



Le edizioni digitali oggi: sviluppi e sfide

Roberto Rosselli Del Turco
Dipartimento di Studi Umanistici
Università di Torino
roberto.rosselidelturco@unito.it

EVT Home: <http://evt.labcd.unipi.it/>
EVT GitHub: <https://github.com/evt-project>

DISH: <https://www.dish.unito.it/>

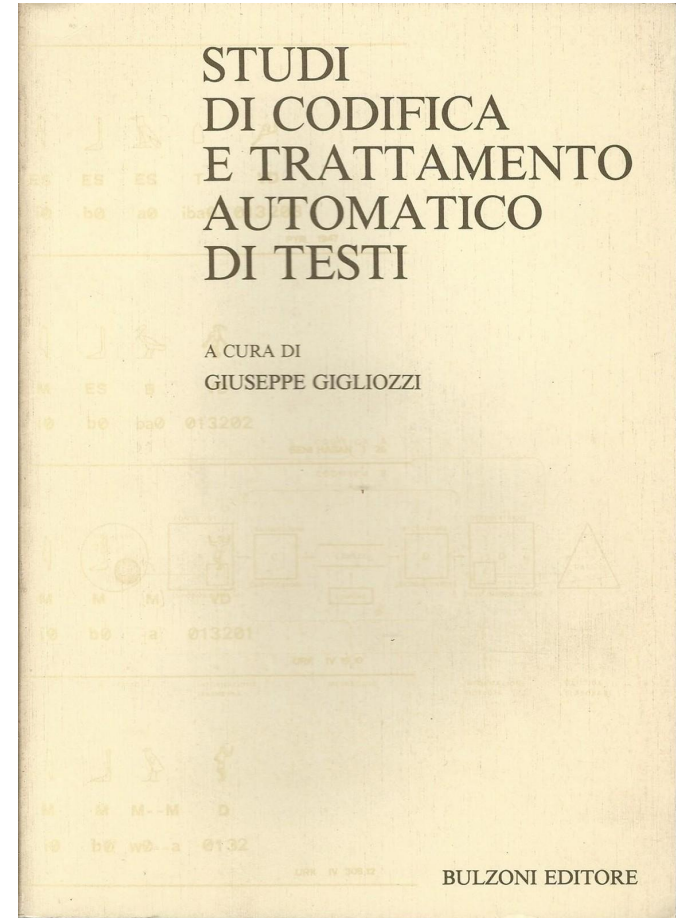
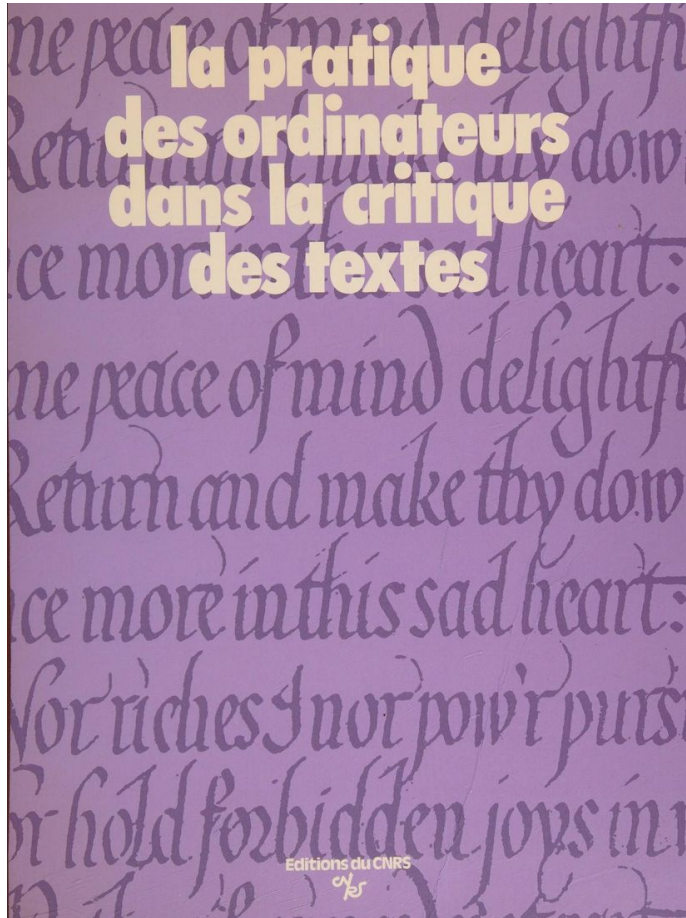




Premessa



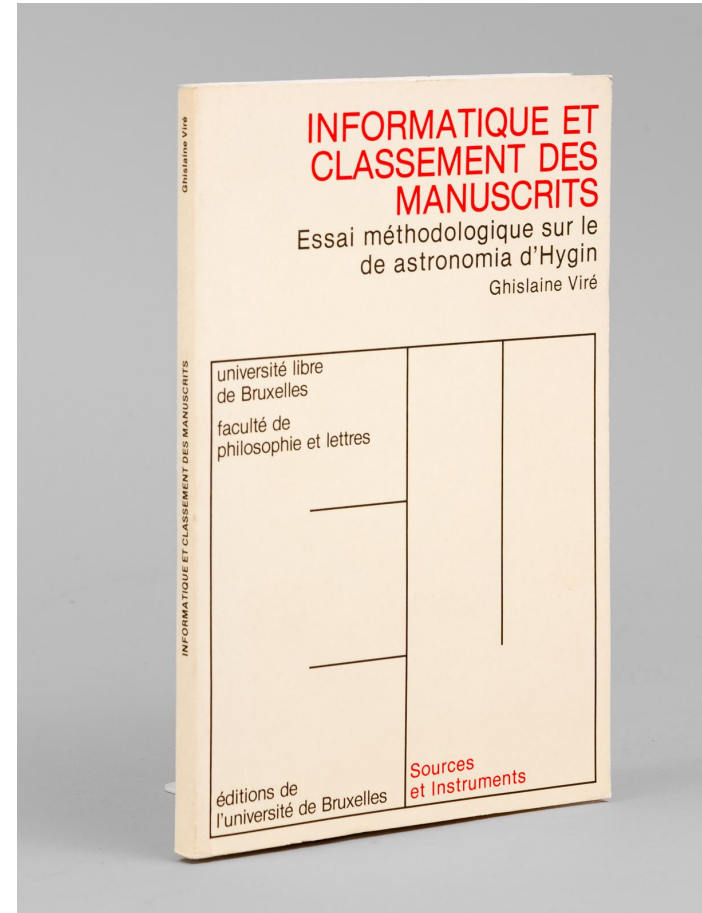
Once upon a time



Once upon a time



ORI.ANDI T. 1986,
Problemi di codifica e
trattamento informatico in
campo filologico



Informatics and New Philology

Roberto Busa SJ

Aloisianum, 21013 Gallarate, Italy

Abstract: This paper considers the impact of informatics and the new technology on the field of philology. A brief overview of developments in the field is given, and the present bottleneck of computational linguistics are considered. Also discussed are some of the challenges posed by natural language processing, and the ways in which we can rise to the challenge.

Key Words: informatics, computer science, philology, lexicography, linguistics.

I.

Let me say at the beginning that the title “Informatics and New Philology” should not be understood as implying that traditional philology is to be forgotten, nor as suggesting that we need to start again from scratch. That is certainly not the case.

The science of philology was not born with the computer. The new philology will be a continuation from its predecessors, but with a quality-leap and new dimensions. If so much was done without the computer, how much more should the young people of today be able to achieve with the computer? Such is the meaning of the title of this paper.

growth. Our generation has not done everything: for the young people there are still immense open spaces. This is what I would like to show.

II.

First of all, I would like to summarize the developments in philological methods that the computer has been responsible for over the past forty years — that is, from 1949 when I first began to grapple with those gadgets!

Working on physical entities which the philologist uses and classifies, the computer has served above all to demarcate the borders between the graphic and the semantic. Furthermore it allows one to recheck and redocument every conclusive affirmation whenever the need arises.

What has perhaps not yet sunk in, is the huge change in dimensions that has occurred: the computer enables us to perform a complete census on texts of tens and hundreds of millions of words. For so long the philologist, being able to work only with pen and paper, had to be happy with a mere sampling, like a bee flitting from flower to flower. We all know how unreliable sampling becomes as

COMPUTER SECTION

Data Bases v. Critical Editions

*Manfred Thaller**

1. Editions and Computers

Historical studies as we know them today, are the result of a long development. Many of the phases of this history of History could have taken place differently, without changing the outcome into something, we would not recognize any more. Indeed, like happiness, history still means very different things to different people — what could scarcely be observed better anywhere else than at such a large-scale meeting of historians as the congress at Kalamazoo. Subjects and methods of historians are so different, that, indeed, some of us may find themselves much more closely related to our colleagues from other departments than to historians from other subdisciplines.

What distinguishes history as an academic discipline from purely speculative writing about the past, is its claim to present the times gone by in a way, which can be criticized not because of a particular literary taste, but

Journal of Computing in Higher Education
Winter 1990, Vol. 1 (2), 3-26.

What is Text, Really?

Steven J. DeRose
praXis

David G. Durand
Brandeis University

Elli Mylonas
Harvard University

Allen H. Renear
Brown University

ABSTRACT

THE WAY IN WHICH TEXT IS represented on a computer affects the kinds of uses to which it can be put by its creator and by subsequent users. The electronic document model currently in use is impoverished and restrictive. The authors argue that text is best represented as an ordered hierarchy of content object (OHCO), because that is what text really is. This model conforms with emerging standards such as SGML and contains within it advantages for the writer, publisher, and researcher. The authors then describe how the hierarchical model can allow future use and reuse of the document as a database, hypertext, or network.

What is Text, Really? TEI and Beyond – An Introduction

Georg Vogeler

The 2019 annual TEI conference in Graz was titled “What is text, really? TEI and beyond.” The development of the TEI and the discussion around it has shown how manifold text(s) can be, and how far-reaching the TEI approach is. Thus, the theme of the conference, “What is text, really?” posed a fundamental question, which goes beyond the pure reference to the seminal paper by DeRose, Durand, Mylonas, and Renear (1990). In 1990 they answered the question somewhat pragmatically by proposing a model for text as an ordered hierarchy of content objects which can easily be formalised with digital technologies, but, as they said later on: text can be much more than that. Encoding can make various aspects of texts explicit, enabling scholars to examine their nature and their relationship with other objects. In this context, the power of TEI rests in its technological



Le Edizioni Digitali Scientifiche (EDS - DSE)



La prima generazione di DSE



DynaText

File Edit Book Collection Journal View Window Help

Collection - Cambridge University Press

The General Prologue - fulltext

The General Prologue - regapp

Double-click the icon to the left to see the unregularized collation at this word.

Regularized collation at: hath

Double-click the icon to the left of a reading to see the text of all witnesses to that reading; single-click on the sigil for any witness to see the text only of that witness.

The base text line here is: 6 Inspired hath in eury holt and heeth

This line OUT in 19 witnesses: Ad 3 Ad 4 Bw Cn Cp Dd Dl Do Gg Gl Ln Mm Ne Nl Ra 2 Ra 3 Ry 1 Sl 1 Sl 2

The lemma for the following word variant or variants is: hath

Agreeing with the lemma:

30 witnesses Bo¹ Bo² Ch Cx¹ Cx² Ds¹ En¹ En³ Fi Ha² Ha³ Ha⁴ Hg Ht Ii La Lc Ld¹ Ld² Ma Mg Ph² Pn Ps Pw Py Ry² Se To¹ Wy

Reading: hath /

Find:

10 That slepen on the nyght with open eye
11 So priketh hem nature / in hir corages
12 Thanne longen folk to goon on pilgrimages
13 And palmeres for to seeken straunge strondes
14 To ferne halwes / kouthes in sondry londes
15 And specially / from eury shyres ende
16 Of Engelond / to Caunterbury they wende
17 The holy blisful martir / for to seke
17-1 To seche the blisful marter that is meke
18 That hem hath holpen whan that they weere seeke
19 Bifel that in that sesoun on a day
20 In Southwerk at the Tabard / as I lay
21 Redy to weenden / on my pilgrymage

Find:

10.15

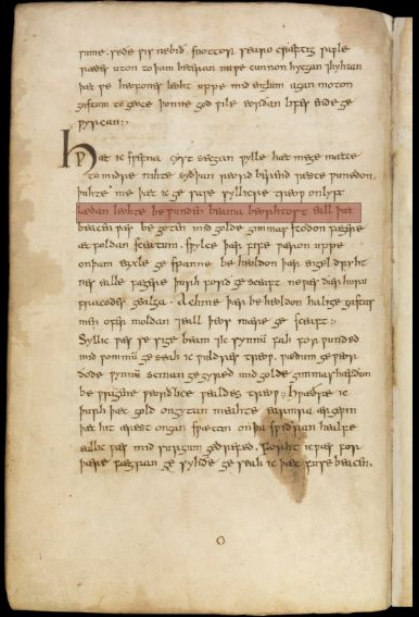
Critical edition of the
Canterbury Tales

La seconda generazione di DSE



The Digital Vercelli Book **BETA**

Thumbs Magnifier HotSpots Image-Text MS Desc DOTR 104v Diplomatic Info



7 **Ð**wæt ic swefna cýrt secgan wylle hætmeges mætte
8 to midre nihte syðþan reord berend reste wunedon.
9 þuhte me þæt ic ge fawe fyllicre treow onlyft
10 **leodan** leohte be wunden beama beorhtort eall þæt
11 beacen wæf be goten mid golde gimmarf stodon fægere
12 æt foldan sceatum . swylce þær fife wæron uppe
13 onþameaxle ge sþannebe heoldon þær engel dryht
14 nef ealle fægere þurh forð ge sceafnewærfðærhuru
15 fracodef gealga .æchine þær beheoldon halige gasta
16 men ofer moldan geall þeow mære ge sceaf :
17 Syllic wæf ge riġe beam ge cynnū fāh for wunded
18 mid wommūge seah ic wuldref treow. wædum geweor
19 ðode wynnū scinan ge gyred mid golde gimmarf hæfdon
20 be wriġene weorðlice wealdes treow : hwæðre ic
21 þurh þæt gold ongytan meahthe earmra ærgewin
22 þæt hit ærest ongan fwætan on þa fwiðran healfē
23 eallic wæf mid furgum gedrefed .Forht icwæf for

33%

Search Q A⁻ A⁺

Powered by EVT

La seconda generazione di DSE



Toc, front matter and back matter



TABLE OF CONTENTS

▶ NAMED ENTITIES

📍 LISTA DEI LUOGHI

👤 LISTA PERSONE

🏛️ ISTITUTI RELIGIOSI

☰ ST

BIBLIOGRAPHY

▶ PROJECT INFO

A B C D E F G H I L M N O P R S T U V W Y Z Ç

▶ Albertus de Buiano ✕

▶ Albertus de Calvo ✕

▶ Albertus ✕

▶ Albertus ✕

▶ Albertus de Cruce ✕

▼ Albertus ✕

More Info Occurrences

📘 Albertus

♂ M

📄 1: Episcopus Lunensis

📄 Alberto, vescovo di Luni, di cui non esistono testimonianze dirette; viene detto successore di Gottifredo II (1129-1156) in un documento del [1181](#) e predecessore di Pietro (1178-1190) in un documento del [1212](#); viene poi citato in un documento del [1230](#). Il Podestà ([PODESTÀ 1895](#), pp. 38-42) lo pone tra l'ultima attestazione di Gottifredo II (1156) e la prima di Andrea II ([1160](#)), ma quest'ultima indicazione non è certa in quanto il prelado vi è indicato dalla sola cifra iniziale.

▶ Albertus quondam Ferri de Lacala ✕

▶ Albertus de Formicoso ✕

▶ Albertus filio Frederici de Carraria ✕

▶ Albertus quondam Gandulphini / de Gandulfino ✕

▶ Albertus quondam Gerardi ✕

▶ Albertus filius Gerardi de Carraria ✕

▶ Albertus quondam domini Guidonis ✕

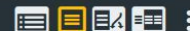
▶ Albertus ✕

▶ Albertus ✕

La seconda generazione di DSE



Edizione Logica Avicennae [experimental encoding]



Div 1 Critical Info

ideo convenientius est prius cognoscere simplices quam compositos.

Cognitio autem^{ay} simplicium^{az} fit duobus modis, quia aut cognoscuntur secundum^{ba} hoc^{bb} quod apti sunt^{bc} ut ex eis fiat^{bf} compositio^{bg} praedicta^{bh}, aut cognoscuntur secundum^{bi} hoc quod sunt^{bj} naturae et res^{bk} quibus accidit hic intellectus, ad similitudinem domus, quae componitur ex lignis et aliis, compositorum cuius opus est cognoscere simplicia domus, scilicet ligna et lateres et lutum, sed ligna et lateres et lutum habent dispositiones^{bl} propter quas^{bm} fiunt apta domui, et eius constructioni, et alias dispositiones praeter has, sicut hoc quod ligna sunt haec^{bn} substantiae in qua fuit anima vegetabilis aut quod natura eorum est calida aut frigida, aut quod comparatio eorum inter ea quae sunt est talis vel talis. Hoc autem scire non est necesse artificii domorum, sed an trabes sit mollis vel dura, aut^{bo} sit^{bp} sana vel putrida aut huiusmodi, necesse est scire domorum artificii. Similiter est doctrina logica: non enim considerat incomplexa harum rerum^{br} secundum quod sunt alicuius duorum modorum esse, scilicet quod est in his sensibilibus^{bs} aut in intellectu, nec etiam in essentia harum rerum ex hoc quod sunt essentiae, sed secundum quod sunt praedicata aut subiecta, et universalis aut particularia, aut huiusmodi quae non accidunt his intellectibus^{bt} nisi ex hoc quod diximus in praemissis.

Ad considerationem autem dictionum ducit nos necessitas: logicus enim, ex hoc quod est logicus, non habet primo occupari circa verba prima, nisi quantum ad loquendum et agendum. Si enim possibile esset logicam discere solo intellectu, ita ut non considerentur aliqua eius nisi soli intellectus, tunc sufficeret. Si etiam^{bu} doctor artis posset revelare id quod est in animo eius alio modo, supersederet semper^{bx} verbis^{by} sed quia necessitas ducit nos ad agendum cum verbis praecipue - non enim potest ratio componere intellectus quin cum illis proferat verba, immo quia cogitatio quasi locutio est inter ipsum hominem et cogitatum suum verbis imaginatis - sequitur ut verba habeant diversas dispositiones propter quas differant dispositiones intentionum quae comitantur eas in

Search Create index Filtri Mappa di variabilità

causam F G M N P R V

- GROUP 1: causam G M N P R V
- GROUP 2: causam F causa U B

intelligi F G M N P R V

- GROUP 1: intelligi G M N P R V
- GROUP 2: intelligi F intelligere B U

Cognitio autem B F G M N P R U V

- GROUP 1: Cognitio autem G M N P R V autem cognitio C
- GROUP 2: Cognitio autem B F U

Apparato Critico Fonti Passi paralleli

doctor artis posset revelare id quod est in animo eius alio modo, supersederet semper

PASSO PARALLELO IN

- Avicenna, Burhān

Riferimento Bibliografico XML

Dicemus quod verbum graece significans intentionem generis, prius apud eos, secundum primam impositionem, significabat aliud, et deinde, per impositionem secundam, translationem est ad significandam intentionem quae apud logicos vocatur «genus».

PASSO PARALLELO IN

- Avicenna, Burhān

Riferimento Bibliografico XML

Multi reprehenderunt,

PASSO PARALLELO IN

- Avicenna, Burhān

Riferimento Bibliografico XML

ut « canis » cum dicitur pro « latrabili », et haec est eius significatio, deinde imponat ad significandum aliud, sicut « pisci »

PASSO PARALLELO IN

- Boethius, In Porph. Isag.

Testo Riferimento Bibliografico XML



Opportunità (e non solo)

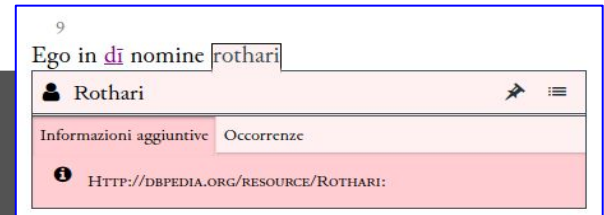


Supporto per LOD in EVT

<teiHeader>

```
<listPerson>
  <person xml:id="Rothari">
    <persName>Rotari</persName>
    <sex>M</sex>
    <birth>606</birth>
    <death>652</death>
    <note>V. anche la voce corrispondente su
      <ref target="http://dbpedia.org/resource/Rothari">
        DBpedia</ref>.
    </note>
  </person>
</listPerson>
```

Partial EVT support
(showing URL)



<body>

```
<persName ref="#Rothari">
  <w>rothari</w></persName>
```

About: [Rotari](#)

An Entity of Type: [animale](#), from Named Graph: <http://dbpedia.org>, within Data Space: [dbpedia.org](#)

Re dei longobardi in Italia (r. 636-652)



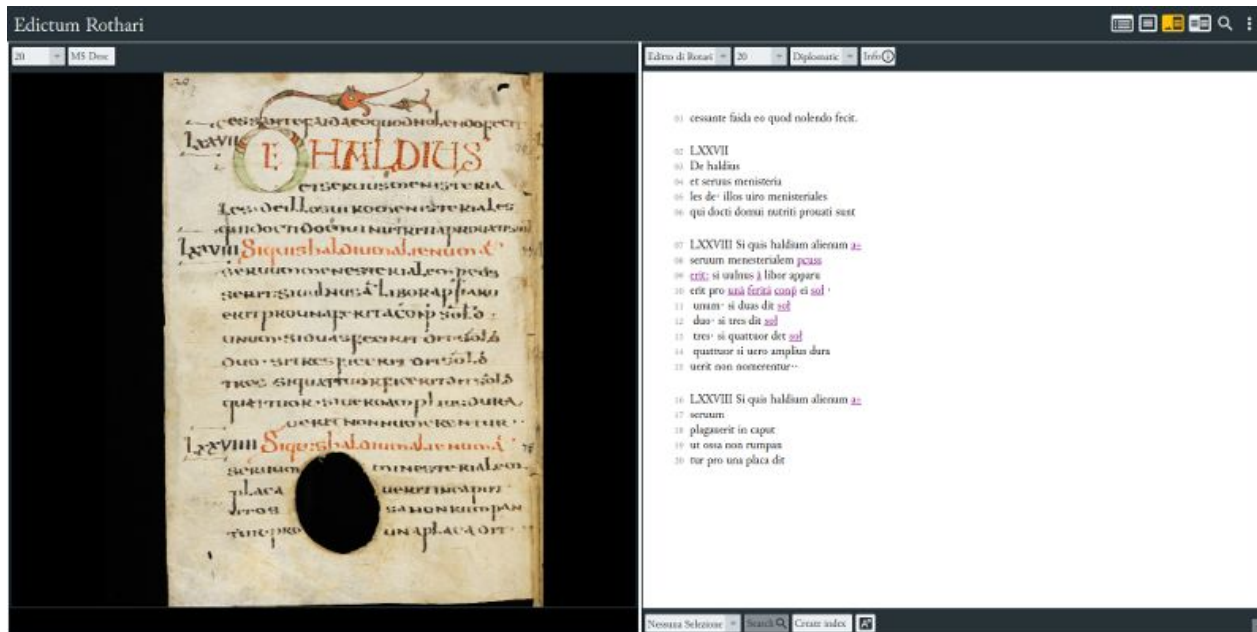
Property	Value
dbo:activeYearsEndYear	<ul style="list-style-type: none">0652-01-01 (xsd:gYear)
dbo:activeYearsStartYear	<ul style="list-style-type: none">0636-01-01 (xsd:gYear)
dbo:child	<ul style="list-style-type: none">dbr:Rodoald
dbo:description	<ul style="list-style-type: none">king of the Lombards (en)re dei longobardi in Italia (r. 636-652) (it)
dbo:predecessor	<ul style="list-style-type: none">dbr:Arioald
dbo:spouse	<ul style="list-style-type: none">dbr:Gundeberga

L'edizione distribuita



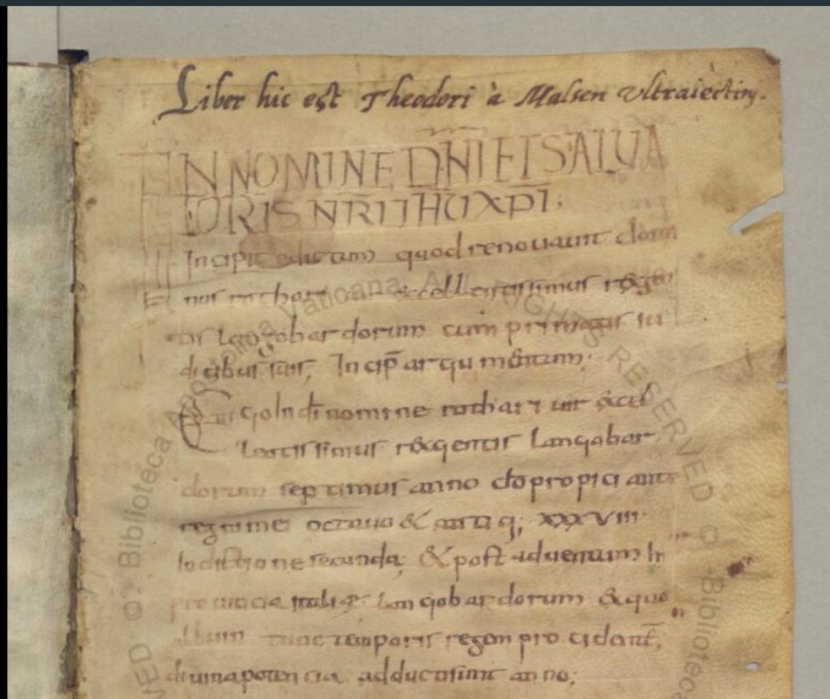
Nuova pagina → intera immagine via IIIF

<pb facs="https://www.e-codices.unifr.ch/loris/csg/csg-0730/csg-0730_020.jp2/full/full/0/default/jpg"/>



1r MS Desc

Editto di Rotari 1r Diplomatic Info



Liber hic est Theodori à Malsen Ultraiectry.


1

- 1 IN NOMINE **DNI** ET SALUA
- 2 TORIS **NRI IHU XPI** ;
- 3 Incipit edictum quod renouauit domi
- 4 nus rothari uir excellentissimus rex gen
- 5 tis langobardorum cum primatis iu
- 6 dicibus suis; **incip** argumentum;

- 7 Ego in **di** nomine rothari uir excel
- 8 lentissimus rex gentis langabar
- 9 dorum septimus anno **do** propiciante
- 10 regni mei octauo **er**atq; XXXVIII
- 11 indictione secunda ; **er** post aduentum in
- 12 prouincia italiæ langobardorum ex quo
- 13 albuin tunc temporis regem **pro**cidentē ;
- 14 diuina potencia adducti sunt anno

Nessuna Selezione Search Create index A°

1r

- 
- in una edizione tradizionale il prodotto finale è statico
 - testo critico e apparato, o trascrizione diplomatica con note paleografiche etc.
 - il lettore fa tutto il *processing* necessario
 - in un'edizione digitale i dati testuali sono elaborati soprattutto per la visualizzazione e la navigazione ipertestuale
 - ma se **text = data** allora posso elaborarlo per altri scopi
 - per visualizzazioni complesse (grafi, network, etc.)
 - per fare calcoli sui dati numerici (tabelle)
 - per ricerche incrociate al fine di ottenere nuove informazioni
 - un documento TEI con codifica ricca è un database da esplorare

Introduction: “Medieval archival sources in the digital world. The challenge of treating and visualizing semi-structured data”

Roberto Rosselli Del Turco

Università di Torino

Enrica Salvatori

Università di Pisa

<http://orcid.org/0000-0002-2933-4341>

Andrea Nanetti

Nanyang Technological University

Marco Giacchetto

Università di Siena

Vera Isabell Schwarz-Ricci

Università degli Studi di Napoli Federico II

Antonella Ambrosio

Università degli Studi di Napoli Federico II

DOI: <https://doi.org/10.6092/issn.2532-8816/12582>

Problems and Issues in the Study of Late Medieval Tax Records: the Sienese "Lira" in the XV Century

Marco Giacchetto

373-385

PDF (Italiano)

HTML (Italiano)

Semi-structured data processing: implementation hypotheses and use cases taken from Old English texts

Roberto Rosselli Del Turco

387-407

PDF (Italiano)

HTML (Italiano)

How to Cite

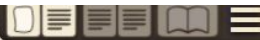
Rosselli Del Turco, R., Salvatori, E., Nanetti, A., Giacchetto, M., Schwarz-Ricci, V. I., & Ambrosio, A. (2021). Introduction: “Medieval archival sources in the digital world. The challenge of treating and visualizing semi-structured data”. *Umanistica Digitale*, 5(10), 289-298. <https://doi.org/10.6092/issn.2532-8816/12582>

More Citation Formats

L'elaborazione di dati semistrutturati



Le denunce della Lira del 1453 **BETA**



Miniature

Lente

Immagine Testo

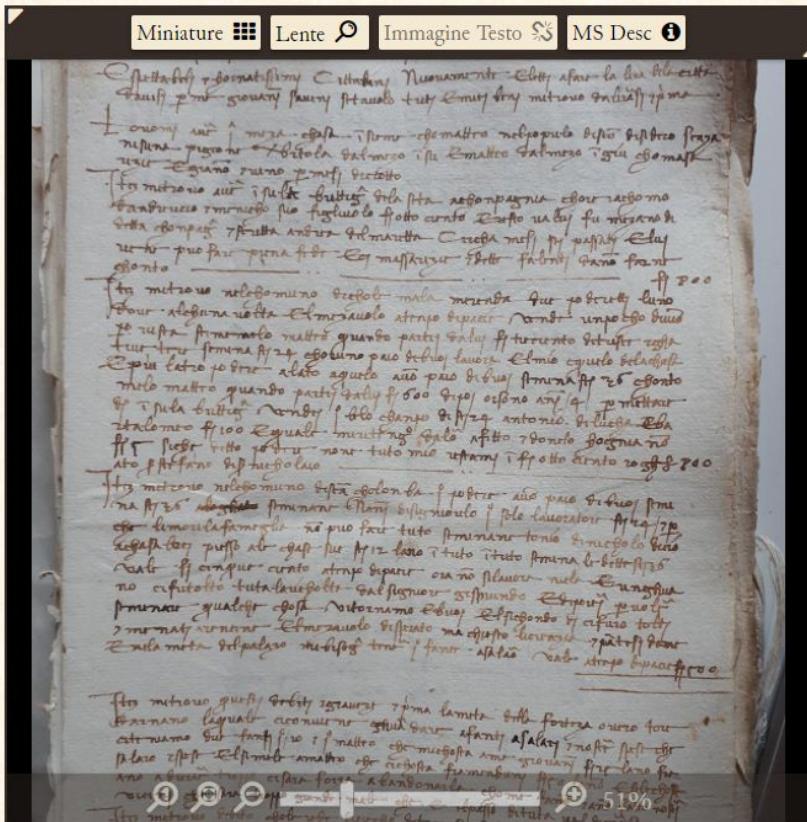
MS Desc

L137

329r

Interpretativa

Info



- 1 Espettabili et hornatissimi cittadini nuovamente eletti a fare la lira dela città davi si per me Giovanni Savini setauolo tuti e' miei beni mi trovo d'alirarsi et prima:
- 2 Truomi avere una meça chasa insieme cho' **Matteo** nel populo di Santo Disidero scaça nisuna pigione abitola dal meço in su e Matteo dal meço in giù cho masericie e granno et vino per mesi diciotto
- 3 Item mi trovo avere in sula buttiga dela seta a chomapagnia chore **Iachomo d'Andreuccio** et Menico suo figliuolo fiorini ottociento, e resto vò lui meçano di detta chompagnia et scritta Andrea del Maretta circha mesi sei passati e lui ve ne può fare piena fede e ci massariçie et dette falenti d'anno farne chonto_____fiorini 800
- 4 Item mi trovo nel chomuno di Chole Malamerenda due poderetti, l'uno dove alchuna volta el mezzaiole a tempo di pacie vende un pocho di vinno però vista stime melo **Matteo** quando partii da lui fiorini trecento de triste et ghative tere semina stiaia 24 choruno paio de buoi lavora el mio, e quello dela chiesa. E più l'altro podere a lato a quello anno paio di buoi semina stiaia 36, chonto melo **Matteo** quando partii da lui fiorini 600. Di poi orsono anni 4 per mettere denari in sula buttiga vendei uno belo campo di stiaia 24 Antonio di Lucha e Bartolomeo fiorini 100 e' quale mi ritengo da loro afito et donelo hogni anno fiorini 5 sichè detto podere non è tutto mio restami in fiorini ottociento rogato ser Stefano di ser Nicholaio_____lire 800
- 5 Item mi trovo nel chomuno di **Santa Cholomba** uno podere, anno paio di buoi

Lista delle persone

Lista dei luoghi

Lista di enti

Ricerca

Liste

Selezione multipla

A⁻ A⁺

L'elaborazione di dati semistrutturati



Quire	S+T		S+P		S+Wynn	
	<i>no lig.</i>	<i>lig.</i>	<i>no lig.</i>	<i>lig.</i>	<i>no lig.</i>	<i>lig.</i>
1	24	17	8	5	2	2
2	21	37	7	9	1	2
3	17	26	6	9	4	0
4	17	30	5	10	3	1
5	15	18	9	4	4	0
6	16	38	15	16	4	5
<hr/>						
7	1	70	0	38	0	8
8	0	49	0	42	0	13
9	1	53	0	29	0	10
10	1	59	0	23	0	10
11	0	43	0	33	0	5
12	0	61	0	29	0	4
<hr/>						
13	2	86	6	43	2	8
14	1	44	0	39	0	6
15	2	54	2	20	1	4
16	0	48	1	15	1	7
17	1	54	0	13	0	1

```
<char xml:id="amacr">
  <charName>LATIN SMALL LETTER A WITH MACRON</charName>
  <mapping type="codepoint">U+0101</mapping>
  <mapping type="diplomatic">ā</mapping>
  <mapping type="normalized">a</mapping>
</char>
<char xml:id="aeligmacr">
  <charName>LATIN SMALL LETTER AE WITH MACRON</charName>
  <mapping type="codepoint">U+01E3</mapping>
  <mapping type="diplomatic">ǣ</mapping>
  <mapping type="normalized">æ</mapping>
</char>
<char xml:id="AEligmacr">
  <charName>LATIN CAPITAL LETTER AE WITH MACRON</charName>
  <mapping type="codepoint">U+01E2</mapping>
  <mapping type="diplomatic">Ā</mapping>
  <mapping type="normalized">Æ</mapping>
</char>
<char xml:id="emacr">
  <charName>LATIN SMALL LETTER E WITH MACRON</charName>
  <mapping type="codepoint">U+0113</mapping>
  <mapping type="diplomatic">ē</mapping>
  <mapping type="normalized">e</mapping>
</char>
<char xml:id="imacr">
  <charName>LATIN SMALL LETTER I WITH MACRON</charName>
  <mapping type="codepoint">U+012B</mapping>
  <mapping type="diplomatic">ī</mapping>
  <mapping type="normalized">i</mapping>
</char>
```

L'elaborazione di dati semistrutturati



NOMINUM LOCORUM QUÆ IN CHRONICO SAXONICO memorantur, EXPLICATIO.

Litera, quæ in hoc Glossario Historicorum nomina designant.

B. Beda	720	H. Hagust
Al. Alfredus M.	890	mes
Af. Afferus	900	R. Ricardus
Et. Eitelwerdus	1109	G. Gervastus
I. Ingulphus	1114	Hovd. Hovd.
F. Florentius	1140	B. Bromton
M. Will. Malnesburienf.	1148	S. Stabbes
Hunt. Henr. Huntingdonienf.	1164	W. Matt.
D. Sim. Dunelmensis	1166	T. Thorn.
E. Eitelredus, Abbas Rievallis		K. Knibton

A Bban-bun. Abben-bun. p. 124, l. 9. 125, 31. 150, 38. 151, 22. 157, 20, 34. 158, 29. 161, 32, 33. 181, 20. 220, 29. i. c. *Abbatia mons five oppidum. Abbandonia, F. Abbandune, E. Abbingdon, B. Abendon, Abyndon, K.* Nomen hoc paulatim induit, postquam Cissa, Occidentalium Saxonum Rex, Monasterium five *Abbatiam* ibi construxerat. Quo autem nomine antea vocabatur, non ita liquido constat.

Lists

- List of person 29
- List of place 23
- List of org 9
- Relations 2

Alfred the Great (King of Wessex) is son of Æthelwulf (King of Wessex)

Alfred the Great was the youngest son of King Æthelwulf and his first wife Osburh, who both died when Alfred was young. Three of Alfred's brothers, Æthelbald, Æthelberht and Æthelred, reigned in turn before him.

Æthelwith (Queen of Mercia) is the wife of Burgred (King of the Mercians)

Æthelwith (Queen of Mercia)

Info Occurrences Relations XMLSource

- PERSNAME: Æthelwith
- SEX: F
- BIRTH: 838
- DEATH: 888
- FLORUIT: 853-888
- OCCUPATION: Queen of Mercia
- IDNO: 2134
 - SOURCE: <https://pase.ac.uk/jsp/DisplayPerson.jsp?personKey=2134>
 - TYPE: PASE
- IDNO: Q3577177
 - SOURCE: <https://www.wikidata.org/wiki/Q3577177>
 - TYPE: WKD

The Epistemological Value of the Computational Turn in Scholarly Editing

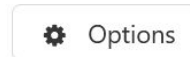
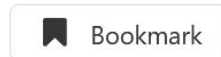
Elisa Cugliana, Andreas Kuczera, Max Grüntgens, Aengus Ward, Joris J. van Zundert, Tara Andrews & Ronald Haentjens Dekker



More download options

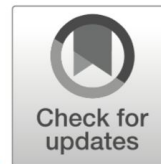
Abstract

In the last decade a decidedly computational turn in scholarly editing is noticeable. Texts are not just encoded anymore. To store and execute scholarly decisions increasingly also program code is applied, which thereby becomes a computational expression of scholarly theory. Theorists and practitioners of such computational approaches have not just put forward scalability and replicability as motivations for computational methods, but also the promise of increased analytical power and on the horizon a changed epistemology. This panel invites some of the most visible proponents of these approaches to reflect on these promises and especially to try to deepen our understanding of the epistemological value that is tied to applying code in the domain of scholarly textual editing.



International Journal of Digital Humanities
<https://doi.org/10.1007/s42803-019-00025-5>

RESEARCH ARTICLE



The 'assertive edition'

On the consequences of digital methods in scholarly editing for historians

Georg Vogeler¹

Published online: 11 May 2019

© The Author(s) 2019

Abstract

The paper describes the special interest among historians in scholarly editing and the resulting editorial practice in contrast to the methods applied by pure philological textual criticism. The interest in historical 'facts' suggests methods the goal of which is to create formal representations of the information conveyed by the text in structured databases. This can be achieved with RDF representations of statements extracted from the text, by automatic information extraction methods, or by hand. The paper suggests the use of embedded RDF representations in TEI markup, following the practice in several recent projects, and it concludes with a proposal for a definition of the 'assertive edition'.

VESPASIANO DA BISTICCI

LETTERE

a knowledge site



LA BASE DI CONOSCENZA

» L'embedded markup XML/TEI

» L'informazione esterna: il file RDF


» I collegamenti con il web of data


» Contributi critici sul tema



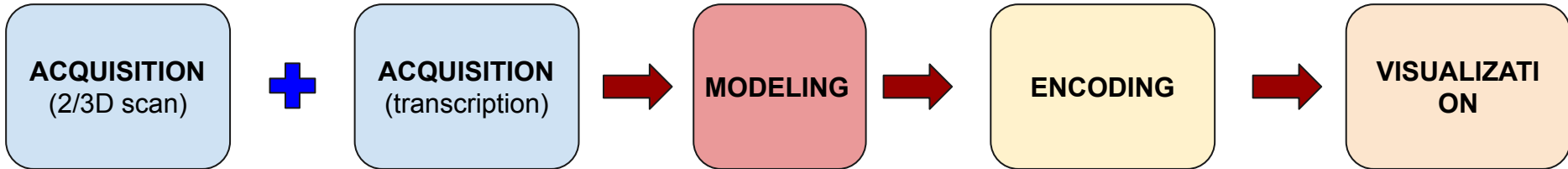
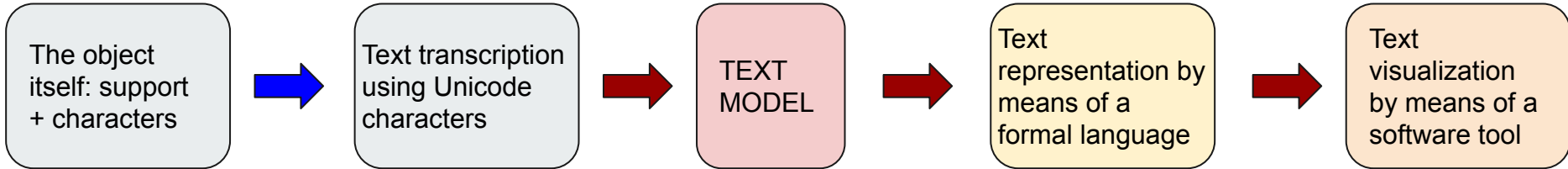
**Opportunità, questioni aperte,
problemi (ir)risolti**



- 
- **filologia digitale e linguistica computazionale:** fanno cose diverse, è possibile l'integrazione in una DSE? servirebbero protocolli per il dialogo
 - **close e distant reading:** filologia è *close reading* per definizione, possibile una integrazione di strumenti di visualizzazione tipici del *distant reading* rivolti all'analisi di fenomeni testuali?
 - **avvento dell'AI:** possibile strumento per la critica del testo
 - trascrizione dei testimoni > reti neurali vs. AI generativa
 - annotazione automatica dei testi
 - suggerimenti per emendazioni / lacune
 - classificazione dei testimoni > stemma codicum (sort of)


- 
- **l'annotazione semantica:** strumenti (XML), modalità di applicazione, flessibilità ed espandibilità degli schemi TEI
 - il concetto di modellizzazione / rappresentazione del testo come premessa per la sua annotazione
 - spesso TEI/XML visto come male necessario (es. FAIR)
 - **fondamenta informatiche:** il web è in un flusso di evoluzione continua, framework di 5-10 anni fa sono già superati
 - per lo sviluppo di strumenti rivolti alla filologia serve una prospettiva di più lungo periodo
 - **accessibilità e conservazione** nel lungo periodo: varie soluzioni possibili, problema dei **PID** per l'edizione distribuita

Abstraction layers



Everybody loves TEI, until they don't



- 
- A decorative horizontal bar with a teal segment on the left and an orange segment on the right.
- gli schemi TEI sono lo standard *de facto* per la creazione di DSE
 - comunità aperta, collaborazione con gli utenti, miglioramento ed evoluzione costanti
 - tuttavia quando si hanno esigenze che vanno un po' oltre quanto offrono gli schemi TEI è necessario un certo sforzo per adattarli in maniera soddisfacente
 - tre esempi in cui ho dovuto creare un modello di codifica personalizzato sulla base di quanto offerto dalla TEI:
 - edizione integrata
 - codifica per la filologia d'autore
 - formule poetiche in inglese antico

```
<listPrefixDef>
  <prefixDef ident="ce34-5" matchPattern="([a-z]+)"
    replacementPattern="CE34-5_ER_diplomatic_sample.xml#$1">
    <p>In the context of this project, private URIs with the prefix "ce34-5" point to
      elements in the project's
      | CE34-5_ER_diplomatic_sample.xml file.</p>
  </prefixDef>

  <prefixDef ident="cs730" matchPattern="([a-z]+)"
    replacementPattern="CS730_ER_diplomatic_sample.xml#$1">
    <p>In the context of this project, private URIs with the prefix "cs730" point to
      elements in the project's
      | CS730_ER_diplomatic_sample.xml file.</p>
  </prefixDef>

  <prefixDef ident="cv188" matchPattern="([a-z]+)"
    replacementPattern="CV188_ER_diplomatic_sample.xml#$1">
    <p>In the context of this project, private URIs with the prefix "cv188" point to
      elements in the project's
      | CV188_ER_diplomatic_sample.xml file.</p>
  </prefixDef>
```

```
<!-- SAMPLE APP FOR INTEGRATED EDITION - SOURCE 1 -->  
<!-- Uso di range() per un singolo testimone; qui produce una duplicazione del <rdg> -->
```

```
<app>  
  <lem><w>credimus</w>  
  | <w>esse</w></lem>  
  <rdg source="cv188:range(CV188_w_002-025,CV188_w_002-026)" wit="#B2">  
  | <w>esse</w>  
  | <w>credimus</w>  
  </rdg>  
  <rdg source="ce34-5:range(CE34-5_w_002-025,CE34-5_w_002-026)" wit="#B3">  
  | <w>esse</w>  
  | <w>credimus</w>  
  </rdg>  
</app>, <w>non</w>  
<w>est</w>
```

```
<!-- SAMPLE APP FOR INTEGRATED EDITION - SOURCE 2 -->  
<!-- Uso di range() per più testimoni nello stesso <rdg> -->
```

```
<app>  
  <lem><w>credimus</w>  
  | <w>esse</w></lem>  
  <rdg source="cv188:range(CV188_w_002-025,CV188_w_002-026) ce34-5:range  
  (CE34-5_w_002-025,CE34-5_w_002-026)"  
  | wit="#B2 #B3">  
  | <w>esse</w>  
  | <w>credimus</w>  
  </rdg>  
</app>, <w>non</w>  
<w>est</w>
```

THE FOUR ENCODING LEVELS

TEXTUAL LEVEL

The critical text defined by the editor

META-TEXTUAL LEVEL

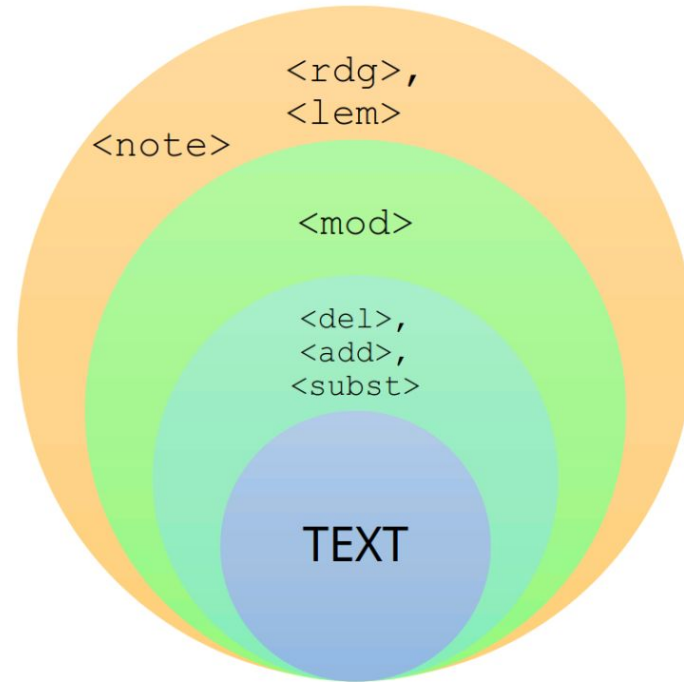
Author's comments and notes

INTRA-TEXTUAL LEVEL

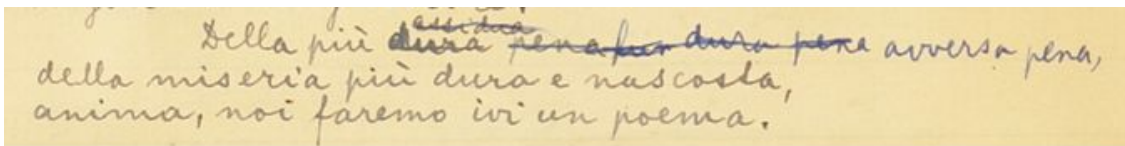
Alternative variants

GENETIC LEVEL

Author's progressive changes to the initial draft



La filologia d'autore digitale



C19, c. 102, da *Verso casa*

Apparato a stampa:

^{1A}dura pena, > ^{2A}assidua
pena, > ^{3B}dura (prima
lun<ga>) pena >
^{4B}avversa pena

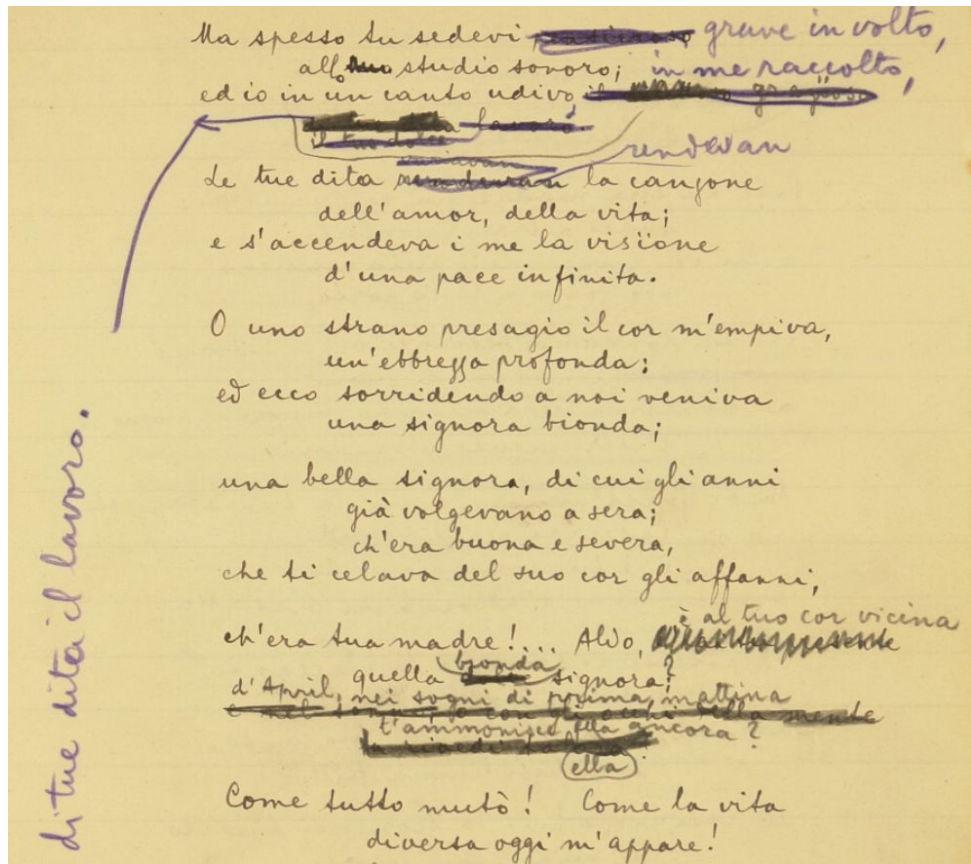
```
<l n="20">Della più <app>  
  <lem varSeq="3">  
    <mod change="#strato-2">  
      <add>avversa pena,</add>  
    </mod>  
  </lem>  
  <rdg varSeq="2">  
    <mod change="#strato-2">  
      <del>  
        <subst>  
          <del instant="true">lun<supplied>ga</supplied></del>  
          <add instant="true">dura</add></subst> pena  
        </del>  
      </mod>  
    </rdg>  
    <rdg varSeq="1">  
      <mod change="#strato-2">  
        <del>  
          <mod change="#strato-0">  
            <subst>  
              <del seq="1">dura</del>  
              <add seq="2">assidua</add>  
            </subst></mod> pena  
          </del>  
        </mod>  
      </rdg>  
    </app></l>
```

3B

1A e
2A

4B

La filologia d'autore digitale



```
<app>  
<lem varSeq="3">  
  <mod change="#strato-5">  
    <del rend="strikethrough">il tuo dolce lavoro.</del>  
    <add place="margin-left">di tue dita il lavoro.</add>  
  </mod>  
</lem>  
<rdg varSeq="2">  
  <mod change="#strato-2">  
    <del rend="strikethrough">di tue dita</del>  
    <add place="below">il tuo dolce</add></mod> lavoro.  
  </rdg>  
<rdg varSeq="1">di tue dita lavoro.</rdg>  
<note>In questa voce di apparato...</note>  
</app>
```

Ma spesso tu sedevi grave in volto

allo studio sonoro;

ed io in un canto udivo, il diletto

di tue dita lavoro.

di tue dita lavoro¹] < il tuo dolce lavoro² < di tue dita il lavoro³

Nota critica

Sequenza correzioni

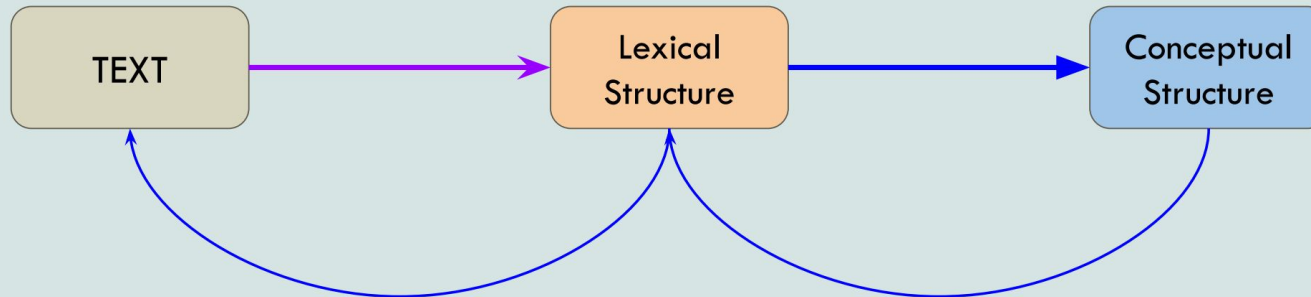
XML

1] Strato 0: di tue dita lavoro

2] Strato 2: di tue dita il tuo dolce lavoro

3] Strato 5: il tuo dolce lavoro di tue dita il lavoro

- TEI schemas offer a variety of elements that can be matched and combined to create a functional encoding model
 - For very specific models, finding the correct combination can be frustrating
- Case in point: despite its short definition (“contains a mathematical or other formula”), TEI **<formula>** is clearly unsuitable for our purpose due to its very limited content model
- Our current encoding model is still incomplete, and very likely to change, but we had to draw on several TEI modules to fulfil our needs



```
<l n="32">
  <seg ref="#witig_dryhten" type="formula" xml:id="DEOR_witig_dryhten">
    <w pos="ADJ">witig</w>
    <w pos="NOUN">dryhten</w>
  </seg> wendeþ geneahhe,</l>

<entry corresp="#adj_noun" xml:id="witig_dryhten">
  <form type="lemma">
    <orth>witig dryhten</orth>
  </form>
  <form type="components">
    <form>
      <orth><w lemmaRef="#https://bosworthtoller.com/36291" xml:id="witig">witig</w></orth>
      <gramGrp><pos>ADJ</pos></gramGrp>
    </form>
    <form>
      <orth><w>dryhten</w></orth>
      <gramGrp><pos>NOUN</pos></gramGrp>
    </form>
  </form>
</entry>
```

```
<item xml:id="eval_phrase">
  <label>EVALUATION PHRASE</label>
  <seg type="formulaPattern" subtype="phrase" function="evaluation" met="halfline">
    <term type="pos" ref="https://universaldependencies.org/u/pos/PRON.html">PRON</term>
    <term type="pos" ref="https://universaldependencies.org/u/pos/VERB.html">VERB</term>
    <term type="pos" ref="https://universaldependencies.org/u/pos/ADJ.html">ADJ</term>
    <term type="pos" ref="https://universaldependencies.org/u/pos/NOUN.html">NOUN</term>
  </seg>
  <fs>
    <f name="semantic_field">
      <vColl>
        <string>appraisal</string>
        <string>eulogy</string>
      </vColl>
    </f>
    <f name="theme"><string>good vs. bad rulers</string></f>
  </fs>
</item>
```



Grazie per l'attenzione!

