

Proposta di tesi

Creazione di grafi di conoscenza a partire da testi e dati strutturati

Referenti: Rossana Damiano, Marco Antonio Stranisci (Unito), Enrico Daga (Open University, KMI, Open University)

Descrizione: La tesi ha come oggetto la creazione di un grafo di conoscenza a partire da dati testuali e dati strutturati (csv, schemi relazionali), con particolare enfasi sull'estrazione da testi e sulla valutazione della qualità dei dati estratti.

In particolare si sperimenterà la possibilità di ri-utilizzare ed estendere per l'estrazione di dati da testi il Cultural Heritage Cookbook (suite di script disponibile pubblicamente che utilizza LLM e Linked Data, <https://pelagios.org/lm-lod-enriching-heritage/intro.html>) usando come dominio un campione di testi del Museo dell'Automobile.

Requisiti: conoscenza di RDF/OWL e SPARQL, elementi di base di NLP