

Calendario del corso

Analisi e trattamento automatizzato del dato testuale

che si terrà dal 3 al 7 maggio 2010 dalle ore 15,00 alle 19,00

docente Dott.ssa Valentina Talucci

1ª Giornata

Introduzione all'analisi testuale: quali testi analizzare e come organizzare il testo; ottenere i primi risultati: lessici di frequenza, concordanze e specificità

- Esercitazione

2ª Giornata

Il trattamento del testo: procedure tipiche di analisi testuale e text mining. il contributo delle risorse linguistiche all'analisi testuale; parsing, normalizzazione, calcolo delle misure lessicoquantitative, estrazione del vocabolario, estrazione dei segmenti ripetuti.

- Esercitazione

3ª Giornata

L'estrazione dell'informazione: individuare parole chiave, tagging tematico e tagging semantico; analisi delle specificità; analisi delle concordanze; confronto con lessici di frequenza, esportazione di file per altri software.

- Esercitazione

4ª Giornata

L'analisi delle corrispondenze lessicali.

- Esercitazione

5ª Giornata

Predisposizione degli output.

- Prova finale