RESEARCH ARTICLE

Exploring place through user-generated content: Using Flickr tags to describe city cores

Livia Hollenstein* and Ross S. Purves

Department of Geography, University of Zurich-Irchel, 8057 Zurich, Switzerland

Abstract: Terms used to describe city centers, such as Downtown, are key concepts in everyday or vernacular language. Here, we explore such language by harvesting georeferenced and tagged metadata associated with 8 million Flickr images and thus consider how large numbers of people name city core areas. The nature of errors and imprecision in tagging and georeferencing are quantified, and automatically generated precision measures appear to mirror errors in the positioning of images. Users seek to ascribe appropriate semantics to images, though bulk-uploading and bulk-tagging may introduce bias. Between 0.5–2% of tags associated with georeferenced images analyzed describe city core areas generically, while 70% of all georeferenced images analyzed include specific place name tags, with place names at the granularity of city names being by far the most common. Using Flickr metadata, it is possible not only to describe the use of the term Downtown across the USA, but also to explore the borders of city center neighborhoods at the level of individual cities, whilst accounting for bias by the use of tag profiles.

Keywords: user-generated content, city core, vernacular geography, gazetteers, georeferenced, tagging, Flickr

1 Introduction

Cities and their centers are concepts with which most people are familiar. But where is the city center? If we wish to find a *city center hotel* in London, how can we formulate such a request through an information system, and if a friend is staying in another city center

---

hotel, will we find ourselves nearby? Is the city center the same place for a resident of London as for a tourist, and do both describe it using similar language? Such questions, which relate to how we name and describe space, are at the core of the study of *vernacular geography* and form the central area of study addressed in this paper.

Vernacular geography encapsulates the spatial knowledge that we use to conceptualize and communicate about space on a day-to-day basis. Importantly, it deals with regions which are typically not represented in formal administrative gazetteers and which are often considered to be vague [19]. In turn, this means that vernacular geography can be seen as an element of naïve geography which is defined by Egenhofer and Mark [7] as "the body of knowledge that people have about the surrounding geographic world." Since vernacular geography concentrates on how people name and delimit space in everyday use, dealing with it is a key component of any attempt to develop information systems which replicate typical cognitive models according to non-expert users' notions of space, rather than the traditionally administrative geography represented in current gazetteers and geographic information systems (GIS). Representations of such notions are becoming increasingly important as the number of information systems based on querying by place name increases, where knowledge by users of administrative place names cannot be presupposed, for example in geographic information retrieval (GIR) [20].

Exploring commonsense notions requires that we have access to descriptions which somehow reflect vernacular usage. One data source with considerable potential in GI-Science is so called *user generated content* (UGC) or, perhaps more specifically, *volunteered geographic information* (VGI) [15]. Both UGC and VGI are essentially data uploaded to the web by individuals, such as YouTube videos, Open Street Map data, or Flickr images and tags. References to location may be stored as (potentially) unambiguous coordinates or references to place names in the form of, for example, tags or titles, if users feel that location is a useful (or practical) way of describing such an information object.

In this study we draw on the abundance of georeferenced photographs in Flickr (http://www.flickr.com), an online photo library, which links absolute, explicit references in the form of (sometimes imprecise and inaccurate) coordinates to user-assigned descriptions of these photographs in the form of tags. The central aim of our study is to explore whether and how we can use the tags associated with Flickr images as a *proxy* for empirical data showing how humans *name* places in order to assist current attempts aimed at imitating this behavior by computer systems.

The remainder of the paper is structured as follows. We first discuss some previous work on place and vernacular regions, before briefly reviewing existing research on the use of tagged data and setting out open questions arising from the literature. We then describe the data and methods used in this research, before presenting and discussing the results of our analysis. The paper is concluded by exploring the extent to which the research questions set out below have been addressed, and by identifying areas of ongoing and future work.

2   Previous work

2.1   Vernacular geography and place

Vagueness is inherent to the way humans conceive and refer to geographic locations. Rather than exact distances and coordinates, we conceptualize geographic space in terms

of hierarchically structured and often ill-defined places. We employ vague spatial relations such as *near Central Station* and talk about places such as *London* without being concerned about their exact boundaries, whilst other place names may correspond to regions which either lack, or have inconsistent legal definitions, such as *Downtown* or the *West End* [21]. This vernacular geography comprises a complex set of places at various scales. Places are a persistent, but not static component of every culture and society, providing a central metaphor in the way humans communicate geographic information [6, 19, 33]. By contrast to typical computational descriptions of locations as a position defined by coordinates, a human centered perspective of place considers it to be a shared frame of reference, corresponding to a collective conception of regions and associated names, typically with vague extents [27]. The vagueness inherent in the human conceptualization of location is based on the quality and limitations of spatial knowledge as well as in the continuous nature of geographic entities. Typically, there will be locations that are clearly agreed upon as being part of a place, perhaps situated towards the center of an associated region. However, there will often be uncertainty and disagreement on locations towards the fringe that might be less characteristic instances of that place.

Geographic regions of urban space, for instance, seem to have collective definitions, which are based on shared conventions in the conceptualizations of residents [9]. Campari [4] showed how the uncertainty of urban boundaries arises from a combination of factors such as administrative, religious, social, and physical artifacts, of which many are per se uncertain. The boundaries of administrative units typically do not physically manifest themselves in urban space, but their existence nonetheless alters the conception of the city. At the same time, administrative artifacts are overlain by a variety of other factors, such as the built structure, land use, social (in)homogeneity, population density, and housing systems. The interplay and influence of these factors goes some way to explaining why residents' designations of their neighborhood may differ fundamentally from administrative definitions, and that urban neighborhoods and their configuration are in constant transition [4]. Such neighborhoods, which can be considered to be sub-places making up the place which is a city, were identified by Davies [6] as a key place type for the end users of mapping products, thus suggesting that there is a real need to develop methods to define and explore such regions.

Approaches to describing urban areas, and the city center in particular, mirror developments in the discipline of geography as a whole and encompass quantitative approaches based around intensity of land use (e.g., the classic concentric zone model from Burgess) or through indices such as the *skyscraper index* [28]. More qualitative behavioral approaches can be seen as counterpoints to such normative assumptions, whereby the underlying cognitive approach is based on the notion that, "... in a sense the city is what people think it is ..." [24]. In his groundbreaking work in this area Lynch [25] made clear that the vagueness or otherwise of boundaries in urban space could be related to their physical manifestation in an environment, that is to say whether the boundaries between neighborhoods are fiat or bona fide [35].

Work to define such vernacular regions in general has taken a variety of forms at a variety of scales. For example, in empirical studies, subjects have been asked to indicate whether locations are found in a particular region and to delineate such regions on a map [27]. Other work has explored how locations are described in texts, typically in relation to other named places whose location is well defined [3]. Recent approaches have explored the notion that vernacular place names are likely to be used in association with

place names listed in gazetteers, and attempted to retrieve such locations from the web, before delineating vernacular regions [21]. These methods suffer from the disadvantage that the identification of references with place names in text is subject to multiple forms of ambiguity (e.g., does Bath refer to a place to bathe or a town in England, and which Springfield is which?).

The advent of large volumes of tagged, georeferenced digital objects provides a new data source for such research. These collections have rapidly grown, with, for example, Flickr containing at the time of writing some 98 million georeferenced images. Thus, researchers have started to explore the use of georeferenced digital objects to delineate vernacular regions, and as a method to suggest potential place names or tags given a set of coordinates (e.g., [1, 16, 23]).

However, to date most research has concentrated on attempting to derive the borders of given vernacular regions, assuming that the names of these regions are known. Furthermore, although user needs studies such as [6] suggest that vernacular place names are an important need in products provided by National Mapping Agencies, we have little evidence for the extent to which vernacular names are actually used in UGC.

2.2 Tagging behavior

Tagging is a classic way for users to annotate UGC, for example in the form of images in Flickr or videos on YouTube. The motivation behind tagging is generally thought to be twofold; apart from the organization of content for personal means, users of tagging systems are driven by the idea of social contribution and the desire to share with others [2]. The vocabulary system in tagging systems is entirely flat and has been claimed to directly reflect the conceptual and linguistic structure of the users and their diverse geographical and cultural backgrounds [39]. The potential and drawbacks of content in tagging systems in terms of information organization and retrieval [14, 39], the generation of ontologies [32], and its suitability to represent the perception of the individual [17], as well as *distributed knowledge* or the *wisdom of the crowd* [36, 38] have been discussed at length. Here, we briefly review work focussing on the spatial component of tagged online image collections.

In a large scale analysis of users' tagging behavior and the information provided through tags, Sigurbörnsson and van Zwol [34] found that 28% of the tags in a random set of 52 million photos from Flickr corresponded to the location type of WordNet [8] categories. A more detailed analysis of geographic tags was presented by Winget [39] who checked the reliability of tags related to images showing volcanoes. The degree to which Flickr users include the hierarchical structure of place name descriptions was measured by comparing the annotations against the Getty Thesaurus of Geographic Names (TGN). Nearly all granularities of geographic terms, ranging from continents to the names of individual volcanoes as well as alternate names, occurred in the Flickr sample. Since data sets such as Flickr appear to be, at least at the moment, dominated by images taken in urban regions [5], then the question arises as to how people describe and name regions within cities through tags in everyday use.

Girardin et al. [13] explored how tourists and residents of Rome left *digital traces* in Flickr through georeferenced photos. The authors identified locations of tourist activity in Rome and compared the spatial distribution of tag semantics, for instance for "ruins," to the actual cityscape. Grothe et al. [16] derived density surfaces from predefined Flickr tags for large-scale geographic vernacular regions such as the Alps, the Black Forest, or

the Rocky Mountains, while other authors have made initial investigations of the use of neighborhood terms in Flickr [23]. Finally, Flickr itself reported on the use of tagged images to derive neighborhood boundaries using alpha shapes[1], which may then be used in automated addition of so called machine-generated tags. This work also points to an important problem in research in this area. Flickr's intention in deriving neighborhood boundaries appears to be to suggest names to users who upload georeferenced images; but of course such machine-generated tags may alter the way in which *individuals* choose to name places in their own tag lists.

Other work with georeferenced tags in Flickr has extracted information by data-driven approaches not dependent on predefined lists of landmarks and places or a manual classification of tags. Rattenbury et al. [30], for instance, presented two automated techniques to successfully extract tags corresponding to place names. Ahern et al. [1] identified representative tags for geographic regions by a *k-means* clustering algorithm. Highly ranked keywords, typically corresponding to place names or landmarks, are displayed on a tag map, referred to as aggregated "psychological map" by the authors. However, to our knowledge, current research has not explored in detail how vernacular names are used within Flickr tags, and explorations of the precision and accuracy of geotagging with respect to place name tags have been limited.

2.3   Research gaps

All of this work suggests that exploring how cities are described through georeferenced, tagged objects may provide an excellent opportunity to explore the potential of UGC in understanding how individuals name regions within cities. Furthermore, the cores of cities (Downtown, City Center, etc.) are seen as classic examples of vague regions, with Thurstain-Godwin and Unwin [37] describing city centers as "... almost archetypal examples of geographic objects with indeterminate boundaries ...", and since Flickr appears to have many more images in urban areas, we investigate in this study tagging of images in urban regions. Although previous studies have explored the nature of tags, and the granularity of place names, these studies have not i) explored the accuracy and precision of georeferencing and tagging; ii) investigated the prominence of vernacular place names in general; nor iii) the use of *generic vernacular names*, as opposed to specific (that is to say the use of terms such as *Downtown* or *City Center* as opposed to *Belltown* or *Aldgate*). These research gaps suggest, in turn, three more specific research questions:

1. How reliable are user-contributed tags as a means of describing geographic space, and how can we deal with issues of precision, accuracy, and bias?
2. How is urban space described using tags, and in particular how are city centers described? How prominent are vernacular place names in such descriptions?
3. How can georeferenced tagged images be used to gain knowledge of the collective understanding of the location and extent of vernacular regions?

Our study focuses on tags, since we are particularly interested in how users wish to describe objects in a flat vocabulary, which is well suited to retrieval with a minimum of processing (unlike, for example, titles or comments) and tagging appears to be a particularly popular way of describing information objects.

---

[1]http://code.flickr.com/blog/2008/10/30/the-shape-of-alpha/

3   Working with Flickr images

3.1   Description of data

In Flickr, images are stored with a wide range of information (i.e., contributing user, image metadata, and time of upload) as well as with information which was optionally contributed by the user. This typically includes title, caption, usage restrictions, and a set of tags. Textual tags retrieved from Flickr can be used without further standardization as they are processed as single strings and letter case is ignored by the database[2]. Spatial references are stored in the form of a distinct geotag expressing a location in coordinates as a latitude and longitude. Geotags may be provided by the photo owner either by means of synchronization with tracklogs from an external GPS, cameras and phones with built-in GPS, or by manually locating photos using a map interface. An accuracy level ranging from 1 (world level) to 16 (street level) is automatically assigned to georeferenced photos, depending on the precision represented by the GPS coordinates or the zoom level of the map used to locate an image. Users may choose to georeference either the scene being photographed, or the photographer's location, though where coordinates are collected automatically using GPS the latter is always the case. In the form of neighborhoods, these differences typically do not appear to be significant, but for very salient landmarks or inaccessible objects (e.g., the Eiffel Tower or the north face of the Eiger) this will not be the case.

   The publicly available Flickr API was used to collect metadata for over 8 million images analyzed in this paper. Image metadata were collected between May 02, 2008 and June 27, 2008 but include accessible photos uploaded by users from any preceding date. Six cities with a variety of cultural and linguistic backgrounds were chosen as examples for quantitative and qualitative evaluations of place tag usage: Zurich, London, Sheffield, Chicago, Seattle, and Sydney. Other than capturing data from a variety of regions, the selection was motivated by the intention to explore if, and how, the choice of tags was influenced by the nature and size of the urban environment. Georeferenced image meta-data were collected by specifying a bounding box corresponding to the administrative region (Table 1) for each of the sample cities without any restriction on tag content. After examination of place tags and city core tags in this first data set, a second data set was created by retrieving georeferenced items geotagged anywhere on the globe and matching a tag restriction corresponding to one of the following generic city core tags: "downtown," "central," "cbd," "innercity," "citycentre," or "citycenter" (reflecting British and American English usage of center/center respectively)[3]. A third data set was collected by sampling photos tagged with at least one of a list of five city toponyms associated with one of three English-speaking regions (the UK, USA, or Australia) but with no restriction with respect to the location or existence of a geotag. Hence, this data set contains both geotagged and non-geotagged photos, and since images were retrieved only using place name tags, these may be ambiguous. Table 1 shows summary statistics for all three data sets.

---

[2]http://www.flickr.com/services/api/misc.tags.html
[3]We present all tags used in searches or scraped from Flickr in lower case and in quotes: "tag."

Data set 1: Bounding boxes		
City bounding box	*Retrieved photos*	*Average tags/photo*
Zurich [8.46E,47.34N–8.58E,47.42N]	47 005	5.4
London [0.63W,51.24N–0.39E,51.73N]	1 061 883	5.4
Sheffield [1.53W,53.36N–1.40W,53.42N]	21 124	5.6
Chicago [87.96W,41.60N–87.45W,42.11N]	389 703	4.9
Seattle [122.47W,47.52N–122.24W,47.75N]	313 796	4.9
Sydney [151.13E,33.94S–151.26E,33.85S]	139 542	5.2
Total/average	1973053	5.2
Data set 2: Global		
Georeferenced tag	*Retrieved photos*	*Average tags/photo*
"downtown"	133 655	11.8
"central"	29 042	15.5
"cbd"	4422	15.4
"innercity"	2359	20.7
"citycentre"	2057	15.7
Total/average	171 535	15.2
Data set 3: Regions		
Tags	*Retrieved photos*	*Average tags/photo*
Great Britain[a]	859 176	8.6
United States[b]	4 140 602	6.0
Australia[c]	1 682 648	6.4
Total/average	6 682 426	7.0

[a] search tags: birmingham AND uk, liverpool AND england, glasgow AND scotland, edinburgh AND scotland, london AND england
[b] search tags: chicago, seattle, boston, miami, houston
[c] search tags: brisbane, sydney, perth, melbourne, adelaide

Table 1: Summary of data retrieved from Flickr. The terms used in retrieving images for data set 3 are shown in a, b, and c. Note that because of greater place name ambiguity, UK place names were collected using both country and city names.

3.2 Characteristics of user-contributed metadata

An initial step in the analysis of any collection of UGC should be to explore data quality. Since our study focuses on location, an evaluation of Flickr data with a special focus on the quality of the geotags was carried out. To gain an impression of user behavior during the creation of spatially relevant metadata, the georeferenced data sets (data set 1) retrieved from Flickr were first investigated with respect to the precision applied in geotagging photos. The diagram in Figure 1 reflects the cumulative frequency of geotag precision levels found in these large samples. The majority of images were found to be associated with relatively high precision varying between street (level 16) and city precision (level 12) with decreasing frequencies towards coarse granularities. An exception is the Sydney sample, with a smaller proportion of detailed geotags. Variation in the application of locational resolution in data from different cities has been observed before [12]. In our case, the deviation for the Sydney sample is simply explained by a difference in the level of detail offered by the backdrop mapping. For maps in the Sydney area, zoom levels more accurate than level 12 were not available at the time of retrieval. Geotags with higher resolution

had to be automatically generated through GPS, or by employing the satellite image interface, which was not the default mapping displayed by the georeferencing interface. Since such changes to interfaces are typically not formally documented, it is important to treat apparent differences with care, as they may reflect changes in systems rather than perceived differences in user behavior.
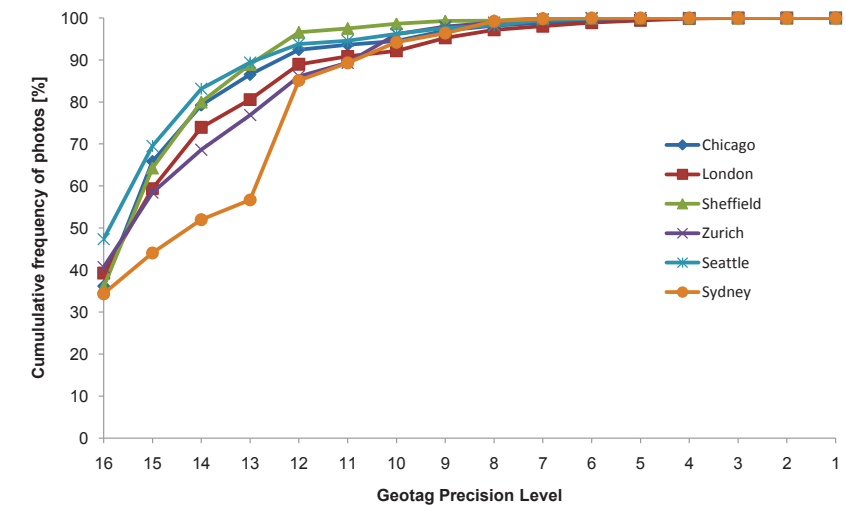

Figure 1: Geotag precision for six city bounding boxes

Having established that users appear to *intend* to precisely geotag images, we next explored how *accurately* and *reliably* users tagged georeferenced images. For this analysis, we used Hyde and Regent's Park in London as exemplars, since these are urban objects with relatively well-defined, unambiguous boundaries. However, in the case of Hyde Park many users seem not to distinguish between the two contiguous regions of Hyde Park and Kensington Gardens (the western part of the park which technically, has been considered separate from Hyde Park since 1728) the entire area was classed as belonging to Hyde Park in the following evaluation. Incidentally, this is a classic example of the administrative and vernacular uses of a place name mismatching, with the potential for confusion.

Figure 2 shows the location of all images tagged with "hydepark" and "regentspark" within London. In the case of Hyde Park, the locations fit well to the shape of the park, with local clusters near Speakers' Corner in the southeast and around the Serpentine— both significant as locations where salient activities or objects can be found. In total, 8398 of 9775 (86%) instances were located either within Hyde Park or on adjacent main roads surrounding the park. The correctly placed instances were uploaded with precision levels of between 10 and 16 with a mean of 14.1 and a standard deviation of 1.4. The precision of the outliers ranges from 3 to 16 with a mean of 11.1 and standard deviation of 3.5. In the case of Regent's Park 628 of 755 (83%) instances are found within the park, with a number of outliers found nearby, for example on Primrose Hill. The correctly placed instances were uploaded with precision levels of between 11 and 16 with a mean of 14.5 and a standard deviation of 1.3. The precision of the outliers ranges from 6 to 16 with a mean of 14.1 and standard deviation of 2.8. These summary statistics suggest in turn that unreliable geotags

appear to be identifiable, at least implicitly, since they are accorded less precision in the georeferencing process.

By exploring Figure 2 we can also identify possible reasons for errors in tagging. Some images have clearly been placed towards the geographical center of London (low precision), whilst others are placed in nearby parks (blunders). For Hyde Park 42% of the outliers stem from only three contributors who posted 196, 184, and 172 items, respectively. One user has obviously automated the geotagging process, as all 196 pictures from a London photoset are placed in the correct location, but have identical tags. The users contributing 172 and 184 photos placed their pictures in the wrong parks—namely Hampstead Heath and St. James's Park. This basic error analysis reveals the influence small numbers of users can have on seemingly very large data sets, especially through processes such as the automation of tagging.

In order to scrutinize the reasons for inaccurate data in more detail, 100 random outliers posted by 100 different users and tagged with "hydepark" were analyzed by hand. The analysis revealed three major reasons for the anomaly between location and tag:

- The image shows the entity tagged, but was taken from a location well outside it. In the case of Hyde Park, seven outlying pictures were taken from aircraft but did show the park, which cannot be considered as incorrectly tagged data.
- The photos are tagged correctly but are misplaced during the georeferencing process. This was the case for 69 items, where 17 were placed in another park.
- The tag choice is apparently incoherent. 22 items tagged "hydepark" did not have an obvious relation to the location.

In the third case, users appear to have added the same taglist to a whole set of photos uploaded simultaneously. Overall, the relationship between misplaced and mistagged items suggests that most users take tagging seriously, but not all of them are willing or able to correctly locate images on a map when georeferencing—in other words people are better at describing what they have seen in terms of semantics than they are at assigning an accurate georeference. Nonetheless, this first assessment allows us to be optimistic about the quality of textual and formal place name tags in user-generated metadata. Furthermore, the specification of the location provided by users seems sufficiently accurate to assist in the investigation of smaller scale geographic places such as urban neighborhoods. It is also clear from this analysis that individual users may make bulk uploads of images with either identical tags or identical locations or both. Since our approach focuses on selecting points for analysis on the basis of tags, we present in Section 3.3.1 a method to explore the influence of prolific users on the metadata pattern of individual tags.

3.3 Analyzing place tags

The examination of quantitative aspects of the use of place names in tagging of images is aimed at gaining insights into how people describe urban spaces. To explore this issue, a list of individual tags and counts was generated for the bounding boxes representing each of the cities under examination (data set 1 in Table 1). Given the unlimited possibilities to encode place information in terms of tags, a manual classification method was employed to identify tags designating a place. For this study, the following conventions were adopted for the identification of place tags:
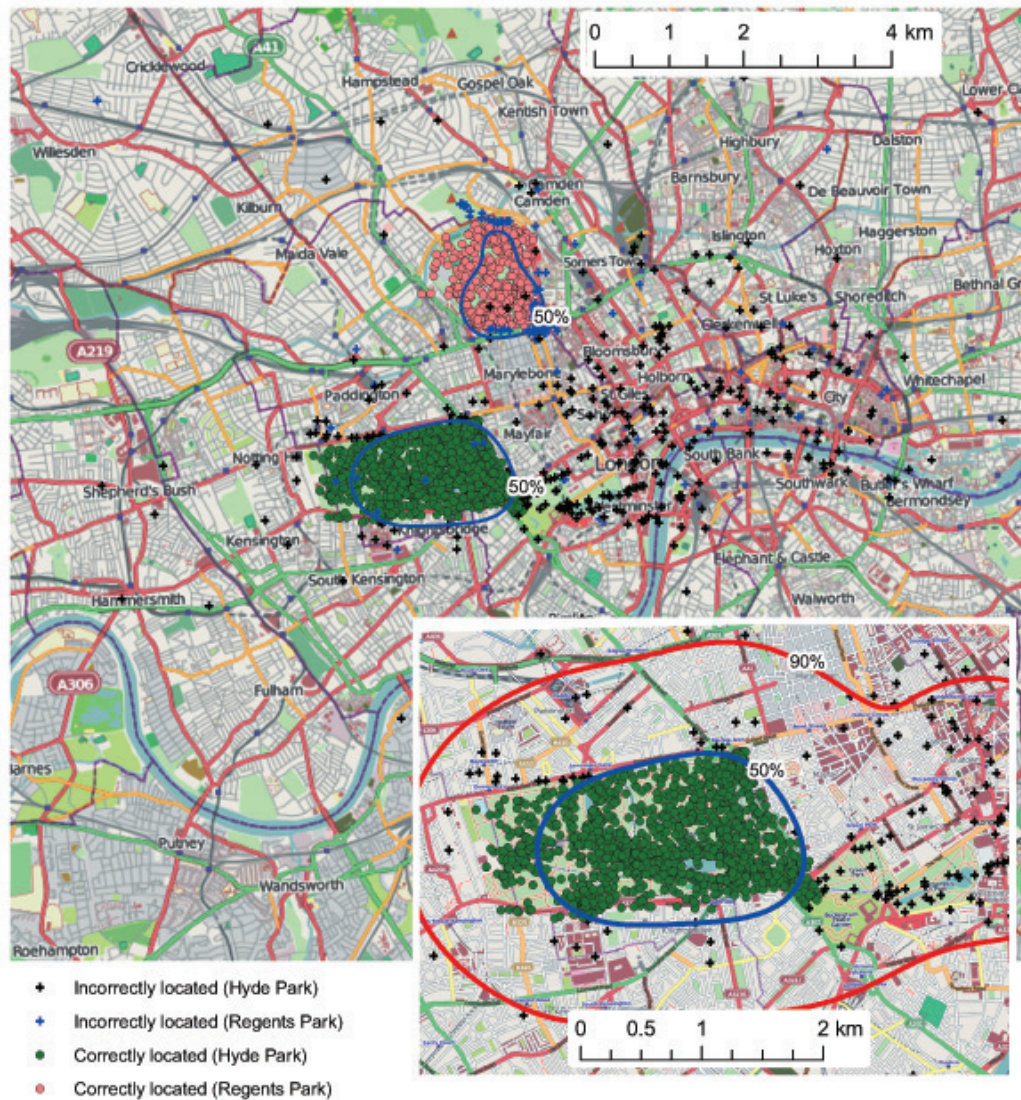
Figure 2: Accuracy of geotagging for images tagged with "hydepark" and "regentspark." Contours show 50% (and for detailed inset 90%) volume surfaces for kernel density surfaces derived from all points as described in Section 3.3.2. Background mapping open-streetmap.org licensed under a Creative Commons Attribution-Share Alike License.

- Names representing places which belong to the hierarchy of continents-states-counties-cities (additionally districts and streets for Zurich) and descriptive terms of location such as city, city center, and neighborhood were regarded as place tags. Place names not related to the city under consideration (e.g., "vienna" in the bounding box

of Zurich) and indications in the form of coordinates (e.g., "geo:lat=47.3722") were
not counted.

- Landmarks and geographic features such as parks, lakes, airports, and locative
  adjectives (e.g., "british") or institutions, buildings, and events (e.g., "university-
  ofchicago") were not considered as place names.
- Interpretable misspellings were included and counted with the correctly spelt version
  of the place, since user intention is relevant [11].
- Different languages as well as compound expressions (e.g., "zurich," "zürich," and
  "zurich2007") were counted and the compound tags counted with the instances of
  the corresponding simple place name tag.
- Compound tags were allocated to all granularity levels represented in the tag, e.g.,
  "bahnhofstrassezurich" was considered to belong to the street and the city level.

Accordingly, the tags for each city in data set 1 were first divided into non-place name
tags and candidate place name tags which were checked against the Geonames gazetteer[4]
and other internet resources. Each identified place tag was labelled according to one of
four granularity levels: continent, country/state, city, or city center concept. The relative
frequency of occurrence with respect to the total sum of tags within a bounding box was
calculated for all identified place name tags.

3.3.1 Measuring the popularity of tags

As the first assessment of georeferenced tags using the example of Hyde Park had shown
that single users may introduce major distortions into the metadata pattern, simple global
tag frequencies were not, alone, considered representative. A method to check for possible
effects of bias resulting from both prolific users (delivering very large numbers of images
within a particular data set) and unprolific users (posting only a few images within the
same data set) was adapted from ongoing work (Jo Wood, pers. comm.). In this method,
the collective popularity of distinct tags within a data set is examined through construction
of tag profile histograms (Figure 3). To construct a tag profile, all tagged photos in a
data set are sorted according to the number of contributions per user, with the most
prolific contributors to the left and decreasingly active contributors to the right of the
histogram. The associated tags are thus divided into bins, where each bin corresponds to
one-hundredth of the total number of photographs in the data set. For each of these bins,
the absolute count of images with a particular tag is then displayed in the histogram. The
histogram represents the ubiquitousness of a tag with respect to the underlying population
of photographs in a region as a whole, and with respect to the number of users who have
contributed that tag.

In order to compare patterns of contribution for different tags and for tags between
different data sets, standard or z-score values are computed to normalize the counts of a
tag per bin. An overall measure of whether a tag is employed with equal frequency among
active and unprolific users can be expressed by the coefficient of variation, defined as the
ratio of the standard deviation to the mean of the bins with respect to the given tag.

Figure 3 shows profiles for tags in London's bounding box: firstly, "london" which is
equally distributed among active and unprolific users and, secondly, "innercity" that is
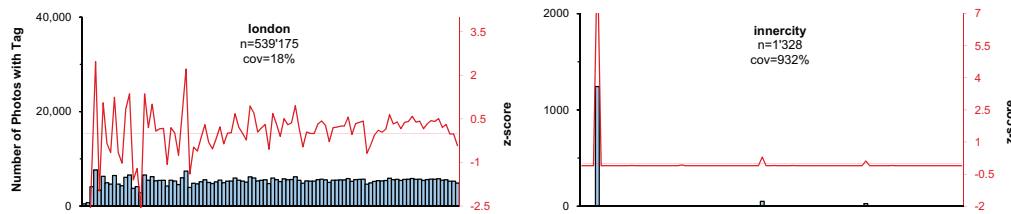predominantly used by prolific users (indeed, the tag is almost certainly used by a *single*

---

[4]http://www.geonames.org/

Figure 3: Tag profiles for "london" and "innercity" (London in data set 1) showing absolute tag counts and associated z-scores. The z-scores are indicated by lines; the histogram shows absolute number of images with this tag ranked by contributor, most prolific contributors to the left.

prolific user), together with the resulting z-scores and coefficients of variation (cov). In general, tags (whether frequent or not) used equally across a population of users, show an erratic pattern and a low coefficient of variation and can be considered ubiquitous with respect to user behavior, while those with high coefficients of variation are indicative of bias. We thus use coefficient of variation, together with visual inspection of histograms as a useful way of exploring tagging behavior and one useful indicator of potential bias.

3.3.2   Visualization of spatial tag distribution

To conceptualize the regions collectively demarcated by multiple users for different place descriptions, vague footprints were derived from the distribution of georeferenced tags associated with vernacular terminology using kernel density estimation (KDE) [29]. KDE methods produce representations of local density estimates from two-dimensional point distributions. The density value is estimated at each observed point by spreading the search radius by a kernel function with defined bandwidth. The critical aspect is the selection of a range of parameters that influence the resulting surface. The choice of the kernel bandwidth, also termed smoothing parameter, and in our case the choice of the threshold value to exclude possible outliers, strongly affect the resulting surfaces and delineations. There is no general agreement on how to approach this problem and previous efforts to approximate vague regions by density estimators have addressed the problem of the kernel parameters in various ways. As the studies were mostly concerned with small numbers of regions, authors have experimentally determined the bandwidth (e.g., [21, 37]) and the threshold point density to generate sharp boundaries is often defined interactively [21]. Data-driven approaches were presented by Henrich and Lüdecke [18] and Grothe and Schaab [16]. Due to the large number of data sets to be processed in this work an experimental choice of kernel parameters was not considered appropriate. Rather, we chose an objective (in so far as it is data-driven), automated choice of the smoothing parameter as implemented in the home range tools (HRT) [31] application for ArcGIS. Gaussian kernel density surfaces were calculated using a standard distribution h_{ref} as the bandwidth parameter, calculated from the mean variance in the x- and y-coordinates, where n is the number of points, as follows:

$$h_{ref} = n^{-1/6}\sqrt{\frac{var_x + var_y}{2}}$$

This method is appropriate if the underlying point pattern is unimodal which is typically the case for georeferenced tags referring to places at the sub-city level. To minimize the effects of bulk uploads, internal clusters, and errors in the Flickr data, all x- and y-multiples and all items geotagged with a precision lower than 9 were removed in a pre-processing step. The outlier problem was addressed by means of volume contours with surfaces thresholded at 90% of volume. Testing of the method using public parks in London revealed that bandwidths so generated are, despite outliers, almost identical to search radii established empirically (for example, 270m for Regent's Park). The peak regions of the density surfaces (defined as the 50% contour) appropriately represent the extent of these well-defined entities (see Figure 2 for examples) and provide an appropriate explorative means to explore the collective definition of vernacular regions in cities.

4 Results and interpretation

4.1 City center concepts at a global level

Table 2 shows the number of images retrieved with tags related to city centers, together with the number of individual users contributing and the proportion of terms contributed by the most prolific users. For example, the most prolific 10% of users contributed 70% of the "downtown" tags and "cbd" tags, whilst a single user is responsible for 61% of the "innercity" tags. Thus, Table 2 suggests that tags describing city center concepts, with the exception of "downtown," are used by a relatively small proportion of users. The preliminary analysis suggests that some city center expressions, such as inner-city, are rarely employed in colloquial language. Nonetheless, tags are common, with for example "central" being used some 29 000 times by around 3000 different users—an order of magnitude more data than collected in traditional survey-based empirical work.

City-center concept	Number of geotagged photos worldwide	Contributing users	Portion of photos by 10% most prolific users
"downtown"	133 655	9364	69.9%
"central"	29 042	3196	71.6%
"cbd"	4422	473	69.7%
"innercity"	2359	104	82.6%
"citycentre"	2057	300	57.1%
"citycenter"	949	199	52.5%

Table 2: Georeferenced images associated with city core tags retrieved from the whole world

The images from all three data sets were also analyzed with respect to the total number of tags assigned. As shown in Figure 4, tag frequencies are very different for images collected using bounding boxes (data set 1, Figure 4a), city center concepts (data set 2, Figure 4b), and city toponyms (data set 3, Figure 4c)

In Figure 4a a sizeable proportion of images have no tags—since these images are retrieved using a bounding box this is possible. The distributions are long tail, with most

images having a relatively small number of tags and the sample sizes appear to be large enough to generate smooth distributions. This is also the case for "downtown" in Figure 4b. However, for "cbd" a relatively small number of users have contributed images so tagged, and the distribution is correspondingly noisy. Notable in Figure 4b are the right-shifted tag distributions, with mean tag counts of 15.2 compared to means of 7.2 and 5.9 for tags retrieved using bounding boxes and city toponyms (Figure 4c) respectively. Thus, generic city center tags appear to mostly be applied in the context of longer-than-average tag lists, implying that users using such terms are more descriptive in general. Figure 4c shows tag frequencies for three regions generated by searching with city toponyms, which have long tail distributions similar to those in Figure 4a.
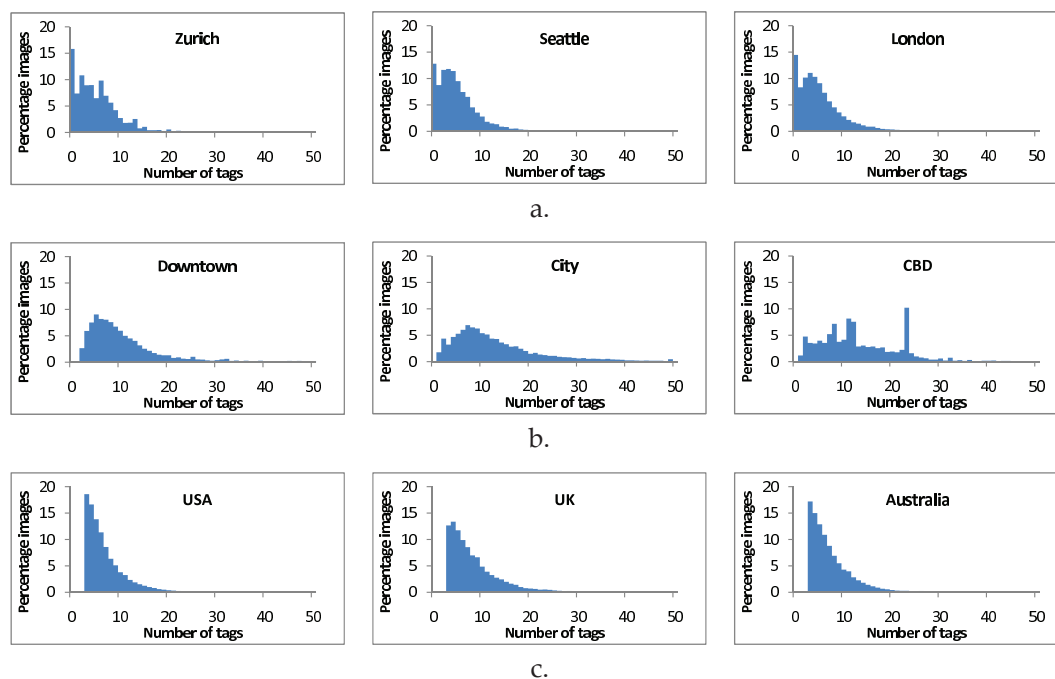

a.

b.

c.

Figure 4: Tag frequencies expressed as percentages for the 3 data sets collected. Note that in data sets 2 and 3 tags are used as search terms, so at least one tag must exist.

Table 3 shows counts of individual tags referring to city core concepts associated with georeferenced images ranked by the countries in which they occurred (data set 2). Once, again, the relative rarity of tags other than "downtown" is clear. The prominence of the United States in three of the four lists (all except the British spelling of "citycentre") is notable, and demonstrates the geographic bias of Flickr as a predominantly US data set, at least at the time of writing. "cbd" is popular in Australia and "downtown" is clearly a predominantly North American term.

Given the predominance of "downtown" in North America, we investigated its spatial distribution with respect to georeferenced images by mapping the density of images tagged with "downtown" in the USA (Figure 5). We first produced a simple map of all occurrences of images tagged with "downtown" in the USA, and represented this as a density surface.

"citycentre"	"citycenter"	"cbd"	"downtown"
United Kingdom (1444)	United States (204)	Australia (2325)	United States (94 070)
Ireland (120)	United Kingdom (136)	China (735)	Canada (17 684)
Netherlands (110)	Germany (96)	United States (500)	Brazil (1652)
Italy (51)	Ireland (72)	Canada (179)	Mexico (793)
Poland (31)	Switzerland (38)	South Africa (163)	Germany (385)

Table 3: Top 5 countries associated with georeferenced images tagged with generic city core terms

This density surface appears to correlate strongly with cities with populations of more than 500 000, which at first glance suggests that "downtown" is a ubiquitous tag in describing North American cities. However, since we have already argued that most Flickr images are urban, the question arises as to whether this is merely a reflection of the overall distribution of Flickr images. To investigate this issue, we generated χ-maps visualizing expected and observed frequencies [40] of images tagged with "downtown" by comparing with a random sample of images with at least one tag drawn from the USA. The χ-maps were produced at resolutions of 100km and 20km, with kernel radii of 150km and 100km respectively. χ was mapped in all cells with more than 3 occurrences of "downtown" as:
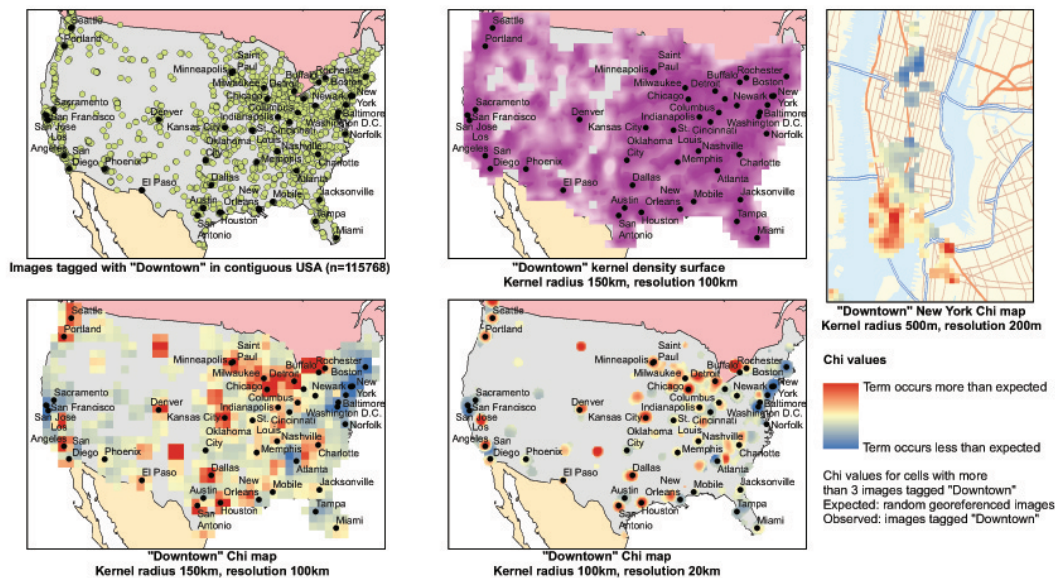
$$\chi = \frac{(obs - exp)}{\sqrt{exp}}$$

Figure 5: Density surface and χ-maps for "downtown" in the USA—cities with populations of more than 500000 shown. Base data © Environmental Systems Research Institute and USGS.

These surfaces allow us to explore qualitatively the distribution of tags with respect to the overall distribution of tagged, georeferenced images in the USA. A number of features are apparent. Firstly, both the 20km and 100km surfaces show similar patterns in general corresponding to the locations of large cities. However, the 100km surface provides a much clearer overview, showing that Downtown appears to be ubiquitous on both the west coast and in the Midwest. By contrast, Downtown appears to be more rarely used on the east coast at this scale. Of course the origin of the term Downtown is in New York, and to investigate this, at first glance surprising, result, we generated a further χ-map in Manhattan with a resolution of 200m. Here we see that Downtown is indeed used more than might be expected in the area traditionally considered to be Downtown as well as at a location in Brooklyn. We assume that this is a location where many pictures of the Downtown skyline are taken. Importantly, these χ-surfaces do not show where high densities of "downtown" tag are found (other than that at least 3 "downtown" tags must occur before we calculate a χ value) but rather locations where the number of images tagged with "downtown" cannot simply be argued to be a function of the underlying distribution. This analysis allows us to explore the use of terms at broad scale within the USA, which appear to be interesting and worthy of further exploration.

4.2   Nature and granularity of place descriptions in tags

To better understand how toponyms and generic city center concepts were used with geo-tagged photos, we analyzed the frequencies of such tags with respect to images collected using city bounding boxes (data set 1, Table 4). In the case of Zurich, where we have good local knowledge, we categorized all identified place name tags as continental, country, canton, city, district, or street level respectively. For the bounding boxes of Anglo-Saxon cities, the 1000 most frequent tags were analyzed and only generic city center concepts were considered at the subcity level of granularity—in other words we did not attempt to identify named neighborhoods or streets within these cities, since we believe detailed local knowledge is required to carry out this task. Even with the advantage of local knowledge, the identification and classification of place tags in the entire tag list of Zurich was challenging. Particularly for tags with low frequencies, issues were identified with respect to quality and idiosyncrasy, as is reflected in the long tail distribution with 56.5% of tags being used only once, a typical property of tagged images [26,34]. The situation was further complicated by the numerous languages used, sometimes within a single compound tag, reflecting the influence of tourists and French and Italian speaking parts of Switzerland, as well as the use of both Swiss and Standard German in Switzerland. By contrast, among the 1000 top-ranked tags analyzed for the Anglo-Saxon cities, only a small number of malformed, misspelled, and idiosyncratic keywords occurred. Nonetheless, due to referent class ambiguity (i.e., does Seattle refer to a football team or a city?) and ambiguity at the place granularity level (i.e., does Zurich refer to the city or the canton?), the numerical values in Table 4 should be considered approximate.

In Zurich, the most commonly occurring keywords out of 14046 distinct tags included "zurich" at rank 1, "zürich" at rank 3, and "zuerich" at rank 12, which were all marked by comparably low coefficients of variation indicating minimal or slight user bias (48%, 71%, and 272%). Together with less popular equivalents of the city toponym (e.g., "züri" in Swiss-German), they account for 18.10% of all the tags employed within the bounding box. 14.36% of tags correspond to a place annotation at the country level, while the continent is

Tag granularity	*London*	*Sheffield*	*Chicago*	*Seattle*	*Sydney*	*Zurich*
Continent Country	13.13%	6.23%	6.25%	4.65%	17.42%	0.86% 14.36%
City	17.72%	12.36%	18.98%	18.98%	14.98%	18.10%
City center	0.70%	0.34%	1.28%	1.28%	0.77%	0.50%
District Street	n/a	n/a	n/a	n/a	n/a	1.22% 0.72%
Total	**31.55**%	**18.93**%	**26.51**%	**16.10**%	**33.17**%	**35.12**%

Table 4: Tag granularity for city bounding boxes (note that district and street were only investigated in Zurich where we had sufficiently detailed local knowledge)

in this case a much less important frame of reference (0.86%). The place tags designating any of the districts, neighborhoods, or post code areas of Zurich sum to only 1.2%. The street level, yielding a portion of 0.72% of tags, appears to be considered to be too detailed for the annotation of photographs.

As for Zurich, the dominant tag in the bounding boxes corresponding to the cities of London, Sheffield, Chicago, Seattle, and Sydney is the official city name itself. Generally, the continent/country level is the second most common spatial reference in users' annotations of photos located in urban space (note that the use of America in a vernacular sense to represent the United States of America makes discriminating between continent and country level difficult in some cases). In the bounding box of Sheffield, for instance, "england," "yorkshire," and "uk" were ranked 2nd, 3rd, and 4th respectively, and in Sydney "australia," "nsw," and "newsouthwales" have the corresponding ranks. Overall, an average of 25% of tags for georeferenced photos within cities correspond to place tags. This fraction is consistent with the 28% of location tags reported by Sigurbörnsson and van Zwol [34]. If we include streets and neighborhood names, as within Zurich's bounding box, the proportion of tags related to place names increases to around 35%. Finally, if we analyze not tags, but images, then around 70% have at least one tag referring to a place name. This implies that users do not consider a georeference in itself to be a sufficient reference to place, and that place names are an important part of the vocabulary related to images.

The fraction of generic city center concepts is, by contrast, low in all samples. Since the sample data were retrieved from the entire extent of the cities, the frequencies of tags at the city level and at the neighborhood level designating the city center is of course limited as a fraction of the total area covered. However, georeferenced items within the bounding boxes were typically highly clustered towards the center, which one might typically expect to be the most photographed part of a city.

In Table 5, the generic place name tags, which are potentially used to refer to the city center, are shown in detail for three cities. As was evident from the examples in Table 4, most of these tags are not very frequent. For example, in Zurich, only "city" appears to be ubiquitous among the tags with all others having high coefficients of variation. In the German language area, city might be used to refer to the city center, but also has a more general connotation.

Within London, the identified terms are either general, such as "city," or ambiguous, such as "centre," a problem that is amplified by the inconsistency of term boundaries of many tags occurring within Flickr. The two most explicit terms, "innercity" and "central-london," occur rarely and have very high coefficients of variation. In London there seems

Zurich

Tag	Occurrence	% of tags	Coeff. of variation
"city"	895	0.353%	195%
"stadt"	132	0.051%	487%
"town"	86	0.033%	654%
"altstadt" & "oldtown"	83	0.019%	406%
"centre/center"	59	0.023%	775% / 677%
"citycenter"	38	0.015%	995 %
"downtown"	8	0.003%	550%
Other	12	0.005%	NA
Total	1313	0.502%	NA

London

Tag	Occurrence	% of tags	Coeff. of variation
"city"	17 698	0.54%	149%
"town"	1975	0.06%	318%
"innercity"	1328	0.04%	932%
"centre"	1013	0.03%	324%
"centrallondon"	910	0.03%	467 %
Total	22 924	0.70%	NA

Chicago

Tag	Occurrence	% of tags	Coeff. of variation
"city"	7689	0.66%	141%
"downtown"	6039	0.56%	138%
"town"	407	0.03%	531%
"center"	373	0.03%	176%
Total	14 508	1.28%	NA

Table 5: Occurrence of generic city core tags for Zurich, London, and Chicago

no widely acknowledged consensus on a means to refer to its central area as a whole. We compared the use of generic city center terms with more specific vernacular neighborhood names, identified by use of a variety of sources. For example, 53 vernacular neighborhood names listed on a London travel website occur 26 623 times altogether, yielding a portion of 0.8% of the tags in the bounding box. By contrast, in Chicago, the generic "downtown" is the most common tag used to refer to the central area. It is 11th in the list of tags within Chicago's bounding box, whereas "cbd" occurs only four times within the entire sample. The Chicago specific appellation of the "theloop," which is commonly said to be used to describe the central business district of Chicago, appeared 4466 times and is 32nd in the overall tag list. It appears that users do not use tag locations generically (i.e., "citycentre"), but rather by using a specific name (i.e., "theloop"), even though the former might be considered more useful with respect to the idea of sharing with many users who might not all be familiar with a particular place. The possible exception here is "downtown," which seems to be used as both a specific tag in North American cities, and possibly as a generic description by Flickr users visiting other cities.

4.3   Deriving boundaries using place tags

The use of Flickr tags to determine region boundaries has been investigated by a number of authors (e.g., [16, 23]). However, these studies have mostly concentrated on larger named regions, or neighborhoods whose names are already known (for example [23]). Here, buttressed by our experiments defining regions with clear boundaries such as Hyde and Regent's Park reasonably accurately, we explore how georeferenced tags can be used to define and compare the usage of different city core areas in the Chicago and London: these were identified by our exploration of data set 1, using data-driven density surfaces as described in Section 3.2.2. We discuss the boundaries thus defined and illustrate the ubiquity of tags in describing these generic and specific regions using tag profiles.

Figure 6: Neighborhoods in and around downtown Chicago. Chicago background mapping openstreetmap.org licensed under a Creative Commons Attribution-Share Alike License; neighborhoods from zillow.com licensed under a Creative Commons Attribution-Share Alike License

Figure 6 shows the region delineated as "theloop," "downtown," and "city" by Flickr tags in Chicago along with neighborhood boundaries provided by Zillow. The delineation of The Loop fits reasonably within the boundaries described by the Zillow data, although the region extends to the north of the Chicago river, which was suggested as a bona fide border to the region (e.g., [25]). This appears to be the result of Flickr users photographing the region from the other side of the river at a prominent vantage point, and indicates two of the weaknesses of the approach: firstly, we assume that tags describe image content located at the image coordinates and, secondly, the methods used in region delineation do not take account of sharp boundaries in urban areas, which, though visible in the original point data are smoothed by the method used to generate region borders. The Downtown area is larger, as might be expected, than The Loop, whilst the region described as City extends yet further, and includes an outlier to the west, which again indicates a vantage point from which users take pictures of the city. These differing vague regions suggest that we can indeed start to explore the geography of vernacular usages of city core terms using Flickr images, especially in North America where data volumes are large. Figure 7 shows

the associated tag profiles, all of which have low coefficients of variation, and demonstrate that these tags are used by a wide range of Flickr users (both prolific and non-prolific). Interestingly, the tag profiles also show that the most prolific users in this area do not use these tags in describing their images.
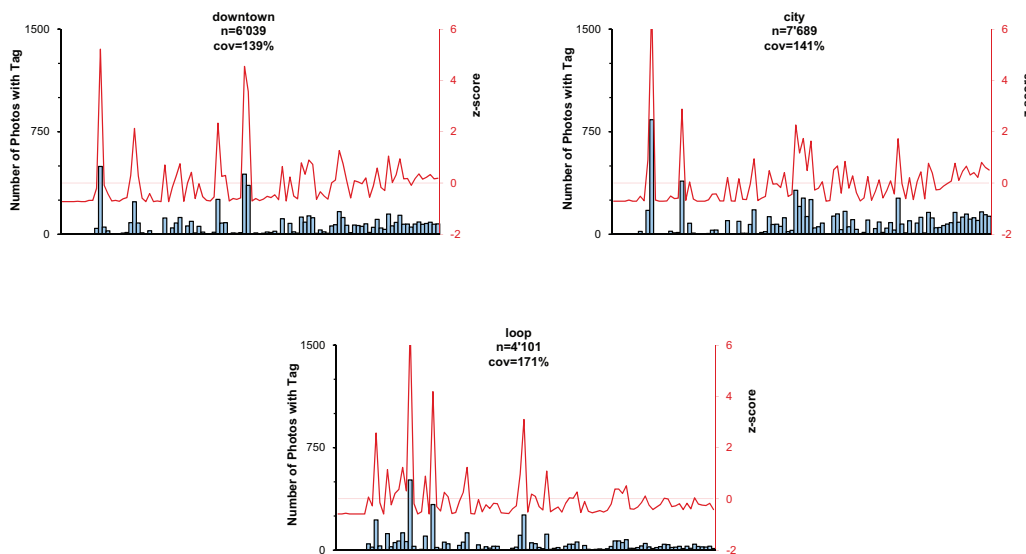

Figure 7: Tag profiles for "downtown," "city," and "theloop" in Chicago

Figure 8 shows three large generic named areas in London, "northlondon," "innercity," and "eastlondon" together with three smaller specific areas, "camden," "mayfair," and "soho." All of the regions appear plausible, but several issues are worthy of attention. Firstly, "northlondon" appears to have a main core nearer to Central London, and then a nearby second core area. This geometry means that the surface is not unimodal, thus violating an assumption made in deriving the kernel bandwidth. Secondly, "eastlondon" appears to actually be defined by Flickr users as east of Central London and north of the Thames. This effect agrees well with the notions of Campari [4] and Lynch [25] with respect to how city neighborhoods can have both bona fide and fiat borders. Finally, the extent of "innercity" appears plausible, extending into "eastlondon" and "northlondon," both areas which might be considered to form part of London's inner city. However, the tag profiles shown in Figure 10 illustrate how these definitions should be treated with caution. The tags for "innercity," with a coefficient of variation of 932%, can be seen in the tag profile to have all been contributed by, most likely, a single prolific contributor. "northlondon" also has a relatively high coefficient of variation (427%) and in the tag profile it is clear that primarily a number of prolific as well as some less prolific users contributed these tags. Finally, "eastlondon" has a lower coefficient of variation (274%) and the tag profile shows that this tag is the most ubiquitous of the three. By contrast to the generic regions, the three specifically named regions all have low coefficients of variation (between 215% and 117%). Such coefficients of variation and the related tag profiles suggest that these regions represent a shared definition of these specifically named places within Flickr.
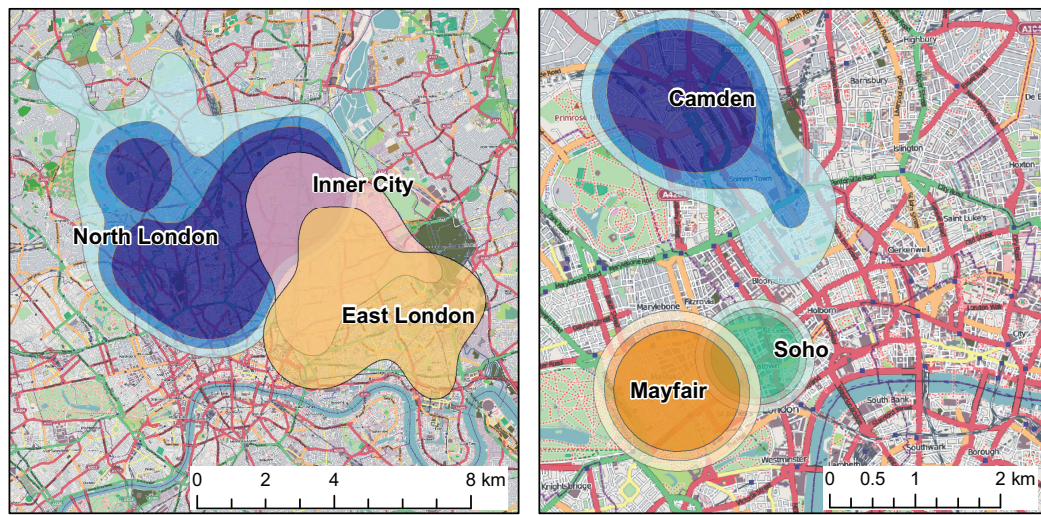
Figure 8: Generic ("northlondon," "innercity," and "eastlondon") and specific ("mayfair," "soho," and "camden") place names in London. For "innercity" and "eastlondon" only 50% contours of the volume surface are shown, for all other regions 50, 60, and 70%. London background mapping openstreetmap.org licensed under a Creative Commons Attribution-Share Alike License.

5 Concluding discussion

This paper opened by setting out a number of examples of the uses of vernacular geography in examples of putative information systems, and argued that such geographies were not often captured by current administrative representations. Furthermore, we argued that user generated content from sources such as Flickr might provide one way of exploring such vernacular geographies, and in particular both *specific* and *generic* use of place names in urban areas where Flickr predominates. Although a number of authors have explored the use of Flickr as a source for vernacular geography, in this paper we set out to address three research questions, which relate not primarily to the extraction of specific regions, but also to the nature and quality of the underlying data and our ability to exploit it to answer broader geographic questions.

The first research question asked "How reliable are user-contributed tags as a means of describing geographic space, and how can we deal with issues of precision, accuracy, and bias?" By analyzing the precision information automatically assigned to geotagged images we firstly observed that the majority of images (Figure 1) are assigned an *accuracy level* of >10 (which is roughly equivalent to city level). An important finding related to these metadata, which is generally relevant to the exploitation of user generated content is however, that they must be treated with caution, since apparent differences in distributions (as was the case for Sydney) may be the result of, sometimes undocumented, changes to proprietary systems. By exploring images tagged in two relatively well delimited regions (parks), we were able to illustrate that, firstly most tagged images ($\approx$86% and $\approx$83% respectively) fell within the region, and these images were associated with a higher precision than
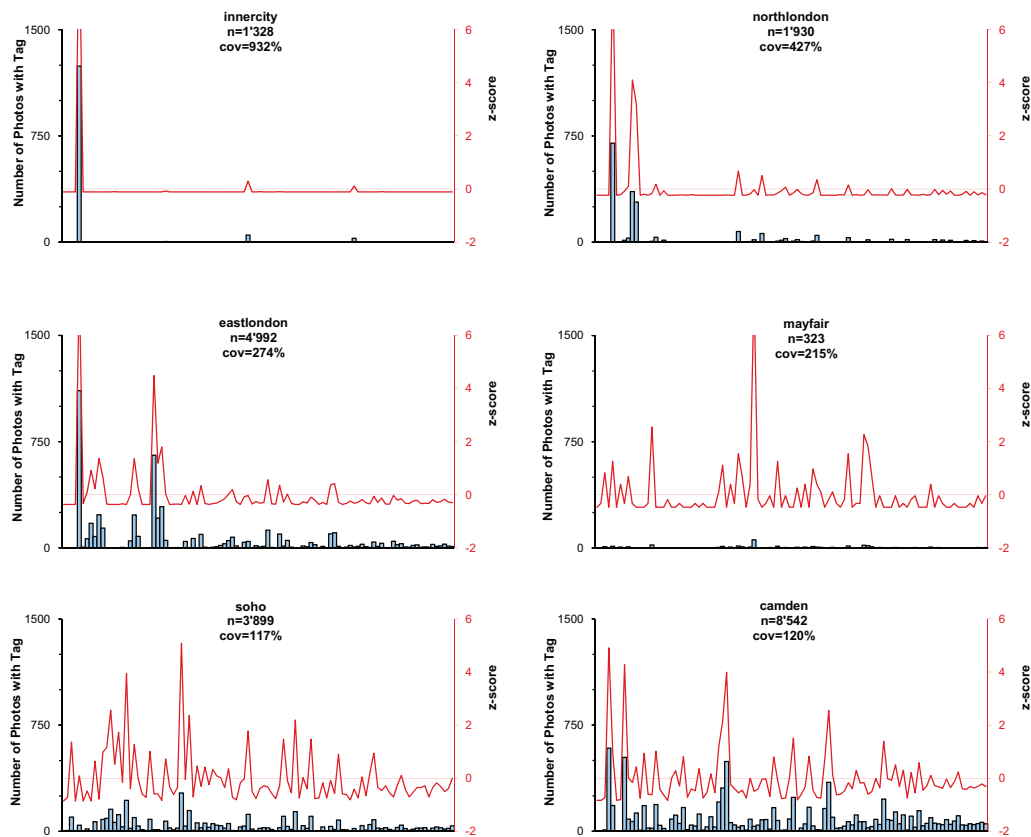
Figure 9: Tag profiles for the six regions shown in Figure 8

those found outside the region. This in turn implies that automatically assigned precision values are a useful filter of falsely placed images. Many falsely assigned images were associated with a few users, and other issues included the location of the photographer versus the object being photographed. Overall, though errors occurred both in semantics (users wrongly naming an object) and geotagging, the precision and accuracy of user generated data appear to be high enough to describe city neighborhoods. Errors resulting from bulk uploads can be dealt with by filtering points with identical coordinates, and by using methods which consider overall point densities, such as kernel density estimation as applied in this paper. We can distinguish between semantic blunders and semantic disagreement—vis the difference between the official and colloquial definitions of Hyde Park—as opposed to the blunder of labelling an image of Regent's Park as Hyde Park, by the use of these density surfaces, provided the tag itself is sufficiently ubiquitous (i.e., are we really capturing the *wisdom of the crowd*). One method of exploring this ubiquity is demonstrated by the use of tag profiles and coefficients of variation, which allowed us to explore how particular tags were used by users with respect to the overall use of tags in a region. Based on our work in this paper, we would suggest that, for large image collections

and histograms with 100 bins, tags with coefficients of variation of less than 300% can be considered (depending on the total number of tags) to be representative. However, inspecting tag profiles is also important, as this reveals further details of the behavior of users. Where coefficients of variation are high, but the overall tag distribution still contains many tags from multiple users, very prolific users could be filtered from the data.

Our second research question asked "How is urban space described using tags, and in particular how are city centers described?" and "How prominent are vernacular place names in such descriptions?" By exploring eight million images collected from Flickr using bounding boxes, city core concepts, and regional place names we established some general characteristics in terms of tagging behavior.

Firstly, georeferenced images have, in general, similar long-tail distributions and mean tag counts to non-georeferenced images retrieved using place names. By investigating the properties of place tag distribution as well as the co-occurrence and frequency of keywords within different cities, it could be shown that the generic city core terms identified in the literature are colloquially used (Tables 3, 4, and 5) but that these tags are employed only by a relatively small group of users with distinctive tagging behavior. "innercity," "central," and "cbd" stem from taglists with a mean of 15.2 keywords per picture. Even the popular "downtown" is associated with photos having on average 11.8 tags. These totals are well above the average tag counts of 7.2 and 5.9 images retrieved using bounding boxes and toponym searches, respectively. Thus, it appears that such generic city core terms are assigned by people describing images in detail. This has important implications for our notion of a collective sense of place, since it may be that the sense we are capturing is one held by a particular group of people, with particular interests and motivations. This is not in itself a fundamental flaw with such data sources, which are still incredibly rich and diverse, but rather an issue which requires further research and consideration. We would argue that, for example, 9000 users tagging more than 100 000 images with "downtown" gives us access to data which are potentially far richer, but much less controlled, than those collected in previous empirical work. Tags also allowed us to explore usage of city core terms in different regions of the world, with "cbd" being more prominent in Asia and Australia, whilst "citycentre" though rare, is mainly used in Europe. Although city core terms are relatively rare (between 0.5 and 2% of georeferenced images from four cities were associated with such tags) the sheer numbers of georeferenced images mean that such tags are associated with large numbers of images.

For "downtown," which appears to be ubiquitous in North America, we have an opportunity to explore patterns of usage at the broader scale. Exploring its distribution through the use of χ-maps (Figure 5) gave us a first insight into a picture that goes beyond that of the underlying image (and typically population density). Analyzing occurrence at different scales gives us clues as to the potential of such methods in analyzing the use of language across space, though it is important to remember the particular characteristics of tags (that is to say flat lists) and their properties.

Chief amongst these tag properties appears to be the importance of both vernacular names specific to individual cities, and more generally, place names, in describing georeferenced images. More than 70% of georeferenced images, and up to 35% of tags, included at least one place name tag of some granularity, reflecting the overall importance of place names in tagging behavior, and tags at the granularity of cities dominate (Table 4). These specific place names are the most frequent keywords used within cities regardless of cultural and linguistic backgrounds. In other work, it has been shown that the city level is

also essential when seeking information. Jones et al. [22] analyzed web queries containing a place name and found that about 84% of the place indications belonged to the city level while only 16% referred to a state/country. The basic geographic level people intuitively think of when describing the location of online items is thus clearly the city name. Perhaps they consider the finer granularity regions within a city as too specific to be searched for by others. Since specific vernacular names appear to be more common than generic names (and thus be associated with less bias, e.g., Figure 10), this in turn suggests that further research is required to automatically identify place names. Since tag lists have no associated structure, then methods which use both semantic and geographic clues are likely to be a promising potential route, but future work should also consider methods which merge data sets with different underlying properties (for example, where precision, in the sense of retrieval, is considered more important by users than recall, as may be the case in online auctions where sellers who falsely attribute themselves to a nearby location may receive negative ratings).

In a third research question we asked "How can georeferenced tagged images be used to gain knowledge of the collective understanding of the location and extent of vernacular regions?" The addition of georeferences to tagged images allows us not only to explore how objects are described, and to use other references to place names to suggest the likely area of interest of an image, but explicitly adds information about the location of images. Using automatically generated metadata related to precision, and by filtering tags that are used in an unusual way (for example very commonly by small numbers of users) it was possible to model collective views of the extents of regions in Chicago and London, illustrating the potential of such techniques to further our understanding of how such regions are understood. However, due to the nature of photography, the data is highly susceptible to internal clustering where a popular photographic viewpoint is found within a region. Thus, there is a need to develop methods to better deal with ambiguity in terms of location— does a tag describe image content or the photographer's location, and if the latter does this change the nature of the surface? Furthermore, current methods for determining regions based on such data typically do not use any ancillary data and, as is the case here, represent regions with smooth borders which may not accord with some of the bona fide divisions found between neighborhoods within cities, although if tag densities are sufficiently high this can be overcome (see for example the boundary of "northlondon" along the Thames in Figure 9).

Unsurprisingly, the performance of the approach was found to be dependent on the data that could be mined from Flickr. It worked more reliably for specific place names, which are more commonly used than generic place tags and typically exhibit well clustered unimodal point patterns. Tags related to neighborhood place names are typically highly spatially auto-correlated, despite the complex nature of cognitive processes and the distributed and uncoordinated process of tagging. This suggests that the average user has a distinct idea of specific places, their location, and their extent. The derivation of spatial footprints from Flickr data is possible using a plethora of methods, including those presented here. We demonstrate that the use of density surfaces and thresholding of the resulting volumes at 50% allows reasonable derivations of the regions associated with tags. These regions can then be represented in an information system as a convex hull or bounding box, which is often sufficient for querying purposes [10]. This research demonstrates for the first time, to our knowledge, that users' overall attitude towards the creation of metadata meets the basic requirements for the generation of footprints for

practical purposes at the sub-city level of granularity. However, it is important to note that such collections still contain significant biases, and that the volume of data does not remove such issues. In the case of Flickr, our results demonstrate that the data set is biased towards American uses of language, and of course the demographic of Flickr users themselves who should not be considered to be representative of society as a whole. Finally, there is a need to develop operational gazetteers and associated methods to take account of representations of vernacular regions which may not only overlap, but have degrees of membership in different locations.

6 Acknowledgements

References

[1] AHERN, S., NAAMAN, M., NAIR, R., AND YANG, J. H. I. World explorer: Visualizing aggregate data from unstructured text in geo-referenced collections. In *Proc. 7th ACM/IEEE-CS Joint Conference on Digital Libraries* (Vancouver, BC, Canada, 2007), ACM Press, pp. 1–10. doi:10.1145/1255175.1255177.

[2] AMES, M., AND NAAMAN, M. Why we tag: Motivations for annotation in mobile and online media. In *CHI '07: Proc. SIGCHI conference on Human Factors in Computing Systems* (San Jose, CA, 2007), ACM Press, pp. 971–980. doi:10.1145/1240624.1240772.

[3] BYRKIT, J. W. Land, sky, and people: The Southwest defined. *Journal of the Southwest 34*, 3 (1992).

[4] CAMPARI, I. Uncertain boundaries in urban space. In *Geographic Objects with Indeterminate Boundaries*, P. A. Burrough and A. U. Frank, Eds., Gisdata 2. Taylor & Francis Ltd, London, UK, 1996, pp. 57–69.

[5] CRANDALL, D. J., BACKSTROM, L., HUTTENLOCHER, D., AND KLEINBERG, J. Mapping the world's photos. In *WWW '09: Proc. 18th international conference on World Wide Web* (Madrid, Spain, 2009), ACM, pp. 761–770. doi:10.1145/1526709.1526812.

[6] DAVIES, C., HOLT, I., GREEN, J., HARDING, J., AND DIAMOND, L. User needs and implications for modelling vague named places. *Spatial Cognition and Computation 9*, 3 (2009), 174–194. doi:10.1080/13875860903121830.

[7] EGENHOFER, M. J., AND MARK, D. M. Naïve geography. In *COSIT '95* (Semmering, Austria, 1995), A. Frank and W. Kuhn, Eds., vol. 988, Springer, pp. 1–15. doi:10.1007/3-540-60392-1_1.

[8] FELLBAUM, C., ET AL. *WordNet: An Electronic Lexical Database*. MIT Press Cambridge, MA, 1998.

[9] FERRARI, G. Boundaries, concepts, language. In *Geographic Objects with Indeterminate Boundaries*, P. A. Burrough and A. U. Frank, Eds., Gisdata 2. Taylor & Francis Ltd, London, UK, 1996, pp. 99–108.

[10] FRONTIERA, P., LARSON, R., AND RADKE, J. A comparison of geometric approaches to assessing spatial similarity for GIR. *International Journal of Geographical Information Science 22*, 3 (2008), 337–360. doi:10.1080/13658810701626293.

[11] GAN, Q., ATTENBERG, J., MARKOWETZ, A., AND SUEL, T. Analysis of geographic queries in a search engine log. In *Proc. First International Workshop on Location and the Web* (Beijing, China, Apr. 2008), 17th International World Wide Web Conference, pp. 49–56. doi:10.1145/1367798.1367806.

[12] GIRARDIN, F., AND BLAT, J. Place this photo on a map: A study of explicit disclosure of location information. In *Late Breaking Result at Ubicomp 2007* (Innsbruck, Austria, Sept. 2007).

[13] GIRARDIN, F., BLAT, J., CALABRESE, F., FIORE, F. D., AND RATTI, C. Digital footprinting: Uncovering tourists with user-generated content. *IEEE Pervasive Computing 7*, 4 (2008), 36–43. doi:10.1109/MPRV.2008.71.

[14] GOLDER, S., AND HUBERMAN, B. A. *The Structure of Collaborative Tagging Systems*. HP Labs technical report, 2005.

[15] GOODCHILD, M. F. Citizens as sensors: The world of volunteered geography. *GeoJournal 69*, 4 (2007), 211–221. doi:10.1007/s10708-007-9111-y.

[16] GROTHE, C., AND SCHAAB, J. Automated footprint generation from geotags with kernel density estimation and support vector machines. *Spatial Cognition and Computation 9*, 3 (2009), 195–211. doi:10.1080/13875860903118307.

[17] GUY, M., AND TONKIN, E. Folksonomies: Tidying up tags? *D-Lib Magazine 12*, 1 (2006).

[18] HENRICH, A., AND LÜDECKE, V. Determining geographic representations for arbitrary concepts at query time. In *Proc. First International Workshop on Location and the Web* (Beijing, China, Apr. 2008), 17th International World Wide Web Conference, pp. 17–24. doi:10.1145/1367798.1367802.

[19] HILL, L. L. *Georeferencing: The Geographic Associations of Information*. MIT Press, 2006. doi:10.1080/00330120701787274.

[20] JONES, C. B., AND PURVES, R. S. Geographical information retrieval. *International Journal of Geographical Information Science 22* (2008), 219–228. doi:10.1080/13658810701626343.

[21] JONES, C. B., PURVES, R. S., CLOUGH, P. D., AND JOHO, H. Modelling vague places with knowledge from the Web. *International Journal of Geographical Information Science 22*, 10 (2008), 1045–1065. doi:10.1080/13658810701850547.

[22] JONES, R., ZHANG, W., REY, B., JHALA, P., AND STIPP, E. Geographic intention and modification in web search. *International Journal of Geographical Information Science 22*, 3 (2008), 229–246. doi:10.1080/13658810701626186.

[23] KESSLER, C., MAU, P., HEUER, J. T., AND BARTOSCHEK, T. Bottom-up gazetteers: Learning from the implicit semantics of geotags. In *GeoS '09: Proc. Third International Conference on GeoSpatial Semantics* (Berlin, 2009), Springer, pp. 83–102. doi:10.1007/978-3-642-10436-7_6.

[24] KING, L. J., AND GOLLEDGE, R. G. *Cities, Space, and Behavior*. Prentice-Hall, Inc., Englewood Cliffs, US, 1978.

[25] LYNCH, K. *The Image of the City*. MIT Press and Harvard University Press, Cambridge, MA, 1960.

[26] MATHES, A. Folksonomies - cooperative classification and communication through shared metadata. Tech. rep., University of Illinois Urbana-Champaign, 2004.

[27] MONTELLO, D. R., GOODCHILD, M. F., GOTTSEGEN, J., AND FOHL, P. Where's downtown? behavioral methods for determining referents of vague spatial queries. *Spatial Cognition and Computation 3*, 2–3 (2003), 185–204. doi:10.1207/S15427633SCC032&3_06.

[28] MURPHY, R. E. *The Central Business District*. Aldine Atherton, Inc., Chicago, Il, US, 1972.

[29] O'SULLIVAN, D., AND UNWIN, D. J. *Geographic Information Analysis*. John Wiley & Sons, Inc., Hoboken, New Jersey, US, 2003.

[30] RATTENBURY, T., AND NAAMAN, M. Methods for extracting place semantics from flickr tags. *ACM Transactions on the Web 3*, 1 (2009). doi:10.1145/1462148.1462149.

[31] RODGERS, A. R., AND CARR, A. P. HRE: The home range extension for ArcView. User's Manual. Tech. rep., Centre for Northern Forest Ecosystem Research, Ontario, July 1998.

[32] SCHMITZ, P. Inducing ontology from Flickr tags. In *Proc. Collaborative Web Tagging Workshop (WWW '06)* (Edinburgh, Scotland, 2006).

[33] SEN, S. Characterizing places in geospatial ontologies: Specifying partial knowldege about their use. In *International Workshop on Computational Models of Place* (Park City, Utah, US, Sept. 2008), S. Winter, W. Kuhn, and A. Krüger, Eds., PLACE'08, The University of Melbourne, pp. 29–44.

[34] SIGURBJÖRNSSON, B., AND VAN ZWOL, R. Flickr tag recommendation based on collective knowledge. In *WWW '08: Proc. 17th international conference on World Wide Web* (Beijing, China, 2008), ACM, pp. 327–336. doi:10.1145/1367497.1367542.

[35] SMITH, B., AND VARZI, A. C. Fiat and bona fide boundaries. *Philosphy and Phenomenological Research 60*, 2 (2000), 401–420.

[36] STEELS, L. Collaborative tagging as distributed cognition. *Pragmatics and Cognition 14*, 2 (2006), 287–292. doi:10.1075/pc.14.2.09ste.

[37] THURSTAIN-GODWIN, M., AND UNWIN, D. J. Defining and delineating the central areas of towns for statistical monitoring using continuous surface representations. Tech. Rep. 18, CASA, Centre for advanced spatial analysis, London, 2000.

[38] WEISS, A. The power of collective intelligence. *netWorker 9*, 3 (2005), 16–23. doi:10. 1145/1086762.1086763.

[39] WINGET, M. User-defined classification on the online photo sharing site Flickr. . . or, how I learned to stop worrying and love the million typing monkeys. In *Advances in Classification Research, Vol. 17: Proc. of the 17th ASIS&T SIG/CR Classification* (Philadelphia, PA, Nov. 2006), J. Furner and J. T. Tennis, Eds.

[40] WOOD, J., DYKES, J., SLINGSBY, A., AND CLARKE, K. Interactive visual exploration of a large spatio-temporal dataset: Reflections on a geovisualization mashup. *IEEE Transactions on Visualization and Computer Graphics 13*, 6 (2007), 1176–1183. doi:10.1109/TVCG.2007.70570.