

Oggetto: **Relazione scientifica finale - Short Term Mobility
2013**

Nel corso della mia visita presso il Laboratoire d'Informatique de Paris-Nord dell'Università Paris 13, svoltasi dall'11 aprile al 2 maggio 2013, ho collaborato con il gruppo di "Representation de la Connaissance et Langage Naturel" (RCLN) allo sviluppo di un metodo per l'interpretazione automatica della semantica dei link ipertestuali.

In particolare dopo un'analisi approfondita del problema sono stati individuati un insieme di algoritmi e risorse da sviluppare e/o riusare.

L'obiettivo finale del lavoro di ricerca è, dato un link ipertestuale e il testo che circonda il link nel documento HTML corrispondente, analizzare automaticamente tale testo (tipicamente una frase) allo scopo di estrarre la semantica del collegamento.

Nelle tre settimane di soggiorno i risultati ottenuti sono i seguenti:

- Un algoritmo e un software che lo implementa per il *pre-processing* di frasi (associate a link) allo scopo di produrne una versione adeguata all'analisi successiva;
- Una risorsa *campione* su cui eseguire esperimenti e test di valutazione, basata su un sottoinsieme di DBpedia (versione RDF di Wikipedia);
- Un *workflow* di componenti software che sono stati integrati per produrre una rappresentazione formale (grafi RDF) di frasi espresse in linguaggio naturale;
- Un insieme di euristiche che consentano l'estrazione di sotto-grafi dai grafi rappresentanti le frasi associate ai link, rilevanti allo scopo di produrre le relazioni semantiche per la classificazione dei link stessi;
- Un prototipo software che supporterà la valutazione dei risultati basata su uno studio con utenti.

La collaborazione è tuttora attiva e il lavoro descritto sopra sta ulteriormente avanzando. Attualmente sono in visita nuovamente presso il LIPN, presso lo stesso gruppo, come ricercatore invitato a spese dell'Università Paris 13 con un programma simile alla STM del CNR.

Data 1 ottobre 2013

Firma

