



[ve]dph



SAPIENZA  
UNIVERSITÀ DI ROMA

Federico Boschetti, Paolo Monella

# IL DIALOGO FRA ESPERTI DI WEB SEMANTICO ED ESPERTI DEL DOMINIO: PROBLEMI E SOLUZIONI

*Workshop Web semantico e antropologia del mito*  
Università per Stranieri di Siena, 25 giugno 2021

# Workshop Proteus 2020

- 04.09.2020
- Tutti i materiali (tagli libro)
  - Schede
- Formalizzazione, DH
- Feedback, proposte

# Workshop Proteus 2021

- Schede → Testi
- Cristiana Franco
  - Più miti consigli
  - Un percorso di navigazione
  - Connessioni tematiche
- Gloria Mugelli
  - Estrarre informazione
    - Citazioni
    - Entità nominate (persone, luoghi...)
    - Parole chiave / temi → Ontologia

# Proteus: gli inizi del progetto

- Simone Beta
- Atlante del mito classico
  - classi: uomini, animali, mondo naturale, dei, parentela, economia...
- Volume cartaceo: troppo materiale
  - 400-600 pagine
  - nonostante taglio del 50%

# Proteus: nuovo progetto

- Quantità
  - Materiali tagliati
- Qualità
  - Dimensione narrativa del mito
  - Aggiunti: iconografia, Nachleben, tema/classe guerra...
- Formalizzazione, DH

# Colori in queste slide

- Rosso
  - Questioni, problemi, cose cui fare attenzione
- Blu
  - Proposte, opportunità

# Voce dell'atlante

- Formalizzare partendo da un file Word?
  - Celle / colori
- Testo: meno strutturato della scheda
  - Cristiana Franco: destrutturare in oggetti?

# Citazioni

- CTS / DTS
  - Problema: i tardo-antichi
- Testo attestazioni
  - Si può linkare a Perseus / [Scaife Viewer](#)
    - (O importare il testo da Perseus automaticamente tramite CTS)



# Entità / Luoghi

- Per essere più granulari: risorse web semantico (Pelagios)

# Entità / Personaggi

- LOD
- Attanti
- Varianti dei personaggi (Pandareo / Pandione)
  - Non è facile identificare il personaggio

# Fluidità / Tempo

– Il tempo del mito?

# Fluidità / Varianti mitiche

## – Trama vs ‘altre versioni’

- una ha uno status speciale?
- oppure tutte le vers. sullo stesso livello?
- conflare versioni diverse dai dizionari di mitologia in una trama unica?

# Temi: categorie (modello)

- Sezione/sotto-sezione → categorie
- F. Boschetti: complessità crescente dei tipi di modello
  - folksonomy
  - list
  - synonymy list
  - taxonomy
  - corpus
  - ontology

# Metodi e tecnologie illustrate

- T2K per estrarre la terminologia e le relazioni fra i termini
- Identificazione delle citazioni
- Ontologie fondazionali e ontologie di dominio

# Come integrare strumenti e risorse?

- Usare soluzioni pronte all'uso
  - studiosi interessati al progetto, che purtroppo non possono dedicare tempo allo sviluppo di soluzioni ad hoc
  - allineare i risultati dei diversi strumenti di analisi automatica

# Validare i risultati automatici a campione

- Paragonare l'estrazione automatica di terminologia con i campioni annotati manualmente e valutare l'accuratezza
- Validare l'accuratezza delle citazioni identificate automaticamente



# Correzione dei risultati automatici

- Quale editor usare per correggere manualmente l'output degli strumenti di analisi?
- Quanto tempo richiede?
- È possibile coinvolgere tirocinanti e collaboratori esterni?

# Annotazione manuale

- Affiancare l'annotazione in stand-off all'annotazione inline
- Quali strumenti di annotazione usare?
  - Euporia
  - Hypothes.is
  - Pundit
  - Altro?

# Granularità delle annotazioni

- Quali sono i livelli di granularità delle annotazioni?
  - termini monorematici
  - termini polirematici
  - blocchi di testo di varia ampiezza

# I testi

- Come risolvere il problema delle licenze?
  - caricare nel sistema i testi integrali, ma mostrare soltanto snippets
  - lavorare con le schede e non con i testi a stampa

# Collegare l'ontologia al testo

- Come collegare l'ontologia al testo in modo efficiente?
  - Espansione di query come in EuporiaRAGT
  - Altro?

# Casi d'uso

- Cosa si aspettano dal motore di ricerca gli esperti di dominio (antropologi del mondo antico, storici, filologi, storici della lingua, studiosi del mondo antico in generale)?
- Come raccogliere i casi d'uso, in modo da passare dall'**analisi dei requisiti** alla **definizione delle specifiche**?

# Ontologie

- Quale tipo d'interazione è necessario fra gli esperti di dominio, la sviluppatrice dell'ontologia di dominio e gli specialisti di ontologie?
- Come organizzare l'eventuale consulenza scientifica?
- Quale potrebbe essere il ruolo di CLARIN?

# Interfaccia d'interrogazione

- Come dovrebbe presentarsi l'interfaccia d'interrogazione?
- Come vengono presentati i risultati?
- Da chi sarà sviluppata?