

## COURSE SYLLABUS

### **Web and Social Networks Search and Analysis**

2324-3-E311PV030

---

#### **Obiettivi**

Il corso si propone di fornire agli studenti i principali concetti alla base della gestione dei dati provenienti dal Web e dai social media, dalla raccolta, alla modellazione, alla successiva analisi. In particolare, gli studenti saranno in grado di recuperare e archiviare dati dal Web sociale, sia attraverso l'uso di API che di *scraping*, di utilizzare rappresentazioni avanzate (sia topologiche che semantiche) e di analizzare e visualizzare strutture di rete complesse e le relative analisi. Una parte del corso si concentrerà in particolare sui concetti di *Web search* e *social search* e sullo studio dei modelli e delle dimensioni di rilevanza più appropriate nel contesto del Web sociale.

#### **Contenuti sintetici**

- Introduzione al Web sociale;
- I dati nel Web e nei social media;
- Rappresentare reti di interazioni complesse: teoria dei grafi e delle reti;
- *Social Network Analysis* (SNA): metriche e algoritmi per l'identificazione di comunità;
- *Social Content Analysis* (SCA): analisi della soggettività e del sentiment, rilevamento dell'ironia, riconoscimento di entità e *linking*;
- *Web e Social search*: i modelli di IR nel Web e nelle reti sociali, le dimensioni di rilevanza, la valutazione dei risultati di ricerca;
- Visualizzazione dei dati reperiti nel Web sociale e analisi di tali dati.

#### **Programma esteso**

##### **1. Introduzione**

- Introduzione al Web (1.0, 2.0), Web semantico e Web sociale e alla terminologia utilizzata;
- Gli “oggetti informativi” nel Web sociale;
- *Web e Social Media Analytics*: definizione e obiettivi, i concetti di auto-presentazione e auto-rivelazione, gli incentivi impliciti ed esplicativi.

## **2. I dati nel Web sociale**

- Le principali piattaforme, le tipologie di dato, le interfacce di programmazione, il processo di *crawling* e di *scraping*.
- Pre-processing e memorizzazione di dati nel Web sociale.
- Cenni di problematiche legate alla raccolta dei dati, sia dal punto di vista legale (il GDPR) sia dal punto di vista tecnologico.

## **3. Rappresentare strutture complesse dei dati online: teoria dei grafi e delle reti**

- Strutture dati elementari e complesse;
- Rappresentazione di reti di interazione mediante grafi (teoria dei grafi, tipologie di reti).

## **4. Social Network Analysis (SNA)**

- *Link analysis, Web link analysis*, principali metriche;
- *Network clustering*: algoritmi di *community detection*;
- Modelli di influenza e contagio nelle reti sociali.

## **5. Social Content Analysis (SCA)**

- Introduzione a concetti di *Natural Language Processing* nell’ambito delle reti sociali;
- Oggettività/soggettività, polarità, emozioni e ironia nelle reti sociali;
- Approcci lessicali e approcci semanticici;
- *Named-Entity Recognition* e *Linking*.

## **6. Web & Social search**

- I principali modelli di IR nel Social Web;
- Le dimensioni di rilevanza;
- La valutazione dei risultati di ricerca.

## **7. Visualizzazione dei dati reperiti dal Web sociale e analisi di tali dati**

- Interfacce grafiche;
- Usabilità e studi per mezzo di utenti.

## **Prerequisiti**

Sono richieste competenze base di algebra lineare e programmazione.

## **Modalità didattica**

- Lezioni frontali;

- Esercitazioni in aula;
- Esercitazioni in laboratorio.

L'insegnamento verrà tenuto in lingua inglese.

## Materiale didattico

Verranno fornite slide, approfondimenti e letture suggerite durante il corso.

### Libri suggeriti:

- Greenlaw, R., & Hepp, E. (2001). *Inline/online: fundamentals of the internet and the world wide web*. McGraw-Hill, Inc.
- Rahman, M. S. (2017). *Basic graph theory* (Vol. 9). Cham: Springer.
- Knoke, D., & Yang, S. (2019). *Social network analysis*. SAGE publications.
- Liu, B. (2020). *Sentiment analysis: Mining opinions, sentiments, and emotions*. Cambridge university press.
- Ledford, J. L. (2015). *Search engine optimization bible* (Vol. 584). John Wiley & Sons.

## Periodo di erogazione dell'insegnamento

Secondo semestre

## Modalità di verifica del profitto e valutazione

### Prova scritta con esercizi e domande aperte

La prova scritta ha come obiettivo la valutazione estensiva ed intensiva delle competenze teoriche e teorico-pratiche acquisite durante il corso.

### Progetto di gruppo (con presentazione orale)

Il progetto ha come obiettivo la valutazione della capacità degli studenti di tradurre in ambiti applicativi reali le competenze acquisite durante il corso, attraverso lo sviluppo e l'impiego di soluzioni tecnologiche per l'analisi ed il reperimento dei dati nel Web sociale.

### Valutazione globale

- La prova scritta viene valutata su una scala da 0 a 24
- Gli studenti devono ottenere una valutazione maggiore o uguale a 12 nella prova scritta.
- Il progetto, con relativa discussione orale, viene valutato su una scala da 0 a 8.
- Il voto finale sarà dato dalla somma del voto ottenuto nell'esame scritto e dal voto relativo al progetto.

## Orario di ricevimento

Su appuntamento: Marco Viviani ([marco.viviani@unimib.it](mailto:marco.viviani@unimib.it))

## **Sustainable Development Goals**

SALUTE E BENESSERE | ISTRUZIONE DI QUALITÁ | PARITÁ DI GENERE

---