

# Neogeography: The Challenge of Channeling Large and Ill-Behaved Data Streams

Mena B. Habib,  
Database group, University of Twente, The Netherlands  
[m.b.habib@ewi.utwente.nl](mailto:m.b.habib@ewi.utwente.nl)

Users are not passive recipients. Not only can they choose the information they want, but also they can even produce the information themselves. The term “Neogeography”, is a special case of the more general web phenomenon of “user-generated content (UGC)”, which has a relation to the geographic features of the earth [1]. UGC refers to various kinds of media content, publicly available, that are produced by end-users. Such contents includes: digital video, blogging, mobile phone photography, and wikis. UGC provides citizens, consumers and students with information and knowledge as its contents tend to be collaborative and encourage sharing and joint production of information, ideas, opinions and knowledge among users. In neogeography, end-users are not only beneficiaries but also contributors of geographic information. Neogeography combines the complex techniques of cartography and GIS and places them within reach of users and developers [2].

In this project, our wide objective is to propose a new portable, domain-independent XML-based technology that involves set of free services that: enable end-users communities to express and share their spatial knowledge using free text; extract specific spatial information from this text; build a database from all the users’ contributions; and make use of this collective knowledge to answer - natural language - users’ questions through a question answering service. Also Users can get this information visualized on geographic mapping application like Google Earth.

The proposed system may help some workers’ committees in developing countries to share their experience through mobile phones. For example, the trucks drivers may provide the system with SMS messages about the traffic situation at particular places at a specific time. Structured information about the place, the time and the situation are extracted from these messages, linked with some other GISs and then stored in a spatial DB. Users can benefit from this system by asking about the best way to go to somewhere by sending an SMS question.

The main core of the project involves extracting information from semi or unstructured text. Unfortunately, Information Extraction (IE) from textual sources is challenging in many ways. Information contained in text is often partial, subject to evolution overtime, in conflict with other sources, and sometimes untrustworthy. Also, resolving spatial vagueness in geographic data is another issue. Finally, IE systems are always built for a specific domain, so it is challenging to have a portable IE system that can be applied on different domains with only minor customization.

## References

1. M. F. Goodchild. “Citizens as sensors: the world of volunteered geography”. *GeoJournal*, Vol. 69, No. 4. Pages 211-221. 2007.
2. Andrew J. Turner. “Introduction To Neogeography”. O’Reilly Media, Inc. 2006.