

ABSTRACT

Digital information with a spatial component is being generated at an astounding rate, from sources such as Flickr Videos, online news, and “tweets” on Twitter. The ability to identify locations in unstructured text and quickly generate a map unlocks valuable information about the context of the locations in the text. Geoparsing, the process of assigning geographic coordinates or other geographic identifiers to unstructured text, extracts this valuable information from text (Nikolajevs and Jekabsons 2013). Existing studies and tools focus on the challenges of location extraction and disambiguation. These studies do not focus on visualizing the extracted locations, and generally use a simple method of displaying each location as a single point on a map. This thesis examines the current geoparsing text-to-map applications, identifies challenges to generate a map from a text document, and defines an approach to display locations with boundaries and relationships between locations on a map. The outcome of this thesis is a semi-automated geoparsing, data integration, and visualization application to convert the locations in text-based news articles to locations on a map. This approach provides an efficient and effective way to display the spatial context of a text document and allow for interpretations of the data that is not readily apparent from the text by itself.