



UNIVERSITÀ DI PISA LINGUISTICA COMPUTAZIONALE II

GIULIA VENTURI

Anno accademico	2022/23
CdS	INFORMATICA UMANISTICA
Codice	513LL
CFU	6

Moduli	Settore/i	Tipo	Ore	Docente/i
LINGUISTICA COMPUTAZIONALE II	L-LIN/01	LEZIONI	36	FELICE DELL'ORLETTA SIMONETTA MONTEMAGNI GIULIA VENTURI

Obiettivi di apprendimento

Conoscenze

Il corso si propone di introdurre lo studente a settori chiave della Linguistica Computazionale caratterizzati da un forte impatto applicativo. In particolare, si articola attorno a due macro-temi, l'annotazione linguistica multi-livello del testo e l'estrazione di conoscenza linguistica da basi documentali annotate, che sono affrontati da una duplice prospettiva, teorica e applicativa.

Capacità

Al termine del corso lo studente saprà utilizzare in modo critico e consapevole strumenti di annotazione linguistica automatica e di estrazione di conoscenza linguistica, identificare le problematiche legate al trattamento di varietà non-standard della lingua e ipotizzare possibili soluzioni.

Prerequisiti (conoscenze iniziali)

Nozioni di base di linguistica computazionale, di linguistica generale e di linguistica italiana.

Indicazioni metodologiche

Durante il corso si alterneranno lezioni frontali, con l'ausilio di slides powerpoint che vengono messe a disposizione degli studenti, ed esercitazioni di laboratorio, sia individuali sia di gruppo (svolte con PC delle aule informatiche e/o PC personali), in cui gli studenti sono invitati a confrontarsi con l'applicazione di strumenti software di annotazione linguistica del testo e di estrazione di conoscenza disponibili come demo online e ad analizzarne criticamente i risultati ottenuti in relazione a diverse varietà d'uso della lingua.

Programma (contenuti dell'insegnamento)

I contenuti del programma potranno subire variazioni e/o integrazioni, che verranno comunicate durante la prima lezione del corso.

I contenuti del corso sono suddivisi in due macro-temi, per ciascuno dei quali segue una lista dei principali argomenti trattati:

- **Annotazione linguistica**
 - annotazione linguistica come processo incrementale; strumenti software per l'annotazione linguistica del testo; schemi di annotazione per l'annotazione morfo-sintattica e sintattica, con particolare attenzione allo schema delle "Universal Dependencies"; costruzione di corpora annotati e valutazione dell'annotazione; adattamento al dominio o altre varietà d'uso della lingua (es. storiche) di strumenti di annotazione;
- **Estrazione di conoscenza linguistica**
 - ricostruzione del profilo linguistico di collezioni di testi; monitoraggio linguistico di diverse tipologie testuali e/o varietà d'uso della lingua; uso dei risultati del monitoraggio linguistico all'interno di diversi scenari applicativi, ad esempio per la classificazione di generi testuali o per l'identificazione della lingua materna di produzioni L2; analisi della leggibilità del testo.

Bibliografia e materiale didattico

L'elenco dei testi d'esame è disponibile alla pagina [Moodle](#) del corso. Gli studenti non frequentanti sono pregati di contattare i docenti per concordare il programma d'esame.



UNIVERSITÀ DI PISA

[Indicazioni per non frequentanti](#)

Contattare i docenti per concordare il programma d'esame.

[Modalità d'esame](#)

Le modalità d'esame verranno comunicate durante la prima lezione del corso.

Ultimo aggiornamento 12/05/2023 14:56