

- 1. Rappresentazione della conoscenza**
- 2. Ontologie**
- 3. Usi delle ontologie**
- 4. Progettazione di un'ontologia**
- 5. Esempio di progettazione di una ontologia**

Rappresentazione della conoscenza

Il problema di rappresentare la conoscenza in funzione della sua archiviazione, gestione, trasferibilità e condivisione ha da sempre interessato l'essere umano ed ha coinvolto, via via, diversi settori e diverse comunità scientifiche e/o professionali (Rivoltella P.C., in Paparella N., a cura di, 2007). La rappresentazione deve limitare quanto più possibile ambiguità interpretative quando la conoscenza viene condivisa internamente a comunità o fra comunità. Se si usa il linguaggio naturale per descriverla, potrebbero generarsi interpretazioni differenziate quando le pratiche nelle comunità sono caratterizzate da conoscenze ed esperienze diverse. Se si tiene anche conto che una descrizione in linguaggio naturale non è facilmente compresa da un elaboratore, diventa essenziale utilizzare opportuni formalismi per la rappresentazione di una conoscenza gestibile anche dalle macchine. In questo contesto si fa particolare riferimento alle modalità di rappresentazione derivate da studi in ambito informatico.

Ontologie

In letteratura sono state date diverse definizioni di cosa si intende con il termine ontologia¹:

- i. "L'insieme dei termini basilari e delle relazioni, che costituiscono il vocabolario di un'area specifica, e delle regole per combinare termini e relazioni per determinare estensioni del vocabolario" (Neches R. et al., 1991);
- ii. "Un'ontologia è una specificazione esplicita di una concettualizzazione" (Gruber T.R., 1993);
- iii. "Un'ontologia è una specificazione formale ed esplicita di una concettualizzazione condivisa" (Borst W.N., 1997);

Dalle definizioni di Borst e Gruber:

- *concettualizzazione*: rappresentazione, in un modello astratto, di un determinato dominio del mondo reale attraverso i concetti (classi) rilevanti che lo caratterizzano;
- *esplicita*: i concetti, le relazioni fra questi, le proprietà del dominio debbono essere definiti in maniera non ambigua;
- *formale*: deve essere espressa in un linguaggio formale comprensibile da una macchina;
- *condivisa*: un'ontologia deve rappresentare conoscenza condivisa da un gruppo.

Con il termine *ontologia* si indica, quindi, una rappresentazione (schema concettuale) esaustiva e rigorosa di un dato dominio; essenzialmente contiene tutte le classi (concetti) rilevanti, le relazioni fra di esse, gli individui (istanze) che popolano la realtà rappresentata e che sono inseriti nelle classi.

Per ogni classe dell'ontologia, quindi per tutti gli individui della classe, vengono definite proprietà (gli attributi) che ne descrivono caratteristiche ed eventuali restrizioni. Le classi possono essere *primitive*, se non derivano da altre presenti nell'ontologia, oppure *derivate* se sono ereditate da altre classi presenti; in questo caso si realizza una *tassonomia*.

¹ In questo contesto si fa riferimento all'ambito informatico ed alle ontologie di dominio. "La distinzione principale fra l'ontologia in senso filosofico e quella in senso informatico è data dall'ambizione (interpretabile come pretesa, in termini di logica formale), da parte dell'ontologia filosofica, di spiegare 'ciò che è' in assoluto, ovvero tutto l'essere, mentre nell'ambito dell'informatica, nel primo decennio del 21. secolo, la creazione di un'ontologia fondante e totale risulta un'impresa titanica. Impresa che tuttavia non può essere a priori esclusa, in termini tecnologici, anche se è sempre più evidente che essa richiede una enorme attività di "negoiazione tra pari" per la conciliazione di moltissime esigenze e punti di vista diversi: primi fra tutti quelli rappresentati dai diversi sistemi semantici associati alle diversità linguistiche. Una cosa, invece, già possibile e praticata è la creazione di molte ontologie, ciascuna limitata a un *dominio* ben preciso e persino ad un ben preciso punto di vista, o scopo, di quel dominio, allora abbiamo qualcosa che può già essere realizzato e può essere utilizzato per molti scopi, come gestire un servizio Web o integrare sistemi diversi. Le ontologie così create potrebbero poi, in caso di necessità, venire mappate le une sulle altre, sfruttando il meccanismo di importazione delle ontologie, in modo da farle interagire senza perdere la complessità e particolarità di ciascuna." [http://it.wikipedia.org/wiki/Ontologia_\(informatica\)](http://it.wikipedia.org/wiki/Ontologia_(informatica))

Gli elementi fondamentali di un'ontologia sono, quindi: un dominio di interesse del mondo reale; un linguaggio comune; un insieme di concetti descritti all'interno dell'ontologia; una eventuale tassonomia che esprime l'ereditarietà tra i vari concetti; un insieme di relazioni non tassonomiche che mettono in relazione i concetti; un insieme di istanze: gli elementi del mondo reale rappresentato (individui); un insieme di proprietà: gli attributi delle istanze; un insieme di restrizioni. Le restrizioni specificano il modo in cui concetti e relazioni possono essere usati, vincolano cioè le classi ed i suoi individui a determinate regole (ad esempio il tipo di relazione tra i concetti, la cardinalità di una relazione, simmetria e transitività tra relazioni, le caratteristiche degli attributi).

Progettazione di un'ontologia

La progettazione e la realizzazione di una ontologia si sviluppano attraverso una analisi negoziata del dominio di interesse, che permette di individuare le classi significative, di organizzarle in eventuali tassonomie e di stabilire le relazioni fra esse.

Nella progettazione e realizzazione occorre sicuramente individuarne il dominio, stabilire quale sarà il suo utilizzo e quali tipi di domande deve soddisfare, chi utilizzerà e chi manterrà l'ontologia. La negoziazione dovrebbe portarci ad individuare una serie di termini significativi che rappresentano i concetti e le relazioni dell'ontologia.

Si percorrono, in genere, modalità *Top-Down*, *Bottom-Up*, miste o, in alternativa, metodologie semi-automatiche come, ad esempio, attraverso l'utilizzo del *software* *OntoLearn*.

Nella modalità *Top-Down* la progettazione si sviluppa a partire dal livello più alto per arrivare, verso il basso, a quello che si ritiene idoneo per gli scopi dell'ontologia. Si procede dai concetti generali a quelli più specifici.

Nella modalità *Bottom-Up* la progettazione si sviluppa a partire dal livello più basso per arrivare, verso l'alto, a quello che rappresenti una generalità adeguata agli scopi dell'ontologia.

Nella modalità mista si selezionano i concetti che si ritengono essenziali per il dominio che si vuole rappresentare, poi si individuano gli eventuali legami gerarchici che legano i concetti verso il basso o verso l'alto.

Vanno quindi definite le relazioni, definiti eventuali vincoli sugli attributi e sulle relazioni, create le istanze.

Nella progettazione semi-automatica (Navigli R., Velardi P., 2005) le soluzioni fornite dalla macchina, attraverso programmi, vengono via via revisionate, integrate e validate da esperti che seguono il processo. Un esempio è fornito da *OntoLearn* (ibidem), un programma che offre un supporto nella fase di progettazione di un'ontologia; deve essere affiancato da esperti e presuppone che il dominio sia già delineato da un insieme di documenti. Il procedimento si sviluppa in 4 fasi e, per ciascuna di esse, operano *OntoLearn* e gli esperti.

- Fase 1: creazione di un Lessico fornendo ad *OntoLearn* un insieme di documenti riferiti al dominio dell'ontologia;
- Fase 2: dare un significato al Lessico; viene creato un glossario sui termini individuati nella fase uno;
- Fase 3: creazione di una tassonomia; i termini individuati nella fase uno vengono sistemati in una organizzazione gerarchica;
- Fase 4: creazione di relazioni che, insieme ai concetti e alla tassonomia, costituiscono l'ontologia.

Ad ogni fase c'è un'interazione fra l'*output*, generato da *OntoLearn* nella fase precedente, ed il gruppo di esperti del dominio e della gestione e rappresentazione della conoscenza. Nella prima fase *OntoLearn* scandisce l'insieme dei documenti, mediante tecniche di NLP², e restituisce una lista di termini significativi per il dominio. Il gruppo di esperti, controllerà la lista di termini, accettando termini che ritiene significativi ed eliminando quelli valutati poco rappresentativi. Nella fase 2 *OntoLearn* riceve in *input* l'elenco revisionato e restituirà lo stesso elenco arricchito con delle definizioni per i termini. Con identico procedimento della fase 1, verranno accettate o eliminate le singole definizioni. Nella successiva fase 3 il Glossario così realizzato sarà passato ad *OntoLearn* che realizzerà una tassonomia individuando una serie di relazioni gerarchiche. Anche questa volta si avrà un intervento umano che provvederà ad accettare o eliminare relazioni. La tassonomia ed il glossario diventano *input* per la fase 4 che arricchirà la tassonomia con relazioni *orizzontali* (non tassonomiche) fra i concetti, restituendo una ontologia che dovrà anch'essa essere valicata. Il prodotto finale è l'ontologia che è pronta per l'uso.

² Processamento del Linguaggio Naturale.

Linguaggi per ontologie

Esiste una rappresentazione grafica dell'ontologia e la relativa versione in un linguaggio; affinché una ontologia possa essere utilizzata dalle macchine deve essere scritta in un linguaggio. Una rappresentazione grafica ha una utilità espositiva dei concetti e delle relazioni presenti nell'ontologia stessa. Se l'ontologia deve essere fruita nel web, come avviene nella maggioranza dei casi, viene generalmente codificata in linguaggio *OWL (Ontology Web Language)*. Un frammento di ontologia, espresso in forma grafica, è presentato in Figura 1.

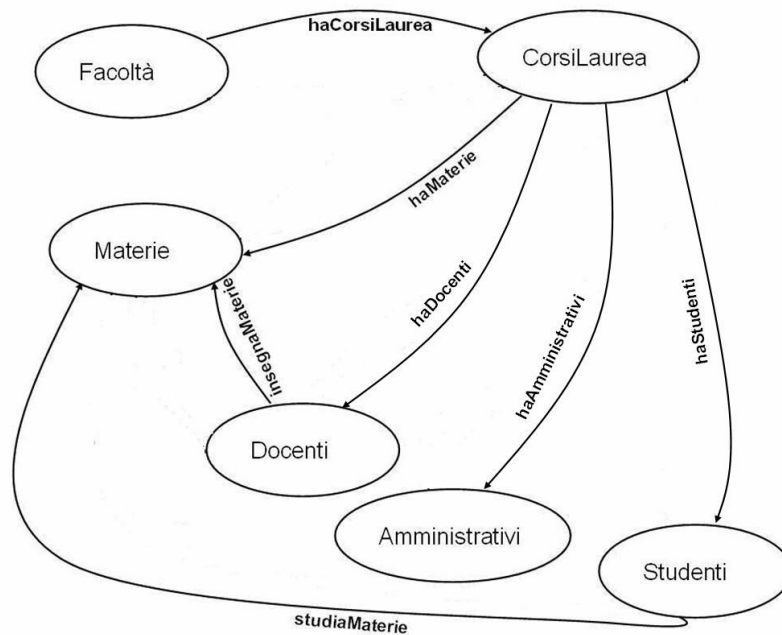


Figura 1-Frammento di ontologia riferita a un segmento del mondo universitario. L'ontologia è utilizzata per tener traccia degli insegnamenti (Materie), dei professori che le insegnano e degli studenti che le studiano. Vengono rappresentate alcune entità e le relazioni che intercorrono tra di esse.

Dal grafico di Figura 1 si nota:

1. la classe *Facoltà* è in relazione con la classe *CorsiLaurea* attraverso *haCorsiLaurea*;
2. la classe *CorsiLaurea* è in relazione con la classe *Materie* attraverso *haMaterie* e con le classi *Docenti*, *Amministrativi*, *Studenti* attraverso *haDocenti*, *haAmministrativi*, *haStudenti*;
3. la classe *Docenti* e la classe *Studenti* sono in relazione con la classe *Materie* attraverso *insegnaMaterie* e *studiaMaterie*.

La classe alla quale si applica la relazione si chiama *dominio (domain)*, la classe alla quale si riferisce la relazione, si chiama *codominio (range)*. Ad esempio, la classe *Facoltà* rappresenta il dominio della relazione *haCorsiLaurea* e la classe *CorsiLaurea* rappresenta il codominio.

Le classi contengono gli individui, chiamati istanze; ad esempio gli individui della classe *Facoltà* sono rappresentati dalle singole Facoltà (Scienze della Formazione, Scienze delle Comunicazioni, Giurisprudenza, ..), quelli di *CorsiLaurea*, dai Corsi di Laurea di ciascuna Facoltà. Agli individui della classe *Facoltà* (dominio) corrispondono gli individui della classe *CorsiLaurea* (codominio). Alla Facoltà Scienze della Formazione, corrispondono i Corsi di Laurea Scienze della Formazione Primaria, Scienze del Turismo, Scienze dell'educazione e della Formazione, .. . Le singole classi hanno delle proprietà (equivalenti agli attributi nella gestione di un database) e tutti gli individui delle classi le assumono. Queste proprietà hanno delle restrizioni che rappresentano o il tipo di dato di appartenenza (intero, con virgola, carattere, ...), oppure delle particolari limitazioni rappresentate

da intervalli di variazione; ad esempio per proprietà di tipo intero, si può imporre che i valori possano variare da 1 a 10, ...

Possono essere applicate delle restrizioni alle relazioni che rappresentano dei vincoli a cui debbono sottostare gli individui delle classi alle quali si applicano le relazioni stesse. Per le ontologie rappresentate in OWL, esistono le seguenti restrizioni: *quantifier restriction*, *cardinality restriction* e *hasValue restriction*³.

In *OWL* la precedente ontologia verrebbe rappresentata nel seguente modo⁴:

```
<?xml version="1.0"?>
<rdf:RDF
.....
.....
1. <owl:Class rdf:ID="Amministrativi"/>
   <owl:Class rdf:ID="Facoltà"/>
   <owl:Class rdf:ID="CorsiLaurea"/>
   <owl:Class rdf:ID="Docenti"/>
   <owl:Class rdf:ID="Studenti"/>
   <owl:Class rdf:ID="Materie"/>
2. <owl:ObjectProperty rdf:ID="haCorsiLaurea">
   <rdfs:range rdf:resource="#CorsiLaurea"/>
   <rdfs:domain rdf:resource="#Facoltà"/>
   </owl:ObjectProperty>
3. <owl:ObjectProperty rdf:ID="haStudenti">
   <rdfs:range rdf:resource="#Studenti"/>
   <rdfs:domain rdf:resource="#CorsiLaurea"/>
   </owl:ObjectProperty>
4. <owl:ObjectProperty rdf:ID="studiaMaterie">
   <rdfs:range rdf:resource="#Materie"/>
   <rdfs:domain rdf:resource="#Studenti"/>
   </owl:ObjectProperty>
5. <owl:ObjectProperty rdf:ID="haAmministrativi">
   <rdfs:range rdf:resource="#Amministrativi"/>
   <rdfs:domain rdf:resource="#CorsiLaurea"/>
   </owl:ObjectProperty>
6. <owl:ObjectProperty rdf:ID="haMaterie">
   <rdfs:range rdf:resource="#Materie"/>
   <rdfs:domain rdf:resource="#CorsiLaurea"/>
   </owl:ObjectProperty>
7. <owl:ObjectProperty rdf:ID="haDocenti">
   <rdfs:domain rdf:resource="#CorsiLaurea"/>
   <rdfs:range rdf:resource="#Docenti"/>
   </owl:ObjectProperty>
8. <owl:ObjectProperty rdf:ID="insegnaMaterie">
   <rdfs:range rdf:resource="#Materie"/>
   <rdfs:domain rdf:resource="#Docenti"/>
   </owl:ObjectProperty>
```

³ Non affrontiamo in questo contesto il linguaggio OWL; l'argomento sarà trattato nel prossimo modulo.

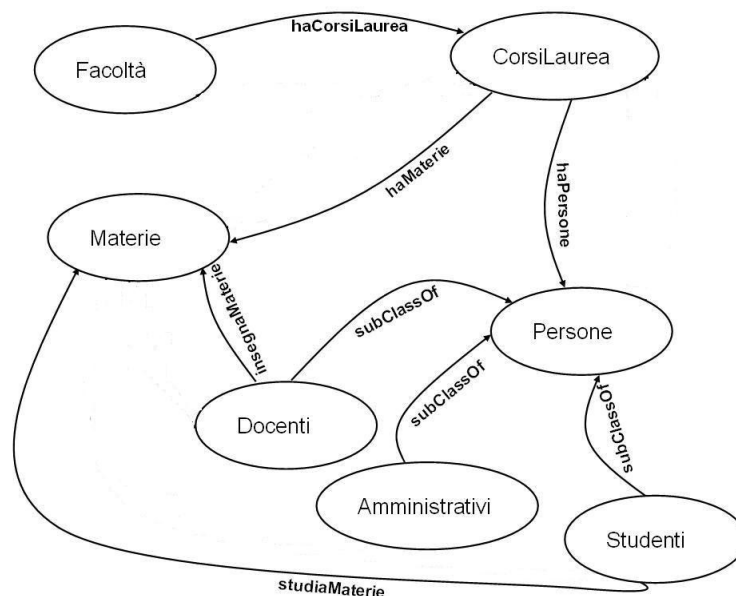
⁴ L'intento, in questo contesto, non è quello di far apprendere il linguaggio *OWL*, ma solo di far capire (intuire) come la rappresentazione grafica di un'ontologia, debba e possa essere 'addomesticata/trasformata' per renderla leggibile ad un elaboratore.

</rdf:RDF>

Si possono *intuire* le modalità con le quali il linguaggio *OWL* rappresenta le **classi** (riga1) e le **relazioni** (righe 2–8). I numeri di riga sono stati inseriti per facilitare la lettura e non vengono memorizzati nel file *OWL*. In riga 2 si dichiara che la relazione⁵ *haCorsiDiLaurea* si applica alla classe *Facoltà* (*domain*) e si riferisce alla classe *CorsiLaurea* (*range*).

Scrivere una ontologia con il linguaggio *OWL* è una operazione abbastanza complessa che si aggiunge alla complessità della progettazione della stessa ontologia. Tuttavia esistono dei programmi che facilitano questo lavoro. *Protégé*⁶ è uno di questi, è un prodotto *free* che facilita la creazione di ontologie attraverso l’inserimento di classi, proprietà, relazioni ed istanze. Il programma permette anche la creazione di rappresentazioni grafiche dell’ontologia.

Ad una analisi più approfondita si può notare che le classi *Docenti*, *Amministrativi*, *Studenti* possono essere considerate come sottoclassi di una classe *Persone* (superclasse), quindi l’ontologia diviene:

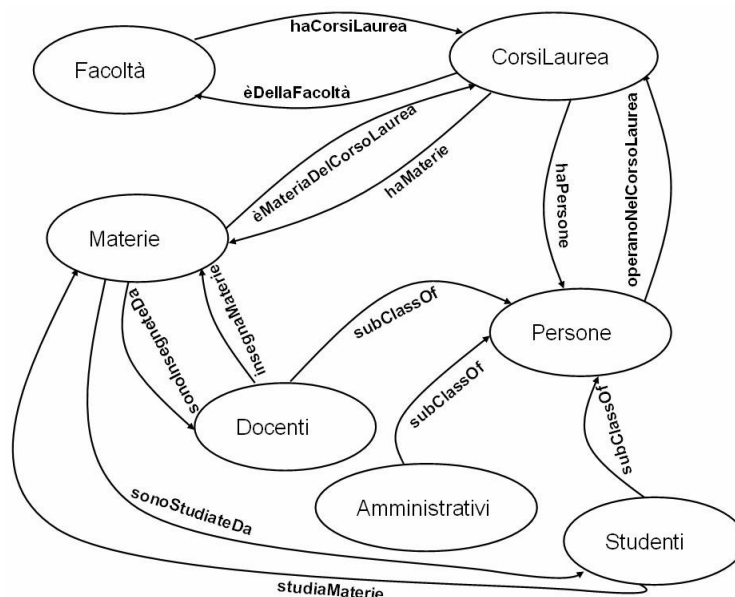


Le sottoclassi ereditano tutte le proprietà (attributi) della superclasse ed in più possono possederne di proprie; la sottoclasse *Studenti* può avere tutte le proprietà di *Persone* (ad esempio: *codice*, *cognome*, *nome*, *url*) ed in più potrebbe avere la proprietà *media*, contenente la media degli esami sostenuti.

Ogni relazione può avere una corrispondente relazione inversa. Se gli individui di una certa classe *A* sono in relazione con quelli di una classe *B*, la relazione inversa metterà in relazione gli individui della classe *B* con quelli della classe *A*. Considerando le inverse, la nostra ontologia diviene infine:

⁵Vedremo in seguito che in *OWL* le proprietà delle classi e le relazioni fra classi si chiamano entrambe proprietà; in particolare le proprietà si chiamano *datatype properties* e le relazioni *object properties*.

⁶URL: <http://protege.stanford.edu/download/release/full/>, verificato il 20 Giugno 2008.



In precedenza abbiamo detto che ciascuna classe è formata da individui (istanze, sono gli elementi che costituiscono la classe); ciascun individuo assume gli attributi delle classi di appartenenza. Occorre quindi definire gli attributi per ciascuna classe; nel nostro esempio potremmo decidere nel seguente modo:

Facoltà: codice, nome, url;

CorsiLaurea: codice, nome, url;

Materie: codice, nome, url;

Persone: codice, cognome, nome;

Docenti, Amministrativi: gli attributi ereditati da Persone;

Studenti: gli attributi ereditati da Persone e l'attributo *media*, contenente la media degli esami sostenuti dallo studente.

Teniamo presente che in OWL vengono prima creati gli attributi e poi questi vengono associati alla varie classi; in questo caso dapprima vengono creati i cinque attributi *codice*, *cognome*, *nome*, *url*, *media* e poi vengono associati alle classi secondo quanto qui sopra stabilito. Nella dispensa relativa alla creazione di un'ontologia con Protégé si vedrà un esempio.

Riepilogando, la costruzione di un'ontologia può avvenire attraverso le seguenti fasi:

i. progettazione e costruzione grafica come in Figura 2⁷;

⁷ Ciò significa:

1. individuare il dominio di interesse,
2. identificare i concetti fondamentali e creare i nodi della rete,
3. costruire legami.

Individuare il dominio di interesse rappresenta la fase preliminare alla progettazione e consiste nell'individuare cosa interessa rappresentare. Punto di partenza è porsi alcune domande fondamentali:

- Qual è il dominio di interesse?
- Quali sono i concetti principali intorno ai quali lavorare?
- Quali conoscenze occorre avere per svilupparli?
- Quali relazioni occorre prevedere?
- Quale utilizzo si prevede di fare della ontologia?

Una volta individuato il dominio, occorre selezionare i concetti fondamentali che lo definiscono. L'impegno va nella direzione di documentarsi e di sintetizzare per punti i contenuti individuati. La valutazione della rilevanza dei concetti può/deve avvenire in modo cooperativo; si inizia con un elenco che potrà essere integrato e corretto durante la realizzazione della rete ontologica. I concetti individuati diventano poi nodi della rete.

- ii. individuazione delle proprietà (attributi) delle varie classi;
- iii. realizzazione della ontologia in OWL usando un programma autore, per noi Protégé⁸:
 - a. creazioni delle classi,
 - b. creazione delle relazioni,
 - c. individuazione del dominio e del codominio di ciascuna relazione,
 - d. definizione di eventuali restrizioni sulle relazioni,
 - e. creazione degli attributi e definizione di eventuali restrizioni,
 - f. associazione degli attributi alle classi,
 - g. inserimento individui.

In conclusione vogliamo puntare l'attenzione su due aspetti: la progettazione iterativa delle ontologie e la differenza con i Database.

Un'ontologia deve riflettere, *nel miglior modo possibile*, la realtà del mondo che cerca di modellare, tenendo, tuttavia, presente che non vi è un metodo corretto a priori per modellare un dominio di conoscenza, che la soluzione migliore dipende dall'applicazione prevista e dalle sue estensioni future, e che il processo di sviluppo di un'ontologia è un'opera continuamente in divenire.

Ricordando i Database possiamo osservare che entrambi (Database e Ontologie) sono *metodologie di rappresentazione della conoscenza*. Con un Database l'attenzione è maggiormente centrata sui dati e sulle loro strutturazioni, in un'Ontologia si cerca di modellare concettualmente il mondo. Ricordando la progettazione di un Database, possiamo pensare lo schema concettuale, prodotto nella progettazione concettuale, come un'ontologia della realtà del problema. Il Database è maggiormente orientato alle istanze (individui), l'Ontologia alle entità e relazioni.

La costruzione dei legami è probabilmente la parte più complessa e comunque più costruttiva di tutto il procedimento. Occorre scegliere fra le connessioni possibili quelle ritenute significative nel contesto che si sta costruendo.

Il processo di collegamento fra i nodi genera continuamente una ristrutturazione della rete; infatti, si può evidenziare l'esigenza di inserire un nuovo concetto che possa permettere il collegamento fra altri due, oppure di suddividerne alcuni, oppure di aggiungerne di nuovi per completare l'ontologia.

Una ontologia può/deve costituire una rappresentazione condivisa di un dominio anch'esso condiviso. In tutte le fasi, la progettazione può/deve svilupparsi cooperativamente.

⁸ Come operare con Protégé verrà illustrato nell'apposita dispensa.

Usi delle ontologie

Molteplici sono i campi nei quali le ontologie sono utilizzate; uno di questi è quello del web semantico. Dapprima analizziamo il web semantico e poi una sua applicazione nel campo delle ricerche.

Questi temi verranno trattati nel successivo modulo.

Esempio di progettazione di una ontologia

Ontologia “Associazione Culturale”⁹

“Una associazione culturale vuole tenere traccia degli eventi che si svolgono nella propria città: in particolare si vuole sapere la data dell’evento, la sua tipologia, e una sua breve descrizione.

Di ogni evento si vuole inoltre sapere il luogo in cui si svolge, specificandone il nome, l’indirizzo, la capienza ed il proprietario.

Si vuole inoltre tener traccia degli artisti che partecipano all’evento ed il loro ingaggio. Un artista può partecipare a più eventi. Di ogni artista si vuole sapere nome, cognome e la sua arte. Inoltre ad ognuno viene assegnato un voto di interesse compreso tra 1 (noioso) a 5 (fenomenale).

Per ogni evento deve infine essere identificato l’organizzatore (nome, indirizzo, telefono); ogni evento ha un solo organizzatore.”

Determinare il dominio e lo scopo dell’ontologia

- Quale è il dominio che deve essere descritto dall’ontologia?
- Per quale scopo creiamo l’ontologia?
- Come vogliamo utilizzarla?
- A quali tipi di domande l’ontologia dovrebbe fornire una risposta?

Elencare i termini principali dell’ontologia

- Quali sono i termini di cui abbiamo necessità di parlare?
- Quali sono le proprietà di questi termini?
- Cosa vogliamo rappresentare con questi termini?

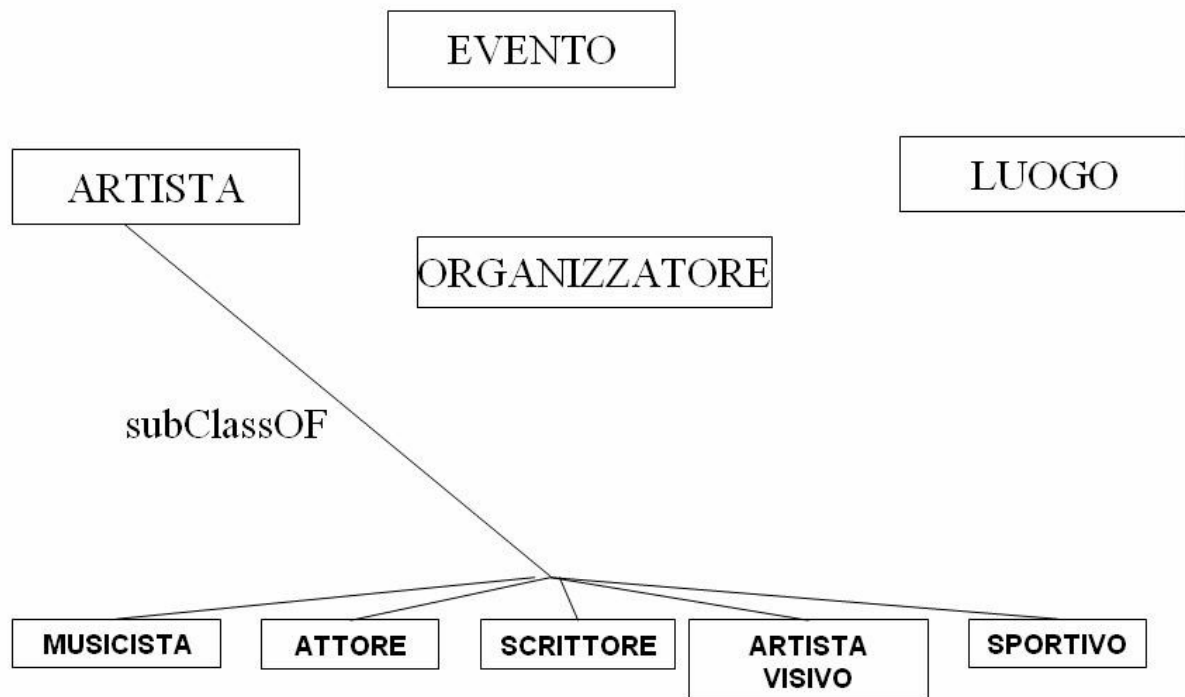
In questa fase è possibile elencare tutti i termini che si ritengono siano di ‘competenza’ dell’ontologia; solo successivamente potranno essere individuati quelli che possono rappresentare i concetti, quelli che possono rappresentare gli attributi, quelli che possono aiutare nella costruzione delle relazioni ed infine quelli che possono essere scartati.

Definire le classi e le sottoclassi

- i concetti (le classi) corrispondono solitamente ai sostantivi utilizzati nel dominio di conoscenza;
- è importante distinguere tra le classi e i relativi nomi;
- le classi rappresentano concetti nel dominio, e non le parole usate per esprimere tali concetti;
- i sinonimi non costituiscono classi diverse.

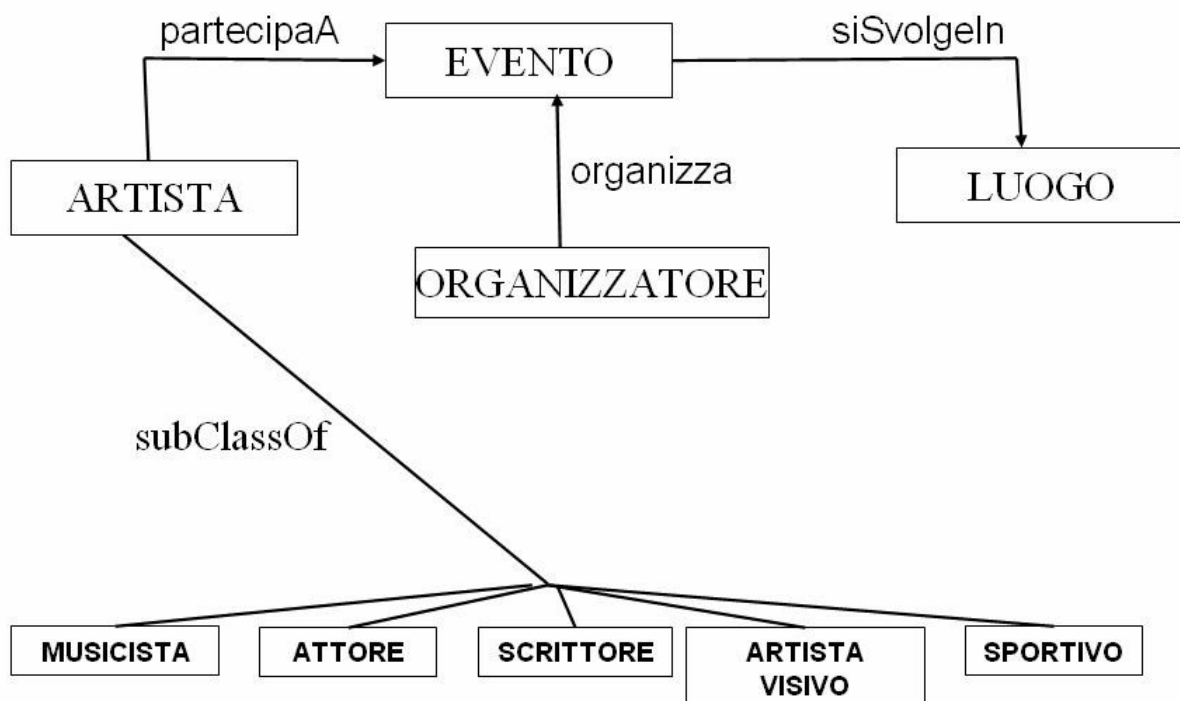
Dai termini individuati in precedenza selezionare quelli che possono individuare le classi; in questo caso potrebbero essere: *artista, evento, luogo, organizzatore*; inoltre *artista* potrebbe avere delle sottoclassi: *attore, musicista, scrittore, ...*.

⁹ Tratto da <http://www.di.unito.it/~carmagno/sw/lez3.ppt>

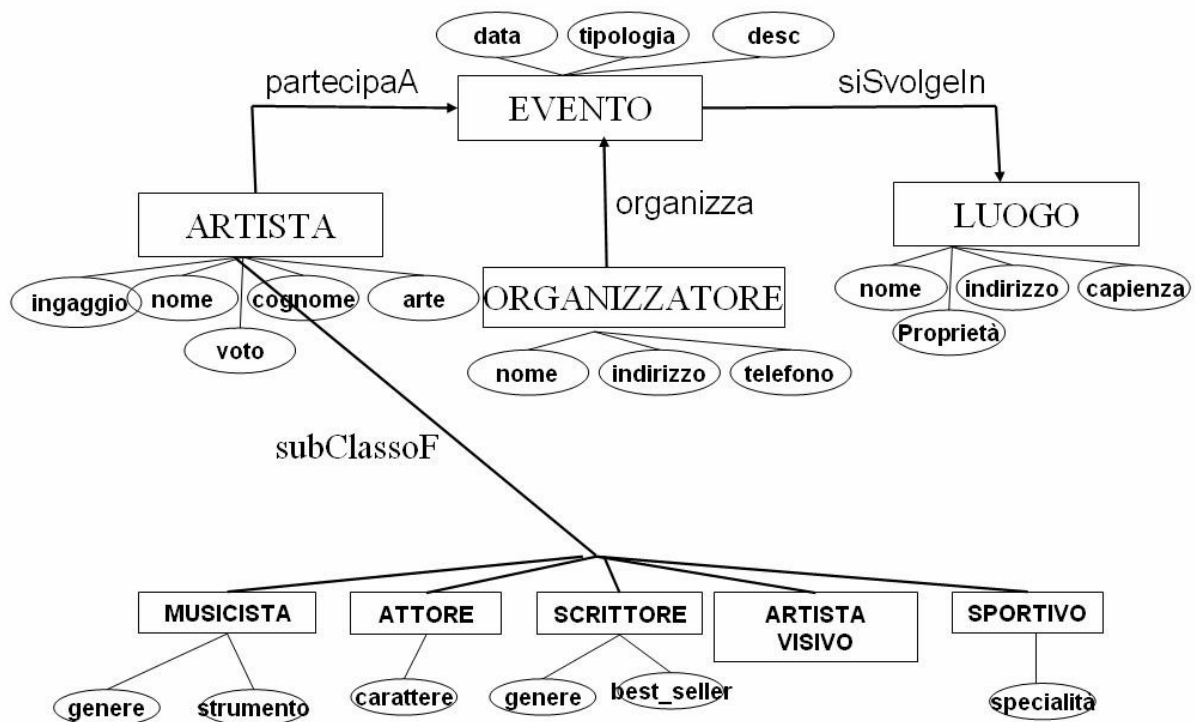


Definire le relazioni fra classi

Definire le relazioni fra le classi, stabilendo a quale classi si applicano (dominio) e a quali si riferiscono (codominio)



Definire le proprietà (attributi) delle classi



Una volta realizzata la rappresentazione grafica dell'ontologia è possibile renderla elaborabile dalla macchina, traducendola, ad esempio, in *OWL*. Per far ciò è possibile usare il programma Protégé.