GIR'18 Workshop Report 12th ACM SIGSPATIAL Workshop on Geographic Information Retrieval Seattle, USA 6^{th} November 2018

Ross S. Purves[1]
[1] University of Zurich, Switzerland
ross.purves@geo.uzh.ch

Christopher B. Jones[2]
[2]Cardiff University, United Kingdom
c.b.jones@cs.cf.ac.uk

(Workshop Co-chairs)

Geographic Information Retrieval (GIR) addresses challenges of gaining access to documents and their content that relate to geographical locations. The field can be regarded as a fusion of Information Retrieval (IR), which is concerned with information in unstructured documents and associated methods of natural language processing, with Geographical Information Systems, which are oriented to structured data, and to a range of spatial analytical and data access methods that are relevant to GIR. This workshop is the twelfth of a series of workshops, the first of which was held in 2004. The topics covered by the workshops have have included methods for recognising and disambiguating references to place names in text (referred to as geo-parsing); determining the geographic scope of documents; developing gazetteers and ontologies to maintain knowledge of toponyms and geographic concepts; spatio-textual indexing methods that combine inverted file methods with those of spatial database indexing; managing vagueness and uncertainty in geographic terminology; extracting geo-spatial facts and events from documents; and evaluating the performance of geo-information retrieval systems.

The 12th GIR workshop was held on 6th November 2018 at the ACM SIGSPATIAL conference in Seattle, USA. Previous workshops have been held either in combination with the SIGIR and CIKM conferences or as stand-alone events in cooperation with ACM SIGSPATIAL. The stand-alone events have all been located in Europe, in particular in Zurich, Paris and last year in Heidelberg.

At GIR'18 there were 8 presentations, of which 2 were full papers and 6 were short papers. We received 12 submissions. The number of people who registered for the workshop was 14, while the typical attendance was slightly less at around 10 to 12.

The workshop was organized around four sessions, the third of which included a discussion session.

The papers cover a wide range of themes. These included the well-established challenge of geo-parsing of which there were two papers. One of these (Hu) presented a website for evaluating geo-parsers, with support for different corpora and various methods of geo-parsing. Akdemir et al applied deep learning with word embeddings (which characterise words by a representation of their contexts) to English language news documents published in India, which are distinctive in their use of English. Two papers were concerned with corpora and geospatial resources, with Yin et al looking at issues of the geographical variation in characteristics of social media and automatically categorising volunteered resources (with a view to automated quality control), while Chesnokova and Purves described methods for building a corpus that represented personal perceptions of landscape. Two papers in the programme were focused on question answering (QA) systems with Punjani et al using DBpedia, the semantic web (linked data) representation of Wikipedia, in combination with the Geonames gazetteer. The other paper by Mai et al was concerned with the creation of a test dataset for QA based on Point of Interest (POI) data from Yelp, that was evaluated using a machine learning method that again demonstrated

an advantage in using text embedding methods (in their case for sentences). The programme also included a position paper by Adams that encouraged us to push the boundaries of GIR to take more account of processes, relations and inferred thematic concepts.