**Clough, P. (2005). Extracting metadata for spatially aware information retrieval on the internet. *Proceedings of the 2005 Workshop on Geographic Information Retrieval*, Germany, 25-30.**

Keywords: Information retrieval, geospatial information retrieval. SPIRIT

The author discusses approaches to extract and use geospatial information in the Spatially-Aware Information Retrieval on the Internet (SPIRIT) project. Clough (2005) states that geospatial information such as addresses, postal codes, hyperlinks and geographic references from documents can be exploited and used in information systems. Clough contends that this use could provide spatial awareness to transport timetables, routing systems for motorists, map-based web sites and location-based services (e.g. Google Local and Yellow Pages).

The author explains that the extraction of geospatial references from documents involves the identification of geographic references (geo-parsing) and the assignment of spatial coordinates to these references (geo-coding). Clough (2005) points out that the approach adopted for geo-parsing and geo-coding is influenced by the following constraints: (a) speed, (b) reliability, (c) flexibility and (d) multilingualism. Because of the aforementioned constraints, the methods used in the SPIRIT project were based on simple approaches.

After conducting the study, Clough (2005) concludes that both user and system evaluation of the SPIRIT prototype have shown promising results. The author explains further that in the case of SPIRIT, further methods for ranking results help to reduce the effects of incorrect markup. Finally, Clough recommends a number of methods that could be used to improve the operations of the systems.