

year	code	course name	ECTS	type	semester	educational activity type	ECTS	hours	faculty
1	F9101Q011	Data semantics	6	elective	Second semester	lecture	4	32	Palmonari Matteo
						practice exercise	2	20	to be defined

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

### Content of the course

1) Data semantics: from conceptual modeling to conceptual data management (interpretation, conceptual models, metadata). 2) Conceptual data management and Knowledge Graphs (KG) (KG as an abstraction; RDF and Semantic Technologies; KG and document management systems; vocabularies, classifications, ontologies). 3) Semantic data modeling (modeling and querying data in RDF and RDFS). 4) Semantic information integration and enrichment (schema and instance matching, similarity measures, mapping, data linking and fusion, from texts to RDF).

### Educational goals

At the end of the course, the student will learn how to: model and query RDF data using or creating suitable vocabularies or ontologies in RDFS; build algorithms for reconciling data using semantic similarity measures; use semantic data models to link and analyze heterogeneous data (tables, texts, and so on)

At the end of the course, the student will be able to: build or enrich a KG starting from one or more tabular data sets, by linking these data to reference KGs; enrich a source data set using data from other KGs (e.g., open data).

### Prerequisites

Relational and NoSQL databases. Basic SQL

### Reference books

Grigoris Antoniou, Paul Groth, Frank van Harmelen, *A Semantic Web Primer*, (Third Edition), MIT press, 2012.

### Teaching methods

Lectures and exercises with a personal computer

### Assessment methods

Final test or project

**Extended program**

1. Data semantics: from conceptual modeling to conceptual data management
  - a. Introduction: semantics and interpretation
  - b. Conceptual modeling with ER and UML
  - c. Metadata and vocabularies
2. Conceptual data management: Knowledge Graphs (KGs) and beyond
  - a. KGs as abstractions
  - b. KGs with RDF and semantic technologies
  - c. KGs and document management systems
  - d. Vocabularies, classifications, ontologies
3. Semantic data modeling
  - a. RDF modeling and querying with SPARQL
  - b. Simple knowledge representation for RDF with RDFS
4. Lab I: Data lifting by mapping tables to KGs
5. Semantic information integration
  - a. Data Reconciliation
  - b. Schema and instance matching
  - c. Similarity measures, uncertainty and decision making
6. Semantic data enrichment
  - a. Data linking
  - b. Data fusion
  - c. Linking texts to KGs
7. Lab I: Data integration and enrichment