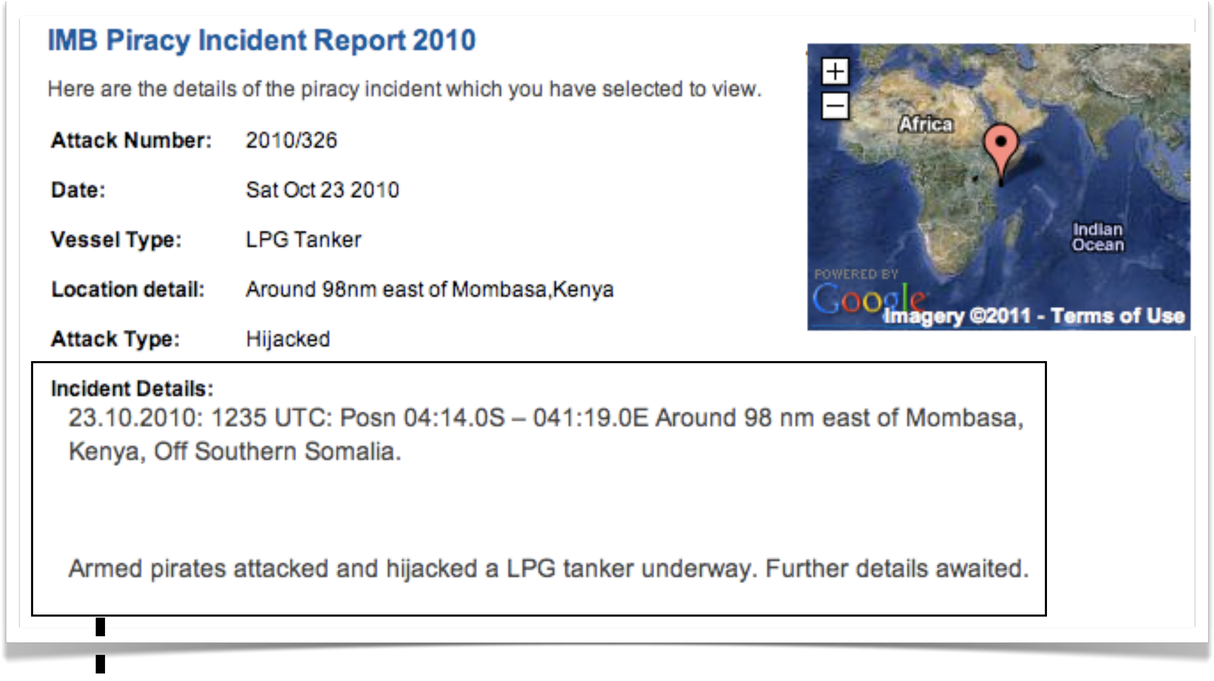
Linked Open Piracy

Willem Robert van Hage, Véronique Malaisé, Marieke van Erp and Guus Schreiber

http://semanticweb.cs.vu.nl/lop/

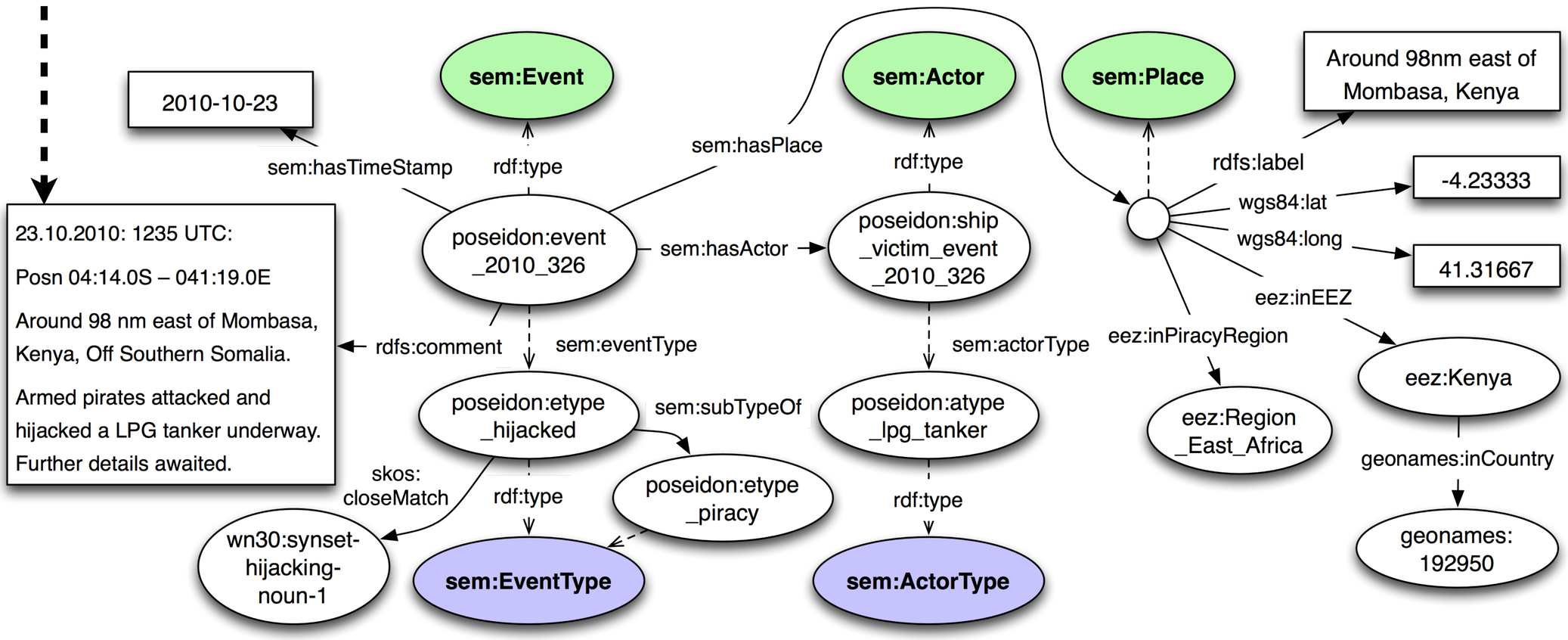


Abstract

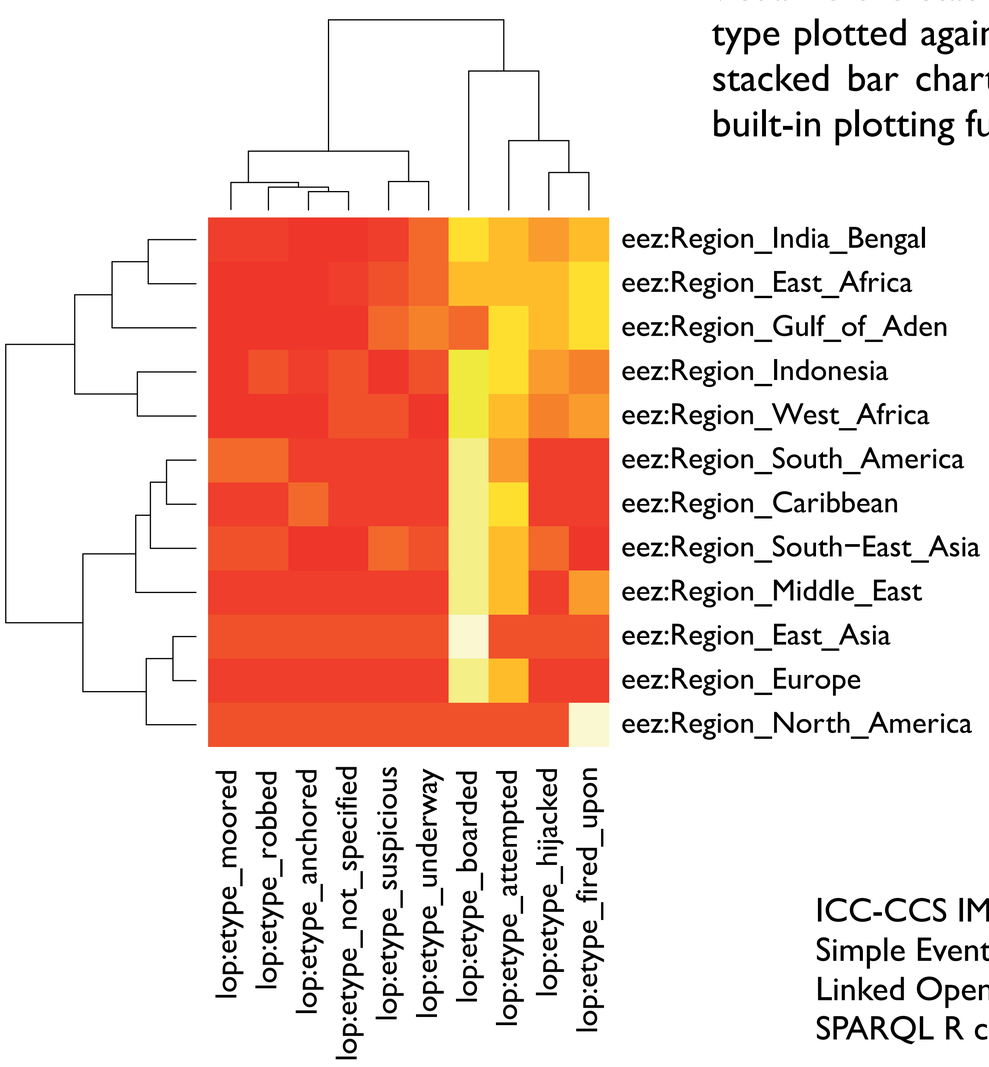
Linked Open Piracy is a Linked Data version of the maritime piracy reports of the International Chamber of Commerce's International Maritime Bureau (ICC-CCS IMB). We create an RDF representation of these event descriptions by screen scraping the ICC-CCS website with XPath and Prolog and representing the events with the Simple Event Model (SEM). The types of the vessels, events, and areas where they take place are mapped to external vocabularies such as Princeton WordNet, FreeBase, and DBpedia. This richer representation of the pirate attacks is easier to process, analyze, and visualize in order to answer domain questions. We host the resulting dataset on a ClioPatria server.

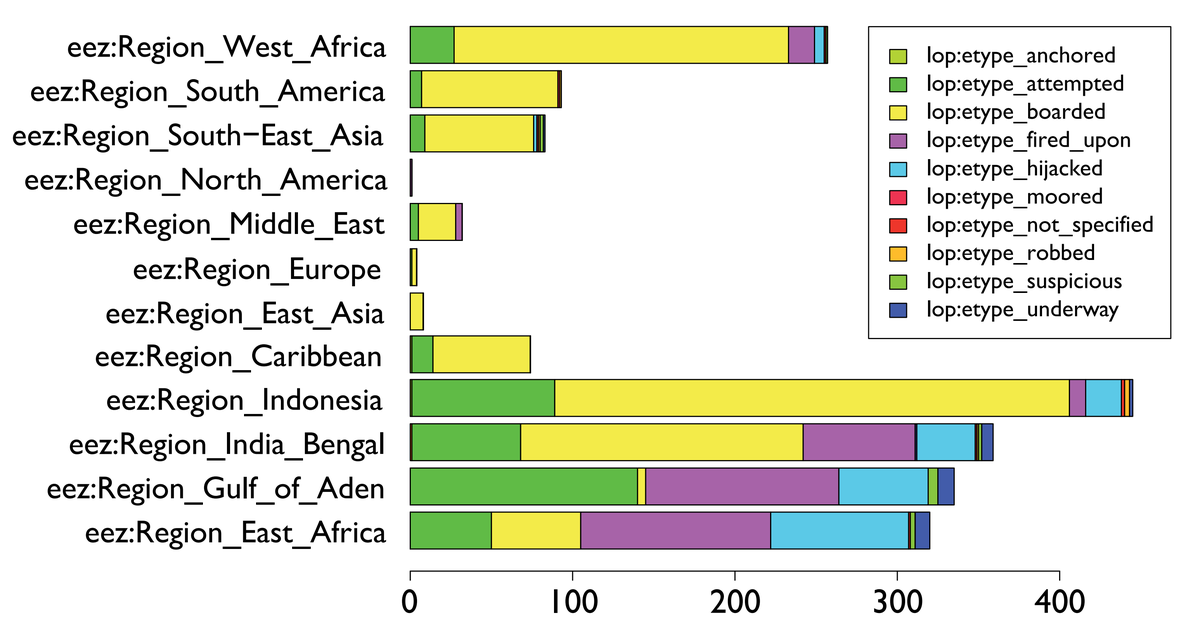
Example XPath expression

//div[contains(@id,"narrations")]/p/text()



The Linked Open Piracy data set now makes it easy to automatically create data plots and visualize the statistics of the events in the data set. Below we show for example the event type plotted against the region as a heat plot with similarity clustering (left) as well as as a stacked bar chart (right). These plots were made with the SPARQL package for R and built-in plotting functions.





ICC-CCS IMB - http://www.icc-ccs.org/piracy-reporting-centre/
Simple Event Model - http://semanticweb.cs.vu.nl/2009/11/sem/
Linked Open Piracy - http://semanticweb.cs.vu.nl/lop/
SPARQL R client - http://cran.r-project.org/web/packages/SPARQL/

VU UNIVERSITY AMSTERDAM