

Ontologia

Semantica lessicale:

È lo studio del significato delle parole e delle loro relazioni

L'**ontologia**, in senso lato, è definibile come lo studio dell'essere (*un ingrediente fondamentale per riuscire ad adempiere al nostro task di rappresentazione della realtà*)

L'ontologia è sempre stato un grande tema di discussione per i filosofi ma negli ultimi anni anche l'**intelligenza artificiale** e la **linguistica computazionale** hanno iniziato a ricercare e a lavorare sul concetto di ontologia al fine di rappresentare informazioni e stralci di realtà.

La definizione filosofica di ontologia differisce leggermente dalla visione della computer science:

Nel primo caso l'essere è studiato nei termini di ciò che c'è perché esiste e non studiandone le caratteristiche dettagliate (uno studio fine a sé stesso insomma)

Nel secondo caso siamo interessati allo sviluppo di una struttura semantica vera e propria, formalizzata e con base logica che sia interrogabile e comprensibile da sistemi automatici.

Un'ontologia permette di rappresentare la struttura della conoscenza

Le ontologie sono uno strumento fondamentale sia per la condivisione della conoscenza (fra le persone e fra gli agenti) che per il riutilizzo dei dati e dell'informazioni stessa.

Perché costruire ontologie

Per la condivisione della comprensione delle entità in un certo dominio

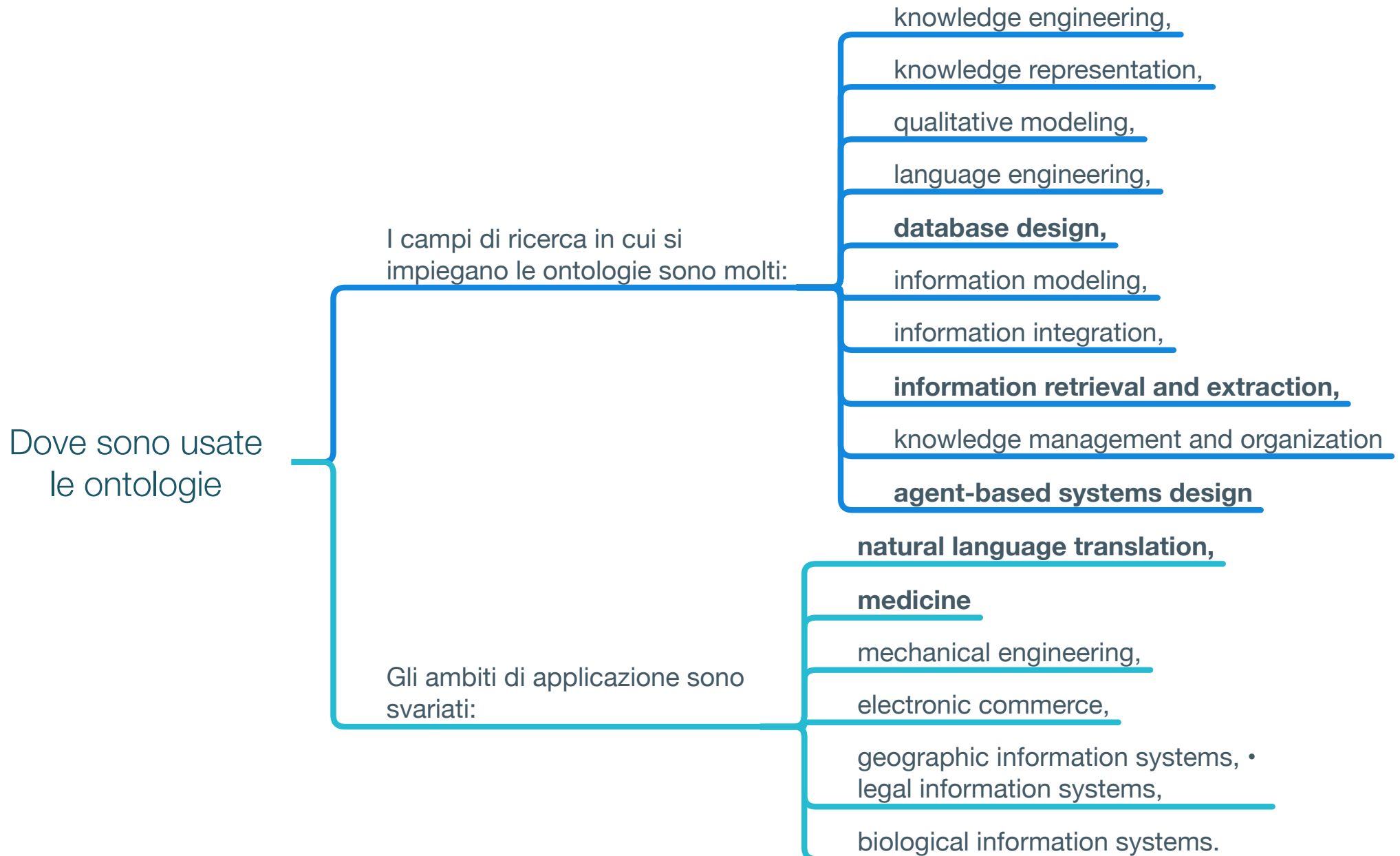
Condivisione tra persone, software

In funzione del riutilizzo dei dati e dell'informazione

Per non ripartire sempre da 0

Introdurre degli standard

Per creare comunità di ricercatori





Cosa hanno
in testa
i computer

Uno dei più grandi traguardi ottenuti dell'intelligenza artificiale sono senza ombra di dubbio i **voice assistant**, quali **Siri** o Cortana, o Alexa. Questi prodotti, che noi riteniamo *'intelligenti'*, *quanto conoscono bene la realtà?*

Se proviamo a chiedere a Siri Cosa ne pensi di Bardonecchia? non otterremo alcuna risposta soddisfacente. Questo poiché Siri non dispone del concetto di Bardonecchia, anzi, non dispone di nessun concetto.

Siri è infatti un agente che è in grado di interagire con grande abilità relativamente alle funzionalità dell'iPhone, ma non certo di 'ragionare' in senso lato.

Non confondiamo dunque l'intelligenza di un agente in senso lato con la rappresentazione della conoscenza: il nostro compito è più complicato di quanto si immagini.



Consideriamo il classico problema di un bicchiere capiente 10cc pieno riempito con 5cc di acqua. Potremmo descrivere la situazione come:

Il bicchiere è mezzo pieno.

Il bicchiere è mezzo vuoto.

Il bicchiere è riempito a metà della sua capacità totale.

Il bicchiere contiene 5cc di acqua.

Ciò che abbiamo appena osservato è il problema del **commitment ontologico**:

A seconda della descrizione che diamo della realtà è come se indossassimo un paio di occhiali da sole che filtrano alcuni aspetti della realtà evidenziandone invece altri.

Il commitment ontologico

Ne deduciamo che tutte le rappresentazioni della realtà sono intrinsecamente imperfette ed approssimate poiché quando si sceglie come rappresentare qualcosa, inevitabilmente si tralasciano alcuni aspetti dell'oggetto che si sta rappresentando.

In questo senso, il commitment ontologico non è soltanto un effetto collaterale della scelta di una metodologia di rappresentazione, ma è bensì l'essenza della rappresentazione stessa.

Definiamo un'ontologia

Ora entriamo più nel dettaglio parlando di cosa sia effettivamente una struttura ontologica.

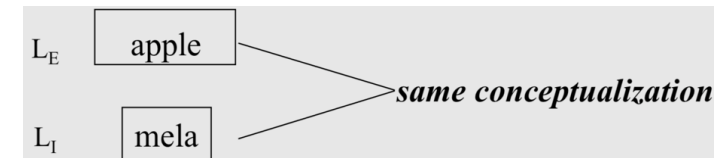
Come prima definizione potremmo dire che un'ontologia definisce un insieme di primitive rappresentazionali con le quali modellare un dominio di conoscenza o di discorso.

Le *primitive rappresentazionali* sono **classi** e **relazioni** e le loro definizioni includono informazioni sul loro *significato* e *vincoli* su come applicarle in maniera logicamente consistente.

Ora vediamo come raffinare questa definizione che è un pò troppo generare al momento

Concettualizzazioni

Le ontologie sono fondate su concettualizzazioni, ovvero strutture formali che rappresentano pezzi di realtà indipendenti dal vocabolario e del contesto in cui vengono utilizzate.



Chiaramente diversi termini (eventualmente di diversi vocabolari) che coinvolgono gli stessi oggetti possono avere la stessa concettualizzazione. Ad *es.* *apple* e *mela* avranno la stessa concettualizzazione.

Ontologia

Studiamo cosa c'è

Semantica

Studiamo come ci si riferisce a ciò che c'è

Ontologia e semantica

È importante non confondere l'ontologia con la semantica:

Nota:

Le due cose, sebbene differenti, sono fra loro collegate. Infatti, parole che hanno determinati ruoli semantici ricopriranno determinate componenti all'interno dell'ontologia.

Definiamo un'ontologia

Ontologia e basi di conoscenza

Le basi di conoscenza sfruttano le ontologie per poter rappresentare porzioni di realtà. Si hanno due grandi componenti:

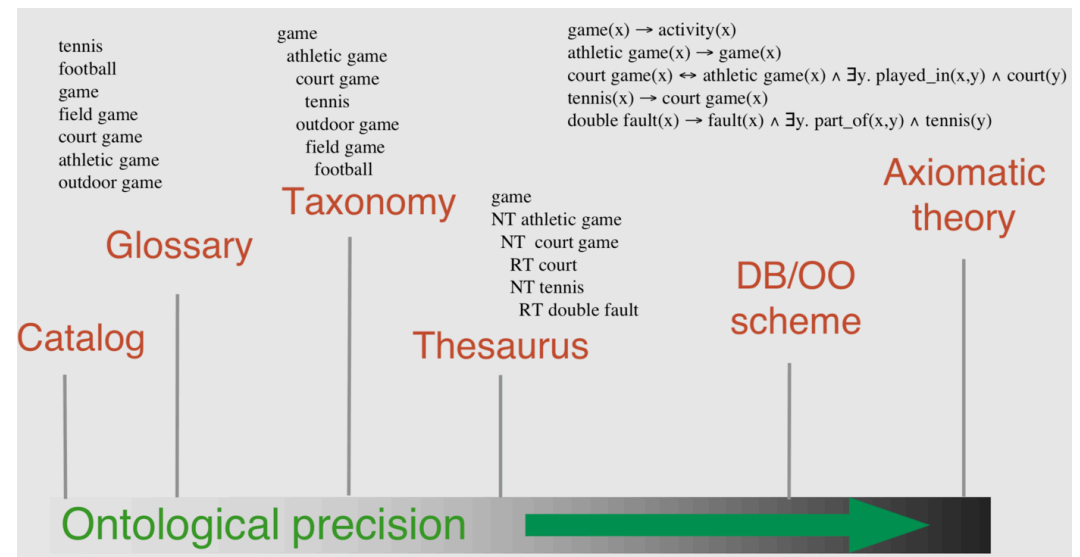
A-BOX *Componente terminologica:* Costituita dall'ontologia, che è indipendente da un preciso stato della realtà ma mantiene informazioni vere in generale.

T-BOX *Componente asserzionale:* che riflette uno specifico stato della realtà.

Es. potremmo dire che il fatto che gli uccelli volino fa parte della componente terminologica mentre il fatto che Tweety, un determinato uccello, stia volando fa parte della componente asserzionale.

Diversi livelli di ontologie

In Figura vediamo le varie tipologie di ontologie che si possono costruire: tanto più un'ontologia è complessa, tanto più sarà precisa ed informativa nel rappresentare la realtà.



I diversi livelli di precisione delle ontologie. Da destra verso sinistra vediamo aumentare la complessità della e quindi la precisione dell'ontologia: partiamo dal mero insieme di termini senza relazioni fra loro (il **catalogo**), per poi passare alla **tassonomia** che istituisce una relazione di gerarchia, al **thesaurus** che fornisce anche informazioni semantiche ed infine alle formule nella **logica del prim'ordine**.

Nota:

Possiamo già anticipare che un approccio che usi la logica del prim'ordine è praticamente impossibile per via della complessità che richiederebbe, studieremo in seguito invece delle tecniche decisamente più usabili per rappresentare la conoscenza.

Ontologia e lessico

Il lessico è lo strumento fondamentale mediante il quale rappresentiamo i concetti.

Fra gli elementi del lessico vi possono essere diverse **relazioni**:

Omonimia:

Due lessemi con stessa forma ma con sensi diversi (*significati diversi*)

Es. A **bank** can hold the investments vs We can go on the right **bank** (*sponda*) of the river.

Significato completamente diverso

Polisemia:

Lo stesso lessema con due sensi diversi (*significati simili*)

Es. A bank can hold the investments vs He got the blood from the bank (*la banca del sangue*).

In questo caso l'etimologia attribuita a bank è la stessa (stesso lessema), ma ovviamente i sensi sono diversi, sebbene correlati.

I due sensi per quando diversi sono correlati

Sinonimia:

Due lessemi con forma diversa ma stesso significato

Es. how **big** is that plane? vs how **large** is plane?

Iponimia e iperonimia:

Due lessemi di cui uno denota una sottoclasse dell'altro.

Es. automobile è un **iponimo** (*più specifico*) di veicolo, mentre veicolo è un **iperonimo** (*più generale*) di automobile.

Meronomia e olonimia:

Due lessemi i cui significati sono uno la parte dell'altro.

Es. dito è **meronimo** di mano e quindi mano è **olonimo** di dito.

Antonimia:

Due lessemi con significato opposto.

Es. ricco e povero.

Ontologia e lessico

Parte delle relazioni scritte nella MM precedente possono essere utili per costruire un'ontologia, istituendo così **relazioni fra concetti**.

Es. la relazione di iponimia è molto simile alla relazione di subclass-of, mentre la meronomia è molto simile alla relazione di has-part.

Purtroppo il lessico porta con sé anche **alcune problematiche**, fra cui:

Overlapping dei significati:

Nelle ontologie le sotto-categorie di una data categoria sono normalmente intese come mutualmente esclusive, ma nel lessico esistono sovrapposizioni parziali di significati.

La sovrapposizione di significati non rappresenta di per sé un problema (si tratta infatti di una sinonimia, semplicemente assoceremo i due lessemi allo stesso concetto), ma l'essere **“quasi sinonimi” è invece difficile da rappresentare.**

Lexical gap:

Si tratta della presenza di categorie ontologiche (concetti) non lessicalizzate in un determinato linguaggio (*una lingua non li ha*).

Tali categorie richiedono delle perifrasi per essere descritte debitamente nelle lingue in cui è presente il lexical gap.