

Describing the space and place of digital cities through volunteered geographic information

Ross Purves¹, Jason Dykes², Alistair Edwardes¹, Livia Hollenstein¹, David Mueller¹ and Jo Wood²

¹Department of Geography, University of Zurich

²School of Information Science, City University

Historically data describing cities have been the preserve of a privileged group of data producers and have therefore typically contained single perspectives on the nature of the cities described, reflecting the particular skills of those responsible for data collection as well as the intended purpose of the data, primarily representing cities as containers of quantitative data. Thus, datasets describing administrative regions typically represent these as having clear, well defined borders, despite a wealth of research showing that people don't typically perceive city spaces in such a way. The advent of large volumes of so-called Volunteered Geographic Information (VGI) produced by anyone with the time and interest to contribute to both moderated and unmoderated collections provides seemingly endless opportunities for exploration of the place and space represented digitally as cities.

In this paper we will present firstly a set of opportunities in the analysis and visualisation of city spaces through VGI, before discussing some of the challenges and threats raised by uninformed and oversimplified portrayal and interpretation of such data. We will illustrate how visualisation can be a powerful tool in not only presenting the results of such analysis, but in exploring the properties of digital cities.

Our experiences have primarily been gathered in the analysis of two large datasets of georeferenced images with, in one case, associated descriptions in the form of sentences, and in the second case, tags with no explicit semantic relationships. We will describe 3 applications of these volunteered data to the description of digital cities:

- 1) Identifying how facets of place are described through the analysis of text describing images, and an exploration of how these facets of place vary in broader geographic space, for example, what qualities are typically ascribed to cities?
- 2) An investigation of how generic city spaces, such as downtown, are used to label images, and the relationship between such descriptions and other toponyms (Figures 1 & 2).
- 3) Explorations of how the spatial relationship "near" is used in conjunction with descriptions of cities and non-urban areas.

In exploring these questions we will also present results demonstrating how this research can be related to previous theoretical work, and most importantly, discuss issues related to bias and data quality and their impacts on our work. Typically, many researchers assume that large volunteered resources (for example, at the time of writing, Flickr contains around 90 million georeferenced images) are so large that sampling problems and bias can be ignored. We show that large collections of VGI do contain significant biases and propose visual and analytic methods for detecting and accounting for such biases. We will also explore how descriptions of the place and space of digital cities vary not only through such biases, but as a function of different cultural influences within Europe.

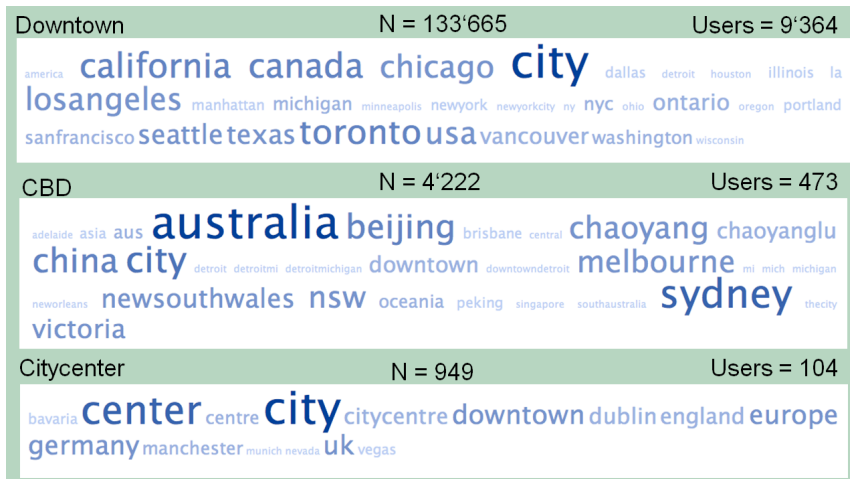


Figure 1: Tag clouds showing specific and generic toponyms co-occurring with generic city core tags “downtown”, “CBD” and “Citycenter” in georeferenced Flickr images (Tag clouds produced using www.tagcrowd.com).

Zurich City

Standard Distance: 1421 m

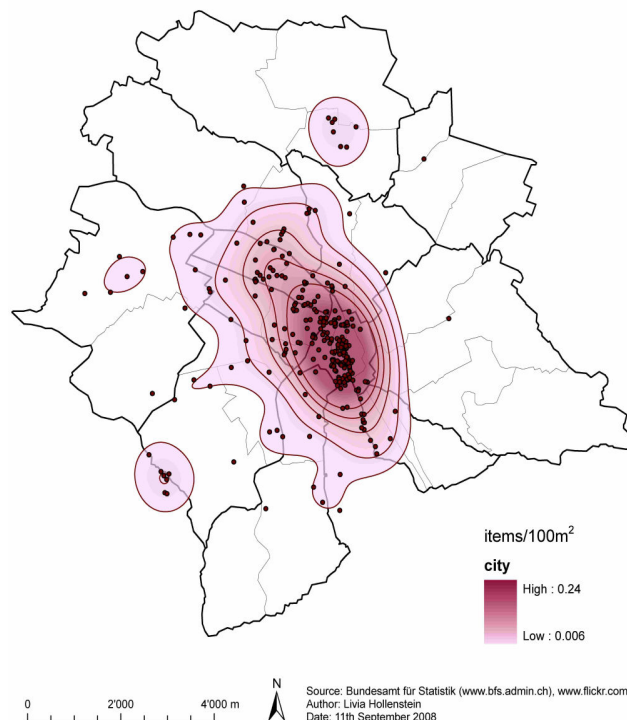


Figure 2: Kernel density representation of Zurich City based on locations of Flickr images. Small cluster to west of image shows a typical bias – here users photograph Zurich City from a local vantage point above the city.