# Progetti IA2 a.a. 2018/2019 (tracce per un solo studente)

**Traccia 3: Analisi geo-spaziale di testi**

Analizzare un testo, con tecnologie a vostra scelta, alla ricerca esclusivamente di informazioni idonee alla sua collocazione spaziale. Fare una ampia disamina di casi possibili. Definire le modalità di riconoscimento differenziate ed implementare una soluzione per almeno 3 differenti tipologie. Porre l’enfasi sugli aspetti di reasoning.

Si potrebbero utilizzare le estensioni geo-spaziali di Ontotext GraphDB per indicizzare l’informazione spaziale, e permettere di recuperare efficientemente eventi “vicini” oppure “dentro una certa area” così come specificato dall’utente. Si potrebbero utilizzare per lo stesso scopo risorse con informazioni geografiche.

Esempi di frammenti di testo utilizzabili per la geo-localizzazione:

* Il Presidente del Consiglio Conte si è recato in Parlamento……
* Il treno Milano-Roma viaggia con 20 minuti di ritardo
* Solo 9 nazioni europee hanno accettato di ospitare parte degli immigrati sbarcati a Malta
* Si sono scoperte tracce di arsenico nel fiume Olona
* I turisti tedeschi preferiscono trascorrere le vacanze in Italia
* La dieta mediterranea è la più salutare
* ……

Lo studente utilizzi testi a sua scelta.

**Traccia 4: Visualizzazione ed integrazione di open data**

Due aspetti chiave dei Big Data sono la combinazione estemporanea dei dati e l’uso di visualizzazioni appropriate per migliorare la comprensione dei dati stessi. Il concetto di open data torna, in questo caso, di grande utilità fornendo la materia prima su cui lavorare. Spesso questi dati open sono forniti esclusivamente in formati tabulari in cui la semantica del dato non è del tutto esplicitata (se non informalmente attraverso le etichette delle colonne, o attraverso documenti di accompagnamento).

Si chiede di utilizzare dati tabulari della pubblica amministrazione italiana per rispondere a domande di interesse generale per esempio per la programmazione politica di nuove realtà ospedaliere. A titolo di esempio: i posti letto nelle strutture sanitarie sono distribuiti uniformemente nel paese (numero di posti letto/per abitante)? Per rispondere a questa domanda, si potrebbero usare i dati sui posti letto forniti dal Ministero della Salute ed i dati sulla popolazione forniti dall’ISTAT.

I dati tabulari utilizzati dovranno essere preventivamente convertiti in RDF, effettuando l’identity resolution. Nella conversione si chiede di usare dove possibile ontologie esistenti e creare collegamenti a data set di uso comune (es. Wikidata, DBpedia, …)

Nel caso di esempio, si può anche usare le visualizzazioni di Ontotext GraphDB per mostrare su una mappa il dato derivato, in modo da permettere di capire meglio la distribuzione delle risorse sanitarie.