

# Migrazione di tecnologie e linguaggi

## Teorie e pratiche per il trattamento di testi nelle scienze umane

**16 ottobre 2014**

ore 10-13

**Felice Dell'Orletta, ILC-CNR**

**La piattaforma T2K:  
dal testo alla conoscenza**

La necessità quotidiana di accedere a grandi quantità di dati per lo più testuali ha dato grande impulso allo sviluppo di tecnologie per l'acquisizione, la classificazione e la gestione automatica del contenuto testuale e al loro sempre più diffuso impiego in una miriade di contesti applicativi.

T2K è una piattaforma web finalizzata all'acquisizione di informazione semanticolessicale da corpora di dominio. Attraverso l'uso combinato di tecniche statistiche e di strumenti avanzati di Trattamento Automatico del Linguaggio, T2K è in grado di analizzare il contenuto dei documenti, estrarre i termini e le entità (nomi, luoghi, persone) potenzialmente più significativi, individuare le relazioni che legano termini e entità, indicizzare il corpus di partenza e ricostruire una mappa multidimensionale della conoscenza di dominio contenuta nella collezione documentale.

Presiede

**Gian Carlo Fedeli, ILIESI-CNR**

Il calendario completo dei seminari è disponibile alla pagina:  
[www.iliesi.cnr.it/seminari\\_formativi](http://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](mailto:pamela.barletta@iliesi.cnr.it)

tel. +39 06 4991 7214

**ILIESI-CNR**

Via Carlo Fea, 2 - 00161 Roma

[www.iliesi.cnr.it](http://www.iliesi.cnr.it)

# 16 ottobre 2014

ore 10-13

**Felice Dell'Orletta**, ILC-CNR

## La piattaforma T2K: dal testo alla conoscenza

La necessità quotidiana di accedere a grandi quantità di dati per lo più testuali ha dato grande impulso allo sviluppo di tecnologie per l'acquisizione, la classificazione e la gestione automatica del contenuto testuale e al loro sempre più diffuso impiego in una miriade di contesti applicativi.

T2K è una piattaforma web finalizzata all'acquisizione di informazione semanticolessicale da corpora di dominio. Attraverso l'uso combinato di tecniche statistiche e di strumenti avanzati di Trattamento Automatico del Linguaggio, T2K è in grado di analizzare il contenuto dei documenti, estrarre i termini e le entità (nomi, luoghi, persone) potenzialmente più significativi, individuare le relazioni che legano termini e entità, indicizzare il corpus di partenza e ricostruire una mappa multidimensionale della conoscenza di dominio contenuta nella collezione documentale.

Presiede

**Gian Carlo Fedeli**, ILIESI-CNR