



*Verso la definizione e  
l'implementazione di una  
piattaforma orientata allo studio  
critico del testo*

**Angelo Mario Del Grosso**

*Istituto di Linguistica Computazionale "A. Zampolli"*

`angelo.delgrosso@ilc.cnr.it`



## *Posizione Attuale*



Ricercatore III livello del Consiglio Nazionale delle Ricerche, in servizio permanente presso l'Istituto di Linguistica Computazionale "A. Zampolli", Area della Ricerca di Pisa

## *Background*



Laureato in Ingegneria Informatica, PhD in Ingegneria dell'Informazione con tesi nell'ambito delle Digital Humanities e del Digital Scholarly Editing presso la Scuola di dottorato L. da Vinci di Pisa

## *Interessi e Attività di ricerca*



Analisi, progettazione e sviluppo di modelli e strumenti software per sistemi di linguistica e filologia digitale/computazionale volti allo studio di testi di tradizione medievale, a stampa e di autori moderni e contemporanei, favorendone lo studio anche attraverso l'allineamento con le rispettive scansioni e traduzioni eseguite in antico o realizzate in epoche successive.

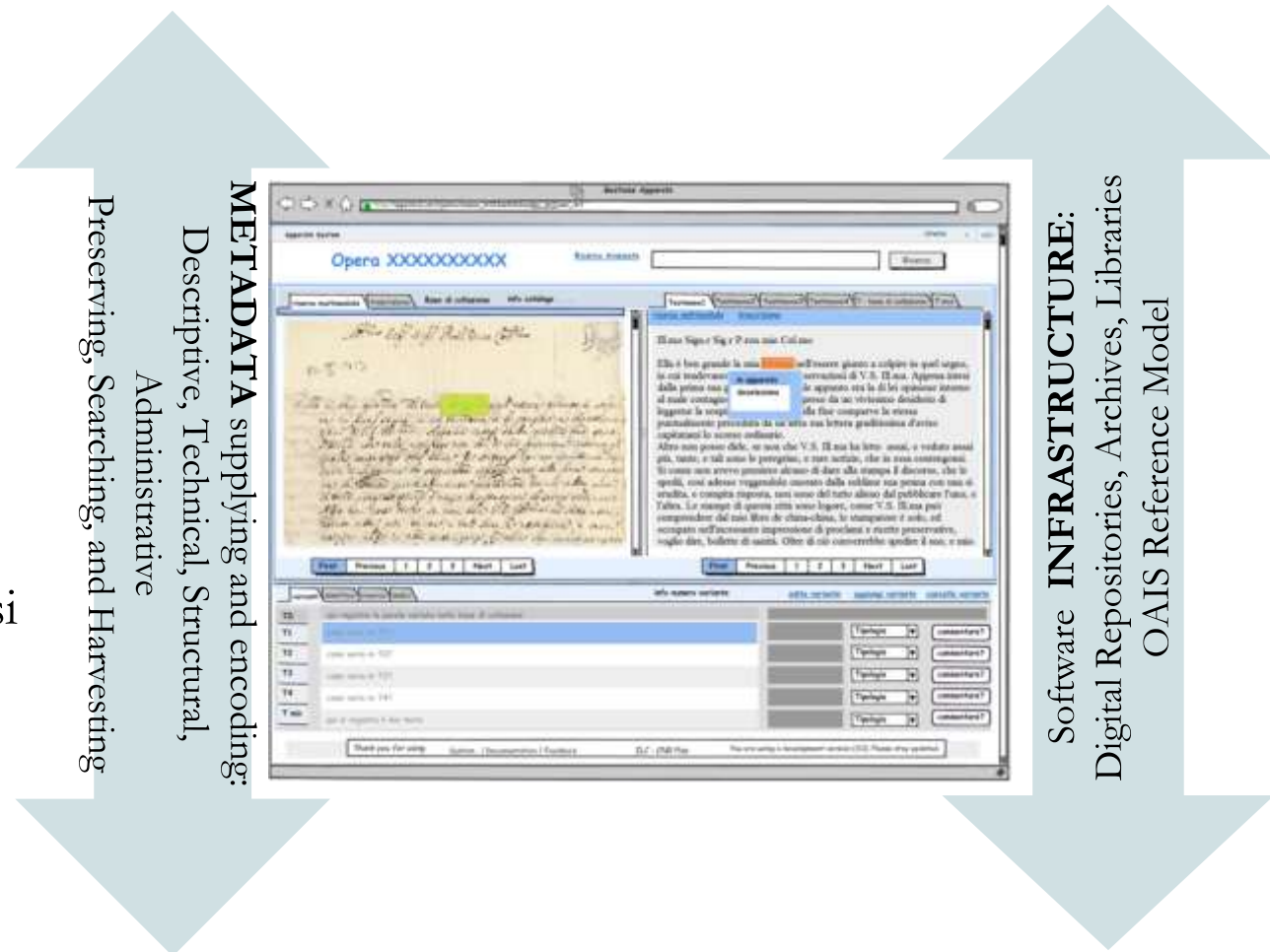
- ❖ Requisiti funzionali nel settore della filologia computazionale
- ❖ Rappresentazione e trattamento digitale del testo e del documento
- ❖ Piattaforma per lo studio scientifico del testo (Filologia Computazionale)
- ❖ Esempi e casi di studio in scenari concreti di progetto
- ❖ Video Demo:
  - L'edizione digitale delle lettere di Bellini
  - Il software LESMU (Lessico della letteratura musicale italiana)

Principali Esigenze

- ✓ Acquisizione
- ✓ Rappresentazione
- ✓ Trascrizione
- ✓ Produzione (Editing)
- ✓ Studio e Traduzione
- ✓ Gestione
- ✓ Conservazione
- ✓ Elaborazione e Analisi
- ✓ Valorizzazione
- ✓ Interrogazione
- ✓ Annotazione
- ✓ Fruizione e Visualizzazione

Entità di Riferimento

Documenti testuali di interesse storico e letterario

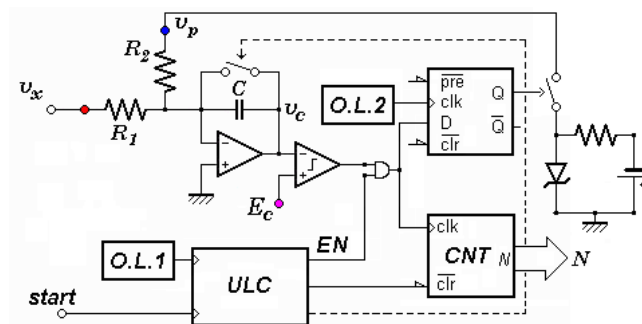


METADATA supplying and encoding:  
Descriptive, Technical, Structural,  
Administrative

Software INFRASTRUCTURE:  
Digital Repositories, Archives, Libraries  
OAIS Reference Model

**Nell'ambito dello studio scientifico del testo** è sentita l'esigenza di sviluppare sistemi che consentano di:

- **navigare in grandi archivi** di *testi* e di *immagini* digitali favorendo il **riconoscimento e l'estrazione** di fenomeni *testuali* e *intertestuali* quali passi paralleli, citazioni esplicite o implicite;
- **il ritrovamento e la classificazione di errori e varianti testuali** in documenti che rappresentano la storia di una stessa opera (es. copie di manoscritti con variazioni);
- produrre **indici e concordanze** sia del *testo* sia delle informazioni di *contesto* e *paratesto* (es. apparato critico, commenti, repertori, bibliografia), aumentando le possibilità di effettuare *scelte editoriali corrette*;
- **annotare** un testo in modo **collaborativo** da parte di una molteplicità di utenti, anche con *diverse specializzazioni* (filologi, linguisti, storici, filosofi, ecc).

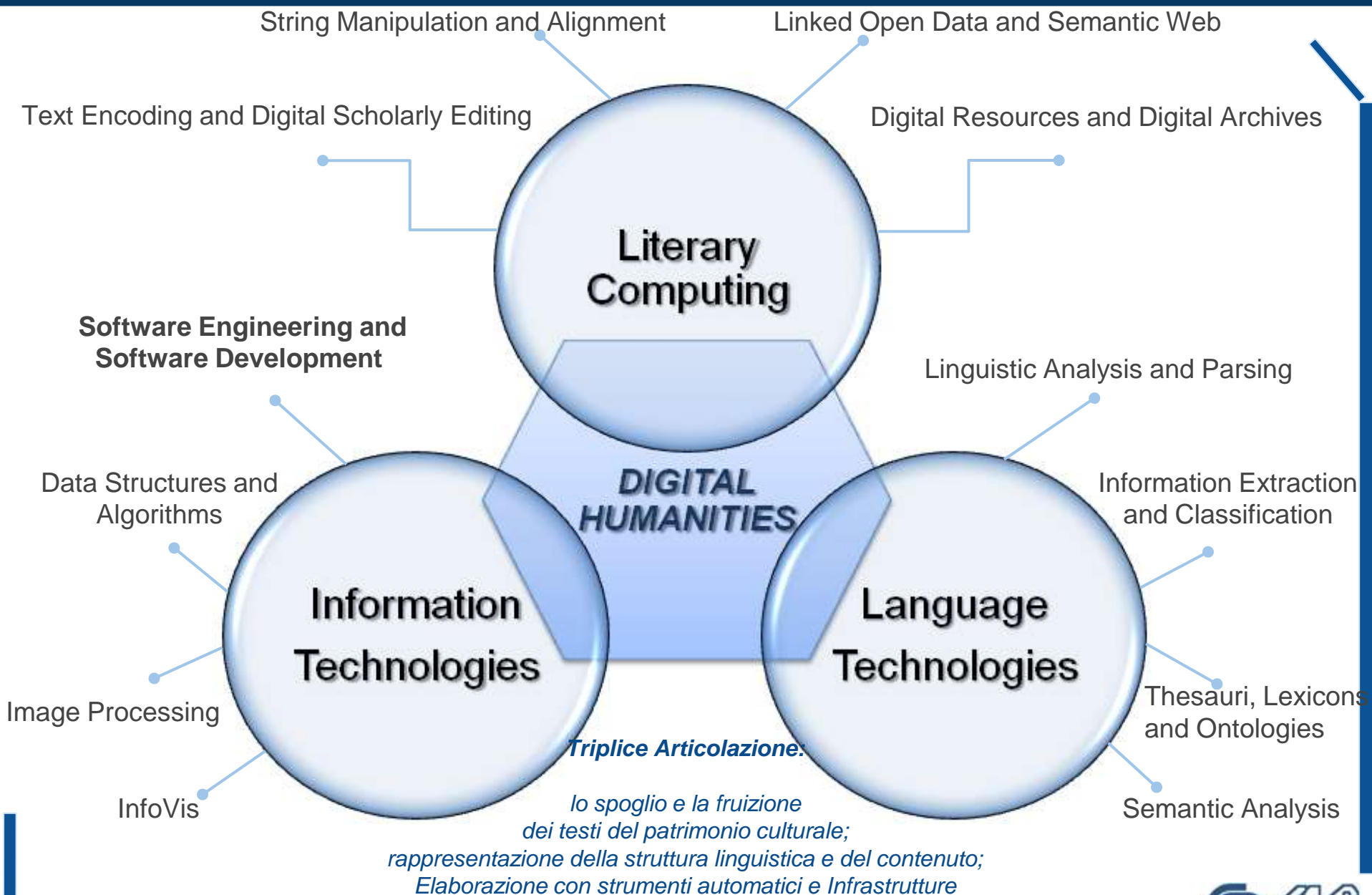


## Sperimentare l'applicazione di tecnologie digitali nell'ambito degli studi umanistici

The screenshot shows a digital interface for text analysis. At the top, a ruler indicates line numbers from 1 to 15. Below it, the text 'Théorie des sonantes' is visible. A specific line is highlighted with a yellow background, containing the text 'lasciato bianca 42r e ha scritto ancora su 43r capovolta'. Below this, there are several numbered annotations in French, such as '1. À quoi se reconnaît un sonantiste?' and '2. nous pouvons examiner les objections contre la conception courante'. The interface also includes checkboxes for 'Sul margine sinistro, scritto trasverso' and 'contre', and a legend for symbols like 'le' and 'qu'.

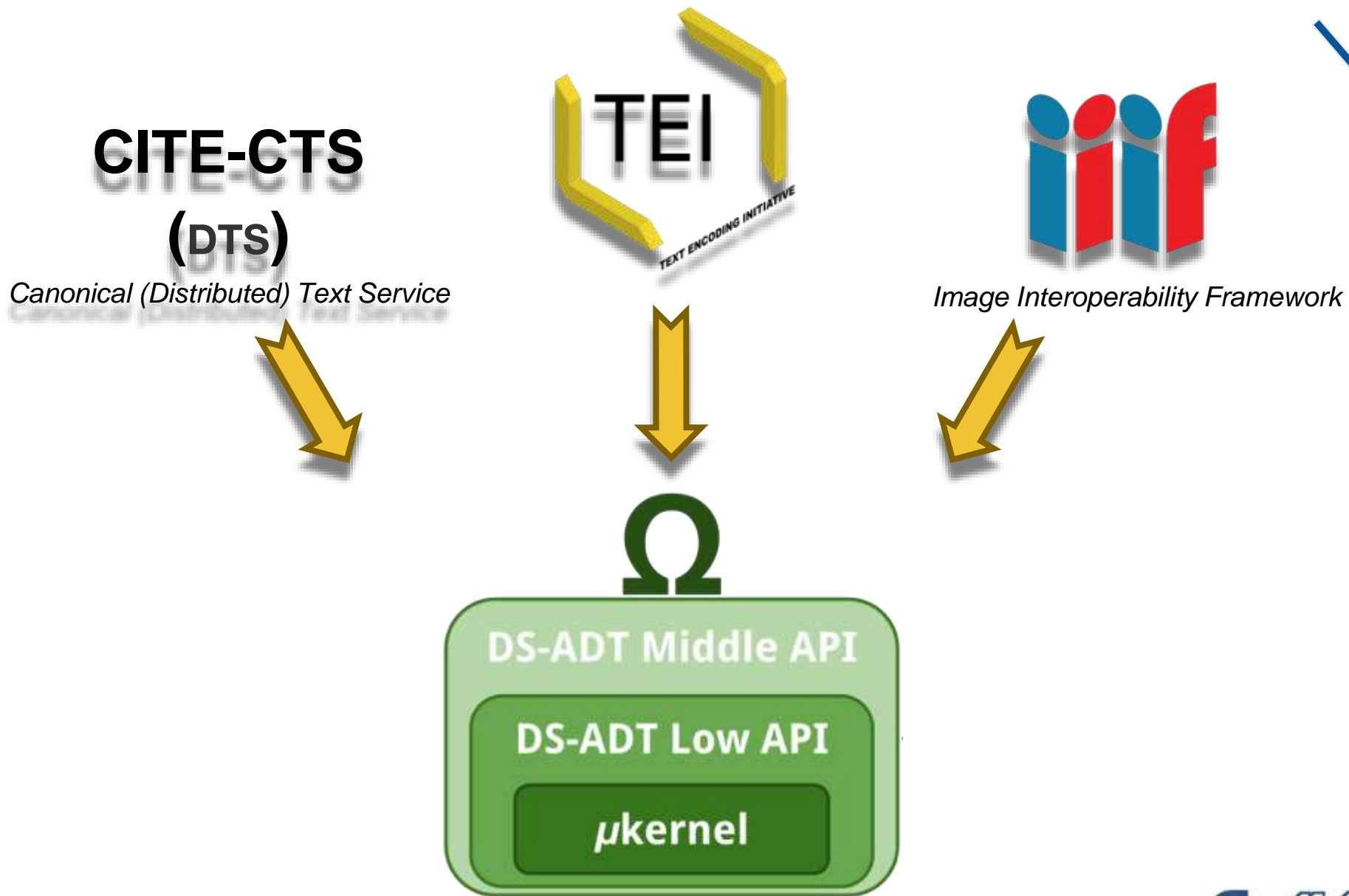
This image shows a handwritten manuscript page with digital annotations. A ruler at the top indicates line numbers. The text is in French and discusses linguistic concepts like 'sonantes' and 'radiaux'. A specific line is highlighted with a yellow background. The text includes phrases like 'Le dernier chapitre nous embarrasse un peu, critique fort courtoise de M. S. de' and 'des radiaux radiaux dixyllabiques'. There are several annotations in French, such as '1. À quoi se reconnaît un sonantiste?' and '2. nous pouvons examiner les objections contre la conception courante'. The page also features a ruler at the top and a legend for symbols like 'le' and 'qu'.

**Modello del Testo:** la rappresentazione formale e strutturata di varie tipologie di dati testuali al fine di descriverli, elaborarli, conservarli, indicizzarli, interpretarli, confrontarli e recuperarli.



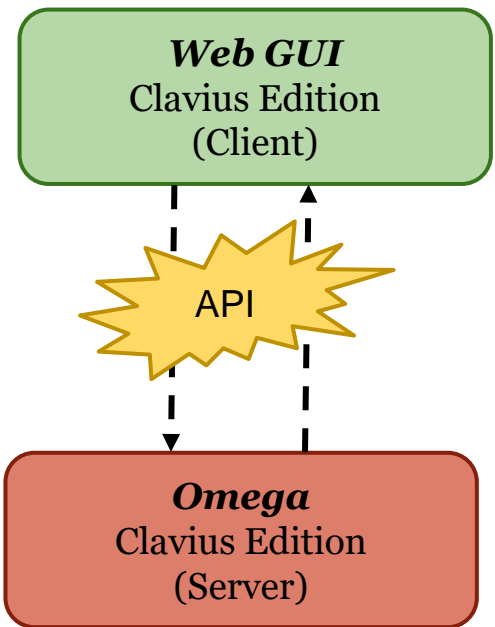






**Obiettivo:** conservare, promuovere e valorizzare la corrispondenza di Christophorus Clavius (1537-1612)  
**Partner:** IIT-CNR, ILC-CNR, APUG **Funding:** Registro.it

Lo sviluppo di **Omega** fornisce alla comunità DH strumenti orientati allo studio dei testi attraverso una architettura modulare e interfacce condivise

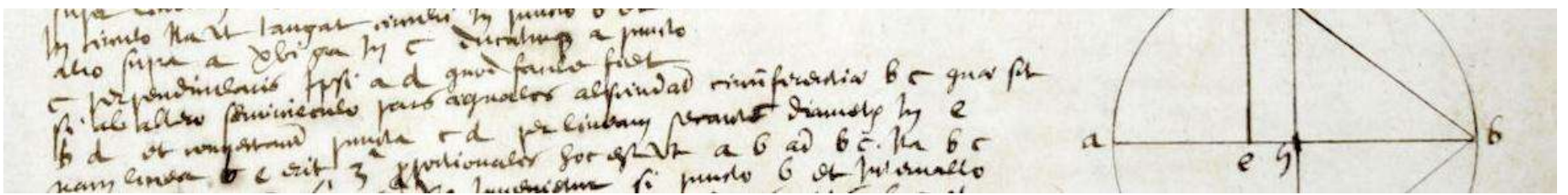


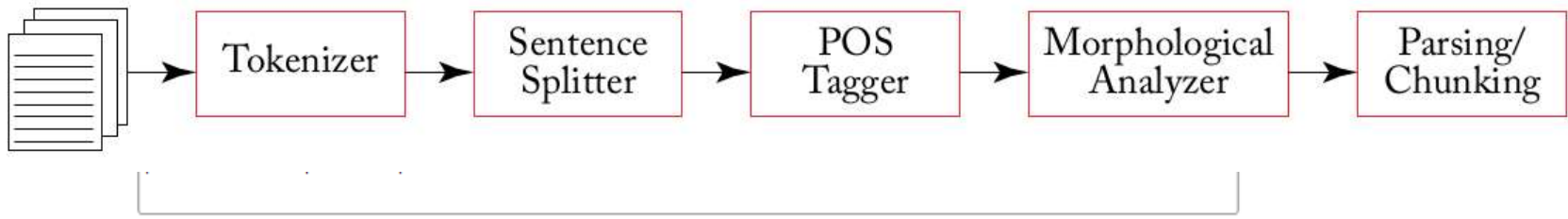
TEA	Search	Lexica
<p>Abscindatur circumferentiae bc quae sit bd et connectantur &lt;puncta&gt;(p) c d per &lt;lineam&gt;(ln) secantem &lt;diametrum&gt;(d) in e nam &lt;linea&gt;(ln) be erit 3a proportionalis hoc est ut ab ad bc ita bc ad be per 8. 6.i 4.a vero inveniatur si &lt;puncto&gt;(p) b et intervallo be sectio &lt;circuli&gt;(c) versus c et d fiat, donec &lt;lineis rectis&gt;(lrr) bc et bd occurrat in &lt;punctis&gt;(p) f et g quae per &lt;lineam&gt;(ln) fg connecti debent secantem &lt;diametrum&gt;(d) in h. tunc enim &lt;linea&gt;(ln) hb quarta proportionalis erit nam ut se habet ab. ad bc. ita bc ad be sed ut bc ad be ita bf hoc est be ad bh per 2am 6.i Igitur quemadmodum ab ad bc. Ita bc ad be et be ad bh eademque arte inveniatur 5.a proportionalis si &lt;puncto&gt;(p) b et intervallo bh portio &lt;circuli&gt;(c) describatur quae secet &lt;lineas&gt;(ln) cb. Et bd et &lt;puncta&gt;(p) sectionum iungantur per &lt;lineam&gt;(ln) secantem &lt;diametrum&gt;(d) interiecta enim inter hanc &lt;lineam&gt;(ln) et &lt;punctum&gt;(p) b. 5a erit proportionalis. Hinc constat datis quibuscumque &lt;linei&gt;(ln) facile inveniri posse duas medias continuas proportionales si hoc problema inventum esset datis duabus &lt;lineis&gt;(ln)//</p>		
<p>[73r]            1. a Datis differentiis diametri rectanguli a lateribus rectum ambientibus et latera et diametrum invenire</p>		

Source

loci

annotations





Proofreading of the analyzed sentences

Navigation controls: 1 2 3 4

Token	Lemma	Part of Speech	Person	Number	Tense	Mood
occupationibus	occupatio	n	-	p	-	-
tuis	tuus	a	-	p	-	-
doctissime	doctus	d	-	-	-	-
clavi	clavi*	d	1	s	r	i
parcens	parco	n	-	s	p	p
,	,	v	-	-	-	-
in	in	t	-	-	-	-
tertia	tertius	a	-	s	-	-
mea	meus	c	-	s	-	-
ad	ad	r	-	-	-	-

Clavius on the Web live example on letter 153:

<http://licodemo:8080/ClaviusSentenceEditing/faces/transcriptionView.xhtml?docid=153>



**Obiettivo:** studiare con tecnologie digitali e approcci propri della filologia computazionale la pseudo teologia di Aristotele testimoniata da documenti arabi confrontandola con le Enneadi di Plotino, opera scritta in Greco.

**Partner:** UNIPI, ILC-CNR, Bochum **Funding:** ERC Ideas, Advanced Grant

The screenshot displays the 'Greek into Arabic Web Application v.0.9.9' interface. It features a navigation bar with options like 'Home', 'Pericopes by Greek', 'Pericopes by Arabic', 'Search', 'Comments Editor', 'Linguistic Analysis', 'Pericopes Manager', and 'Computational Lexicon'. The main content area is split into two columns: 'Greek pericope IV 3, 20.21-24' on the left and 'Arabic pericope II, 43.5-43.8' on the right. The Greek text is in a serif font, and the Arabic text is in a sans-serif font. Both columns have a vertical scrollbar. The application is logged in as 'simone'.

Greek into Arabic Application for aligning parallel textual fragments:

[http://g2a.ilc.cnr.it:8080/Teologia\\_Wapp/Home.xhtml?centerPage=teologia](http://g2a.ilc.cnr.it:8080/Teologia_Wapp/Home.xhtml?centerPage=teologia)



La funzionalità di ricerca avanzata permette di interrogare il testo attraverso un lessico greco-arabo  
 Un utente esperto chiede al sistema di visualizzare i luoghi del testo impostando vari filtri cross-linguistici

Greek into Arabic Web Application v.0.9.9

Home Pericopes by Greek Pericopes by Arabic Search Comments Editor Linguistic Analysis Pericopes Manager Computational Lexicon Logged in as simone Logout

**Greek Search**

Type	Word	Pos	Operator
Lemma	κόσμος	Any	And
Lemma	νοητός	Any	
Form	Insert word ...	Any	

Search Clear Form

**Arabic Search**

Type	Word	Pos	Operator
Lemma	عالم	Any	And
Lemma	عقلی	Any	
Form		Any	

Search Clear Form

**Combined Search**

And Search

---

**Search Results**

Greek Pericope Info	Greek Pericope	Arabic Pericope	Arabic Pericope Info
IV 8, 1:46-51	"Ἡ τε οὖν ψυχὴ ἡ τοῦ παντός τούτου χάριν εἰς αὐτὸ παρά τοῦ θεοῦ ἐπέμφθη , ἣ τε ἐκάστου ἡμῶν , πρὸς τὸ τέλος αὐτὸ εἶναι - ἐπειδὴ ἔδει , ὅσα ἐν νοητῷ κόσμῳ , τὰ αὐτὰ ταῦτα γένη ζῶων καὶ ἐν τῷ αἰσθητῷ ὑπάρχειν .	فهذه العلة أرسل الباري النفس إلى هذا العالم وأسكنها فيه ، ثم أرسل أنفسنا فسكنت في أبداننا ليكون هذا العالم تأملًا كاملًا ونبلاً يكون دون العالم العقلي في التمام والكمال ، لأنه كان ينبغي أن يكون في العالم الحسي من أجناس الحيوان ما في العالم العقلي .	I, 25.3-25.6
V 8, 1:1-4	Ἐπειδὴ φαμεν τὸν ἐν θεῷ τοῦ νοητοῦ κόσμου γεγενημένον καὶ τὸ τοῦ ἀληθοῦς νοῦ κατανοήσαντα κάλλος τούτων δυνήσεσθαι καὶ τὸν τούτου πατέρα καὶ τὸν ἐπέκεινα νοῦ εἰς ἔννοιαν βαλέσθαι ,	في شرف عالم العقل وحسنه ونقول : إن من قدر على خلق يده وتسكرين حواسه ووساوسه وحركاته كما وصفه صاحب الرموز من نفسه - قدر أيضًا في فكرته على الرجوع إلى ذاته والصعود بعقله إلى العالم العقلي فيرى حسنه وبهائه ، فإنه يقوى على أن يعرف شرف العقل ونوره وبهائه ، وأن يعرف قدر تلك الشئ ، الذي هو فوق العقل ، وهو نور الأتوار وحسن كل حسن وبهائه كل بهاء .	IV, 56.3-56.8



**Obiettivo:** Traduzione del Talmud Babilonese dal testo originale in ebraico/aramaico al testo in italiano, attraverso funzionalità computazionali per il supporto alla traduzione

**Partner:** UCEI, ILC-CNR, MIUR **Funding:** MIUR

Albero delle Fonti

Tabella delle Traduzioni e Filtri

Note, Glossari, Suggerimenti e Analisi Linguistica

The screenshot shows a web application interface for Talmud translation. It is divided into several panels:

- Resources:** A tree view on the left showing a hierarchy of sources, including Mishnah and Ghemara sections.
- Filter:** A central table displaying search results. It has columns for filters, text, and translation. The table shows results for 'MISHNA' and 'GHEMARA' sections.
- References:** A panel on the right containing notes, glossaries, and a linguistic analysis table.

The linguistic analysis table in the 'References' panel is as follows:

Token	Lemma	Part of speech
il	il	DET:art
primo	primo	ADJ
il	il	PRE
dei	dei	NOM
e	essere	VER:pres
il	il	DET:art
capodanno	capodanno	NOM
per	per	PRE
la	il	DET:art
decima	decimadecimo	NOM
dei	dei	PRE:det
bestiame	bestiame	NOM
		SENT

La funzionalità di allineamento parola per parola permette di costruire un lessico tri-lingue ebraico-aramaico-italiano attraverso un sistema stocastico di collegamento tra parole

AliEval

Review

Stats

## Aligned Textual Fragment

BACK

משעה שהכהנים נכנסים לאכול בתרומתן

dall' ora in cui i kohanim entrano a mangiare la loro terumà

משעה	dall'	ora	in	cui	kohanim	entrano	a	mangiare	la	loro	terumà	משעה
שהכהנים	dall'	ora	in	cui	kohanim	entrano	a	mangiare	la	loro	terumà	שהכהנים
נכנסים	dall'	ora	in	cui	kohanim	entrano	a	mangiare	la	loro	terumà	נכנסים
לאכול	dall'	ora	in	cui	kohanim	entrano	a	mangiare	la	loro	terumà	לאכול
בתרומתן	dall'	ora	in	cui	kohanim	entrano	a	mangiare	la	loro	terumà	בתרומתן

SAVE ALIGNMENT!



**Obiettivo:** Realizzare una edizione digitale degli scritti saussuriani. Funzionalità essenziale è l'annotazione e il commento. Gli studiosi selezionano un frammento di risorsa e lo annota sfruttando un sistema di classificazione.

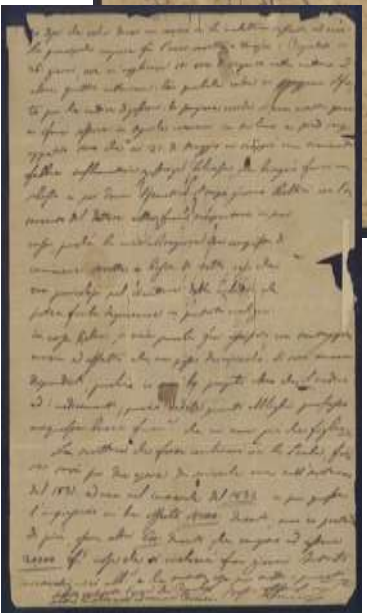
**Partner:** UNICAL, ILC-CNR, UNISA **Funding:** PRIN2008

The interface is divided into several functional areas:

- Lista Commenti (Comments List):** A table on the left showing snippets and their classifications.
 

Snippet	Type	View
connuet et	[a]	
elles, qui mais	[bi]	
extrêmement	[bi]	
infiniment	[bi]	
être que charmé...	[va]	
de voir	[bi]	
toutes les fois...	[bi]	
se trouve	[bi]	
enfermera	[bi]	
ainsi	[re]	
- Trascrizione (Transcription):** A central text area containing a paragraph of French text with various annotations (checkmarks and brackets) indicating specific linguistic or structural features.
- Immagine (Image):** A thumbnail of a handwritten manuscript page on the right, with a magnifying glass icon for viewing details.
- Text & Image Selection:** A central control area with buttons for 'text selection' and 'image selection'.
- Area Testo (Text Area):** A large text input field at the bottom, containing a rich text editor toolbar and a 'Biffure' (crossed-out) button.
- Classificazione (Classification):** A dropdown menu at the bottom right of the text area, currently showing 'Biffure'.





BelliniInRete

```
<code></code>
```



Edizione digitale Lettere di Bellini

Lettere Bellini 1.4 - L11.4.1 - Interpretative - Info

- 1 • vi dissi che volea darvi un cenno su la malattia sofferta ed ecco:
- 2 La principale ragione fu l'aver scritto a **Venezia** i **Capuleti** <sup>(1)</sup> in
- 3 26 giorni, ove m'applicava 10 ore di seguito nella mattina ed
- 4 altre quattro nella sera. la qualche volta mi puzzava il fia
- 5 to per le cattive digestioni: la stagione orrida si era unita pure
- 6 a farmi soffrire: in Aprile ritornai in **Milano** e stiedi senza
- 7 appetito sino che ai 21 di Maggio mi scoppiò una tremenda
- 8 febbre infiammatoria gastrica biliosa che bisognò farmi un
- 9 salasso e poi darmi l'emetico: il terzo giorno **Pollini** con l' in
- 10 tervento del dottore volle farmi trasportare in sua
- 11 casa, perchè la mia abitazione era composta di
- 12 camere strette e basse di tetto, cosa che
- 13 era pericolosa pel carattere della malattia che
- 14 potea farla degenerare in putrida maligna:
- 15 in casa **Pollini** , in una parola fui assistito con tanta pre
- 16 mura ed affetto che non posso descriverlo. Si sono ancora
- 17 **dispensati**, perchè io non ho pagato altro che il medico
- 18 ed i medicamenti; perciò vedete quanti obblighi professo
- 19 a questa buona fam **iglia** che mi ama più che figlio.
- 20 La scrittura che forse combinerò con la **Scale**, for
- 21 se sarà per due opere <sup>(2)</sup> da scriverle una nell'autunno
- 22 del 1831 ed una nel carnevale del 1832 e per queste
- 23 l' impressario <sup>(3)</sup> m'ha offerto 4000 ducati; ma io pretendo
- 24 di più, ossia altri 600 ducati, che vengono ad essere
- 25 20000 f ranchi cosa che si risolverà fra giorni. Intanto
- 26 ricevetevi i miei abh racci e le tante cose per tutti i parenti
  
- 27 Vostro affezionatissimo nipote
- 28 **Vincenzo**
- 29 Subito completata l' opera <sup>(4)</sup> dei **Capuleti** <sup>(1)</sup>
- 30 spedirò le solite copie ed una per **Carcaci** . •

Multiple Selection
Search Q
Create index
AV



**LesMu** è una banca dati informatica contenente circa **22.500 schede lessicografiche**, per un totale di oltre **3.600.000 parole**. Ogni scheda è suddivisa in **29 campi** diversi analizzando **800 opere**. Il programma di interrogazione è il **DBT**, realizzato da Eugenio Picchi presso l'ILC-CNR alla fine degli anni '90.

*Contenuti a cura di Fiamma Nicolodi e Paolo Trovato con la collaborazione di Renato Di Benedetto. Redazione: Luca Aversano e Fabio Rossi*

The screenshot shows the 'DBT per Windows' application window. The menu bar includes 'File', 'Ricerca', 'Contesti', 'Famiglia', 'Opzioni', 'Varie', 'Finestre', and 'Guida'. The toolbar contains various icons for search and navigation. The main window is divided into two panes:

- Quadro (Left Pane):** A list of 22 numbered items, each with a field name in curly braces and a term. The items are:
  - 7\* {Lemma} sortita
  - 7 {Accez} sortita
  - 7 {Forma} sortita
  - 3 {InLoc} sortita
  - 6 {Defin} sortita
  - 1 { Note} sortita
  - 7 {Altri} sortita
  - 35 { Cont} sortita
  - 1 {Locuz} sortita di tono
  - 9 {Lemma} duetto
  - 9 {Accez} duetto
  - 1 {Titol} duetto
  - 9 {Forma} duetto
  - 5 {InLoc} duetto
  - 10 {Defin} duetto
  - 8 {Altri} duetto
  - 417 { Cont} duetto
  - 1 {Forma} arietta
  - 9 {Alter} arietta
  - 4 {InLoc} arietta
  - 3 {Altri} arietta
  - 45 { Cont} arietta
- [Cnt] - sortita (Right Pane):** A list of 7 numbered items, each with a field name in curly braces and a definition. The items are:
  - 1795B/sortita.32191 {Lemma}[L] sortita.32191 {Locuz} sortita di tono {Defin} cambio di tonalità, modulazione
  - 1800B/sortita.12696 {Lemma} sortita.12696 {Defin} fig., passaggio di carattere solistico
  - 1828B/sortita.14058 {Lemma} sortita.14058 {Defin} brano solistico che segna la prima apparizione in scena di un personaggio d'opera
  - 1828B/sortita.14063 {Lemma} sortita.14063 {Defin} brano solistico che segna la prima apparizione in scena di un personaggio d'opera
  - 1833B/sortita.53454 {Lemma}[L] sortita.53454 {Locuz} aria di sortita {Defin} brano solistico che segna la prima apparizione in scena di un personaggio d'opera
  - 1841B/sortita.13794 {Lemma} sortita.13794 {Defin} gen., prima apparizione in scena di un personaggio
  - 1880B/sortita.09243 {Lemma} sortita.09243 {Defin} brano solistico che segna la prima apparizione in scena di un personaggio d'opera



A.M. Del Grosso, E. Capizzi, S. Cristofaro, G. Seminara, D. Spampinato, Promoting Bellini's legacy and the Italian Opera by scholarly digital editing his own correspondence. In Poster papers of TEI2019 Book of Abstracts <https://graz-2019.tei-c.org/program/book-of-abstracts/>

A.M. Del Grosso, E. Capizzi, S. Cristofaro, M.R. De Luca, E. Giovannetti, S. Marchi, G., Seminara, D. Spampinato, Bellini's Correspondence: a Digital Scholarly Edition for a Multimedia Museum. In *Umanistica Digitale* n.7, 2019 (in pubblicazione).

A.M. Del Grosso, S. Cristofaro, M.R. De Luca, E. Giovannetti, S. Marchi, G. Seminara, D. Spampinato, Le lettere di Bellini: dalla Carta al Web. In Spampinato Daria (a cura di), *AIUCD2018 - Book of Abstracts*. AIUCD. ISBN 9788894253528. <http://amsacta.unibo.it/id/eprint/5997>. In: *Quaderni di Umanistica Digitale*, 2018, pp. 60-64.

D. Spampinato, Rapporto scientifico secondo anno progetto Museo Virtuale della Musica BellinInRete, 2019.

D. Spampinato, Rapporto scientifico secondo al progetto Museo Virtuale della Musica BellinInRete, 2018.

D. Spampinato, Rapporto scientifico primo trimestre progetto Museo Virtuale della Musica BellinInRete, 2017.

S. Cristofaro, D. Spampinato, OntoBellini: towards an RDA based ontology for Vincenzo Bellini's cultural heritage. In Adrien Barton, Selja Sappälä, Daniele Porello (a cura di), *JOWO2019 Proceedings*. CEUR Workshop Proceedings (in pubblicazione).



*Verso la definizione e l'implementazione di una  
piattaforma orientata allo studio critico del testo*

*GRAZIE per l'ATTENZIONE*

**Angelo Mario Del Grosso**

*Istituto di Linguistica Computazionale "A. Zampolli"*

`angelo.delgrosso@ilc.cnr.it`

