Do Humans Have Conceptual Models About Geographic Objects? A User Study

Ahmet Aker*
Department of Computer Science, University of Sheffield, 211 Portobello, S1 4DP, Sheffield, United Kingdom. E-mail: a.aker@dcs.shef.ac.uk

Laura Plaza
Dpto. de Ingeniería del Software e Inteligencia Artificial, Universidad Complutense de Madrid, C/ Prof. José García Santesmases s/n, 28040 Madrid, Spain. E-mail: lplazam@fdi.ucm.es

Elena Lloret
Dpto. de Lenguajes y Sistemas Informáticos, Universidad de Alicante, Apdo. de correos, 99, E-03008 Alicante, Spain. E-mail: elloret@dlsi.ua.es

Robert Gaizauskas
Department of Computer Science, University of Sheffield, 211 Portobello, S1 4DP, Sheffield, United Kingdom. E-mail: R.Gaizauskas@sheffield.ac.uk

In this article, we investigate what sorts of information humans request about geographical objects of the same type. For example, *Edinburgh Castle* and *Bodiam Castle* are two objects of the same type: "castle." The question is whether specific information is requested for the object type "castle" and how this information differs for objects of other types (e.g., church, museum, or lake). We aim to answer this question using an online survey. In the survey, we showed 184 participants 200 images pertaining to urban and rural objects and asked them to write questions for which they would like to know the answers when seeing those objects. Our analysis of the 6,169 questions collected in the survey shows that humans have shared ideas of what to ask about geographical objects. When the object types resemble each other (e.g., church and temple), the requested information is similar for the objects of these types. Otherwise, the information is specific to an object type. Our results may be very useful in guiding Natural Language Processing tasks involving automatic generation of templates for image descriptions and their assessment, as well as image indexing and organization.

Introduction

In everyday life, we see and categorize things in our built or natural environment. For instance, if we look at different castles, we see that each of them has a different style, a different look, and a different size; some of them are newer than the others; and so on, but still we are able to categorize all of them as "castles."

For categorization purposes, visual attributes such as the style and size of a castle might be enough; however, to separately report about each castle, we need more attributes whose values make each one distinguishable from others.

Knowing what sets of attributes people use to describe geographical objects has several applications. Automatic text summarization applied to generation of descriptions for geographical objects would be improved by incorporating human preferences into the output summaries. Template-based summarization methods, for example, have been used successfully for news-event summarization (McKeown & Radev, 1995; Radev & McKeown, 1998). The same strategy can be applied to the task of generation of descriptions for geographical objects (e.g., for purposes of generating descriptions for images of these objects). Templates can be derived from the set of attributes relevant to humans to bias the summarization system toward the text units containing the values for these attributes. Furthermore, the attributes could be used in a guided summarization task as organized by the Document Understanding Conferences (DUC)[1] and the Text Analysis Conference (TAC).[2] In addition, the attributes can be used to evaluate the output of an automatic summarization system by testing whether

[1]http://duc.nist.gov/
[2]http://www.nist.gov/tac/

highly relevant attributes are covered in the output summary. Furthermore, the attribute lists present a guide for what to index for images pertaining to locations. If information highly relevant for humans is used for indexing, this could lead to better retrieval and organization of those images. In information extraction, automatic generation of domain templates has been widely investigated (Banko, Cafarella, Soderland, Broadhead, & Etzioni, 2007; Banko & Etzioni, 2008; Filatova, Hatzivassiloglou, & McKeown, 2006; Li, Jiang, & Wang, 2010; Sudo, Sekine, & Grishman, 2003). The set of attributes could be used to generate such templates for geographic objects and/or to evaluate the output of such systems.

For these applications, it is relevant to know whether there is a general set of attributes that people use to describe any geographical object or, given that humans categorize objects into types (e.g., church, museum, lake, etc.), if there are sets of attributes specific to single object types or perhaps shared between object types of similar function (e.g., church and temple).

In this article, we aim to address these questions. Specifically, we analyze three research questions: When seeing a geographic object, what is the set of attributes related to the object for which people would like to know the values? Is this set of attributes specific to a particular object type (e.g., church) or shared between different object types? Do human interests correlate to what can be found in real documents?

Our goal is to answer these questions using an online survey conducted on Amazon's Mechanical Turk. In the experiment, we showed participants different images of objects from around the world. Our set of images contains only images of static features of the built or natural landscape (i.e., objects with persistent geocoordinates, e.g., buildings and mountains, and not images of objects which move about in such landscapes, e.g., people, cars, clouds, etc.). We asked the participants to write questions for which they would like to know the answers when seeing each image. We collected 7,644 questions from 184 participants. The results of our analysis reveal that people share ideas regarding what to ask about geographical objects in general. When the object types resemble each other, the requested information is similar. In contrast, if the object types are distinct enough, then this information is specific to each object type.

The remainder of the article is organized as follows. We first introduce related work. We next describe Mechanical Turk, present the experimental setting of our survey, and outline the preprocessing of the data. We later present our analysis and report, and then discuss the results. Finally, we extract the main conclusions of the study and outline future work.

Related Work

Cognitive psychology has offered several theories and substantial empirical evidence for the existence of categories or concepts and explanation of what constitutes them (Eysenck & Keane, 2005). Theories agree that concepts are characterized by sets of attributes, although they differ in whether a set of attributes is necessary and sufficient to define a concept (defining-attributes theories) or whether the concepts are more fuzzy in their specification in terms of attributes (prototype theories), so that some objects are more representative of a concept than others. If humans use concepts to organize the knowledge about the world, then we assume that they will have ways to describe these concepts in natural language. Our aim is to analyze this hypothesis because it will be complementary to reasoning theories (Gordon, 2004; Gordon, Bejan, & Sagae, 2011; Gordon & Hobbs, 2011), thus providing some insights about how humans perceive and communicate the information.

To our knowledge, similar analyses of what concepts or information types human associate with geographical objects have not been reported previously. However, earlier work has studied image-related captions or descriptions to understand how people describe images by looking at query logs or existing image descriptions (Armitage & Enser, 1997; Balasubramanian, Diekema, & Goodrum, 2004; Choi & Rasmussen, 2002; Greisdorf & O'Connor, 2002; Hollink, Schreiber, Wielinga, & Worring, 2004; Jörgensen, 1996, 1998). These studies were aimed at understanding how people index images in general and were not specific to geographic objects. Our work also is somewhat similar to the one described in Rosch (1999), who investigated what attributes people associate with objects such as fish, tree, fruit, and so on whereas our focus is on geographic objects. Finally, other studies have analyzed the impact of images in memory, focusing on the differences between men and women when describing a seen image (Marks, 1973), which is outside the scope of our work.

Experimental Setting

Mechanical Turk

Mechanical Turk (MTurk) is a crowdsourcing service run by Amazon. It allows users (also called "requesters") to upload tasks and obtain results within a very short time. The tasks are performed by MTurk workers. Each worker has the ability to scroll existing tasks (also called "human intelligence task," or "HITs") and complete them for a fee offered by the requester.

The general advantage of MTurk, the ability to collect results in a time- and cost-efficient manner, makes it a suitable platform for conducting this experiment. MTurk has been widely used for language processing and information retrieval tasks (Aker, El-Haj, Albakour, & Kruschwitz, 2012; Dakka & Ipeirotis, 2008; El-Haj, Kruschwitz, & Fox, 2010; Kaisser, Hearst, & Lowe, 2008; Kittur, Chi, & Suh, 2008; Snow, O'Connor, Jurafsky, & Ng, 2008; Yang et al., 2009). Several studies have shown that the quality of results produced by MTurk workers is comparable to that of traditionally employed experiment participants (Alonso &

Question 1:
Question 2:
Question 3:
Question 4:
Question 5:
Question 6:
Question 7:
Question 8:
Question 9:
Question 10:

FIG. 1. Design of the online MTurk experiment. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

Mizzaro, 2009; Snow et al., 2008; Su, Pavlov, Chow, & Baker, 2007).

Experiment Design

Our experimental design is similar to the one described by Filatova et al. (2006), who were interested in knowing what information people expect to know when they read an article about an event. The authors used four different events (terrorist attack, earthquake, presidential election, and airplane crash). They asked 10 humans to provide questions for which they would like to get the answers when they read an article about each of those events. We set up our experiment in a similar way and asked humans to provide questions about places for which they would like to know the answers when they see the place. We obtained the questions via MTurk workers. In the experiment, we showed the workers an image pertaining to a particular object or place (e.g., the *Eiffel Tower* in Paris, as shown in Figure 1). We also presented the worker with the name of the place (*Eiffel Tower*) and its object type (*tower*). The workers were asked to take the role of a tourist and provide 10 questions for which they would like to know the answers when they see the place shown in the image.

The experiment design illustrated in Figure 1 allows us to address some of the quality issues associated with conducting experiments on MTurk. Related work has reported problems with spammers and unethical workers, who produce incomplete or absurd output (Feng, Besana, & Zajac, 2009; Kazai, 2011; Mason & Watts, 2010). In our experiment, for example, text fields were provided for writing down each question. In this way, we were able to control the quality of the questions and reject all absurd input such as strings of arbitrary characters, and so on. Aker et al. (2012) also reported that text fields are a good design selection to detect unwilling or less careful workers.

TABLE 1. Object types with their number of questions: object type (count, percentage).

Top-10
university (225, 2.94), lake (225, 2.94), cemetery (216, 2.82), cave (216, 2.82), park (216, 2.82), church (216, 2.82), gallery (216, 2.82), canal (216, 2.82), cathedral (216, 2.82), square (216, 2.82), palace (216, 2.82), opera house (216, 2.82), river (216, 2.82), mountain (207, 2.7), beach (207, 2.7), volcano (207, 2.7), zoo (207, 2.7), museum (207, 2.7), temple (207, 2.7), residence (207, 2.7), ski resort (207, 2.7), garden (207, 2.7), glacier (199, 2.6), peak (198, 2.59), monument (198, 2.59), building (198, 2.59), aquarium, (198, 2.59), boulevard (189, 2.47), basilica (189, 2.47), railway station (180, 2.35), house (180, 2.35), abbey (180, 2.35), tower (180, 2.35), stadium (180, 2.35), island (173, 2.26), avenue (171, 2.23), village (135, 1.76), parliament (81, 1.06), railway (81, 1.06), prison (45, 0.59)

Workers also were required to provide the complete set of 10 questions. This was not always fulfilled; however, we also accepted lists containing less than 10 questions, provided that the questions were of "good quality."

We showed images picturing places of 40 different object types randomly chosen from our entire set of 107 object types; this set is described in Aker and Gaizauskas (2010). From these 40 object types, 25 are urban and 15 rural types:

- **Urban types:** abbey, aquarium, avenue, basilica, boulevard, building, cathedral, cemetery, church, gallery, house, monument, museum, opera house, palace, parliament, prison, railway, railway station, residence, square, stadium, temple, tower, university, and zoo.
- **Rural types:** beach, canal, cave, garden, glacier, island, lake, mountain, park, peak, river, ski resort, village, and volcano.

For each object type, five different places were shown. For example, for the object type *tower*, the images of *Eiffel Tower*, *Flag Tower of Hanoi*, *BT Tower*, and *Munttoren*, *Bettisons Folly* were shown. These places (towers) were manually selected from Wikipedia.[3] Each image was shown to 5 different workers. Note that we did not want to experiment with specific places (i.e., *Eiffel Tower*), but we wanted each worker to pose questions on the same five places to collect a great variety of questions for the same place and to study the correlation between different users; that is the why we manually chose the images within Wikipedia.

We ran the experiment for 4 weeks. In total, we collected 7,644 questions for 187 different places (We did not get results for 13 of 200 images.) The questions came from 184 different workers. The expected number of questions was 10,000 (40 Types × 5 Objects × 5 Workers × 10 Questions). However, there are some objects for which we only have questions from a few workers. Table 1 shows the total number of questions collected for each object type. There were 4,815 questions for the urban types and 2,829 questions for rural ones. We think that objects of urban types

such as a church are more familiar to humans than are objects of rural types and that they are thus able to generate more questions for the urban objects than for the rural ones. We believe that this might be one reason why the number of questions for the urban objects is much greater than for the rural objects.

Question Preprocessing

We manually analyzed all the questions to assess their quality. Approximately 2% of the questions were empty because not all the workers wrote 10 questions for each object. Moreover, some questions were related only to the image itself rather than to the place shown in the image (e.g., "When the picture is taken?" "How many flowers you found in the image?" "Is there a bus in the picture?"). In addition to these, some questions presented nonresolved references, so it was impossible to know to which object they refer (e.g., "What language do *they* [emphasis added] speak?"). Finally, there were questions which bore no relation at all with the object in the image (e.g., "How is the manager?"). These questions did not address the task, which was to ask questions about the *object* shown in the image, not about the image itself or related information. Therefore, we categorized all these questions as "noise," making up 19% of the entire question set (1,479 of 7,648 questions).

We categorized the remaining 81% of the questions (6,169 of 7,648) by the *attribute* for which the worker was seeking the value with his or her question. An attribute is an abstract grouping of similar questions. We regard two or more questions as similar if their answers refer to the same information type. For instance, we regard the questions "where is Garwood glacier?" and "where exactly is Edmonton?" as similar because both aim for answers related to the information type *location*. We manually established different abstract grouping categories, and each attribute was named according to the information type to which it refers (e.g., location). Table 2 shows question examples for the top-10 attributes (i.e., the 10 attributes which have the most questions), defined next:

- **Visiting:** Sentences containing information about, for example, visiting times, prices, and so on.
- **Location:** Sentences containing information about where the object is located.
- **Foundationyear:** Sentences containing information about when a building was built, when an organization was established, and so on.
- **Surrounding:** Sentences containing information about what other objects are close to the main object.
- **History:** Sentences containing information about interesting past events related to the main object.
- **Size:** Sentences containing information about the spatial dimensions of the main object, such as the height of a tower, the length of a mountain, and so on.
- **Design:** Sentences containing information about the style, structure of a building, or even the technology used for its construction.

TABLE 2. Top-10 attributes with related questions.

Visiting	where i can buy the ticket? is this tower available to be visited the whole year? when is the best time to visit? how to get there?
Location	where is garwood glacier? where exactly is edmonton? where it's located?
Foundationyear	when was it build? which year was this zoo opened? when it was established?
Surrounding	what are the landmarks found nearby seima palace? what are the nearby places to visit? what are some nearby attractions?
Features	are there any waterfalls in the park? what does the zoo house?
History	what is the history of it? any history related to george mason university? what role did the palace play in the history of france?
Size	how big is the complex? how big is the temple? how big is it, meaning what are the dimensions?
Design	what is the architectural structure style? what style of architecture is this house built in? is it constructed by ancient technology?
Mostattraction	what is the speciality of this abbey? what is the best in this island? what are the extraordinary facts about this place?
Naming	was it named after a person? how did the abbey get its name? how the name for this boulevard came?

- **Mostattraction:** Sentences containing information about the feature or attribute of the main object which makes people want to visit it.
- **Naming:** Sentences containing information about the origin of the name object name or about alternative names for it.
- **Features:** Sentences containing information about other outstanding qualities or characteristics of the main object.

After this analysis, we obtained a total of 146 attributes; however, 95 of them contained less than five questions, so we ignored these attributes in further analysis.[4] We analyzed the remaining set of 51 attributes (see Table 2), each of which has at least five questions related to it.

Analysis and Results

In this section, we carry out the analysis with respect to the three proposed research questions. Each subsection is devoted to a different question. Note that the analysis for some research drew upon data from other questions that were studied.

RQ1: Is there a set of attributes that people generally associate with geographic objects?

In Table 3, the attributes along with the number of questions for each attribute from all object types are given. As the table shows, some attributes in the left column ("Top-10") are very frequently addressed by the participants. These attributes cover the majority of the questions. More than 65% of the questions can be categorized by these 10 attributes. This means that people do share ideas as to what types of information are required about a place, and the set of top-10 attributes captures these information types.

Table 4 lists the most frequent eight attributes for each object type. From this table, we can see that the top-10 attributes are present for most of the object types, indicating that the same type of information is relevant for several object types. However, although these attributes occur for many object types, their *popularity* is not the same in all the object types with which they are associated.

We define the popularity of an attribute for an object type as the number of questions categorized under this attribute

---

[4]The total number of questions discarded with these 95 categories is 150.

TABLE 3. Attributes used to categorize the questions. For each category, we give its number of questions and the percentage value of that number over the total number of categorized questions: attributeName(count, percentage)

Top-10	Below top-10
visiting (1068, 17.31), location (861, 13.96), foundationyear (465, 7.54), surrounding (426, 6.91), features (357, 5.79), history (239, 3.87), size (214, 3.47), design (211, 3.42), mostattraction (165, 2.67), naming (153, 2.48)	purpose (136, 2.20), istouristattraction (132, 2.14), height (132, 2.14), visitors (126, 2.04), founder (113, 1.83), owner (109, 1.77), events (104, 1.69), type (102, 1.65), status (86, 1.39), habitants (74, 1.20), designer (69,1.12), constructioninfo (63, 1.02), temperature (58, 0.94), length (56, 0.91), eruptioninfo (42, 0.68), comparison (42, 0.68), capacity (38, 0.62), depth (36, 0.58), maintainance (35, 0.57), arts (27, 0.44), studentsinfo (26, 0.42), subjectsofferred (25, 0.41), gravesinfo (23, 0.37), preacher (22, 0.36), firstdiscovery (20, 0.32), parkinginfo (20, 0.32), religioninfo (19, 0.31), destination (16, 0.26), origin (15, 0.24), workers (12, 0.19), width (12,0.19), travellers (11, 0.18), travelcost (9, 0.15), waterinfo (9, 0.15), travelling (9, 0.15), firstclimber (9, 0.15), tributaries (8, 0.13), prayertime (6, 0.10), personinmonument (6, 0.10), snowinformation (6, 0.10), firstfounder (5, 0.08)

for that particular object type. Attribute popularity indicates how important the type of information represented by the attribute is for the particular object type. For instance, the attribute *visiting* is the most popular attribute (i.e., has the maximum number of 52 questions) for the object type *museum* (see Table 4). This indicates that it is very important for people to know how much the entry to the museum costs or when the museum opens. This information is more relevant than, for instance, knowing when the museum was built. However, if we look at the object type *house*, we can

TABLE 4. List of attributes (eight most frequent) for each object type. The numbers within the brackets attached to each attribute indicate how many questions were categorized by that attribute.

Object type	Attributes
Abbey	foundationyear (21), location (20), visiting (16), history (14), design (10), founder (7), height (5), mostattraction (4)
Aquarium	visiting (63), features (30), location (19), foundationyear (10), size (7), visitors (6), events (5), surrounding (5)
Avenue	location (28), surrounding (16), visiting (11), foundationyear (8), features (8), type (6), parkinginfo (5), visitors (5)
Basilica	visiting (25), location (23), foundationyear (21), surrounding (16), founder (7), visitors (7), history (7), purpose (6)
Beach	visiting (44), location (24), surrounding (23), features (18), visitors (12), mostattraction (7), naming (4), depth (4)
Boulevard	surrounding (23), location (17), visiting (17), features (13), naming (13), foundationyear (7), length (6), visitors (5)
Building	location (22), design (22), visiting (18), surrounding (10), foundationyear (10), height (9), constructioninfo (9), owner (9)
Canal	visiting (23), location (21), surrounding (18), purpose (16), length (12), foundationyear (12), depth (9), istouristattraction (7)
Cathedral	location (30), visiting (23), foundationyear (16), history (12), founder (9), size (8), design (8), preacher (6)
Cave	visiting (46), location (25), surrounding (15), features (15), firstdiscovery (13), history (12), naming (7), mostattraction (6)
Cemetery	location (30), visiting (29), gravesinfo (23), foundationyear (18), istouristattraction (11), surrounding (8), size (7), naming (5)
Church	foundationyear (26), location (22), visiting (18), preacher (12), events (9), history (9), design (9), features (7)
Gallery	visiting (43), location (23), arts (21), foundationyear (16), surrounding (10), design (8), events (5), mostattraction (5)
Garden	visiting (34), features (23), location (20), foundationyear (11), size (9), constructioninfo (6), surrounding (6), owner (5)
Glacier	location (24), visiting (24), status (13), foundationyear (11), size (10), history (9), surrounding (8), naming (8)
House	design (27), location (22), foundationyear (16), owner (13), visiting (11), history (10), istouristattraction (9), founder (9)
Island	visiting (21), location (19), habitants (19), features (15), size (13), surrounding (11), temperature (6), mostattraction (5)
Lake	visiting (61), surrounding (32), location (29), size (12), depth (10), features (7), mostattraction (6), purpose (6)
Monument	foundationyear (23), location (20), visiting (14), surrounding (13), history (12), purpose (11), designer (11), height (9)
Mountain	visiting (35), location (23), surrounding (18), height (18), features (14), naming (8), temperature (6), size (5)
Museum	visiting (52), location (22), foundationyear (22), mostattraction (10), features (9), history (8), arts (5), owner (5)
Opera House	visiting (33), foundationyear (20), location (17), events (17), mostattraction (10), surrounding (10), designer (9), owner (8)
Palace	location (26), history (23), foundationyear (21), visiting (17), size (12), surrounding (12), design (12), owner (8)
Park	visiting (41), features (39), location (26), size (12), surrounding (10), foundationyear (9), mostattraction (5), events (4)
Parliament	visiting (13), location (7), foundationyear (7), surrounding (6), design (5), history (4), mostattraction (3), type (3)
Peak	visiting (58), location (27), height (20), surrounding (18), temperature (10), features (6), history (6), istouristattraction (5)
Prison	foundationyear (7), design (5), history (4), capacity (4), location (3), prisonersinfo (3), type (2), size (2)
Railway	location (7), travelcost (7), travelers (6), traveling (4), technique (3), foundationyear (3), status (2), owner (2)
Railway Station	foundationyear (22), location (19), visiting (7), naming (5), travelers (5), history (5), design (5), type (4)
Residence	design (28), location (21), foundationyear (15), visiting (13), habitants (10), purpose (8), history (8), size (7)
River	visiting (25), location (19), length (18), surrounding (16), purpose (15), origin (12), status (9), features (8)
Ski Resort	visiting (40), features (26), location (18), surrounding (8), height (7), events (6), size (6), temperature (5)
Square	visiting (28), location (26), surrounding (22), history (17), foundationyear (14), naming (13), mostattraction (8), istouristattraction (8)
Stadium	location (30), events (25), visiting (23), capacity (15), surrounding (11), foundationyear (11), type (6), size (6)
Temple	location (28), visiting (22), foundationyear (17), design (17), visitors (12), surrounding (8), religioninfo (8), size (7)
Tower	visiting (24), location (21), foundationyear (18), design (16), purpose (15), height (14), features (6), designer (5)
University	location (32), studentsinfo (26), subjectsofferred (25), visiting (11), foundationyear (10), size (9), maintainance (8), history (6)
Village	habitants (18), surrounding (17), visiting (16), location (11), features (10), mostattraction (5), size (5), naming (4)
Volcano	eruptioninfo (42), visiting (26), location (21), surrounding (20), height (15), status (14), foundationyear (7), naming (6)
Zoo	visiting (41), features (34), location (19), foundationyear (14), size (10), mostattraction (8), surrounding (5), owner (5)

see that the same attribute *visiting* is not the most popular. It occurs at Position 5 with only 11 questions. For this object type, instead, people seem to be most interested in knowing the *design* of the house.

The remaining 35% of the questions are spread over the remaining 41 attributes as shown in the right column ("Below top-10") of Table 3. From this list, we can observe that there are some specific attributes that are only present in a particular object type or in few object types. For instance, as shown in Table 4, the attribute *eruptioninfo is* associated only with the object type *volcano*. This attribute contains questions related to the eruption of different *volcanoes*: "How often this volcano erupts in fuji?" "Has it erupted before?" "When was the last time it erupted?" From Table 4, we can see that this attribute also is the most popular for the *volcano* object type while information related to *visiting*, *location*, and so on,

which is generally most frequently asked for, comes after *eruptioninfo*.

From this analysis, we can conclude that even though object types share the top-10 attributes, the differences in attribute popularity are an indication that these are not equally important for all object types. Therefore, the question arises as to whether each object type has an associated specific set of attributes, or whether there are similar object types that can be grouped according to which information is required to describe them. This is our second research question, which is analyzed next.

RQ2: Is each set of attributes specific to a particular object type or shared between several object types?

To address this question, we compare different object types and groups of object types and investigate the degree of similarity between them. If two object types are similar,

TABLE 5. Ten object-type pairs from urban and rural areas with high correlation in all attribute sets, with high correlation in the top-10 attributes, and with low correlation, respectively.

Always high correlation	High correlation in top-10 and low correlation in below- top-10 attributes	Always low correlation
church-temple, abbey-temple, building-tower, building-residence, abbey-basilica, house-residence, abbey-church, basilica-temple, basilica-cathedral, museum-operahouse, house-palace, gallery-museum	Urban areas house-prison, museum-railway, basilica-railway, monument-prison, university-cathedral, prison-palace, prison-cathedral, cemetery-aquarium, railwaystation-cathedral, cemetery-parliament	prison-aquarium, monument-railway, railway-cathedral, university-railway, museum-prison, temple-railway, railway-operahouse, parliament-railway, abbey-railway, tower-railway
mountain-volcano, mountain-skiresort, mountain-peak, peak-volcano, peak-glacier, mountain-glacier, glacier-skiresort, park-garden, lake-canal, canal-river	Rural areas village-river, lake-island, canal-volcano, island-river, mountain-river, peak-lake, island-garden, canal-glacier, skiresort-garden, mountain-canal	lake-volcano, canal-island, lake-glacier, glacier-river, canal-village, lake-garden, village-garden, glacier-village, island-glacier, park-cave

then they will not only share a set of attributes but the attributes within this set also will have similar popularity ranking.

We use Kendall's τ rank correlation coefficient as a metric, which indicates similarity between object types while considering the attribute popularity ranking in their attribute sets.[5] Kendall's τ correlation coefficient is equal or close to 1 if two events are highly correlated in rankings, and close to 0 if there exists little or no correlation between the ranks of the two events.

Attributes are ranked according to their popularity (i.e., the number of questions contained under each attribute). If the attributes of two different object types have similar rankings, this is an indication that these attributes are of similar importance for both object types. Kendall's τ for these object types will return a high correlation for the pair, and we will refer to such object types as *similar*. A difference in attribute rankings between different object types indicates that these attribute sets are more specific to one object type and not shared by the other(s). In this case, Kendall's τ will return a low correlation, and we will refer to such object types as *dissimilar*.

Within this research question, we can go further and analyze two more questions that directly derive from this. On one hand, we first report comparisons between single object types. On the other hand, we report similarity between these object-type higher level groups. Both questions are discussed next.

RQ3: How similar are single object types?

Table 5 shows three groups of object-type pairs in rural and urban areas, respectively. In the first column (*Always high correlation*), the object-type pairs whose correlation coefficient is higher than the mean for *all* attributes are shown. The second column (*High correlation in top-10 and low correlation in below top-10 attributes*) shows the object-type pairs whose correlation is higher than the mean in the *top-10* attributes, but drops below the mean for the *below top-10*

attributes. Finally, in the third column (*Always low correlation*), the pairs that are correlated with the coefficient lower than the mean for *all* attributes are presented. On average, the urban object types correlate with 0.55 (*Mdn* = 0.57), and the rural types correlate with 0.53 (*Mdn* = 0.53). The mean and median correlation coefficients lie close to each other, which indicates that the distribution of highly correlating object types is similar to that of object types with low correlation. From Table 5, we can see that high correlation in attributes is always present when the object types have, for instance, the same look, design, or purpose.

For the urban areas, we have object types such as *church*, *basilica*, *abbey*, *cathedral*, and *temple*, which correlate with a coefficient higher than the mean of 0.55 when *all* attributes are considered. A similar picture can be drawn for other urban object types such as *house-residence*, *building-residence*, or *museum-operahouse*. In rural areas, object types related to mountainous areas (e.g., *glacier*, *mountain*, *peak*, *volcano*, *ski resort*, etc.) are correlated above the mean of 0.53. The same is valid for water bodies such as *canal*, *lake*, *river*, and so on. All these object types clearly share features such as look, design, purpose, and so on. The correlation coefficient is always low for object types that do not share these aspects, such as the ones shown in the third column of Table 5. This indicates that aspects used for categorizing objects into object types also play a role in deciding which object types are similar for purposes of describing them.

However, the second column of the table also highlights the importance of shared ideas of what information is relevant for describing geographic objects for object-type similarity (RQ1). The object types shown in the second column of the table are correlated above the mean for *top-10* attributes, and their correlation drops below the mean when *below top-10* attributes are used, leading to a new research subquestion (RQ4).

Based on these results, we can derive groups of similar object types. Such groups can be built from object types shown in column 1 of Table 5, and we can investigate whether such groups of object types are similar.

RQ4: How similar are groups of similar object types?

---

[5]We discarded attributes which do not occur in both lists of attributes, which makes approximately 2% of the questions used for categorization.

TABLE 6. High-level groups and subgroups of object types.

Urban objects	Religious places (e.g., basilica, abbey, church)
	Buildings (e.g., palace, stadium, university)
	Cultural attractions (e.g., museum, opera house, gallery)
Rural objects	Water bodies (e.g., beach, canal, lake)
	Mountainous areas (e.g., ski resort, volcano, peak)

TABLE 7. Correlation coefficients between groups of object types.

Categories	Area pair	Kendall's τ
Groups in urban		
All	(religious places)-(buildings)	0.71
Top-10	(religious places)-(buildings)	0.85
Bellow-top-10	(religious places)-(buildings)	0.63
All	(religious places)-(cultural attractions)	0.71
Top-10	(religious places)-(cultural attractions)	0.71
Bellow-top-10	(religious places)-(cultural attractions)	0.71
All	(buildings)-(cultural attractions)	0.67
Top-10	(buildings)-(cultural attractions)	0.44
Bellow-top-10	(buildings)-(cultural attractions)	0.59
Groups in rural		
All	(water bodies)-(mountainous areas)	0.47
Top-10	(water bodies)-(mountainous areas)	0.74
Bellow-top-10	(water bodies)-(mountainous areas)	0.29
Groups in urban against groups in rural		
All	(religious places)-(mountainous areas)	0.41
Top-10	(religious places)-(mountainous areas)	0.40
Bellow-top-10	(religious places)-(mountainous areas)	0.18
All	(buildings)-(mountainous areas)	0.41
Top-10	(buildings)-(mountainous areas)	0.40
Bellow-top-10	(buildings)-(mountainous areas)	0.18
All	(cultural attractions)-(mountainous areas)	0.39
Top-10	(cultural attractions)-(mountainous areas)	0.25
Bellow-top-10	(cultural attractions)-(mountainous areas)	0.18
All	(religious places)-(water bodies)	0.38
Top-10	(religious places)-(water bodies)	0.41
Bellow-top-10	(religious places)-(water bodies)	0.17
All	(buildings)-(water bodies)	0.46
Top-10	(buildings)-(water bodies)	0.25
Bellow-top-10	(buildings)-(water bodies)	0.32
All	(cultural attractions)-(water bodies)	0.39
Top-10	(cultural attractions)-(water bodies)	0.30
Bellow-top-10	(cultural attractions)-(water bodies)	0.20

Based on Table 5, we group the always-high correlating object types. To build a group from two object types *A* and *B*, we merge their attributes. The set of these higher level categories can be seen in Table 6.[6] For the attributes that exist in both object types, we sum their counts from both types and keep only one occurrence of the attributes. The attributes then are ranked according to their popularity.

The correlation between groups of object types from the same group is given in Table 7. In this table, we report comparisons based on three different sets of attributes: *all*, *top-10*, and below *top-10*. *All* is the entire set of attributes shown in Table 3 (attributes from columns top-10 and below top-10). The *top-10* attributes include those that are shared between most of the object types and thus are more likely to render similarity. The *below top-10* set contains all remaining attributes, which are rather object-type specific, as previously assumed.

Urban object types are grouped into the following three categories: *religious places* (church, basilica, cathedral, abbey, temple), *cultural attractions* (museum, opera house, gallery), and *buildings* (tower, house, building, residence, palace). From Table 7, we see that a high correlation exists between each of these groups when *all* attributes are taken into consideration. *Religious places* and *buildings* correlate with $\tau = 0.71$. The correlation between *religious places* and *cultural attractions* is $\tau = 0.71$, and between *cultural attractions* and *buildings*, $\tau = 0.67$. When the attribute set is split into *top-10* most frequent attributes and the remainder, we see that *religious places* correlate with *buildings* more highly in the *top-10* attributes than in the remainder while their correlation with *cultural attractions* is spread evenly between both attribute sets. *Cultural attractions* and *buildings* are correlated most highly when the entire set of attributes (*all*) is taken into consideration. For both the *top-10* attributes and the remainder, the correlation is lower than for the full set of attributes, but the correlation in the *below top-10* set of attributes is higher than in the *top-10* attributes.

Based on Table 5, the rural object types are grouped into two groups: *water bodies* and *mountainous areas*. These two groups correlate moderately with $\tau = 0.47$. However, the results in Table 7 show a high correlation ($\tau = 0.74$) in *top-10* attributes while the remaining attributes are not correlated ($\tau = 0.29$).

---

[6]Although there exist different clustering hierarchies for geographical places such as the one used in the CORINE Land Cover project (http://www.eea.europa.eu/publications/COR0-landcover), we generated our clusters of object types based on their look, design, and functionality.

From these results, we can conclude that groups of urban object types are generally similar. However, for some group pairs (e.g., *religious places–buildings*), the similarity is substantially higher in more frequently asked information types than in those less frequently asked. In other cases, we see a reverse picture (e.g., *cultural attractions–buildings*) or an equal similarity in both attribute sets (e.g., *religious places–cultural attractions*). Rural object-type groups are only similar when the 10 most frequent attributes are considered. This means that for these groups, the frequently asked information types are of similar importance for all object types within these groups while the remaining information types are object-type specific. This may be due to the fact that for the urban type, we collected 4,815 questions whereas the number of questions for the rural object types was only 2,829. Due to this difference, the correlation between the information users are interested in for the rural objects is lower than that for the urban objects. To analyze whether both object types correlate with each other, we also compared the urban groups with rural ones. From Table 7, we can see that there is, in general, moderate or low correlation between these groups. This indicates that people have different ideas of what attributes to associate with urban and rural

TABLE 8. Percentage of documents in the set of Wikipedia articles that contain the eight most frequent attributes from Table 3.

Object type	Attributes
Beach	location (100%), features (100%), visiting (100%), surrounding (60%), mostattraction (40%), naming (40%), visitors (10%), depth (10%)
Cathedral	location (100%), design (100%), history (70%), foundationyear (50%), visiting (40%), preacher (30%), founder (20%), size (10%)
Cave	location (100%), naming (100%), firstdiscovery (80%), features (70%), visiting (70%), history (60%), mostattraction (50%), surrounding (0%)
Gallery	location (100%), events (100%), arts (90%), foundationyear (80%), visiting (30%), surrounding (10%), design (10%), mostattraction (0%)
Island	location (100%), size (70%), features (50%), habitants (50%), visiting (50%), surrounding (20%), mostattraction (10%), temperature (0%)
Lake	location (100%), features (100%), size (80%), surrounding (70%), visiting (60%), depth (40%), mostattraction (20%), purpose (20%)
Mountain	naming (90%), visiting (90%), height (90%), location (80%), surrounding (70%), size (50%), features (10%), temperature (0%)
Museum	location (100%), features (70%), history (70%), foundationyear (70%), arts (50%), visiting (50%), mostattraction (20%), owner (10%)
Opera House	location (100%), foundationyear (100%), events (70%), designer (70%), owner (70%), surrounding (20%), visiting (20%), mostattraction (0%)
Palace	location (100%), history (100%), visiting (80%), design (70%), foundationyear (40%), owner (30%), size (30%), surrounding (20%)
Park	features (100%), location (90%), size (90%), foundationyear (80%), visiting (70%), mostattraction (40%), surrounding (20%), events (10%)
River	location (100%), length (70%), surrounding (70%), origin (60%), features (60%), status (0%), visiting (0%), purpose (0%)
Stadium	location (100%), capacity (100%), type (100%), foundationyear (80%), events (60%), surrounding (20%), size (20%), visiting (0%)
University	location (100%), history (90%), studentsinfo (70%), size (60%), foundationyear (50%), subjectsofferred (20%), visiting (10%), maintainance (0%)
Volcano	location (100%), surrounding (70%), status (60%), foundationyear (60%), eruptioninfo (50%), height (50%), visiting (30%), naming (30%)

object types. We think that one of the reasons for this low correlation is that humans do use the visual and functional aspects of objects to associate attributes with them. Since urban objects (mainly human-made constructions, e.g., buildings, bridges, do not look like or have similar functions as the rural objects (mainly natural formations, e.g., mountains, rivers), the list of attributes that humans associate with them differ substantially; thus, when compared with each other, they achieve very low correlations.

RQ5: Do human interests correlate to what can be found in existing text documents?

To carry out the analysis, we first randomly selected 15 object types from both the urban and rural categories. Specifically, seven objects were selected from the urban category (*cathedral*, *gallery*, *museum*, *opera*, *palace*, *stadium*, and *university*) whereas eight were chosen from the rural category (*beach*, *cave*, *island*, *lake*, *mountain*, *park*, *river*, and *volcano*). For each object type, we collected 10 Wikipedia articles about objects of this type.

Once the articles were retrieved, three annotators analyzed the sentences of the first paragraph of each article.[7] The objective of the analysis was to determine if the attributes identified as relevant by the MTurk workers also could be found in Wikipedia articles. The most relevant attributes identified by MTurk workers were shown in Table 4, where the eight most frequently addressed attributes for each object type were listed. For each sentence from the first paragraph of Wikipedia articles, it was

[7]We chose to analyze the first paragraph only instead of the entire Wikipedia article to reduce the labor needed for the experiment. Our random analyses of Wikipedia articles showed that the most relevant information is contained in the first paragraph, provided it is of a considerable length. Therefore, we only select Wikipedia articles if their first paragraph is at least 10 sentences or longer.

determined whether that sentence describes one or more attributes from this list, and if so, which one(s). The result of this analysis is shown in Table 8.

The table shows that most of the attributes can be identified in the first paragraph of Wikipedia articles. Approximately half of the attributes for each object type may be identified in the first paragraph of Wikipedia articles. The most common attribute is *location*, which is found in almost all the objects that were analyzed. For urban types, historic and designing information, including the year of foundation, is typically reported. For rural objects, we usually find data about the dimensions of the object and other information concerning the surroundings of the object. However, we also observed that some very specific information for several object types may not always be found. For instance, *temperature* is a detail difficult to find in Wikipedia articles related to *islands* or *mountains*, though it may be interesting according to the MTurk workers. Sometimes this is due to the fact that only the first paragraph in the Wikipedia article is analyzed, but it is usually because the information is not available in the article. Nonetheless, it also may be that the selected Wikipedia articles do not contain such information, but it may be included in other articles concerning the same object type that have not been chosen. To analyze this, we would have to investigate a larger corpus of articles for each object type, which is beyond the scope of this article. Moreover, note that the order of frequencies of the attributes for each object varies with respect to the manual analysis. While the MTurk workers gave preference to the visiting attribute for these selected objects (see Table 4), in Wikipedia articles, we found that this attribute, despite also being relevant, is not always among the two most frequent ones. One reason may be that this kind of information varies over time and must be constantly updated. In Wikipedia articles, we observe that the type of information more frequent for each

TABLE 9. Kendall's τ correlation coefficient between Wikipedia and Mechanical Turk results for each object type.

Object type	Kendall's τ
Beach	0.67
Cathedral	0.37
Cave	0.12
Gallery	0.33
Island	0.56
Lake	0.52
Mountain	0.47
Museum	−0.07
Opera House	0.11
Palace	0.48
Park	0.79
River	−0.07
Stadium	0.22
University	0.14
Volcano	0.37

object type varies depending on the object type itself. For instance, the most frequent attributes for the object type *gallery* are *arts*. This means that articles about galleries focus more on the features of the exhibitions they show rather than on any other type of information (e.g., concerning the entry fees, opening hours, etc.). In contrast, we have found that most of the information in the first paragraph of the Wikipedia articles can be categorized under the most frequent attributes in Table 4. We only found a few exceptions, such as the capacity and dimensions for the object type *stadium* or the characteristics of sea wildlife for the object type *beach*.

After completing this analysis, we also were interested in knowing whether the popularity of the attributes obtained within Wikipedia articles correlates with the popularity of the attributes obtained through the MTurk workers. To do this, we again compute Kendall's τ correlation. Table 9 shows the results of computing the Kendall's τ correlation coefficient between the Wikipedia and the MTurk results for each of the object types analyzed.

The results obtained show acceptable correlation coefficients for most of the object types. Note that the correlation for rural objects is higher, being the object type *park*, the one which has the highest correlation between MTurk workers and Wikipedia articles (τ = 0.79). The lowest correlation can be seen in the *river* and *museum* object types. On average, the correlation for all 15 object types is τ = 0.33, indicating a moderate correlation in the popularity of the attributes. Although the correlation is moderate, our analysis shows that humans have conceptual models about objects of specific types which they use to guide themselves in seeking information or writing about objects of particular types. This was not trivial or obvious when we wanted to test our initial hypothesis since the MTurk workers who participated in the survey may have completely different backgrounds or nationalities that were unknown to us. Consequently, each person may have his or her own interests regarding the sort

of information that he or she would like to recall from an object type. In light of these results, it would be interesting to analyze whether this also applies to other languages and domains.

Discussion

Our analyses have shown that people taking the role of a tourist do share ideas as to what information is relevant for describing geographic objects, and we have identified a set of information types that reflects these ideas. However, although we could say that there clearly are several types of information which people like to know when seeing any geographic object, it also was clear that not every type of information is equally relevant for each object type. These findings prompted the question whether each object type has a specific set of information associated with it or whether object types can be grouped according to how they are described.

We found that some object types are similar and not only share the information types associated with them but also the importance ranking of these types. Such similar objects were mostly objects of like purpose, look, or design (e.g., churches and temples, mountains and peaks, etc.). This indicates that sets of features that people use to categorize objects also play a role in deciding which information should be used to describe these objects. However, some object types that do not share these features are still similar in terms of which information is required to describe them. In these cases, people will refer to the shared set of frequently requested attributes that we identified.

An attempt to group similar object types and investigate similarity between the resulting groups revealed that groups of urban object types are similar. All these objects are objects of the built landscape that can be visited as tourist attractions. For this reason, people taking the role of a tourist are interested in information related to visiting, location, foundation year, and so on. This is not the case for rural object-type groups or for pairs of urban and rural object type groups, as was reflected in our results by the dissimilarity between these object-type groups.

Moreover, we have studied to what extent existing documents (Wikipedia articles) contain the information in which tourists are interested and showed that even though most of the general information is found in the articles, other very specific information for several object types may not always be found (e.g., entry fees or opening hours). Obviously, this specific touristic information is not the focus of Wikipedia articles, and it is more likely to be found in other kind of documents such as travel guides and tourist-information websites and forums.

These results can be used to automatically generate information extraction (IE) style templates (Banko et al., 2007; Banko & Etzioni, 2008; Filatova et al., 2006; Li et al., 2010; Sekine, 2006; Sudo et al., 2003). These templates are descriptions of types of information relevant to a specific domain such as different events reported in the news (e.g.,

terrorist attacks, plane crashes, etc.) or person-specific information. Currently, there exist no templates for descriptions of geographic objects, and the questions collected in the survey could be used to automatically build such templates.

Wikipedia articles also could be categorized by object types to create object-type corpora. Each object-type corpus would contain articles about objects of that object type. For instance, the church object-type corpus would contain all the Wikipedia church articles. Then, from these object-type corpora, different types of information could be extracted, such as IE templates.

This would contribute to populating knowledge databases with geographic information, which could be used to index images pertaining to geographic objects. If information highly relevant for humans is used for indexing, this could lead to better retrieval and organization of those images. In addition, templates can be used in template-based automatic summarization (e.g., for purposes of automatically generating descriptions for images showing geographic objects). Furthermore, the highly relevant information types could be used in a guided summarization task as organized by the DUC and the TAC. In this task, the automated summaries could aim to answer the questions grouped by the attributes relevant for geographic objects.

Furthermore, our results can guide the decision as to which information types should be included in templates. For creating the knowledge database, the full set of attributes is needed to capture the maximal amount of information about object types. However, for indexing and summarization, it is important to know the information types most relevant to a particular object type or groups of similar object types. In this work, we show these information types and indicate which object types can be grouped together.

Evaluation of such automatically created summaries also would profit from the findings reported in this article. Each automatically generated summary about an object could be presented to a user along with a set of questions related to the frequently asked attributes about the object type. The user would be asked to indicate whether the summary answers those questions. This would be similar to the summary content units (SCUs) presented within the pyramid method (Nenkova & Passonneau, 2004). In the pyramid method, however, the SCUs are extracted from human-generated model summaries and presented to the evaluator along with automated summaries. The evaluator is asked whether the automated summary contains that SCU. In our case, the process of extraction of SCUs from reference summaries would be redundant, as the SCUs would be replaced by the questions related to frequently occurring attributes.

Conclusions and Future Work

In this article, we investigated which information types (i.e., attributes) humans associate with geographic objects from urban and rural landscapes. We identified a set of attributes that were relevant for any geographic object type, but also found that an appreciable proportion of attributes are object-type specific. Even in the set of shared information types, not all information types were equally important for each object type. Based on the importance ranking of information types for each object type, we were able to identify similar object types. Furthermore, we showed that urban object types can be grouped together according to what information is used to describe them. We also showed that some of the attributes MTurk workers identified as relevant for the various object types can be found in existing documents such as Wikipedia articles, but others (especially those concerning very specific traveling information) are not likely to be found in this kind of document. We believe that our results can guide tasks involving generation of image descriptions. Furthermore, they can aid assessment of such descriptions as well as image indexing and organization. In addition, these results can be useful for understanding human reasoning (Gordon, Bejan, & Sagae, 2011) and providing some insights for the future development of automatic tools capable of performing reasoning.

Having shown that there may be a conceptual model of geographic objects, our short-term goal is to study and analyze whether such models also can be applied to other languages such as Spanish or German. One limitation encountered in this study is the manual selection of images to be shown to the annotators. Our aim in the future is to broaden this experiment by randomly selecting different images pertaining to object of the same type to verify our findings, and to generalize the experiment, we will not restrict ourselves to Wikipedia documents but also add general documents about tourist places (e.g., information from tourist websites). Furthermore, we aim to perform the future experiments by selecting images pertaining to objects of diverse types rather than restricting ourselves to images pertaining to only locations. In this way, we will be able to collect conceptual models for different types and reuse them for describing images that show, for example, a beach where a person goes for a walk with a dog. We believe that a combination of image features and conceptual models can help to automatically describe such images. The descriptions also can be evaluated by the conceptual models (i.e., to determine if the generated descriptions entail what is said in the conceptual models).

Finally, we believe that our findings can be used as a baseline approach to compare any automatic technique aiming to collect similar pieces of information about locations. The automatic approaches could derive such information from existing resources such as Wikipedia (see Aker & Gaizauskas, 2010) and compare their results with those reported in this article.

References

Aker, A., El-Haj, M., Albakour, D., & Kruschwitz, U. (2012). Assessing crowdsourcing quality through objective tasks. Proceedings of the Language Resources and Evaluation Conference (pp. 1456–1461). Istanbul, Turkey.

Aker, A., & Gaizauskas, R. (2010). Generating image descriptions using dependency relational patterns. Proceedings of the 48th annual meeting of the Association for Computational Linguistics (pp. 1250–1258), Uppsala, Sweden.

Alonso, O., & Mizzaro, S. (2009). Can we get rid of TREC assessors? Using Mechanical Turk for relevance assessment. Proceedings of Special Interest Group on Information Retrieval (SIGIR '09), The Future of IR Evaluation Workshop (pp. 15–16). Boston, MA.

Armitage, L., & Enser, P. (1997). Analysis of user need in image archives. Journal of Information Science, 23(4), 287–299.

Balasubramanian, N., Diekema, A., & Goodrum, A. (2004). Analysis of user image descriptions and automatic image indexing vocabularies: An exploratory study. Proceedings of the International Workshop on Multi-disciplinary Image, Video, and Audio Retrieval and Mining. Quebec, Canada.

Banko, M., Cafarella, M., Soderland, S., Broadhead, M., & Etzioni, O. (2007). Open information extraction from the web. Proceedings of the International Joint Conference on Artificial Intelligence (pp. 2670–2676), Hyderabad, India.

Banko, M., & Etzioni, O. (2008). The tradeoffs between open and traditional relation extraction. Proceedings of the 46th annual meeting of the Association for Computational Linguistics: Human Language Technologies (pp. 28–36), Columbus, OH.

Choi, Y., & Rasmussen, E. (2002). Users' relevance criteria in image retrieval in American history. Information Processing & Management, 38(5), 695–726.

Dakka, W., & Ipeirotis, P.G. (2008). Automatic extraction of useful facet hierarchies from text databases. Proceedings of the Institute of Electrical and Electronics Engineers 24th International Conference on Data Engineering (pp. 466–475), Washington, DC.

El-Haj, M., Kruschwitz, U., & Fox, C. (2010). Using Mechanical Turk to create a corpus of Arabic summaries. Proceedings of the Language Resources and Evaluation Conference Workshop on Semitic Languages (pp. 36–39), Valletta, Malta.

Eysenck, M.W., & Keane, M.T. (2005). Cognitive psychology: A student's handbook. Erlbaum/Psychology Press, United Kingdom.

Feng, D., Besana, S., & Zajac, R. (2009). Acquiring high quality non-expert knowledge from on-demand workforce. Proceedings of the Workshop on The People's Web Meets NLP: Collaboratively Constructed Semantic Resources (People's Web '09) (pp. 51–56), Morristown, NJ.

Filatova, E., Hatzivassiloglou, V., & McKeown, K. (2006). Automatic creation of domain templates. Proceedings of the Joint Conference of the International Committee on Computational Linguistics and the Association for Computational Linguistics on Main Conference Poster Sessions (pp. 207–214), Sydney, Australia.

Gordon, A.S., (2004). Strategy representation: An analysis of planning knowledge. Erlbaum/Psychology Press, United Kingdom.

Gordon, A.S., Bejan, C.A. & Sagae, K. (2011). Commonsense causal reasoning using millions of personal stories. Proceedings of the 25th AAAI Conference on Artificial Intelligence, San Francisco, CA.

Gordon, A.S., & Hobbs, J. (2011). A commonsense theory of mind–body interaction. Association for the Advancement of Artificial Intelligence Spring Symposium: Logical Formalizations of Commonsense Reasoning. March 21–23, Stanford, CA.

Greisdorf, H., & O'Connor, B. (2002). Modelling what users see when they look at images: A cognitive viewpoint. Journal of Documentation, 58(1), 6–29.

Hollink, L., Schreiber, A., Wielinga, B., & Worring, M. (2004). Classification of user image descriptions. International Journal of Human–Computer Studies, 61(5), 601–626.

Jörgensen, C. (1996). Indexing images: Testing an image description template. Proceedings of the annual meeting of the American Society for Information Science, 33, 209–213.

Jörgensen, C. (1998). Attributes of images in describing tasks. Information Processing & Management, 34(2–3), 161–174.

Kaisser, M., Hearst, M.A., & Lowe, J.B. (2008). Improving search results quality by customizing summary lengths. Proceedings of the 48th annual meeting of the Association for Computational Linguistics: Human Language Technologies (pp. 701–709), Columbus, OH.

Kazai, G. (2011). In search of quality in crowdsourcing for search engine evaluation. Advances in Information Retrieval—33rd European Conference on IR Research (ECIR'11) Lecture Notes in Computer Science: Vol. 6611 (pp. 165–176). Amsterdam: Springer.

Kittur, A., Chi, E.H., & Suh, B. (2008). Crowdsourcing user studies with Mechanical Turk. Proceedings of the 26th annual Special Interest Group on Computer–Human Interaction Conference on Human Factors in Computing Systems (pp. 453–456), New York, NY.

Li, P., Jiang, J., & Wang, Y. (2010). Generating templates of entity summaries with an entity-aspect model and pattern mining. Proceedings of the 48th annual meeting of the Association for Computational Linguistics (pp. 640–649).

Marks, D.F. (1973). Visual imagery differences in the recall of pictures. British Journal of Psychology, 64, 17–24.

Mason, W., & Watts, D.J. (2010). Financial incentives and the "performance of crowds." Association for Computing Machinery's Special Interest Group on Knowledge Discovery and Data Mining (SIGKDD) Explorations Newsletter, 11, 100–108.

McKeown, K., & Radev, D. (1995). Generating summaries of multiple news articles. Proceedings of the 18th annual International Association for Computing Machinery Special Interest Group on Information Retrieval (ACM SIGIR) Conference on Research and Development in Information Retrieval (pp. 74–82).

Nenkova, A., & Passonneau, R. (2004). Evaluating content selection in summarization: The Pyramid method. Proceedings of the Human Language Technology/North American Chapter of the Association for Computational Linguistics Conference (pp. 145–152), Boston, MA.

Radev, D., & McKeown, K. (1998). Generating natural language summaries from multiple on-line sources. Computational Linguistics, 24(3), 470–500.

Rosch, E. (1999). Principles of categorization. Cambridge, MA: MIT Press, Concepts: Core Readings.

Sekine, S. (2006). On-demand information extraction. Proceedings of the Joint Conference of the International Committee on Computational Linguistics and the Association for Computational Linguistics on Main Conference Poster Sessions (pp. 731–738), Sydney, Australia.

Snow, R., O'Connor, B., Jurafsky, D., & Ng, A.Y. (2008). Cheap and fast, but is it good? Evaluating non-expert annotations for natural language tasks. Proceedings of the Conference on Empirical Methods in Natural Language Processing (pp. 254–263), Morristown, NJ.

Su, Q., Pavlov, D., Chow, J.H., & Baker, W.C. (2007). Internet-scale collection of human-reviewed data. Proceedings of the 16th International World Wide Web Conference (pp. 231–240), New York, NY.

Sudo, K., Sekine, S., & Grishman, R. (2003). An improved extraction pattern representation model for automatic IE pattern acquisition. Proceedings of the 41st annual meeting of the Association for Computational Linguistics (pp. 224–231), Sapporo, Japan.

Yang, Y., Bansal, N., Dakka, W., Ipeirotis, P., Koudas, N., & Papadias, D. (2009). Query by document. Proceedings of the Second ACM International Conference on Web Search and Data Mining (pp. 34–43), New York, NY.