



University of Rome “Tor Vergata”

Knowledge Engineering

Introduzione al corso

Armando Stellato <stellato@uniroma2.it>

Semantica Esplicita delle informazioni

Il Web non deve più limitarsi a ospitare testo e altri contenuti multimediali all'interno dei quali cercare le informazioni che desideriamo.

Queste devono essere rese esplicite! al fine di poter interrogare il Web stesso come una immensa base dati, trovando esattamente quel che cerchiamo

The screenshot shows a Google search interface. The search bar contains the text "spaghetti cacio e pepe". Below the search bar, there are tabs for "Tutti", "Video", "Immagini", "Notizie", "Maps", "Altro", "Impostazioni", and "Strumenti". The search results show "Circa 1.440.000 risultati (0,56 secondi)". The first result is titled "Ricetta Spaghetti Cacio e Pepe - La Ricetta di GialloZafferano" with a URL "https://ricette.giallozafferano.it/Spaghetti-Cacio-e-Pepe.html". The result includes a star rating of 4.5, a total time of 20 minutes, and a small image of the dish. Annotations include: a green circle around the image labeled "Picture"; a yellow circle around the star rating labeled "Rating"; a red circle around the "20 min" label labeled "Total time"; and a blue oval around the text "Gli Spaghetti cacio e pepe sono uno dei piatti forti della tradizione romana: Pecorino grattugiato e grani di pepe, una ricetta veloce e saporita!" labeled "Dish description".

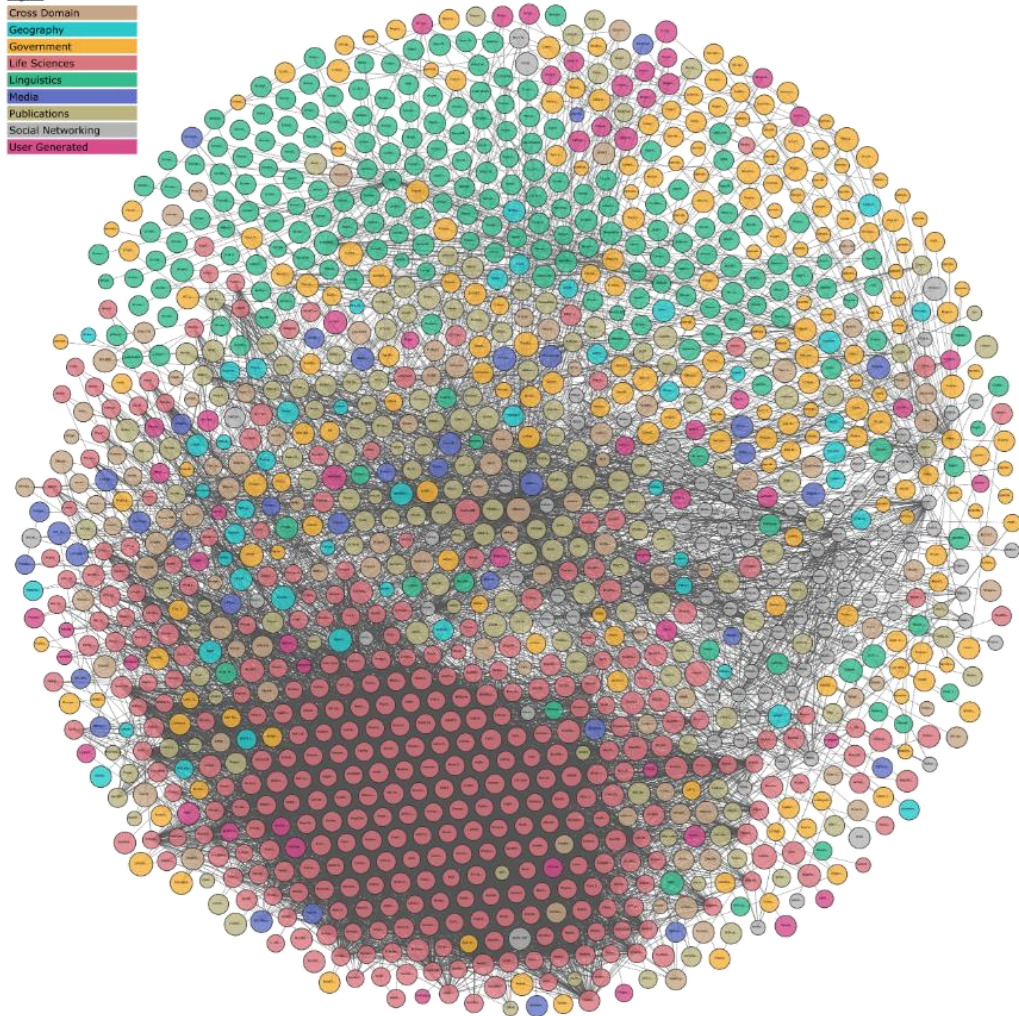
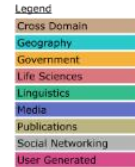
The Linked Open Data Cloud

Non è la morte nera!

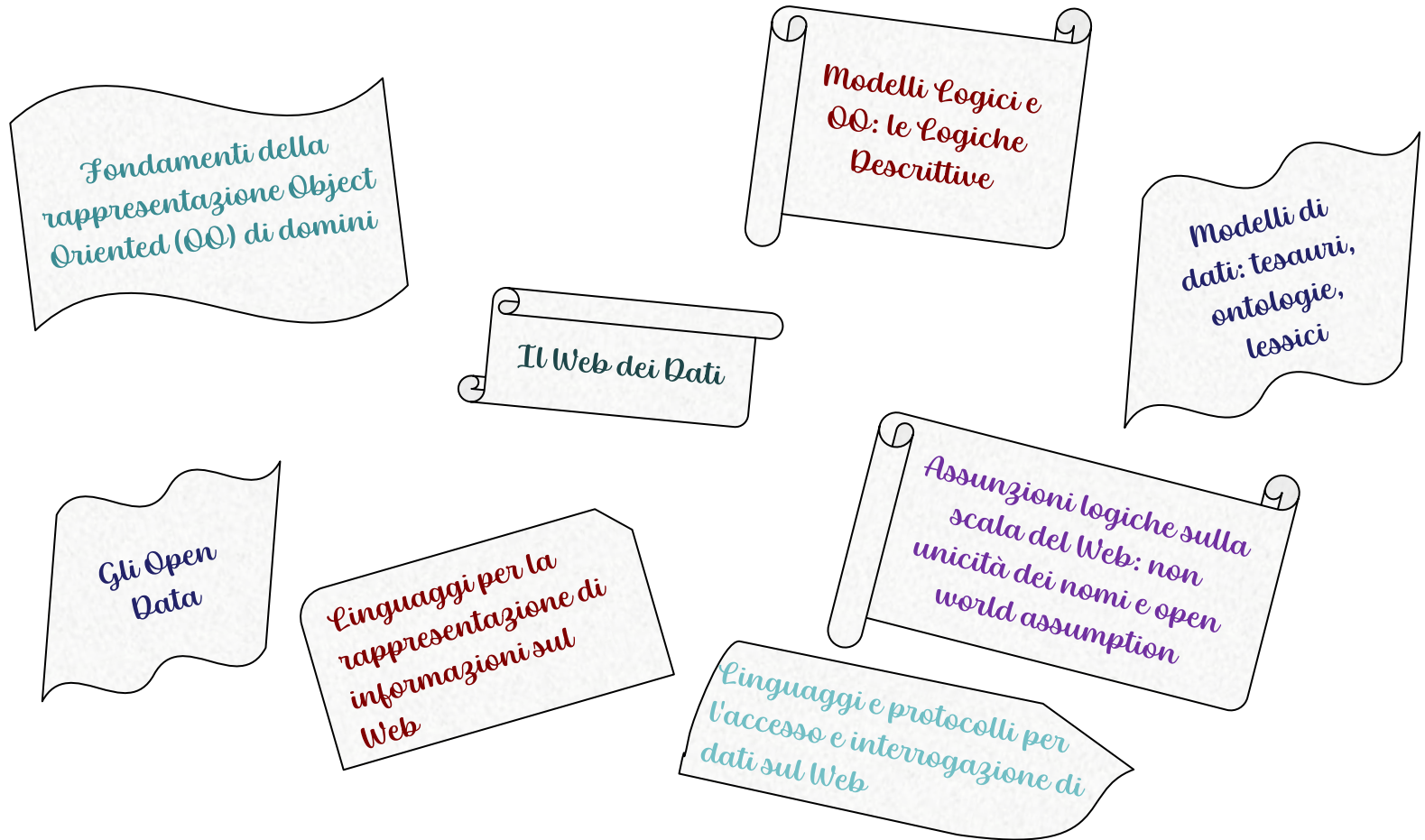
è la:

Linked Open Data Cloud: centinaia di dataset pubblici in rete collegati tra loro

Oggi migliaia di organizzazioni pubbliche e private pubblicano i loro dati in modo aperto e interrogabile



Temi del corso



Organizzazione del corso

Titolo del corso: Knowledge Engineering

Durata del corso: 1 semestre – 9 crediti

Obiettivi del corso

- Riesamina dei fondamenti di rappresentazione della conoscenza e inferenza logica appresi in altri corsi della laurea triennale e ulteriore approfondimento delle conoscenze su questi temi.
- Padronanza del paradigma di gestione e accesso ai dati distribuito nel web.
- Progettazione di modelli della conoscenza come ontologie, tesauri e lessici per il Web dei Dati.
- Pubblicazione, accesso e esplorazione di dati nel web. Tecniche di coordinamento semantico basato su metadati.

Modalità d'Esame

- Prova scritta e al calcolatore.
 - Esercizi di inferenza logica usando le logiche descrittive
 - Esercizi sull'interrogazione di repository aperti
- Prova Orale
- Progetto applicativo, con realizzazione di un modello dei dati, una base di conoscenza e una applicazione basata su di essi

Programma del corso

- Introduzione ai dati aperti. I 5 livelli di qualità dei dati aperti fissati dalla W3C. Licenze sui dati: licenze libere/aperte. Dati, metadati, vocabolari dei dati, ontologie, tesauri: ambiguità terminologiche in uno scenario inerentemente intersettoriale. Il valore economico dei dati. Open Data per il dominio dei Patrimoni Culturali.
- Il Web Semantico: una introduzione. Il Web Semantico e gli Open Data: Linked Open Data e il Web. Modelli dei dati e della conoscenza per il Web.
- Inferenza Logica: limiti di computabilità e nuove prospettive per la ricerca e il ragionamento in ipotesi di mondo aperto. Il supporto logico-inferenziale ai linguaggi per il Web.
- Multilingualità e diffusione: aspetti denotazionali, proprietà terminologiche e modelli di lessico avanzato. Interazione lessico-semantica.
- SPARQL: un linguaggio di interrogazione dei dati RDF & assieme un protocollo di accesso ai dati aperti
- Metadati: importanza dei metadata e nel loro ruolo per l'autonomia delle macchine nel web
- SHACL: vincoli sintattici per i modelli dei dati: un passo indietro rispetto agli approcci inferenziali?
- LOD Cloud: la "nuvola" dei Linked Open Data nel Web, un bootstrap del Semantic Web, ma anche una contraddizione in essere
- Dataset repositories e portali; Tecnologie per i Linked Open Data: Triple stores. RDF API & Middleware. Ontology Editors, Thesauri Editors, differenti paradigmi e dimensioni d'uso
- ETL (Extract, Transform, Load) in chiave RDF: triplificazione di sorgenti legacy e/o non strutturate.
- Pubblicazione dei dati aperti: standard & best practices