FEMAREpViz: Automatic Information Extraction for Geo-temporal Visualization

During emergencies, response personnel are often overwhelmed by vast amounts of information. Processing data manually is not feasible when decision makers are under time pressure. Hence, automatically extracting useful pieces of information from vast sources of textual data is vital for such scenarios. Our system of text processing and visualization consists of three main components, each designed as a standalone Web service to ensure system flexibility and interoperability.

Named entities and entity relations are extracted using FactXtractor to represent the original text. FactXtractor processes text documents using an open source text processing platform (GATE). The output of FactXtractor is a concept map formatted in OWL, which can be visualized with ConceptVISTA.

The primary task for GeoTagger is to resolve ambiguous place names listed in the original text. GeoTagger uses GNIS for determining U.S. locations, NGA GEOnet for determining locations outside the U.S., and Google Earth for representing all locations.

FEMAREpViz is a visualization generation Web service that demonstrates application of our methods to FEMA National Situation Updates. The daily FEMA National Situation Updates contain information from a variety of sources including federal agencies, state and local governments, and news reports. Situation reports generally include location names to indicate where incidents happened, as well as persons or organizations involved in the incidents. FEMAREpViz fetches the daily reports from the FEMA website, and then the reports are processed using FactXtractor and GeoTagger. Processed reports are stored in a repository and can be retrieved by a Web interface. The output is a KML document that provides dynamic updates and interactive visualization, which can be presented with applications such as Google Earth. The system provides an intuitive way to browse and visualize FEMA situation updates.

Since there is no computational approach that can achieve human level accuracy for complex information extraction tasks, visualization could bridge the gap between a fully automated system and manual data processing. Our extraction and visualization system can benefit users who need to analyze massive sets of geo-temporal information in an efficient manner. There are many applications that can be built on top of our system, such as emergency situation pattern analysis and real-time emergency updates monitoring.

Research described here is supported (in part) from the National Visualization and Analytics Center (NVAC), a U.S. Department of Homeland Security Program, under the auspices of the Northeast Regional Visualization and Analytics Center (NEVAC). NEVAC is operated by the Pacific Northwest National Laboratory (PNNL), a U.S. Department of Energy Office of Science laboratory.