

Migrazione di tecnologie e linguaggi

Teorie e pratiche per il trattamento di testi nelle scienze umane

11 dicembre 2014

ore 10-13

Aldo Gangemi, ISTC

Il *machine reading* come metodo per collegare gli archivi testuali al Web Semantico

Il *machine reading* è un paradigma di estrazione dell'informazione (*Open Information Extraction*) che si applica in generale (cioè non è vincolato a un dominio) e non ha bisogno di esempi su cui addestrare un classificatore (cioè non è supervisionato). Recentemente, la ricerca sul Web Semantico e l'ingegneria della conoscenza ha cominciato a ibridarsi con gli strumenti e le risorse sviluppate dalla comunità di trattamento delle lingue naturali, sviluppando nuovi strumenti che aggiungono al paradigma del *machine reading* anche la capacità di estrarre l'informazione dal testo, formalizzandola ed eseguendo inferenze automatiche. Si parla dunque di *estrazione della conoscenza* (*Open Knowledge Extraction*). Il seminario descriverà questa evoluzione della ricerca e come gli strumenti attuali possono essere usati per la produzione di dati strutturati e collegati a dati esistenti (*Linked data*) a partire da archivi testuali. Il problema di ragionare automaticamente su una grande quantità di conoscenza estratta automaticamente da testi fra loro collegati in serie e/o dinamicamente sarà anche oggetto di discussione.

Presiede

Irene Pedretti, Archivio Storico PUG, Roma

Il calendario completo dei seminari è disponibile alla pagina:
www.iliesi.cnr.it/seminari_formativi

Su richiesta viene rilasciato un attestato di frequenza

CNR, Aula Bisogno
Roma, Piazzale Aldo Moro, 7

Seminari tecnici formativi 2014

Progetto scientifico
Cristina Marras, Ada Russo

Segreteria

Pamela Barletta

pamela.barletta@iliesi.cnr.it

tel. +39 06 4991 7214

ILIESI-CNR

Via Carlo Fea, 2 - 00161 Roma

www.iliesi.cnr.it