

Percorso di Eccellenza

Matteo Mancanelli

1711823

2018

Il percorso di eccellenza è stato svolto con la supervisione del Prof. Riccardo Rosati e si è focalizzato in particolar modo sullo studio e l'approfondimento del Semantic Web, tanto da un punto di vista teorico quanto dalla prospettiva delle applicazioni pratiche e della creazione di semplici programmi implementativi. In un primo momento, infatti, sono stati analizzati gli assunti di base che delineano questo ambito e i principi concettuali che lo caratterizzano; successivamente si è proceduto con la concretizzazione delle conoscenze acquisite, in modo tale da consolidarne la comprensione ed indagarne i risvolti nel reale sviluppo di software operativi.

L'obiettivo prefissato per il percorso è stato quello di esaminare un contesto distante dalle attività svolte regolarmente nei corsi proposti, ma che nello stesso tempo trova diversi legami con tali attività e permette quindi una proficua integrazione. Una introduttiva sperimentazione nel Semantic Web produce perciò il duplice risultato di ampliare una visione sempre più trasversale e di gettare un rapido sguardo in alcuni dei settori più attuali, che vengono tuttora sviluppati in ambito accademico e non solo.

1 Attività di tipo teorico

La prima attività svolta consisteva nell'analisi dei paradigmi e degli strumenti che costituiscono ad oggi l'architettura del Semantic Web. Il mezzo utilizzato per il raggiungimento di tale scopo è stato lo studio di un libro di riferimento dell'ambito descritto, *Semantic Web. Tra ontologie e Open Data*, scritto dai professori Di Noia, De Virgilio, Di Sciascio e Donini. Questo volume ha fornito una introduzione efficace e graduale dell'argomento, in grado di presentare ai lettori una panoramica completa degli aspetti più rilevanti per comprendere le motivazioni alla base e le tecnologie già standardizzate. Si percorrono così informazioni legate agli innovativi risultati ottenuti nel corso degli anni, peculiari del Semantic Web, ma anche alle relazioni che intercorrono fra tali risultati e i più tradizionali settori dei linguaggi per il Web, delle basi di dati e della logica, di cui si apprende nel corso delle attività didattiche.

A quanto detto si aggiunge poi un ulteriore approfondimento degli stessi temi, realizzato con l'impiego di nuovo materiale relativo alle lezioni dei corsi dedicati all'argomento per la laurea magistrale. In questo modo è stato possibile avere una prima visione di come può essere inserito il settore del Semantic Web nel percorso formativo volto alla preparazione degli studenti per una maggiore consapevolezza e competenza nei campi dell'ingegneria informatica e, in parte, dell'intelligenza artificiale. Parte degli approfondimenti descritti provengono direttamente dai documenti messi a disposizione dallo stesso professor Rosati, che illustrano più in dettaglio le tecniche della rappresentazione della conoscenza e dell'ambito preso in considerazione. Questo si è rivelato di grande utilità anche per la possibilità di osservare le prime esercitazioni ed applicazioni concrete della materia trattata.

I linguaggi che costituiscono la struttura fondamentale del Semantic Web sono volti, in parte, alla descrizione delle relazioni che intercorrono fra un insieme di informazioni e alla interrogazione di queste nuove basi di conoscenza. Sono perciò realizzati per soddisfare gli obiettivi di fornire semantica ai dati e di permettere alle macchine una più efficace elaborazione degli stessi, basata sul loro significato e sulla loro corretta interpretazione. Fra le principali conoscenze acquisite in questa prima fase emergono quindi i più celebri ed utilizzati strumenti quali RDF, che permette di enunciare relazioni fra individui, OWL, che consente di creare ontologie e stabilire relazioni fra predicati, ed infine SPARQL, per l'esecuzione di query sui dataset costituiti da insiemi di triple RDF.

2 Attività di tipo progettuale

Come detto, allo studio di tipo teorico-concettuale deve necessariamente essere associata una analisi altrettanto estesa e rilevante per individuare le implicazioni degli argomenti affrontati ed i possibili impieghi nella scrittura di codice applicativo. In questa ottica, la seconda parte del percorso di eccellenza è stata nuovamente dedicata al Semantic Web, ma questa volta lo scopo risiedeva nel provare ad utilizzare gli strumenti descritti in precedenza per la scrittura di semplici programmi eseguibili. Tali programmi sono prodotti per evidenziare i limiti e le potenzialità delle tecnologie citate e per verificare i modi in cui queste possano essere integrate con i più celebri modelli tradizionali ed i maggiori linguaggi di programmazione.

In particolare sono stati ideati semplici software con il supporto dei linguaggi appresi nel corso di studio, come Python e Java, e delle loro librerie pensate per una migliore e più efficace implementazione dei nuovi paradigmi. In un primo momento sono quindi state suddivise le attività principali e testate in differenti contesti per osservare quali potessero offrire maggiori possibilità. Le attività eseguite nei programmi consistevano soprattutto nella strutturazione dei contenuti informativi in triple RDF, nell'interrogazione

degli stessi e nella creazione di ontologie per la rappresentazione dei domini di interesse.

Per la realizzazione di quanto detto, si è fatto largo uso di strumenti versatili come Jena, framework di Java per la manipolazione dei modelli e dei file che codificano i dataset sfruttati, e Protégé, un editor visuale per l'organizzazione delle classi e delle proprietà che mettono in relazione le une con le altre. È stato così possibile provare in prima persona come definire un dominio ontologico, come eseguire il parsing e la scrittura di documenti OWL, come estendere questi documenti con informazioni ed annotazioni di vario genere. Parte integrante ed essenziale è stata l'estrazione dei differenti dati presi in considerazione con l'esecuzione di query SPARQL, che si discosta dal più popolare SQL tanto nella struttura sintattica quanto nel modello concettuale che ne rappresenta alla stesso tempo presupposto e punto di arrivo.