

Stefano Minozzi
(Università degli Studi di Verona)

Filologia ed esegesi: il ruolo delle tecniche computazionali di analisi lessicale e semantica

Seminario di Studi
Progetti e Prospettive di Filologia Digitale
Verona, 4 – 5 ottobre 2007

Tecniche di analisi computazionale e “*filologia informatica*”

- Filologia digitale:
 - il calcolatore:
 - *strumento di organizzazione* del processo di edizione
 - indispensabile per il *display* dell’oggetto “edizione digitale”
 - tecniche di analisi computerizzata sono ancillari al processo di edizione
- Filologia computazionale:
 - il calcolatore:
 - sostituisce l’intervento umano in alcuni momenti del processo di edizione (*recensio, eliminatio codicum descriptorum, collatio*)
 - Indispensabile per l’esecuzione *algoritmi*
 - tecniche di analisi computerizzata sono parte integrante del processo di edizione

Prima dell'analisi: il ruolo della *codifica*

- codifica: momento condiviso tra filologia digitale e filologia computazionale
- si passa da
 - preoccupazione legata alla rappresentazione (display)
 - a
 - codifica come fatto interpretativo
- rappresentazione del significante → rappresentazione/interpretazione dell'intenzionalità del segno

Prima dell'analisi: il ruolo del mark-up

- Insiste sulla **codifica**, intendendo il termine nell'accezione informatica (codifica binaria)
- È un livello di **codifica**, intendendo la il termine nell'accezione semiotica
- È un atto interpretativo che sviluppa/aggiunge sul testo strutture di significato
- *Fa superare i limiti della codifica binaria, rendendo possibile la computazione di elementi di significazione non altrimenti rappresentabili al calcolatore*

Un segno, o representamen, è qualcosa che sta a qualcuno per qualcosa sotto qualche rispetto o capacità [...]

Definisco un Segno come qualcosa che da un lato è determinato da un oggetto e dall'altro determina un'idea nella mente di una persona, in modo tale che quest'ultima determinazione, che io chiamo interpretante del segno, è con ciò stesso mediatamente determinata da quell'oggetto

da C. Peirce, *Collected Papers*, trad. it. Massimo Bonfantini, Einaudi, 1980

Problemi che si presentano all'analisi computazionale

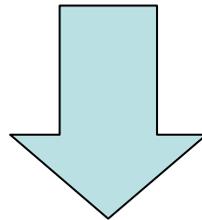
- inferenza vs elaborazione → il computer opera solo sui significanti
- “Formato testo” è “stringa di caratteri” → rappresentazione inadeguata della *testualità*
- Testo elettronico: diventa un luogo da mappare

. . . In quell'Impero, l'Arte della Cartografia giunse a una tal Perfezione che la Mappa di una sola Provincia occupava tutta una Città, e la Mappa dell'Impero tutta una Provincia. Col tempo, queste Mappe smisurate non bastarono più. I Collegi dei Cartografi fecero una Mappa dell'Impero che aveva l'Immensità dell'Impero e coincideva perfettamente con esso. Ma le Generazioni Seguenti, meno portate allo Studio della Cartografia, pensarono che questa Mappa enorme era inutile e non senza Empietà la abbandonarono alle Inclemenze del Sole e degli Inverni.

da **Jorge Luis Borges**, *Storia universale dell'infamia*, 1954,
tr. it. Il saggiatore, Milano 1961

Tre livelli di interrogabilità e di analisi del testo

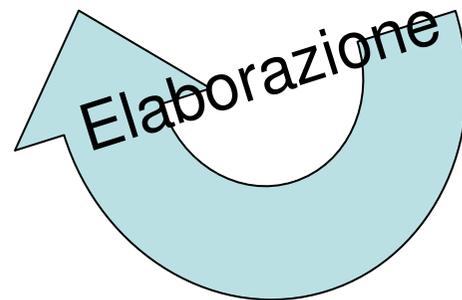
- testuale → *text-retrieval*
 - lessicale
 - semantica
- } *text-mining*

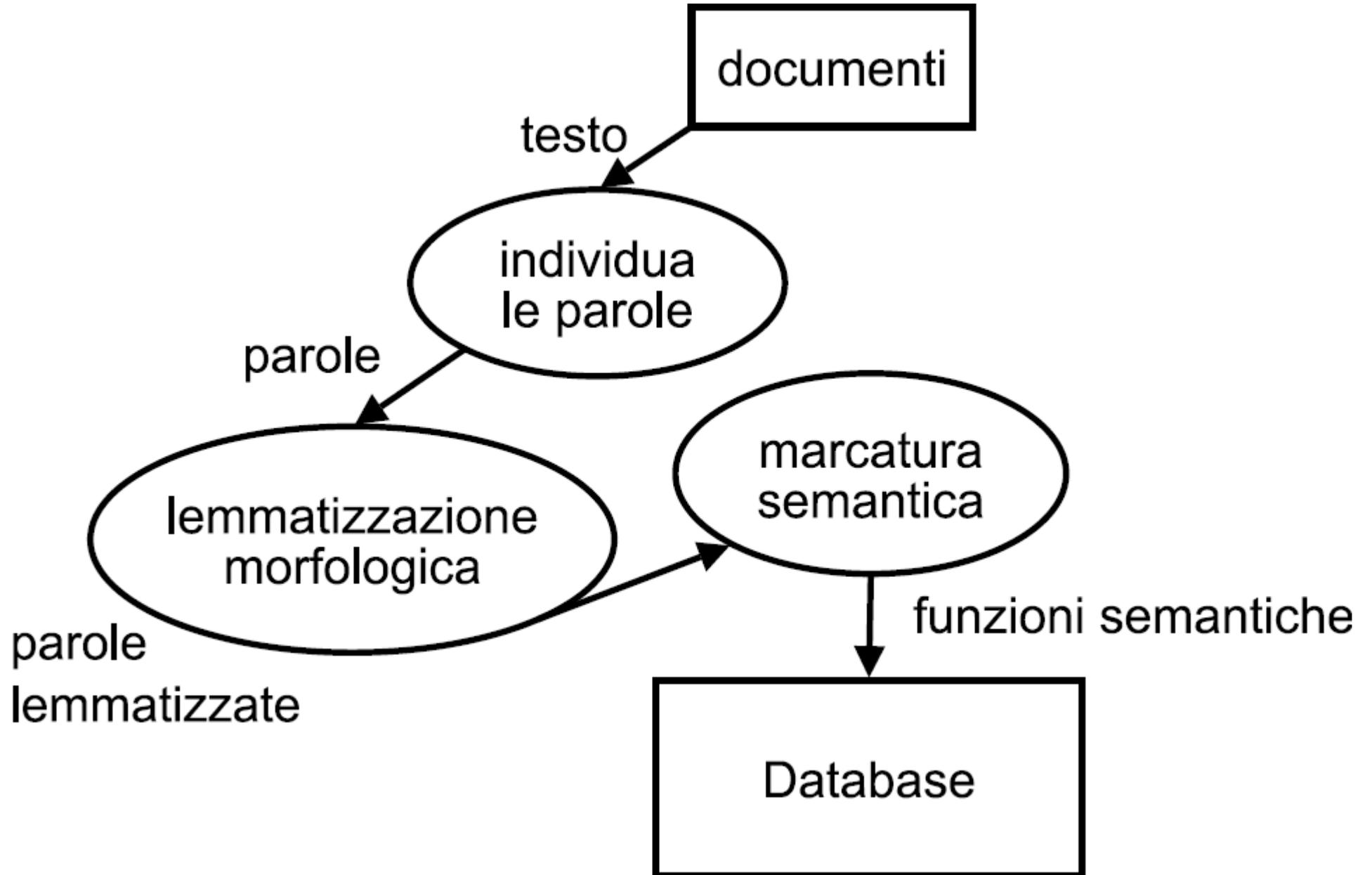


Lexicographie prédictionairique (Quemada)

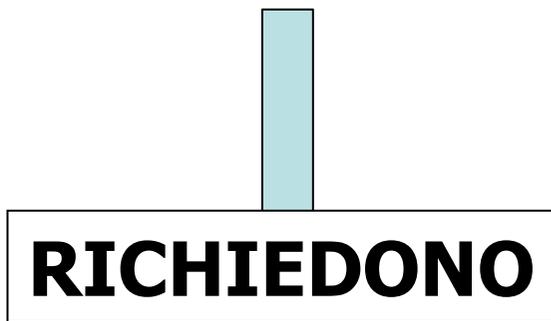


TESTO MARCATO

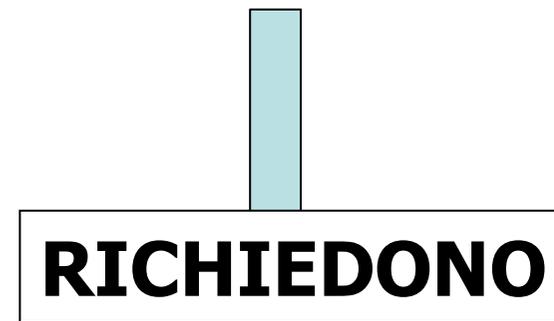




Problemi di identificazione:
omografi ed ***eterografi***



disambiguazione



assimilazione

Struttura di uno strumento di analisi

- **COMPETENZA:** insieme di processi di elaborazione, decisione e controllo che permettono l'esecuzione di un compito (*task*)
- **MEMORIA:** dati esterni utilizzabili dai processi ed esterne ai dati che si stanno elaborando