

insegnamento	modulo	cfu	settore scientifico disciplinare	tipo	semestre	tipo attività didattica	cfu	ore	docente
F9101Q003 Data management and visualization	Data management (6 cfu)	12	ING-INF/05	obbligatorio	Primo semestre	lezione	4	32	Maurino Andrea
						esercitazione	1	10	Rula Anisa
						laboratorio	1	9	Bianchi Federico
	Data visualization (6 cfu)					lezione	2	16	Schettini Raimondo
	lezione						2	16	Cabitzza Federico
	laboratorio						2	24	Uboldi Giorgio Roberto

CV docenti: <http://www.unimib.it/go/176181440>)

Modulo Data management

Contenuti

- Il ciclo di vita dei Big Data
- Introduzione ai Modelli e alle architetture NoSQL (CAP theorem, document, graph, key value, columnar)
- Distribuzione dei dati (replica, frammentazione)
- Cenni su qualità dei dati
- Cenni sui sistemi informativi geografici
- Architecture per big data analysis

Obiettivi formativi

Conoscenza e comprensione

Alla fine del corso lo studente apprenderà come scegliere, progettare e interrogare un database anche di tipo non relazionale, in base alle proprie esigenze applicative

Capacità di applicare conoscenza e comprensione

Lo studente sarà in grado di usare database management system di tipo non relazionale al fine di acquisire, memorizzare e gestire dati semi/non strutturati.

Prerequisiti

- modello relazionale dei dati

Testi di riferimento

- G. Harrison (2015) Next Generation Databases, APress

Metodi didattici Lezioni e esercitazioni in aula e su computer

Modalità di verifica dell'apprendimento
progetto applicativo

Programma esteso

1. Il ciclo di vita dei dati
 - 1.1. Acquisizione
 - 1.2. Storage
 - 1.3. Integration
 2. Modelli NoSQL
 - 2.1. Cap theorem
 - 2.2. Document based system
 - 2.3. graph db
 - 2.4. Sistemi key value e columnar
 3. Distribuzione dei dati
 - 3.1. Replica dei dati
 - 3.2. Frammentazione
 4. Cenni sulla qualità dei dati
 5. Cenni sui sistemi informativi geografici
 6. Architetture per big data analysis
 - 6.1. Modello Map Reduce
 - 6.2. Rassegna dei principali componenti (Hive, Spark, Flink, Impala....)
-

Modulo Data Visualization

Contenuti

Il corso riguarda metodi, tecniche e strumenti di visualizzazione dati (data visualization) e progettazione visuale (information design) per la realizzazione di sistemi che permettano l'analisi interattiva dei dati e l'ottimizzazione flessibile del reporting (sia in ambito organizzativo che di data journalism). A tal scopo in questo corso saranno presentate strategie di visualizzazione di dati sul Web attraverso infografiche e cruscotti (dashboard) che risultino informativi e comprensibili e che possano essere realizzate senza competenze avanzate di programmazione attraverso l'utilizzo di vari strumenti che variano dalle piattaforme software commerciali più comuni (e.g., Tableau) ai diversi pacchetti open source accessibili su Web (Javascript, HTML5, etc.). Una componente importante del corso riguarderà la progettazione iterativa e quindi l'acquisizione di metodiche e tecniche per la valutazione della qualità delle infografiche suddette e la applicazione di tale tecniche per il miglioramento continuo dei sistemi di visualizzazione dati. Nelle ore di laboratorio gli studenti potranno inoltre acquisire le competenze necessarie per lo svolgimento di un progetto applicativo concreto e di complessità realistica, consistente nella realizzazione di un Web report contenente grafici e diagrammi animati e interattivi su temi di interesse comune e pubblica utilità.

Obiettivi formativi

Alla fine del corso lo studente avrà acquisito competenze relative alle attività di analisi, sviluppo e valutazione della qualità di infografiche complesse e interattive.

Prerequisiti Nessuno

Testi di riferimento

- Munzner, T. (2014). Visualization analysis and design. CRC Press.
- Ware, C. (2012). Information visualization: perception for design. Elsevier.

Articoli scientifici e dispense fornite dai docenti.

Metodi didattici Lezioni frontali con il supporto di slideware, discussione di casi pratici nel forum del corso, discussione di progetti.

Modalità di verifica dell'apprendimento

Un progetto in cui applicare metodi e tecniche per realizzare e valutare una infografica complessa o un Web report dotato di una serie di infografiche correlate tra loro. Esame scritto e un orale (discussione progetto).

Programma esteso

- Introduzione alla Visualizzazione.
- Percezione umana ed elaborazione delle informazioni
- Tipi di dato
- Percezione grafica (abilità di interpretare la codifica visuale – grafica – delle informazioni e quindi decodificare l'informazione presente in grafici e diagrammi):
 - Identificazione dei segnali
 - Stima della grandezza (magnitudo)
 - Elaborazione visuale pre-attentiva
 - Uso di codifiche visuali multiple
 - Raggruppamento per Gestalt
- Il colore nella rappresentazione dell'informazione
- Esempi e casi di studio
- Sistemi di gestione del colore
- Visualizzazione e fruizione delle immagini fotografiche
- Trasformazione di dati in fonti di conoscenza attraverso la rappresentazione visuale.
- Requisiti ed euristiche per visualizzazioni di qualità.
- Diagrammi e visualizzazioni standard: pertinenza e appropriatezza.
- Strumenti avanzati e innovativi per data visualization e analisi quantitative avanzate.
- La valutazione della qualità delle visualizzazioni.
 - a. Valutazione qualitativa: esperta ed euristica;
 - b. Valutazione quantitativa: task utente; tecniche di statistica inferenziale.
 - c. Questionari psicometrici validati e loro analisi e comprensione.
- Laboratori in cui verranno acquisite competenze pratiche per:
 - a. L'estrazione di dati non strutturati dal web (import.io, kimono, etc.)
 - b. La gestione e manipolazione dei dati in formato tabellare (google spreadsheet, excel, etc.)
 - c. L'esplorazione e la visualizzazione statica di dati (RAWGraphs, Gephi, illustrator, etc.)
 - d. L'esplorazione e la visualizzazione interattiva di dati (Tableau Public, Carto)
 - e. La progettazione di una narrazione "data-driven" nel contesto del data journalism.

year	code and course name	module	ECTS	type	semester	educational activity type	ECTS	hours	faculty
1	F9101Q003 Data management and visualization	Data management (6 ECTS)	12	mandatory	First semester	lecture	4	32	Maurino Andrea
						practice exercise	1	10	Rula Anisa
						laboratory	1	9	Bianchi Federico
		lecture				2	16	Schettini Raimondo	
		lecture				2	16	Cabitzza Federico	
		laboratory				2	24	Uboldi Giorgio Roberto	
		Data visualization (6 ECTS)							

CV: <http://www.unimib.it/go/176181440>

Data management

Content of the course

- Data life cycle
- NoSQL models
- Data distribution
- Data quality
- Geographical information systems
- Architectures for big data analysis

Educational goals

At the end of the module students will be able to select, design and query a database (relational or not) according to their application needs

Students will be able to use a NoSql database management system to acquire, memorize and query semi structured data

Prerequisites

Relational model

Reference books

- G. Harrison (2015) Next Generation Databases, Apress.

Teaching methods Lectures and exercise in room and on PC

Assessment methods projects

Extended program

1. Data life cycle
1. Acquisition
2. Storage
3. Integration
2. NoSQL models
 - Cap theorem
 - Document based system
 - Graph db
 - key value and columnar models
2. Data distribution
 - Replication
 - Fragmentation
3. Data quality
4. Geographical information systems
5. Architectures for big data analysis
 - Map Reduce
 - Main components (Hive, Spark, Flink, Impala....)

Data Visualization

Content of the course

The course covers the methods, techniques and tools of data visualization and visual design by which to design, implement and evaluate systems that enable the interactive analysis of data and the flexible optimization of reporting (both in an organizational domain and in data journalism). To this aim, in this course strategies will be presented for the visualization of Web data through infographics and dashboards that are both informative and understandable, and that could be implemented without advanced programming skills using various instruments that range from the most common commercial software platforms (e.g., tableau) to the several open source packages accessible on the Web (JavaScript, HTML5, etc.). An important component of the course will cover the iterative design and then the acquisition of methods and techniques to assess the quality of these infographics and their concrete application to the continuous improvement of data visualization systems. In the laboratory hours, the students will also acquire the skills necessary to carry out a concrete application project of realistic complexity, which regards the production of a Web report with graphics and animated and interactive charts on topics of common interest and public utility.

Educational goals

At the end of the course students will have acquired skills in analysis, development and evaluation of the quality of complex and interactive infographics.

Master of science in Data Science a.a. 2017-18

Prerequisites

Nobody

Reference books

Münzner, T. (2014). Visualization analysis and design. CRC Press.

Ware, C. (2012). Information visualization: perception for design. Elsevier.

Scientific articles and handouts provided by the teachers.

Teaching methods

Lectures with the support of slideware, discussion of practical cases through the forum, discussion of practical home-work projects.

Assessment methods

A project in which to apply methods and techniques to create and evaluate a complex infographic or a Web report with a series of infographics related to each other. Written exam and an oral discussion of the project.

Extended program

- Introduction to Visualization.
- Human Perception and Information Processing
- Data types
- Graphical perception (the ability of viewers to interpret visual
- (graphical) encodings of information and thereby decode information in graphs):
 - a. Signal Detection
 - b. Magnitude Estimation
 - c. Pre-Attentive Visual Processing
 - d. Using Multiple Visual Encodings
 - e. Gestalt Grouping
- Color for information display
- Examples and case studies
- Color management systems
- Picture visualization and fruition
- Data Transformation into sources of knowledge through visual representation.
- Requirements and heuristics for high-quality visualizations.
- Charts and standard views: relevance and appropriateness.
- Advanced and innovative tools for data visualization and advanced quantitative analysis.
- The evaluation of the quality of visualizations and infographics.
 - o Qualitative assessment: expert and heuristic;
 - o Quantitative assessment: user tasks; inferential statistical techniques.
 - o Validated psychometric questionnaires and their analysis and understanding.
- Workshops in which students will acquired practical skills to:
 - o extract unstructured data from web (import.io, kimono, etc.)
 - o manage and manipulate data in tabular format (google spreadsheet, excel, etc.)

Master of science in Data Science

a.a. 2017-18

- explore and present static data (RAWGraphs, Gephi, illustrator, etc.)
- explore and build interactive data visualizations (Tableau Public, Carto)
- design a "data-driven" narrative in a data journalism context.