Title: City Search: An evolution of search to incorporate city data

Length of tutorial: half-day

Abstract: Today plenty of data is emerging from various city systems. Beyond the classical Web resources, large amounts of data are retrieved from sensors, devices, social networks, governmental applications, or service networks. In such a diversity of information, answering specific information needs of city inhabitants requires holistic IR techniques, capable of harnessing different types of city data and turned it into actionable insights to answer different queries. This tutorial will present deep insights, challenges, opportunities and techniques to make heterogeneous city data searchable and show how emerging IR techniques models can be employed to retrieve relevant information for the citizens.

Intended audience: (introductory / intermediate) The tutorial is aimed at all practitioners and researchers familiar with a basic IR and data management knowledge and interested in emerging big data technologies for the processing of urban data, in order to improve data consumption and enable advance semantic and contextual search based on user needs and tasks.

Tutorial Description: With the idea on mind of giving more visibility to the challenges and opportunities involved in consuming emerging city data, this is a follow-up of the tutorial presented recently at SIGIR 20131. The previous tutorial attracted a lot of interest both from academia and industry with a number of 24 attendees. This tutorial will address this emerging topic of City Search, and present it as an evolution of the current search technologies, such as Web search and local search, as an adaption to an environment of World cities. In this respect, it distinguishes itself significantly in terms of the complex information needs of the city users, the way search relevance is modeled according to those needs, and diversity of the information sources that are available in a city context different from the Web. As city data becomes more and more available every day and the current information needs of the users in a city context cannot be properly answered with the state-of-the-art search technologies, the topic poses a lot of research challenges and industry potential to be investigated in detail and with a high potential of industrial impact. Thus it is one of the main goals of the tutorial to increase the awareness of the search community on this emerging important area of search and also enumerate the opportunities to attract expertise and interest from different communities.

Outline: As urban data introduces a novel hybrid scenario that requires the synergy of disciplines such as IR, data management, semantic search, local search, social media analysis, geospatial search and so on, the content of the tutorial will span a number of relevant areas each of which has a significant contribution to realize the big vision of City Search. As outline, the tutorial consists of three main parts as the following:

- Part I - Beyond Local Search: This part will introduce the overall vision of City Search and discuss how it differs from its counterparts. It will involve three main sections:
  - A Planet of Smarter Cities
  - City Data and Information
  - Making City Search Smarter
- Part II – Managing City Data: This part of the tutorial will focus on a detailed introduction of the currently available city data sources and how the vast amount of city data can be better managed to enable search. This is an important pre-processing step for City Search and also the main source of its evolution as an emerging search process. It will consists of the following subsections:
  - Semantic processing and lifting of Urban Data

1 SIGIR Tutorial “Searching in the City of Knowledge”: [http://www.dublinked.ie/?q=searchinginthecityofknowledge](http://www.dublinked.ie/?q=searchinginthecityofknowledge)
2 The slides of the previous tutorial can be reached at: [http://researcher.watson.ibm.com/researcher/view_project.php?id=4892](http://researcher.watson.ibm.com/researcher/view_project.php?id=4892)
• Querio City: Consuming & querying governmental data
• Reasonable City: Semantic reasoning over city data to enable insight
• Social City: Social media and city event dynamics
• Care City: Contextual information and access to integrated city healthcare
• Streaming city data

Part III – Searching City Data: This last part of the tutorial will focus on the different steps of the City Search process and review the existing techniques from a number of different fields to enable the search in the city context. Mainly it will cover the related techniques grouped according to its functionality in the following steps of the process:
• Indexing City Data
• Query Processing
• Retrieval and Ranking
• Summary and Outlook

Contact information:
Veli Bicer: VELIBICE@ie.ibm.com
Vanessa Lopez: vanlopez@ie.ibm.com

Smarter Cities Technology Centre
IBM Research Ireland
Damastown Industrial Estate, Dublin.

Brief biography:

Veli Bicer is a researcher at Smarter Cities Technology Center of IBM Research in Dublin in 2012. His research interests include semantic data management, semantic search, software engineering and statistical relational learning. He obtained his PhD from Karlsruhe Institute of Technology, Karlsruhe, Germany and B.Sc. and M.Sc. degrees in computer engineering from Middle East Technical University, Ankara, Turkey. He has more than 30 publications in major international conferences, journals and is the co-editor of the recent book “Modern Software Engineering Concepts and Practices”.

Vanessa Lopez joined IBM Research Ireland as a research engineer in January 2012. Prior to joining IBM, she was a research associate at the Knowledge Media Institute (The Open University, UK). She has been working on the topic of searching and querying heterogeneous semantic data for over eight years, in particular on question answering, ranking and merging. She has published her work in more than 30 conferences and journals. Her current research interests are to investigate how Linked Data technologies can be used to interpret, capture and integrate live and open data from cities.