Representing, extracting, mapping, and interpreting movement references in text

Qualitative geographic references in text from field reports, audio transcriptions, human generated route directions, weblogs, and other sources provide potentially important information about movement of entities (people, vehicles, weapons, etc) and about their underlying spatial behaviors. These qualitative geographic references can only be interpreted if put in appropriate context. Human analysts are often able to interpret even vague and imprecise geographic references by inferring the correct context. They also reduce ambiguity and increase interpretation success by making use of more precise artifacts such as maps and images. But there are orders of magnitude more potentially relevant text sources than there are analysts available to extract the explicit and implicit geographic references manually. While progress has been made on geographic information retrieval from text, the progress has been relatively slow and has focused primarily on extracting and disambiguating place names. While that is an important and sometimes hard task, place name extraction is just a small part of the challenge.

Objectives

GeoCAM research focuses on two specific goals:

1. Build the conceptual/theoretical/data model framework needed to represent, extract, map, and interpret geographic accounts of movement found in text.
2. Apply the framework to creating methods, tools, and a geovisual analytics workspace to accomplish these goals.

The RouteSketcher interactively displays parsed routes on a map. This tool supports connection of statements of movement extracted from text directions, plotting of identifiable entities on a map, and exploration of the text-map connections.

The Document Tagger allows users to tag documents with semantic knowledge. The tool supports drag and drop functionality for tag schema reordering as well as AutoMotion-Tagger, a JAVA web service for classifying and extracting motion data from documents.

This study is supported by a grant from the NGA University Research Initiative (NURI) program. The views and conclusions contained in this document are those of the author(s) and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of the National Geospatial-Intelligence Agency (NGA) or the U.S. Government.