purves
University of Zurich
rsp@geo.unizh.ch
http://www.geo.unizh.ch/~rsp/gir06

Chris Jones
Cardiff University
c.b.jones@cs.cf.ac.uk

1 Workshop motivation

August 2006 saw the third in a series of workshops held to discuss the state of the art in Geographic Information Retrieval held in conjunction with SIGIR'06. The original call for papers for the workshop suggested the workshop should discuss further progress within the field and potential future research strands. Topics considered relevant for the workshop included the following:

- architectures for geographic search engines;
- spatial indexing of documents and other media resources;
- extraction of geographical context from documents and geo-datasets;
- geographical annotation techniques for geo-referenced media;
- design, construction, maintenance and access methods for geographical ontologies, gazetteers and geographical thesauri;
- geographical query interfaces for the web and geo-spatial libraries;
- visualising the results of geographic searches; and
- relevance ranking for geographical search.

A total of 25 papers were submitted for the workshop and after a rapid and comprehensive review process by our program committee 16 of these were selected for presentation in Seattle. In a departure from previous formats it was also decided on this occasion to invite a keynote speaker to address future challenges in the field of GIR. John Frank, of MetaCarta kindly agreed to attend the workshop and present a keynote to the assembled masses.

2 Workshop summary

This year's workshop had the highest attendance yet with 45 researchers from academia, funding bodies and industry squeezing into a seminar room for a stimulating and useful discussion of where we are and where we are going with GIR.

Individual papers from the workshop are available for downloading at <http://www.geo.unizh.ch/~rsp/gir06>.

The workshop was divided into 4 sessions, starting with a discussion of *Place names and geoparsing* where a range of topics centered around the identification, extraction and grounding of place names within texts was presented. In the second session, *Query processing and spatial language*, a set of papers which addressed issues related to query formulation and geographic representation of documents was presented. These papers included the analysis of geographic queries to existing search engines, methods to deal with location in query, fuzzy spatial reasoning in search and the geographical serving area of web resources. This session featured papers from both Yahoo! and Microsoft Research, indicating the importance of GIR to commercial providers of search.

Following a brown bag lunch we kicked off with a stimulating keynote from John Frank of MetaCarta. John's keynote, entitled "*Cartographic Information Retrieval*", focused on the fact that whether or not that was our intention, GIR inherently involved important *cartographic* challenges. John posed the following questions, which from the perspective of a company directly involved with GIR, he considered worthy research challenges where current techniques fell short:

- How should IR label relief representation?
- How should IR align with geometric layers?
- How should IR generalize across scales?
- What is the value of a georef?

His presentation was richly annotated with maps and examples and provoked a stimulating discussion of possibilities for future research directions.

The following session, entitled *Access methods and integration* included a variety of papers addressing issues related to the association of documents and parts of documents with simple and complex descriptions of locations through geographic ontologies, techniques for the distributed storage of such information in hash tables and, importantly, methods for ranking the results of geographic search. Ray Larson rounded the session off with a discussion on the current state of GeoCLEF, The CLEF Cross-Language Geographic Information Retrieval Track.

In the final session of the day, *GIR Systems*, a number of prototypes of fully featured GIR systems were presented. These included a system implementing geographic search for person photograph collections and the development of a context-aware geographic search system. An further example of geographic search illustrated the influence of so-called implicit locations, and a novel application tested the possibilities of retrieving similar travel routes by chunking GPS routes into textual descriptions.

In the evening the workshop attendees adjourned to more comfortable surroundings to carry on our discussions and experience Greek cuisine Seattle style. This year's workshop would not have been possible without the help of the SIGIR local and workshop organizers, our program committee, John Frank's enthusiastic and stimulating response to our invitation to give a keynote and most of all the interest and active participation of all the attendees in presenting papers or participating in discussions. Thank you all!