Sensors **2015**, *15*, 1-x manuscripts; doi:10.3390/s150x0000x

OPEN ACCESS

sensors

ISSN 1424-8220

www.mdpi.com/journal/sensors

Article

A Geospatial Semantic Enrichment and Query Service for Geotagged Photographs

Andrew Ennis [1,]*, Chris Nugent [1], Philip Morrow [2], Liming Chen [3], George Ioannidis [4], Alexandru Stan [4] and Preslav Rachev [4]

[1] School of Computing and Mathematics, University of Ulster, Northern Ireland; E-Mails: ennis-a1@email.ulster.ac.uk; cd.nugent@ulster.ac.uk

[2] School of Computing and Information Engineering, University of Ulster, Northern Ireland; E-Mail: pj.morrow@ulster.ac.uk

[3] School of Computer Science and Informatics, De Montfort University; E-Mail: liming.chen@dmu.ac.uk

[4] IN2 Search Interfaces Development Ltd.; E-Mails: gi@in-two.com; as@in-two.com; pr@in-two.com

[†] These authors contributed equally to this work.

* Author to whom correspondence should be addressed; E-Mail: ennis-a1@email.ulster.ac.uk

Academic Editor:

Received: / Accepted: / Published:

Abstract: With the increasing abundance of technologies and smart devices, filled with a multitude of sensors for sensing the environment around them, such as a GPS receiver, information creation and consumption has become effortless. This is particularly the case for photographs with vast amounts being created and shared every day. Nevertheless, it still remains a challenge to discover the "right" information for the appropriate purpose. This paper describes an approach to create semantic geospatial metadata for photographs which can facilitate photo search and discovery. To achieve this we develop a semantic geospatial data model by which we can enrich a photo's geospatial metadata extracted from several geospatial data sources based on the raw low-level geo-metadata from a smartphone photograph. We also describe our method and implementation for searching and querying the semantic geospatial metadata repository to enable a user or third party system to find the information they are looking for.

Keywords: Geospatial; Semantic; Enrichment; Ontology; Photograph, API

1. Introduction

With the increasing abundance of technologies and smart devices for creating and consuming media, such as smartphones, tablets and smart digital cameras, all of which contain a multitude of sensors for sensing the environment around them, it has become effortless to create vast amounts of information. This is particularly the case for photographs, with vast amounts being created and shared every day. In 2012 the number of smartphones globally exceeded 1 billion and this is expected to double by 2015 [1]. Gartner predicts that by 2015 80% of mobile handsets sold globally will be smartphones and these will outweigh PCs as the most common device to access the web [2]. The ease with which media can be captured and uploaded online results in vast amounts of information being created and stored online daily. There are also an increasing number of online information sources and tools being made publicly available, such as DBpedia [3], Flickr [4] and YouTube [5]. With this information deluge it has become increasingly time-consuming to decipher actionable information upon which informed decision making can be based. This is particularly the case for multimedia content, such as photographs and videos where a means to better organize, categorize and make searchable the generated media is required. Users are subsequently suffering from information overload and struggling to discriminate relevant from irrelevant information. To solve this problem there is a need to have more detailed and useful metadata attached, to facilitate and improve organization and categorization to enable relevant searches. During media capture, limited metadata is attached to the media. Recently, with the increased use of smart digital cameras, Global Positioning System (GPS) coordinates embedded in the metadata can be used for searching and categorization. In addition by analyzing the image pixels it is possible to determine if the photograph was taken indoors or outdoors and to determine the photograph's location based on the photometric effects of shadows [6]. There is, however, a specific lack of semantic geospatial information in the form of metadata. This lack of semantic geospatial metadata restricts the searching capability to the limited existing non-semantic metadata, such as the GPS coordinates in regards to geospatial metadata. Our system however addresses this issue by enriching photographs with semantic geospatial metadata, which in turn allows the system to categorize and make the photographs easily searchable, based on the added semantic geospatial metadata.

The remainder of this paper is organized as follows. Section 2 discusses related work; Section 3 describes our geospatial data model and the Semantic Web Rule Language (SWRL) rules used for inferring further information and geospatial relationships. It also describes the implementation of the data model; Section 4 describes the searching and querying of the semantic geospatial data repository; Section 5 discusses the system validation method; Section 6 discusses the conclusions and future work.

2. Related Work

The majority of related work in the media enrichment research area involves automatic image annotation based on extracted semantic features from the image pixel data.

Ballan *et al* discuss an approach to video annotation and retrieval using ontologies and rule learning [7]. Their approach uses semantic concept classifiers and SWRL to determine what concepts are in the video and then generate annotations based on these concepts. Relevant to work in the current paper is their use of ontologies and rules to automatically determine appropriate annotations. However, their method uses a predefined set of trained concepts to search for in a given image. This limits the annotations to only those that the system has been trained to recognize. In contrast the proposed approach in this paper to geospatial semantic annotation uses the information and concepts extracted from various publicly available datasets to then construct an ontology that enables further semantic inferences to be made based on a set of rules.

Bannour and Hudelot discuss the use of ontologies for image annotation and high-light that in order to enable complete high-level semantic annotations the use of several knowledge sources is needed and inferences must be made across the knowledge sources [8].

Yi discusses the use of ontologies for the fusion of multiple geographic data sources and entity matching [9]. The work describes what information is considered relevant and useful in an ontology based on the different points of view of the different research communities. Given that this makes the process of fusing several ontologies together challenging, the paper proposes a graph model based ontology method to facilitate fusion of multiple geographic ontologies. The approach matches entities, however it lacks the ability to determine geospatial relationships between the entities which is a particular media requirement for our geospatial model.

Taking into consideration related work it is clear that there are still many gaps in media enrichment which have not been addressed, in particular the area of geospatial semantic enrichment of media. This paper proposes to address these issues by developing a way to fuse several geospatial data sources together and model this data such that it can easily be used for searching and retrieval of the media.

3. Geospatial Data Modeling

In our previous work we developed a system to extracted geospatial information from multiple geospatial data sources, including Geonames, DBpedia, Google Places and Open street maps [10], [11]. The types of information extracted from these datasets are latitude, longitude, place name, place feature type and feature description, city, country, elevation, etc. The system, developed in our previous work, uses the GPS coordinates in a photograph to query the datasets and extract the different items of information and fuse the result sets together. This fusion process combines duplicates that may appear in the multiple datasets or combine two places that are similar enough to be assumed the same place, based on given criteria [10], [11]. The criteria of the fusion process involves initially checking if any of the extracted POIs have a reference to each other, such as a “sameAs” attribute. This is in the form of a “sameAs” attribute with a URI to the other POI. If it does have an associated “sameAs” attribute that links to the second POI, then we trust this and classify them as the same. If no “sameAs” attribute is found then we check to see how far apart the two POIs are from each other. The distance threshold will vary depending on the feature type that the POIs have. If a POI has a feature type of “statue” then the spatial footprint will be relatively small. Whereas a POI with a feature type of “building” will have a relatively large spatial footprint. This also means that the POIs have to have the same feature type if they are to be classified as being the same.

In the following sections we discuss the approach taken to model the data extracted from several geospatial data sources and how to automatically infer the relationships between the data and the media.

3.1. Data Model

Once the geospatial information has been extracted and fused, then it needs to be analyzed and interpreted so that an understanding of how the data relates to the media can be established in an automatic way. To do this the data needs to be modeled such that inferences can be made across the data, thus enabling relationships between the extracted data and the media to be determined. For example storing the data and relationships in a semantic description format such as Resource Description Framework (RDF). The model needs to be able to store links back to the original data source in order to enable verification and provenance of the data. Given that some of the data sources are part of the linked data cloud, this link to the original data source can be used by applications as an entry point to the linked cloud. Applications can then further follow the linked cloud links and gather more information beyond the scope of the extracted geospatial data.

For our data model to handle all the criteria described above we choose to model the data in an ontology. This has many benefits and extra features which enables our system to reason over the extracted data and infer relationships between the data. To handle the data model we modified our system architecture, as shown in Figure 1, so that in the Point Of Interest (POI) Handler layer we have the RDF Enrichment Model Generator component, which creates the sets of triples representing the extracted objects and their attributes such as latitude, longitude, place name and place feature (building, statue, etc) [10], [11]. A POI is a significant location that may be of interest and in relation to our system is usually a building, statue or business location. These triples represent subject-predicate-object. For example *POIA hasName Belfast City Hall* and *Photograph isLookingAt POIA*. During the semantic metadata repository population phase, we calculate predefined attributes, such as the distance to each POI from the photograph. This pre-calculation enables faster processing in our semantic geospatial metadata repository later on, particularly when searching. We also then run SWRL rules against the data stored in the semantic metadata repository to determine further information through inferences. Such as what POIs the photograph is looking at or the direction relationship each POI has to the photograph. This then enables searching such as show all POIs to the south of a given photograph, or show all photographs to the east of a given POI.

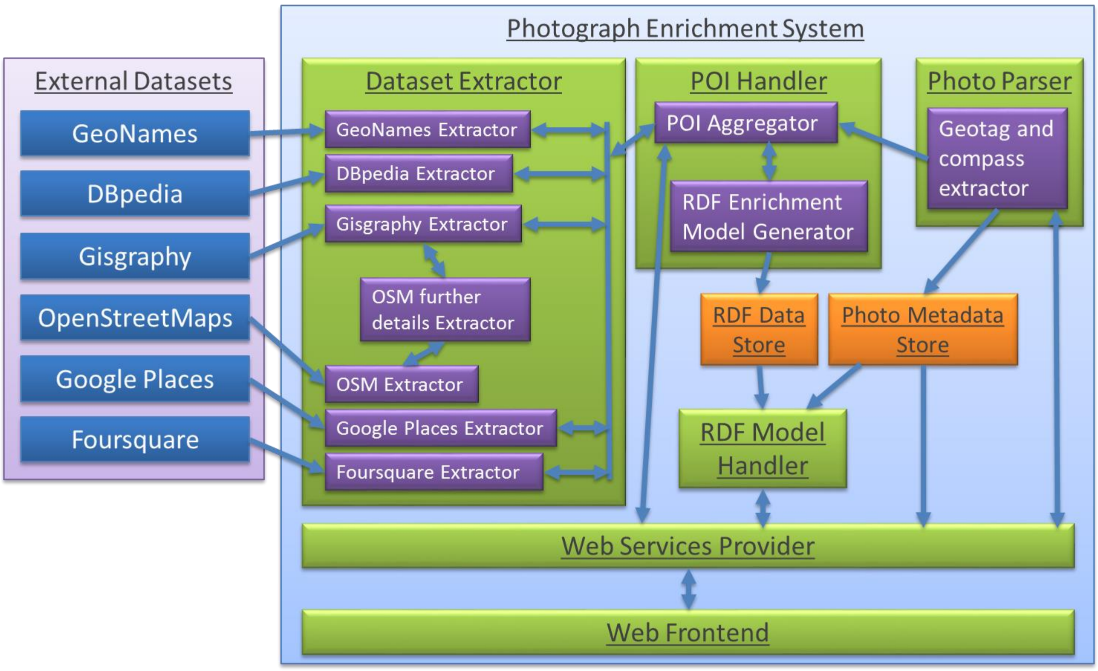

Figure 1 - Diagram showing our system architecture with the external datasets on the left and our photograph enrichment system on the right. It shows each component of our system and how each component interconnects.

3.2. Implementation of Data Model

The semantic geospatial metadata of each photograph will be generated by instantiating the ontological geospatial data model, which will be assigned a unique id so that it can be identified. The id is a 256 bit SHA hash generated from the latitude, longitude and compass of the photograph. This ensures that every photograph will have its own unique id but also ensures that we do not process the same photograph twice. In other words the same place and direction, as we will have already computed the geospatial model for that location. Below is an example of the semantic geospatial metadata in RDF XML format created for a photograph.

```
<rdf:Description rdf:about="http://www.ulster.ac.uk/#0196ebd4...bb2aa0bd">
<rdf:type rdf:resource="http://www.ulster.ac.uk/#photograph"/>
<p:latitude rdf:datatype="http://www.w3.org/2001/XMLSchema#double">
      51.5184661</p:latitude>
<p:longitude rdf:datatype="http://www.w3.org/2001/XMLSchema#double">
      -0.1258781</p:longitude>
<p:hasclosestpoi rdf:resource="http://www.ulster.ac.uk/#3998196a...18076dfa"/>
<p:iseastof rdf:resource="http://www.ulster.ac.uk/#e31e748f...84ab3631"/>
<p:issouthof rdf:resource="http://www.ulster.ac.uk/#4642bdb2...2757bf39"/>
<p:isnorthof rdf:resource="http://www.ulster.ac.uk/#f1ba3f3b...dc7acde8"/>
<p:iswestof rdf:resource="http://www.ulster.ac.uk/#bbd8d6d3...03ba2bde"/>
</rdf:Description>
```

Figure 2 - Snippet of RDF XML created for a photograph by our system. We have abbreviated the 256 bit SHA hash for improved presentation clarity.

For each POI extracted from the datasets we create a "poi" node which is identified with a 256 bit SHA hash. This SHA hash is created from the place name, latitude and longitude. This ensures that each POI node has a unique identifier that each attribute can be linked to. It also enables us to identify if the POI has already been added to the semantic geospatial metadata repository, therefore stopping it from being added twice. Below is a short example of some of the triples that we create for a given POI.

```
<poi> mp:isapoi poi: cbaeaa66…67840a46
poi:85279…284574 mp:name "Buger King Donegall Street"
poi:85279…284574 mp:lat "54.59740006476453"
poi:85279…284574 mp:lon "-5.930394107437821"
poi:85279…284574 mp:address "51-59 Donegall street"
poi:85279…284574 mp:country pl:42635…284605
poi:85279…284574 mp:source "foursquare"
poi:85279…284574 mp:sourceid "4cd8240a2a87a1434240ab09"
poi:85279…284574 mp:foursquareverifiedplace "false"
poi:85279…284574 mp:feature "Fast Food Restaurant"
<photo> mp:isaphoto photo:0196ebd4...bb2aa0bd
photo:0196ebd4...bb2aa0bd mp:islookingat poi: cbaeaa66…67840a46
```

Figure 3 - Snippet of RDF, for a POI, created by our system. We have abbreviated the 256 bit SHA hash for improved presentation clarity.

We also model the administration hierarchy of each of the POIs, such as the city, country and continent that the POI is located within. We do this by using the associated city or country information extracted from the POIs. We then search against the Geonames database to obtain the remaining hierarchical information. For example: if we have the city name, we can query for the country that this city is located in from the Geonames database. We can then in turn query the continent that contains the country. We cannot, however, solely rely on Geonames to provide an accurate country or city name because its data is collected from a variety of sources and errors can appear. To address this we cross-reference with the other datasets. For example if we obtain results from Google Places for the photograph, then we obtain country and/or city information and check if there is a match with the other country information from the other POIs. If a new country is introduced into the semantic geospatial metadata repository that currently does not exist, then we have a conflict and it needs to be resolved. We do this by looking not only at the frequency counts of the country names, but also by looking at higher or lower levels of the administration hierarchy. To determine the most common country from the POIs we use a simple frequency-voting algorithm on the country name, where the country with the most votes from the POIs, is the most likely country name for the given location.

3.3. SWRL Rules

Semantic Web Rule Language (SWRL) is a language used to construct inference rules, which are run against triples to try and infer more semantic relationships. This is particularly useful for our system because it enables our system to infer relationships between a given photograph and many POIs that are not explicitly there in the extracted information. However they can be inferred using the rules and the extracted data.

We apply SWRL rules to the semantic geospatial data repository for collected photographs to infer further information and relationships, in particular that of the POIs to the photograph. Due to the complexity of geospatial calculations we developed several custom functions, called built-ins, for example to calculate compass bearing between two GPS points, distance between two GPS coordinates and POI name similarity. These built-ins then supply values back to the rule which can be used for comparisons or in the result of the rule. An example is shown in Figure 4, where one of the rules is to calculate if the POI is in the direction the photograph was taken. This coupled with the distance from the photograph, means that when our system queries the semantic geospatial metadata repository to determine what the photograph is looking at, these values are already calculated and so the query is much more simplified and also computationally simpler. The result from the rules gets added back into the semantic geospatial metadata repository and so is adding additional semantic context to the photograph, such as how far away the POIs are and in what direction, etc.

```
poi(?p) ∧ photograph(?photo) ∧ latitude(?p, ?poilat) ∧ longi-tude(?p, ?poilon) ∧
latitude(?photo, ?photolat) ∧ longi-tude(?photo, ?photolon) ∧ compass(?photo,
?photocompass) ∧ p1:IsLookingInDirection(?poilat, ?poilon, ?photolat, ?photolon,
?photocompass, 20) → islookingindirection(?photo, ?p)
```

Figure 4 - Example SWRL rule used by our system to determine what POIs the photograph is looking in the direction of.

4. Querying and Searching the Semantic Geospatial Metadata Repository

To make use of our geospatial data model we have implemented a web service that enables third party systems or a user to query and search the semantic geospatial metadata repository for the pieces of information that are of most interest to them. To achieve this searching capability, we store the geospatial annotations for photographs in a semantic repository. Virtuoso was chosen for our semantic repository because it is a popular semantic repository. Our reasoning for using a semantic repository was because our model is based on triples, which describe a semantic meaning of the data. Therefor virtuoso enables us to retain this model and combine the models from many photographs together but importantly we can develop a web service on top to abstract away the underlying structure and provide a simple API for querying the semantic geospatial metadata.

The web service provides two main services. The first enables a photograph to be enriched and the enrichment metadata to be stored for later querying. This is achieved by supplying the GPS coordinates of the photograph and an optional compass heading. This is enough for our system to query the external datasets and create the enrichment semantic geospatial metadat. For example, if we wished to enrich a photograph at the coordinates 54.597N, -5.930E and a compass of 174° and we want the semantic geospatial metadata to be returned, the query for this service would be as follows:

```
http://.../api/enrich/?location=54.5970788,-5.9301246&compass=174
&returnmodel=true
```

In the above *location* is the GPS coordinates of the photograph. *Compass*, is optional, being the heading the photograph was taken at, relative to true north. *Returnmodel*, is optional, being a boolean flag indicating if the model should be returned to the calling application.

The second of the services provided enables querying of the semantic geospatial enrichment repository. This enables a user to specify a location of interest, an optional compass heading relative to true north, optional distance in meters, optional features and an optional free text string.

To facilitate a user in knowing what features are available a service is provided that returns all the features that are currently stored in the repository. This forces a user to search on only the types of data that are actually in the repository, therefore helping to direct them to relevant information. For example if we wished to search for hotel with a free text term inn, and within 500 meters of a photograph at the GPS point 54.597N, -5.930E and a compass heading of 174°, query for this service would be as follows:

```
http://.../api/search/?location=54.5970788,-5.9301246&compass=174&distance=500
&features=hotel&searchtext=inn
```

The above API call could be the scenario where a user is using a photograph to search for a hotel. The photograph could be of a famous monument and they are looking for a hotel within walking distance of the monument. Our search algorithm can also be used in reverse, in other words to find photographs in a particular location, with particular features.

A result set is returned from the search API call containing the details of the POIs that match the supplied criteria. This information can then be displayed to the user, in a user friendly and readable way, to give them a greater understanding about the photograph and where it was taken. An example result would be as follows:


```
{"pois":[{"LatLon":{"lat":54.59272,"lon":-5.93025},"PlaceName":"Holiday Inn
Belfast","Features":[{"FeatureURI":"","FeatureName":"hotel","FeatureDesc":"a building
providing lodging and/or meals for the public","Source":"geonames"}],
"DataSource":"geoname","CountryName":"United Kingdom","CountryCode":"GB",
"IsIn":"Belfast City","City":"Belfast","Elevation":0.0,
"ID":"d7d61b95e2fd3c635b626107ef2bcb24ff6f4fd71e9c834b4b732f5ca8682af1"}]}
```


Figure 5 - JSON object returned from API search query.

Figure 5 shows the JSON object returned from the API query. It contains one POI and its attributes such as place name, its features and feature description, etc. This can then be used by a smartphone app, for example, to display the relevant information back to the user.

Information about related streets and open areas or parks can also be returned if the search criteria have selected this type of information to be returned.

4.1. Integration

IN2 Search Interfaces Development Ltd. provide a number of services that enable organizing, searching and displaying multimedia. Two of their services, Followtheplace and Citypulse, focus on photographs and places. Screenshots of these services are shown in Figure 6 and Figure 7.

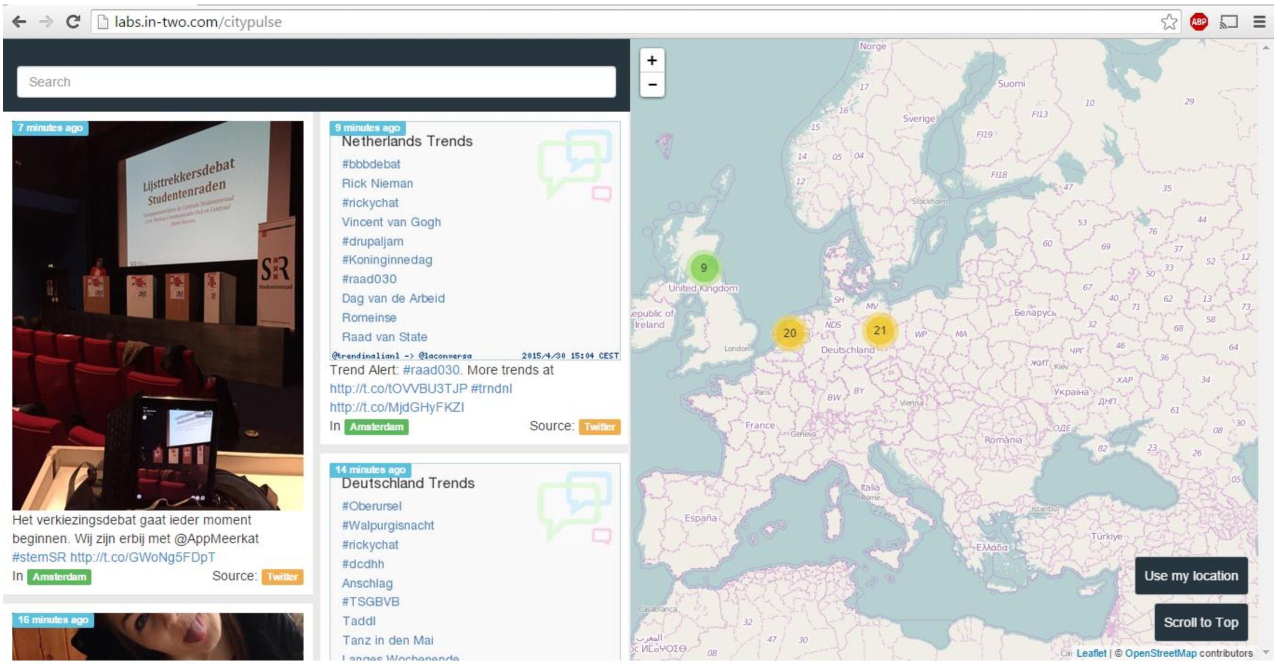

Figure 6 - Screenshot of Citypulse app showing the user interface.

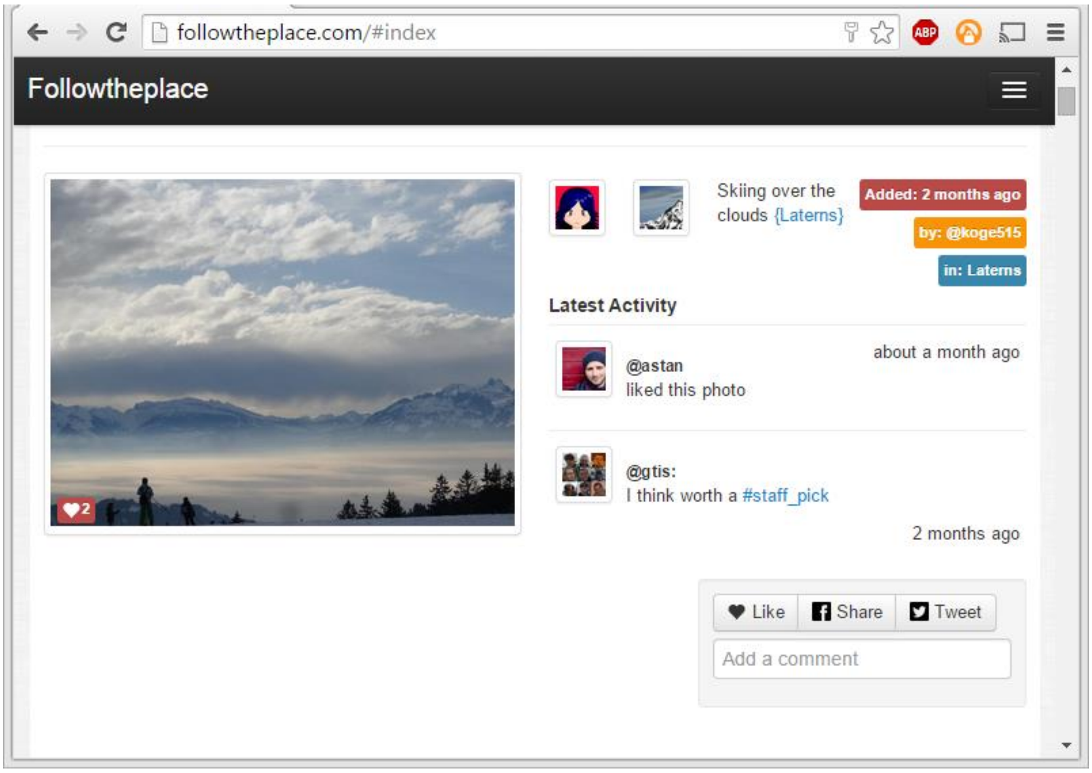

Figure 7 - Screenshot of Followtheplace app showing the user interface.

IN2 are using the API we developed by integrating it with these services to enhance the searching and categorization of digital media that these two services provide. Our API is being integrated into the backend system to enrich the search indexing and the categorization of the photographs with our semantic geospatial metadata. This integration will therefor allow a user to search on the semantic geospatial enrichments to better find photographs and information that is of interest to them. The semantic geospatial metadata will also be used for suggesting similar photographs to the user, by matching photographs on geospatial features such as both are looking at the same statue.

The integration phase is currently underway as of writing this paper. We intend to use this integration with the IN2 products to demonstrate a real world use case of our research and collect usage statistics of how users of the product make use of the geospatial enrichments.

5. System Validation Approach

To verify that our system is producing correct and relevant semantic geospatial annotations, we selected to run a subset of geotagged photographs from the Yahoo Flickr Creative Commons 100M dataset [12]. As per the Yahoo Data Sharing Agreement, this dataset has solely been used for the academic research purposes of producing publishable results to validate our research.

This dataset has been compiled by Yahoo from Flickr's collection of photographs and videos. The dataset does not contain the original photograph pixel data, but is comma separated file containing the metadata of the photograph, such as the geotag coordinates and user created tags. One of the metadata attributes is a URL to the original photograph in Flickr, which can be downloaded if necessary. Since we are only interested in geotagged photographs, we take a subset of photographs that meet the following criteria: must be geotagged, must have user tags, optionally have machine tags and must have a description.

Our justification for this selection criteria is as follows:

- Geotagged: our system requires GPS coordinates. Although this dataset does not contain the compass heading, our system will still work without it.
- User tags: these are what we will compare our systems tags against. A string similarity will be used as the metric.
- Machine tags: if these are available we will compare our systems tags against them. A string similarity will be used as the metric.
- Description: a comparison of the words in the description will also be carried out.

Since our system uses online APIs, there are query restrictions imposed by the providers of the APIs, so we initially intend to further reduce the photograph set to 2500 photographs for the first run. We will then increase the dataset to 5000 photographs in a second run, to see if we get similar statistics. 5000 photographs will produce a sufficient number of results to generate statistics which will show if our system is accurate at determining sensible geospatial annotations for a geotagged photograph.

The statistics that we are intending to gather are what percentage of photographs our system correctly determines geospatial annotations for and the accuracy of the geospatial annotations. This will be achieved by doing a string comparison between the geospatial annotations created by our system and the user tags, machine tags, and description from the Flickr dataset.

We will also look at the precision of our system at determining the main POI the photograph is looking at and on top of this determine the precision of the features our system assigns to the photograph. The precision of the place names assigned by our system will also be compared with the user tags, machine tags and description from the dataset to determine the accuracy of our system in determining where the photograph has been taken. This will include the accuracy of the hierarchy of places, for the photographs were the tags have this data. For example: Donegall Square is in Belfast City, Belfast City is in Northern Ireland, Northern Ireland is in the UK, the UK is in Europe.

6. Conclusions / Future Work

This paper develops a semantic geospatial data model which provides a way to enrich photographs with rich semantic metadata extracted from several datasets. In addition, we develop SWRL rules to infer relationships between the extracted information and a photograph. We also described the development of an API to enable querying and searching of the semantic geospatial enrichment metadata associated to a photograph. The API demonstrates that the semantic geospatial enrichment metadata can be used to find and discover geospatial information about a photograph. The API also demonstrates the ability to search and discover photographs based on geospatial features.

We also discussed how we are using the Yahoo Flickr Creative Commons 100M dataset to gather statistical results that will validate if our system is producing reasonably accurate and precise geospatial enrichment results.

Our future work involves completing the validation testing, as described in our system validation approach, and analyzing and publishing the results.

Acknowledgments

Main text paragraph.

Author Contributions (ONLY for Research Articles)

For research articles with more than one author, authors are asked to prepare a short, one paragraph statement giving the individual contribution of each co-author to the reported research and writing of the paper.

Conflicts of Interest

The authors declare no conflict of interest.

References and Notes

[1] J. Yang, "Smartphones in Use Surpass 1 Billion, Will Double by 2015," 2012. [Online]. Available: http://www.bloomberg.com/news/2012-10-17/smartphones-in-use-surpass-1-billion-will-double-by-2015.html. [Accessed: 29-May-2014].

[2] C. Pettey, "Gartner Identifies the Top 10 Strategic Technology Trends for 2013," 2012. [Online]. Available: http://www.gartner.com/newsroom/id/2209615. [Accessed: 15-May-2013].

[3] DBpedia, "DBpedia," 2013. [Online]. Available: http://dbpedia.org/About. [Accessed: 07-May-2013].

[4] Flickr, "Flickr," 2013. [Online]. Available: http://www.flickr.com/.

[5] YouTube LLC, "YouTube," 2013. [Online]. Available: http://www.youtube.com/. [Accessed: 16-May-2013].

[6] Y.-T. Zheng, Z.-J. Zha, and T.-S. Chua, "Research and applications on georeferenced multimedia: a survey," *Multimed. Tools Appl.*, vol. 51, no. 1, pp. 77–98, Nov. 2010.

[7] L. Ballan, M. Bertini, A. Del Bimbo, and G. Serra, "Video annotation and retrieval using ontologies and rule learning," *IEEE Multimed.*, vol. 20, no. 17, pp. 2–10, 2010.

[8] M. Fernández, I. Cantador, V. López, D. Vallet, P. Castells, and E. Motta, "Semantically enhanced Information Retrieval: An ontology-based approach," *Web Semant. Sci. Serv. Agents World Wide Web*, vol. 9, no. 4, pp. 434–452, Dec. 2011.

[9] S. Yi, "Learning ontologies for geographic entity matching and multi-sources data fusion," *21st Int. Conf. Geoinformatics*, pp. 1–5, 2013.

[10] A. Ennis, L. Chen, C. Nugent, G. Ioannidis, and A. Stan, "A System for Real-Time High-Level Geo-Information Extraction and Fusion for Geocoded Photos," *Proc. Int. Conf. Adv. Mob. Comput. Multimed. - MoMM '13*, pp. 75–84, 2013.

[11] A. Ennis, L. Chen, C. Nugent, G. Ioannidis, and A. Stan, "High-level geospatial information discovery and fusion for geocoded multimedia," *Int. J. Pervasive Comput. Commun.*, vol. 9, no. 4, pp. 367–382, 2013.

[12] Yahoo, "Yahoo! Webscope - Yahoo! Flickr Creative Commons 100M - YFCC-100M." [Online]. Available: http://research.yahoo.com/Academic_Relations. [Accessed: 24-Mar-2015].