**Informatica umanistica (Linguistica computazionale)**

Proff. Guido Milanese, Alessandra Lombardi

**Primo modulo: Prof. Guido Milanese**

***OBIETTIVO DEL CORSO***

Il corso si propone di fornire le competenze di base (1) per analizzare testi di varia natura con l’ausilio del computer e (2) per la scrittura scientifica strutturata.

***RISULTATI DI APPRENDIMENTO ATTESI***

Terminato il corso, ci si attende che lo studente sia in grado di:

1. comprendere il rapporto tra informatica, linguistica e filologia;
2. essere in grado di invididuare il tipo di strumento adatto per le varie tipologie di analisi linguistiche e letterarie.

***PROGRAMMA DEL CORSO***

1. Introduzione storica alla filologia e alla linguistica computazionale.
2. Concetti di hardware, software: definizioni di «informatica umanistica».
3. Software libero e proprietario.
4. Digitalizzazione e OCR.
5. Concetto di «marcatura» e di testo digitale (XML, TEI, Markdown).
6. Banche dati testuali e bibliografiche.
7. Il testo analizzato al computer (livelli di analisi).
8. Scrittura strutturata con LaTeX / LyX, Markdown, BibTeX.

## Strumenti informatici che verranno presentati:

* programmi di OCR, indicizzazione e concordanze
* programmi per lemmatizzazione e analisi morfologica
* LaTeX e programmi collegati (LyX, BibTeX e Jabref in particolare).
* Markdown
* Strumenti per l’analisi filologica

***BIBLIOGRAFIA***

Oltre agli appunti del corso: G. Milanese, *Filologia, letteratura, computer*, Milano, Vita e Pensiero 2020 [Acquista da V&P](https://librerie.unicatt.it/scheda-libro/guido-milanese/filologia-letteratura-computer-9788834337516-549756.html)

***DIDATTICA DEL CORSO***

Le lezioni si svolgeranno con l’uso del computer. Chi dispone di un computer mobile (notebook, netbook, tablet…) è invitato a servirsene durante le lezioni.

***METODI E CRITERI DI VALUTAZIONE***

Occorre mostrare la propria capacità di utilizzare direttamente a computer *uno almeno* degli strumenti illustrati a lezione, e di rispondere a domande sulla parte teorica del corso.

***PREREQUISITI***

Il corso ha un carattere introduttivo e non richiede competenze pregresse. La pratica d’uso del computer è sufficiente.

***AVVERTENZE***

Orario e luogo di ricevimento degli studenti sono sempre comunicati con avviso sulle pagine Blackboard e di avvisi del docente rese disponibili in piattaforma. Nei periodi di lezione l’orario normale è dopo la lezione in via Trieste 17, II piano. Per contatti a distanza, utilizzare l’indirizzo di e-mail guido.milanese@unicatt.it seguendo le indicazioni per i messaggi di mail offerte nella pagina web del docente. Si prega di leggere anche le “FAQ”.

**Secondo modulo: Lessicografia computazionale**

**(Prof.ssa Alessandra Lombardi)**

***OBIETTIVO DEL CORSO***

Il modulo si prefigge di introdurre i concetti-chiave della lessicografia, tracciandone l’evoluzione (dal cartaceo al digitale) e descrivendone i principali ambiti di applicazione. Attraverso una presentazione ragionata delle diverse tipologie di risorse lessicali online attualmente disponibili, delle problematiche e delle prospettive di sviluppo ad esse correlate, ci si propone di fornire gli strumenti interpretativi e operativi necessari a un utilizzo critico e consapevole delle stesse. Particolare attenzione sarà riservata ai sistemi che consentono l’acquisizione automatica, il trattamento e la rappresentazione sistematica di combinazioni lessicali ricorrenti (collocazioni) estratte da corpora.

***RISULTATI DI APPRENDIMENTO ATTESI***

Al termine dell’insegnamento lo studente sarà in grado di:

1. descrivere in modo pertinente e terminologicamente corretto gli aspetti macro- e microstrutturali delle risorse lessicografiche tradizionali, nonché le principali funzionalità delle risorse digitali più utilizzate ai fini dell’apprendimento linguistico o per la traduzione (con riferimento all’italiano e alle lingue straniere studiate);
2. spiegare, almeno nelle sue linee fondamentali, l’evoluzione della teoria e prassi lessicografica nel corso del tempo e con riferimento a diverse tradizioni linguistico-culturali;
3. produrre riflessioni critiche sulle possibilità di ricerca, consultazione e fruizione offerte da diverse tipologie di risorse lessicali online;
4. sviluppare strategie mirate ed efficaci per la ricerca di informazioni lessicali in rete e per la costruzione di una conoscenza lessicale adeguata allo svolgimento di diversi compiti di mediazione linguistica.

***PROGRAMMA DEL CORSO***

1. La lessicografia: fondamenti teorici e applicativi.

2. Evoluzione del dizionario: dalle origini all’era digitale.

3. Strumenti di lessicografia computazionale: tipologie a confronto

a. risorse lessico-ontologiche (basi di conoscenza): thesauri, glossari e ontologie informatiche;

b. dizionari macchina (es. di applicazioni: lessici di frequenza; sistemi di traduzione automatica e assistita);

c. prodotti di lessicografia elettronica:

dizionari informatizzati

dizionari elettronici

dizionari collaborativi

aggregatori di fonti lessicografiche;

4. L’impatto dei corpora sulla prassi lessicografica: *corpus-based dictionaries; corpus query tools* (l’acquisizione di informazioni lessicali attraverso l’analisi di corpora: contesti, concordanze, *word sketches*).

***BIBLIOGRAFIA***

a. Manuale di riferimento: Granger S. & Paquot M. (2013), *Electronic Lexicography,* Oxford University Press

b. Per la parte relativa alla storia della lessicografia: Hanks, P. 2012, “*Lexicography from Earliest Times to the Present*” (lightly edited version of a chapter first published in Keith Allan (ed., 2013), The Oxford Handbook of the History of Linguistics. Oxford: Oxford University Press, p. 503-536.)

http://www.patrickhanks.com/uploads/5/1/4/9/5149363/2012d-lexicography\_from\_earliest\_times.pdf

c. altri 3 saggi monografici, di recente pubblicazione, per l’approfondimento delle tematiche trattate a lezione, i cui titoli verranno comunicati alla fine delle lezioni e resi disponibili su Blackboard.

d. Materiale integrativo (consigliato per tutti, ma obbligatorio per gli studenti *non frequentanti*): *video-relazione* di Lars Trap-Jensen al XVIII EURALEX International Congress *Lexicography in global contexts* (Ljubljana 2018): “Lexicography between NLP and linguistics: aspects of theory and practice”:

http://videolectures.net/euralex2018\_trap\_jensen\_lexicography/

***DIDATTICA DEL CORSO***

Lezioni frontali ed esercitazioni pratiche in laboratorio informatico. Tutti i materiali didattici utilizzati a lezione saranno resi disponibili su Blackboard.

***METODO DI VALUTAZIONE***

L’esame finale terrà conto delle prestazioni riferite a tre diverse aree di valutazione:

1. padronanza delle nozioni di base (lo studente deve dimostrare di aver acquisito i concetti-chiave della lessicografia, di avere cognizione della sua evoluzione storica, dei fondamenti teorici e delle principali applicazioni computazionali, nonché delle problematiche e prospettive di sviluppo relative alla produzione e alla fruizione di risorse lessicografiche online);

2. capacità di riconoscere e descrivere, in modo critico e argomentato, scopi, struttura e funzionalità di una risorsa lessicografica digitale proposta dalla docente (e selezionata fra quelle illustrate durante il corso);

3. presentazione in sede d’esame, e relativo commento ragionato, di un progetto individuale fra quelli suggeriti a lezione (da elaborare secondo i criteri metodologici indicati dalla docente e resi disponibili su Blackboard, da inviare via mail 15 giorni prima dell’appello stabilito).

Essendo il corso suddiviso in due moduli, il voto conseguito per la parte di Lessicografia computazionale (prof.ssa Lombardi) costituirà, nella misura del 50%, parte della valutazione finale attribuita dal prof. Milanese.

***AVVERTENZE E PREREQUISITI***

Ai fini di una proficua frequenza delle lezioni non sono richiesti prerequisiti particolari se non un interesse proattivo per l’ampliamento delle conoscenze lessicali e delle competenze di metodo nell’ambito della cd. *digital literacy.*

*Orario e luogo di ricevimento degli studenti*

Durante il periodo di lezione la prof.ssa Lombardi riceve gli studenti nel suo studio secondo l’orario indicato nella pagina docente (eventuali variazioni saranno comunicate nella sezione Avvisi della stessa). Per qualsiasi informazione la si può contattare tramite la mail istituzionale: alessandra.lombardi@unicatt.it.