

# LA MEMORIA DIGITALE

FORME DEL TESTO E ORGANIZZAZIONE DELLA CONOSCENZA  
ATTI DEL XII CONVEGNO ANNUALE AIUCD

UNIVERSITÀ DI SIENA, 5-7 GIUGNO 2023

ASSOCIAZIONE PER  
L'INFORMATICA UMANISTICA  
E LA CULTURA DIGITALE



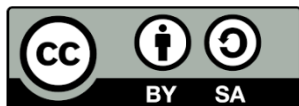
ISBN 978-88-942535-7-3

Copyright ©2023 AIUCD  
Associazione per l'Informatica Umanistica e la Cultura Digitale



Il presente volume e tutti i contributi sono rilasciati sotto licenza  
Creative Commons Attribution Share-Alike 4.0 International license ([CC-BY-SA 4.0](https://creativecommons.org/licenses/by-sa/4.0/)).  
Ogni altro diritto rimane in capo ai singoli autori.

This volume and all contributions are released under the  
Creative Commons Attribution Share-Alike 4.0 International license ([CC-BY-SA 4.0](https://creativecommons.org/licenses/by-sa/4.0/)).  
All other rights retained by the legal owners.



A cura di: Carbé, Emmanuela ; Lo Piccolo, Gabriele ; Valenti, Alessia; Stella, Francesco (2023). La memoria digitale: forme del testo e organizzazione della conoscenza. Atti del XII Convegno Annuale AIUCD, Siena: Università degli Studi di Siena

Ogni link citato era attivo al 22 maggio 2023, salvo ove diversamente indicato.  
All links have been visited on 22th May 2023, unless otherwise indicated

Si prega di notificare all'editore ogni omissione o errore si riscontri, al fine di provvedere alla rettifica.  
Please notify the publisher of any omissions or errors found, in order to rectify them.

**aiucd.segreteria [at] aiucd.org**

I contributi pubblicati nel presente volume hanno ottenuto il parere favorevole da parte di valutatori esperti della materia, attraverso un processo di revisione anonima mediante *double-blind peer review* sotto la responsabilità del Comitato Scientifico di AIUCD 2023.

All the papers published in this volume have received favourable reviews by experts in the field of DH, through an anonymous double-blind peer review process under the responsibility of the AIUCD 2023 Scientific Committee.

Il programma della conferenza AIUCD 2023 è disponibile online

The AIUCD 2023 Conference Program is available online

<http://www.aiucd2023.unisi.it>

### **Comitato Scientifico**

Nicola Barbuti  
Marina Buzzoni  
Emmanuela Carbé (co-chair)  
Fabio Ciraci  
Fabio Ciotti  
Angelo Mario Del Grosso  
Maurizio Lana  
Monica Monachini  
Paolo Monella  
Roberto Rosseelli Del Turco  
Gino Roncaglia  
Francesco Stella (chair)  
Francesca Tomasi

### **Comitato organizzativo**

Elisabetta Bartoli  
Paola Bellomi  
Monica Bianchini  
Silvia Calamai  
Riccardo Castellana  
Rosalba Nodari  
Antonio Rizzo  
Enrico Zanini

**Segreteria del Convegno:** Gabriele Lo Piccolo, Francesca Pietrini;

Giulia Bassi, Silvia Cappa, Chiara Cauzzi, Martina Corti, Elena Crocicchia, Anna Guadagnoli, Giada Giannetti, Bogdan Groza, Alessia Luvisotto, Paola Mocella, Pietro Orlandi, Martina Paccara, Elisa Petri, Maria Grazia Schiaroli

Si ringrazia l'Ufficio stampa, comunicazione istituzionale e stampa digitale, l'Ufficio Ricerca, Biblioteche, Internazionalizzazione e Terza Missione e il Supporto eventi culturali e convegnistici dell'Università di Siena

Supporto tecnico: MCM Service, Santa Chiara Lab, Presidio San Niccolò

### **Enti organizzatori**

AIUCD;

Università degli Studi di Siena: Dipartimento di Filologia e critica delle letterature antiche e moderne (DFCLAM), Centro interuniversitario di Studi Comparati I Deug-Su, Centro Interdipartimentale di Ricerca Franco Fortini in "Storia della tradizione culturale del Novecento", Santa Chiara Lab, in collaborazione con i Dipartimenti di Ingegneria dell'Informazione e Scienze Matematiche (DIISM), di Scienze sociali, politiche e cognitive (DISPOC) e di Scienze storiche e dei beni culturali (DSSBC);

CLARIN-IT.

**Con il patrocinio di:** Journal of the Text Encoding Initiative

## **Chair di track**

### **Archivi, edizioni digitali, organizzazione della conoscenza**

Marina Buzzoni, Paolo Monella, Roberto Rosselli Del Turco

### **Analisi computazionale dei testi**

Fabio Ciotti, Rachele Sprugnoli

### **Intelligenza Artificiale e modelli applicati ai beni culturali**

Monica Bianchini, Federico Boschetti

### **Preservazione della memoria e del patrimonio digitale**

Nicola Barbuti, Maurizio Lana

### **Workshop**

Francesco Stella, Emmanuela Carbé

## **Lista dei revisori**

Stefano Allegrezza, Cristiano Amendola, Paolo Andreini, Laura Antonietti, Luca Bandirali, Sofia Baroncini, Elisabetta Bartoli, Stefano Bazzaco, Andrea Bellandi, Paola Bellomi, Benedetta Bessi, Andrea Bolioli, Luca Bombardieri, Simone Bonechi, Alice Borgna, Flavia Bruni, Paolo Buono, Dino Buzzetti †, Silvia Calamai, Anna Cappellotto, Giuliana Capriolo, Vittore Casarosa, Riccardo Castellana, Paola Castellucci, Simona Chiodo, Fabio Ciraci, Elisa Corrà, Elisa Cugliana, Fabio Cusimano, Christian D'Agata, Elisa D'Argenio, Vincenza D'Urso, Stefano Dall'Aglio, Marilena Daquino, Angelo Mario Del Grosso, Antonio Di Silvestro, Diego Mantoan, Dominique Brunato, Dominique Longrée, Edmondo Grassi, Elena Spadini, Giulia Fabbris, Pierluigi Feliciati, Paolo Fioretti, Franz Fischer, Greta Franzini, Francesca Frontini, Daniele Fusi, Simone Giusti, Marco Grasso, Fabiana Guernaccini, Alessandro Iannella, Benedetta Iavarone, Alessandro Lenci, Eleonora Litta, Agnese Macchiarelli, Marco Maggini, Elisabetta Magnanti, Francesco Mambrini, Tiziana Mancinelli, Anna Maria Marras, Cristina Marras, Luca Martinelli, Stefano Melacci, Federico Meschini, Alessio Miaschi, Andrea Micheletti, Giovanni Morrone, Rosalba Nodari, Giuseppe Palazzolo, Niccolò Pancino, Fiammetta Papi, Enrico Pasini, Marco Passarotti, Giulia Pedonese, Igor Pizzirusso, Federico Ponchio, Francesca Pratesi, Alessia Lucia Prete, Simone Rebor, Giulia Renda, Gino Roncaglia, Irene Russo, Enrica Salvatori, Eva Sassolini, Daniele Silvi, Daria Spampinato, Linda Spinazze', Francesco Stella, Matteo Tiezzi, Francesca Tomasi, Sara Tonelli, Gennaro Vessio, Paul Gabriele Weston, Michelangelo Zaccarello, Patrizia Zambrano, Marco Zappatore, Andrea Zugarini

**LA MEMORIA DIGITALE**  
**XII CONVEGNO ANNUALE AIUCD**  
**Siena, 5-7 giugno 2023**

**Indice**

<i>Prefazione</i>	IV
<i>In ricordo di Dino</i>	VII
<b>1. Archivi, edizioni digitali, organizzazione della conoscenza</b>	<b>1</b>
Creare e interrogare i modelli: un duplice valore conoscitivo. Il caso dei pareri di lettura einaudiani di Natalia Ginzburg	2
Exploring iconographical and iconological content in semantic data through art historians' interpretations	9
L'applicazione del riconoscimento testi neurale per la realizzazione di ristampe digitali	15
PATH, un thesaurus per le arti performative. Osservazioni empiriche e criticità d'uso	24
I dati della ricerca: gestione, diffusione e preservazione nei Next generation repositories	28
ODI and BACODI: a study on <i>Destini incrociati</i> by Italo Calvino with Semantic Web Technologies	35
L'edizione di un graphic novel medievale: il <i>Romanzo di Tristan</i> in München, BSB, Cgm 51	41
I <i>Capricci d'un Re</i> digitale	48
Harmonizing and modelling a bibliographic ontology of imperial diplomas within the OntoVE Knowledge Base	51
Mappare il museo in IIIF. Una combinazione di deep zoom e VR360 per la Galleria Borghese di Roma	57
H2IOSC: Humanities and Heritage Open Science Cloud	63
Śivadharmā Database CMS. HTML and graph as a starting point for digital editions	65
L'approccio human-centered nella progettazione di un ecosistema digitale che massimizzi la diffusione del materiale diaristico preservandone la singolarità	71
Un futuro digitale per gli archivi de 'Gli Anni della Cupola'	78
<i>Ianus 2.0</i> . Corrottele filologiche, congetture digitali	85
Verso una «volontà d'archivio» digitale	92
L'applicazione del digitale alla valorizzazione della lingua come patrimonio culturale immateriale	99
Digital <i>Accrocchio</i> : a computational image searching tool for social history	104
Quali prospettive per ItaDraCor? Risorse e strumenti per la codifica di testi drammatici in lingua italiana	108
Metodologie di indagine per una nuova piattaforma: Digital Scholarly Editions Platform and Aligned Translation	111
Making Korean Perseus Digital Library: The Case Study of 'Dokdo Digital Library'	114
Constructing an Old English WordNet: The Case of <i>Guilt</i>	122
A digital diplomatic edition of the 10th-century charters of Lucca for Latin corpus linguistics	124
On why and how we should build a catalogue of software products for digital scholarly editing	130
Database Allegorico Dantesco	134
Tecnologie digitali per la realizzazione e adattamento di contenuti scolastici facilitati e semplificati	139
<i>Proteus</i> : marcatura di testi per una lettura antropologica del mito greco e romano	145
Seeing the Unseen: an approach for visualising situated statements in the Cultural Heritage domain	149
<i>Soavissime ariette</i> : impiego di repertori digitali per lo studio della tradizione delle arie d'opera	154
<i>Musisque Deoque</i> : la nuova organizzazione dei testimoni manoscritti	159
Storytelling with Linked Open Data	164
From ontology design to user-centred interfaces for music heritage	168

Un modello open data per l'interazione utente: la digital library Open Memory Apulia	173
Nel laboratorio di <i>Poeti d'Italia in lingua latina</i> : lavori in corso e prospettive future	177
Metadata e concetti, una nuova intermediazione delle fonti. Quali prospettive per la ricerca e la diffusione della storia?	181
Metodi di armonizzazione per codifiche TEI cooperative: l'edizione Bellini Digital Correspondence	185
LINEA: per un portale della letteratura italiana del Novecento	191
'Voce dei Margini': Le annotazioni di Giorgio Bassani	197
Scholarly edition of born-digital literary sources: a case study	204
Per un atlante delle biblioteche medievali: il progetto Mirabile-Atlas e una nuova prospettiva sulla circolazione della cultura manoscritta	208
Dai metadati ai dati, dai contesti ai contenuti: aumentare la descrizione archivistica	214
ITINERA: progetto, database e prospettive	220
Per l'edizione digitale delle lettere familiari di Verga: un esperimento di analisi semantica	224
Ease the collaboration making Scholarly Editions: the GreekSchools case study	230

## **2. Analisi computazionale dei testi** **233**

Zoo-KG: A knowledge graph of zoological data	234
<i>Do AIs dream of electric comics?</i> Generative AI models, digital memory, and creativity	237
ChatGPT: un Pappagallo Stocastico può essere di aiuto a un Vero Ricercatore (Umanistico)?	245
Un atlante digitale per i borghi delle aree interne	251
To what extent are lemmatisation and annotation relevant for deep learning assignments and textual motifs detection? The case-study of Peter Damian's letters (11 <sup>th</sup> century)	254

## **3. Intelligenza artificiale e modelli applicati ai beni culturali** **260**

<i>Falso movimento</i> (Moretti 2022): Memories and Translations	261
Verso la creazione della versione italiana di SPARSAR	265
Unsupervised Information Extraction from Medieval Charters	275
The challenges of sense granularity in word sense disambiguation for Latin	279
Machine Learning Algorithm for the Scansion of Old Saxon Poetry	284
Diachronic and diatopic word embeddings from British historical newspapers	289
GPT-3 vs. Delta. Applying stylometry to large language models	292
EcoLet. Ecotemario della Letteratura italiana	298
Risorse linguistiche per lo studio dei <i>Promessi Sposi</i>	301

## **4. Preservazione della memoria e del patrimonio digitale** **304**

Uso e riuso di dati per la valorizzazione del patrimonio culturale italiano online: buone pratiche per l'interoperabilità	305
"Dalla civiltà dell'archivio alla civiltà dell'algoritmo?"	309
<i>Breaking the Canvas</i> A Serious Game to challenge stereotypes in Visual Culture through Fine Arts	315
CoMontology. Towards An Ontology for the Chronology of Mediterranean archaeologies: a model for the digital memory	322
Dematerializzazione e sostenibilità: prospettive ideali e problemi concreti	327
New Media Art: indagini conservative sulle eredità immateriali dell'arte di fine millennio	332
The challenges of preservation: digitizing graffiti in the urban landscape	338
The Lions of Venice Arsenal: a neuroscientific approach to Cultural Heritage	343
L'immaginazione tra 0 e 1. Alcune considerazioni su memoria e oblio nell'era digitale	347
A methodological study on the projects developed by the Cologne Center for Humanities (CCeH)	351

Preservare il <i>Tesoro</i> . Per un nuovo sistema di interrogazione dei corpora dell'Opera del Vocabolario Italiano	358
La galassia MQDQ tra innovazione e tradizione	365
Memorizzare o ricordare? La memoria come fatto emozionale	371
GECA 3.0 – A new tool for cataloguing and enjoying cultural heritage	373
Libri, memoria e territorio: la digital library di Carlo Alberto	380
Un modello di catalogazione digitale delle grammatiche bilingue di spagnolo per italofofi	382
Un'ontologia per la classificazione semantica dei dizionari storici: l'esempio del LEI Digitale	385
How many oral archives are in your home? Piloting a new Tuscan census in the Gra.fo Reloaded project	389
Le mafie e le Digital Humanities	395
I Confini della Lunigiana: un progetto di Digital Public History	401
L'immaginario francese della Puglia: una collezione digitale del 'Petit Tour' dell'altra Italia	404
<b>5. Workshop</b>	<b>409</b>
Workshop sulle edizioni digitali: preparazione con codifica XML TEI e visualizzazione con il software EVT	410
Trascrivere, analizzare e pubblicare documenti d'archivio: Transkribus e l'intelligenza artificiale al servizio dello storico	416
Codificare (a)l'Infinito	421

# Constructing an Old English WordNet: The Case of *Guilt*

Anas Fahad Khan<sup>1</sup>, Michele Cavallaro<sup>2</sup>, Rafael Cruz González<sup>3</sup>, Javier Díaz-Vera<sup>4</sup>,  
Francesca Frontini<sup>5</sup>, Francisco Javier Minaya Gómez<sup>6</sup>

<sup>1</sup> Istituto di Linguistica Computazionale ‘Antonio Zampolli’ (CNR-ILC), Italy – fahad.khan@ilc.cnr.it

<sup>2</sup> University of Siena, Italy – michele.cavallaro@student.unisi.it

<sup>3</sup> Universidad de Castilla-La Mancha, Spain – rafael.cruz@uclm.es

<sup>4</sup> Universidad de Castilla-La Mancha, Spain - javierenrique.diaz@uclm.es

<sup>5</sup> Istituto di Linguistica Computazionale ‘Antonio Zampolli’ (CNR-ILC), Italy – francesca.frontini@ilc.cnr.it

<sup>6</sup> Universidad de Castilla-La Mancha, Spain – francisco.minaya@uclm.es

## ABSTRACT

In this paper, we look at the manual construction of a lexicon of emotion terms in Old English organised as a wordnet lexicon and based on a pre-existing dataset which categorises emotion terms on the basis of cognitive criteria. This is part of a wider initiative to construct a wordnet for Old English which includes both manually and (semi-)automatically constructed parts. We will describe our workflow which foresees the use of lemmas, senses and definitions taken from and in many cases adapted from a legacy Old English lexical resource, ClarkHall’s AngloSaxon Dictionary, and organised into synsets by matching the definitions of OE terms with the synset definitions found in the Open English WordNet. We argue that this is a useful way of disseminating research in lexical semantics and philology taking as an example the semantic field of GUILT. In our poster presentation we will give a visual description of the whole field; in the present submission we only focus on the most ‘core’ nominal synsets relating to GUILT.

## KEYWORDS

Old English; wordnet; lexicography; guilt

## 1. INTRODUCTION

The work described in this submission is part of a wider project which aims towards the creation of a wordnet for the Old English<sup>1</sup> language (OE), the *Old English WordNet* (OldEWN), a project which is described in detail in previous works [3,4]. The OldEWN is being constructed semi-automatically<sup>2</sup> on the basis of the information contained in a public domain Old English dictionary, J. R. Clark-Hall’s *Concise Anglo-Saxon Dictionary* (1916) (CH) [1]. The entries in this latter work are short and easy to process; something which distinguishes it both from the earlier and much more comprehensive Old English dictionary Bosworth and Toller’s *Anglo-Saxon Dictionary* [2], which contains extensive information on the distribution of senses and a considerable amount of sense nesting, and the contemporary Dictionary of Old English (DOE)<sup>3</sup> generally considered the most authoritative dictionary of the language. This paper focuses on the incorporation of a pre-existing dataset describing emotion terms in OE into OldEWN. More precisely, this dataset categorises entries in the OE lexicon of emotions on the basis of the different figurative semantic shifts, if any, which occurred historically between different senses of the same word (this is part of a wider research program in Cognitive Linguistics described in). One of the outcomes of this work of incorporating this dataset into our wordnet has been the development of an extended version of the standard Global Wordnet Association Schema for wordnets<sup>4</sup> as adapted to the exigencies of studying historic languages from a diachronic perspective within a cognitive linguistics inspired framework<sup>5</sup>. Another outcome of the work is to demonstrate the usefulness of the WN organisational approach and schema (with appropriate extensions such as ours) in making datasets such as these in historical lexical semantics more interoperable and potentially easier to disseminate on the basis of principles that are fairly intuitive and easily grasped (notion of synsets with relations like hyponymy between them). In what follows we will show this with the example of OE terms related to the notion of guilt. First, we will describe our workflow, remarking upon the potential of this task for didactic ends. Then we will show some of the OE synsets which we derived for the semantic field of GUILT.

<sup>1</sup> By Old English here we refer to a language spoken in the British Isles from around the fifth century AD up until around the 12th century. It is the earliest ancestor of the English language for which we have written evidence.

<sup>2</sup> We foresee an initial phase of automatic assignment of synsets to word senses, followed by a subsequent phase of correction by Old English experts via a specially designed platform.

<sup>3</sup> <https://doe.artsci.utoronto.ca/>

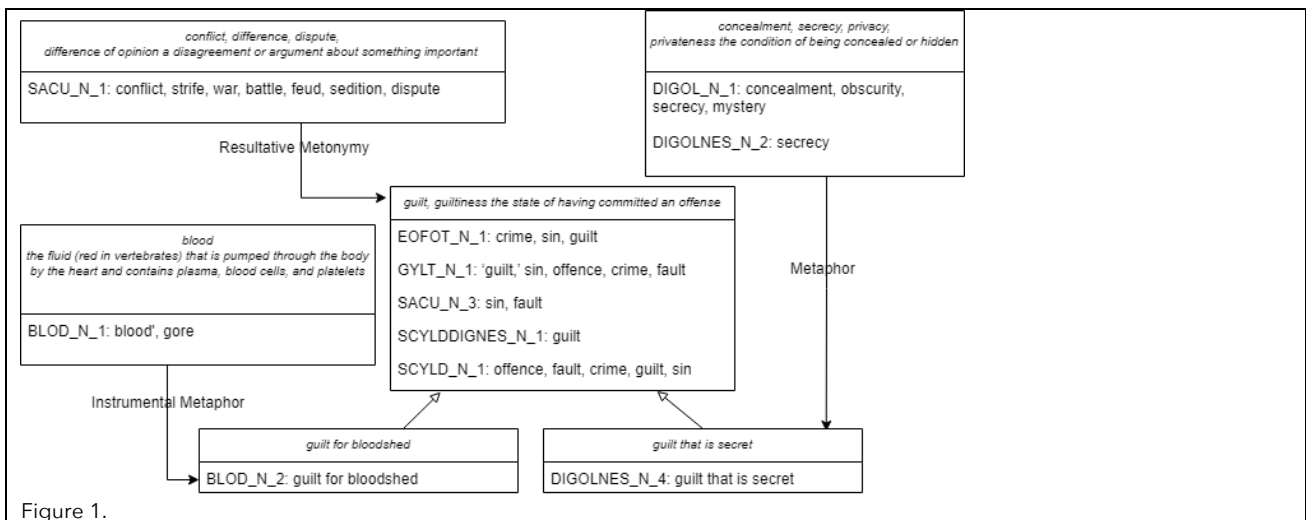
<sup>4</sup> <https://globalwordnet.github.io/schemas/>

<sup>5</sup> <https://github.com/anasfkhan81/OldEnglish/blob/main/WN-IELMF-0.DTD>



## 2. WORKFLOW AND EXAMPLE

The task of converting the original Diaz-Vera (DV) dataset with its description of the OE emotion lexicon into a wordnet style lexicon was part of a collaboration between three institutions<sup>6</sup> and involved interns from the University of Siena working under the supervision of experts from CNR-ILC and UCLM. In addition to contributing to the creation of the OldEWN<sup>7</sup>, this task also serves the didactic purpose of introducing the interns both to wordnets and to the Old English lexicon (and more broadly to language resources for ancient languages) as well as to working with lexicographic resources. The task begins with the alignment of the lemmas and descriptions in the original DV dataset (this dataset is described in more detail in [3]) with those in the Clark Hall (CH) dictionary<sup>8</sup>, followed by the enrichment of the CH definitions from the Bosworth Toller dictionary when this is required by the data; in some cases senses are added based on other scholarly sources. Since this part of the wordnet is based on a specialist dataset we include some very rare senses in our synsets; these specialised senses will however be marked as such in the final version of the resource when it is made available. The OE word definitions are subsequently compared with the synset definitions in the Open English wordnet<sup>9</sup> (OEWN), an open source version of the Princeton WordNet. Each definition is matched to its closest synset in case that one exists. Where there exists no such match, the annotator must look for an OEWN hypernym (this usually tends to be quite straightforward). Although this task is currently fully manual, it has given us significant insight into how the CH can best be leveraged for building OE synsets automatically and the many challenges such a task faces. We end this submission with a diagram representing part of the semantic field of GUILT in OE on the basis of the DV dataset structured as a wordnet. Here we show the more ‘core’ terms in the OE guilt vocabulary, in this case nouns such as *eofot*, *gylt*, *sacu*, and *scyld* in their various different senses. Note the extra relationships which we have added to the basic repository of wordnet relations proposed by the GWA. These allow us to signal diachronic sense shift relationships between the senses in synsets and to categorise the kind of figurative (metaphoric, metonymic) shift involved (e.g., from secrecy to a secret guilt).



In our poster we will present a map of the entire semantic field taking into consideration different parts of speech.

## REFERENCES

- [1] Clark-Hall, J. R. A concise Anglo-Saxon dictionary: for the use of students. Swan Sonnenschein & Company, 1894.
- [2] Bosworth, J., & Toller, T. N. An Anglo-Saxon Dictionary, Oxford, 1889.
- [3] Khan, A. F., Gómez, F. J. M., González, R. C., Diakoff, H., Diaz-Vera, J. E., McCrae, J. P., O’Loughlin C., Short W.M., & Stolk. S. Towards the Construction of a WordNet for Old English. In Proceedings of the Thirteenth Language Resources and Evaluation Conference, pages 3934–3941, Marseille, France. European Language Resources Association, 2022.
- [4] Khan, A. F., McCrae, J. P., Gómez, F. J. M., González, R. C., & Diaz-Vera, J. E. Some Considerations in the Construction of a Historical Language WordNet. Global WordNet Conference, 2023.

<sup>6</sup> The Istituto di Linguistica Computazionale (CNR-ILC), the University of Siena, and Universidad de Castilla-la Mancha (UCLM)

<sup>7</sup> The part of the wordnet dealing with emotions will eventually be incorporated into the rest of the OldEWN.

<sup>8</sup> The CH lemmas are already the basis of the lemma list of the contemporary and authoritative Dictionary of Old English.

<sup>9</sup> <https://en-word.net/>

ASSOCIAZIONE per  
l'INFORMATICA UMANISTICA  
e la CULTURA DIGITALE



UNIVERSITÀ  
DI SIENA 1240



SANTA CHIARA LAB  
Università di Siena 1240

CENTRO DI STUDI COMPARATI  
IDEUG-SU



Centro di Ricerca  
Franco Fortini



Con il patrocinio di  
TEI JOURNAL OF THE TEXT  
ENCODING INITIATIVE