



Università
di Catania

ASSOCIAZIONE per
l'INFORMATICA UMANISTICA
e la CULTURA DIGITALE

Consiglio Nazionale
delle Ricerche

ME.TE. DIGITALI

MEDITERRANEO IN RETE TRA TESTI E CONTESTI

ATTI DEL XIII CONVEGNO ANNUALE
AIUCD 2024



28 - 30 MAGGIO
MONASTERO DEI BENEDETTINI
P.ZZA DANTE, 32 CATANIA

ISBN 978-88-942535-8-0



Copyright ©2024 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). 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). All other rights retained by the legal owners.

A cura di: Di Silvestro Antonio; Spampinato Daria (2024). Me.Te. Digitali. Mediterraneo in rete tra testi e contesti, Proceedings del XIII Convegno Annuale AIUCD, Catania 28-30 maggio 2024, Università di Catania.

Editing: Denise Bruno; Christian D'Agata; Laura Mazzagufò; Francesca Prado; Eliana Vitale; Alessandro Zammataro.

Ultimo accesso agli URL in data 15 maggio 2024.

Si prega di notificare all'editore ogni omissione o errore si riscontri: [aiucd.segreteria \[at\] aiucd.org](mailto:aiucd.segreteria@aiucd.org)
Please notify the publisher of any omissions or errors found: [aiucd.segreteria \[at\] aiucd.org](mailto:aiucd.segreteria@aiucd.org)

Il programma della conferenza AIUCD 2024 è disponibile online <https://aiucd2024.unict.it/programma/>
The AIUCD 2024 Conference Program is available online <https://aiucd2024.unict.it/programma/>

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 di Programma di AIUCD 2024.

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 2024 Program Committee.

Chair

Antonio Di Silvestro (Università di Catania)

Daria Spampinato (CNR Istituto di Scienze e Tecnologie della Cognizione)

Comitato di programma / Program committee

Emmanuela Carbé (Università di Siena)

Massimo Cultraro (CNR Istituto di Scienze del Patrimonio Culturale)

Christian D'Agata (Università di Catania)

Antonio Di Silvestro (Università di Catania)

Greta Franzini (Eurac Research)

Maurizio Lana (Università del Piemonte Orientale)

Cristina Marras (CNR Istituto del Lessico intellettuale europeo e Storia delle Idee)

Marco Mazzone (Università di Catania)

Ouafae Nahli (CNR Istituto di Linguistica Computazionale "Antonio Zampolli")

Marianna Nicolosi-Asmundo (Università di Catania)

Marina Paino (Università di Catania)

Giuseppe Palazzolo (Università di Catania)

Jonathan Prag (University of Oxford Merton College)

Daria Spampinato (CNR Istituto di Scienze e Tecnologie della Cognizione)

Rachele Sprugnoli (Università di Parma)

Francesco Stella (Università di Siena)

Segreteria scientifica / Scientific Secretariat

Liborio Barbarino (Università di Catania)

Denise Bruno (Università di Catania)

Giulia Cacciatore (Università di Catania)

Giuseppe Canzoneri (Università di Catania)

Elisa Conti (Università di Catania)

Milena Giuffrida (Università di Catania)

Miryam Grasso (Università di Catania)

Francesca Prado (Università di Catania)

Emilio M. Sanfilippo (CNR Istituto di Scienze e Tecnologie della Cognizione)

Eliana Vitale (Università di Catania)

Alessandro Zammataro (Università di Catania)

Comunicazione istituzionale: Claudia Cantale (Università di Catania) e Area Per la Comunicazione dell'Università di Catania (ACOM).

Institutional communication: Claudia Cantale (University of Catania) and the Area for Communication of the University of Catania (ACOM)

Supporto tecnico: Rosario Agrò, Area della Terza Missione dell'Università di Catania, per la consulenza e la progettazione grafica dei materiali informativi del convegno.

Technical support: Rosario Agrò, Third Mission Area of the University of Catania, for advice and graphic design of the conference information materials.

Enti organizzatori / Organisers

AIUCD; Università di Catania: Dipartimento di Scienze Umanistiche; CNR Istituto di Scienze e Tecnologie della Cognizione; CINUM: Centro di Informatica Umanistica dell'Università di Catania.

Supporter

CLARIN-IT; Neperia Group; Storage; programma Piaceri 2020-2022, Linea 1; Parmalat-Sole.

Chair di area/ Track chair

Le culture digitali nel Mediterraneo

Cristina Marras (CNR Istituto del Lessico intellettuale europeo e Storia delle Idee)

Paola Moscati (CNR Istituto di Scienze del Patrimonio Culturale)

Archivi ed edizioni digitali

Christian D'Agata (Università di Catania)

Greta Franzini (Eurac Research)

Analisi computazionale dei testi

Angelo Mario Del Grosso (CNR Istituto di Linguistica Computazionale "Antonio Zampolli")

Simone Reborra (Università di Verona)

Ontologie e Semantic Web

Marianna Nicolosi Asmundo (Università di Catania)

Francesca Tomasi (Università di Bologna)

Preservazione della memoria e del patrimonio digitale

Fabio Ciraci (Università del Salento)

Anna Maria Marras (Università di Torino)

Lista dei revisori /List of reviewers

Maristella Agosti (Università di Padova), **Stefano Allegrezza** (Università di Bologna), **Chiara Alzetta** (CNR Istituto di Linguistica Computazionale "Antonio Zampolli"), **Liborio Barbarino** (Università di Catania), **Nicola Barbuti** (Università di Bari Aldo Moro), **Stefano Bazzaco** (Università di Verona), **Benedetta Bessi** (Università Ca' Foscari di Venezia), **Andrea Bolioli** (ricercatore indipendente), **Paolo Bonora** (Università di Bologna), **Federico Boschetti** (CNR Istituto di Linguistica Computazionale "Antonio Zampolli"), **Dominique Brunato** (CNR Istituto di Linguistica Computazionale "Antonio Zampolli"), **Flavia Bruni** (Università Gabriele D'Annunzio di Chieti-Pescara), **Marina Buzzoni** (Università Ca' Foscari di Venezia), **Alberto Campagnolo** (Université Catholique de Louvain/KULeuven), **Anna Cappellotto** (Università di Verona), **Emmanuela Carbé** (Università di Siena), **Vittore Casarosa** (CNR Istituto di Scienza e Tecnologie dell'Informazione "A. Faedo" – Università di Pisa), **Fabio Ciotti** (Università di Roma "Tor Vergata"), **Fabio Ciraci** (Università del Salento), **Elisa Conti** (Università di Catania), **Salvatore Cristofaro** (CNR Istituto per il Lessico Intellettuale Europeo e Storia delle Idee), **Christian D'Agata** (Università di Catania), **Elisa D'Argenio** (HUN-REN Hungarian Research Centre for Linguistics), **Mauro De Bari** (Università di Bari Aldo Moro), **Riccardo Del Gratta** (CNR Istituto di Linguistica Computazionale "Antonio Zampolli"), **Angelo Mario Del Grosso** (CNR Istituto di Linguistica Computazionale "Antonio Zampolli"), **Saulo Delle Donne** (Università del Salento), **Giorgio Maria Di Nunzio** (Università di Padova), **Antonio Di Silvestro** (Università di Catania), **Filippo Diara** (Università di Torino), **Giulia Fabbris** (Università Ca' Foscari di Venezia), **Riccardo Fedriga** (Università di Bologna), **Franz Fischer** (Università Ca' Foscari di Venezia), **Greta Franzini** (Eurac Research), **Francesca Frontini** (CNR Istituto di Linguistica Computazionale "Antonio Zampolli"), **Daniele Fusi** (Stuttgart University & VeDPH – Università Ca' Foscari di Venezia), **Carola Gatto** (Università del Salento), **Lucia Giagnolini** (Università di Bologna), **Emiliano Giovannetti** (CNR Istituto di Linguistica Computazionale "Antonio Zampolli"), **Milena Giuffrida** (Università di Catania), **Edmondo Grassi** (Università San Raffaele di Roma), **Miryam Grasso** (Università di Catania), **Alessandro Iannella** (Università di Cagliari - Università di Pisa – Università di Torino), **Paola Italia** (Università di Bologna), **Maurizio Lana** (Università del Piemonte Orientale), **Pietro Maria Liuzzo** (Bibliotheca Hertziana), **Dominique Longrée** (Université de Liège), **Francesco Mambrini** (Università Cattolica del Sacro Cuore di Milano), **Tiziana Mancinelli** (Istituto Italiano di Studi Germanici), **Anna Maria Marras** (Università di Torino), **Cristina Marras** (CNR Istituto del Lessico intellettuale europeo e Storia delle Idee), **Federico Meschini** (Università della Toscana), **Alessio Miaschi** (CNR Istituto di Linguistica Computazionale "Antonio Zampolli"), **Paolo Monella** (Università Sapienza di Roma), **Ouafae Nahli** (CNR Istituto di Linguistica Computazionale "Antonio Zampolli"), **Marianna Nicolosi-Asmundo** (Università di Catania), **Giuseppe Palazzolo** (Università di Catania), **Valentina Pasqual** (Università di Bologna), **Gianluca Pavani** (Università di Roma "Tor Vergata"), **Giulia Pedonese** (CNR Istituto di Linguistica Computazionale "Antonio Zampolli"), **Jonathan Prag** (University of Oxford Merton College), **Simone Reborra** (Università di Verona), **Giulia Renda** (Università di Bologna), **Roberto Rosselli Del Turco** (Università di Torino), **Enrica Salvatori** (Università di Pisa), **Emilio M. Sanfilippo** (CNR Istituto di Scienze e Tecnologie della Cognizione), **Eva Sassolini** (CNR Istituto di Linguistica Computazionale "Antonio Zampolli"), **Pietro Sichera** (CNR Istituto di Scienze e Tecnologie della Cognizione), **Daniele Silvi** (Università di Roma "Tor Vergata"), **Elena Spadini** (University of Basel), **Daria Spampinato** (CNR Istituto di Scienze e Tecnologie della Cognizione), **Linda Spinazzè** (Università Ca' Foscari di Venezia), **Rachele Sprugnoli** (Università di Parma), **Francesco Stella** (Università di Siena), **Cecilia Tamagnini** (Università di Bologna), **Timothy Tambassi** (Università Ca' Foscari di Venezia), **Francesca Tomasi** (Università di Bologna), **Marco Venuti** (Università di Catania), **Fabio Vitali** (Università di Bologna).

SOMMARIO

PREFAZIONE	I
<i>Antonio Di Silvestro, Daria Spampinato</i>	
RELAZIONI DEI KEYNOTE SPEAKERS	VI
INFORMATICA UMANISTICA, INFOCRAZIA, E INTELLIGENZE ARTIFICIALI	VII
<i>Giuseppe Savoca</i>	
THE MEDIEVAL MEDITERRANEAN IN...DATA? INTERPRETATION, CONJECTURE AND DIGITAL METHODS	XIII
<i>Tara L. Andrews</i>	
RELAZIONI DEGLI INVITED SPEAKERS	XXI
RICOSTRUZIONE DEL TESTO E BANCHE DATI. LA FILOGIA DIGITALE ALLA PROVA DELL'ESEGESI ANTICA DELLA <i>COMMEDIA</i>	XXII
<i>Vittorio Celotto, Andrea Mazzucchi</i>	
IL MEDITERRANEO: UN MARE DI OPPORTUNITÀ E SFIDE	XXX
<i>Salvatore Capasso</i>	
MEDITERRANEO TRA TESTI E CONTESTI	1
COMBINING GENERATIVE AI AND ARCHAEOLOGY TO BUILD DATA-DRIVEN STORIES	2
<i>Francesca Buscemi, Angelica Lo Duca</i>	
DIGITAL PRESERVATION E SOSTENIBILITÀ AMBIENTALE	9
<i>Adele Gorini</i>	
IL PROGETTO DIGITALE SU ELISA CHIMENTI (LAI-ALEEF): LA PROBLEMATICITÀ DI UN PROFILO MEDITERRANEO TRA RETI E FRONTIERE	15
<i>Ada Desideri, Bianca Vallarano</i>	
MEDITERRANEO VERDE FANTASTICO: UN REPERTORIO DIGITALE ATTRAVERSO LA COLLEZIONE BOTANICA DI ULISSE ALDROVANDI (1522-1605)	21
<i>Sara Obbiso</i>	
PER LA DIGITALIZZAZIONE DEL PATRIMONIO LINGUISTICO E CULTURALE ITALIANO IN EGITTO: I PERIODICI ITALIANI (1892- 1940)	26
<i>Wafaa El Beih</i>	
PIATTAFORME WIKI PER L'INSEGNAMENTO UMANISTICO: SPERIMENTAZIONI IN CORSO NEL LICEO DE COSMI DI PALERMO	32
<i>Antonino Fiorino, Paolo Monella, Francesca Saieva, Antonella Sorci</i>	
THE ORGANIZATION AND MANAGEMENT OF THE MAGIC PROJECT FOR ANCIENT MANUSCRIPTS DIGITIZATION: CONNECTIONS BETWEEN MEDITERRANEAN CULTURES	38
<i>Stefania Conte, Andrea Mazzucchi, Guido Russo, Augusto Tortora, Giorgia Tortora</i>	
TOWARDS A RESEMANTISATION OF THE CONCEPT OF MODELLING IN DIGITAL HUMANITIES	43
<i>Cristina Marras, Arianna Ciula, Øyvind Eide, Patrick Sahle</i>	
VALORIZZARE UN ARCHIVIO 'MEDITERRANEO': STUDI PER UN'EDIZIONE CRITICA DIGITALE DELLE OPERE DI GIOVANNI COMISSO	48
<i>Marco Borrelli</i>	
ARCHIVI E MUSEI DIGITALI PER IL PATRIMONIO CULTURALE	55
A WORKFLOW FOR GLAM METADATA CROSSWALK	56
<i>Arianna Moretti, Ivan Heibi, Silvio Peroni</i>	
DIGITALIZZAZIONE E MODELLAZIONE DELLA <i>DRAMMATURGIA</i> DI LEONE ALLACCI	63
<i>Luca Giovannini, Giorgia Gallucci</i>	

FROM DATA COMPLEXITY TO USER SIMPLICITY: A FRAMEWORK FOR LINKED OPEN DATA RECONCILIATION AND SERENDIPITOUS DISCOVERY	67
<i>Marco Grasso, Giulia Renda, Marilena Daquino</i>	
GIUSEPPE CHIARINI: UN'OPERA INEDITA	73
<i>Elena Almangano, Mirko Castaldi, Eleonora De Longis, Daniele Pasqualetti</i>	
HERITRACE: TRACING EVOLUTION AND BRIDGING DATA FOR STREAMLINED CURATORIAL WORK IN THE GLAM DOMAIN	78
<i>Arcangelo Massari, Silvio Peroni</i>	
LIBRI E BIBLIOTECHE TRA MUSEABILITÀ E MUSEALIZZAZIONE DIGITALE: SOGNO O REALTÀ?	84
<i>Nicola Barbuti, Mauro De Bari</i>	
LISTENING2PAINTING: AN AUDIO AUGMENTED REALITY APPROACH FOR ARTS	89
<i>Nicola Orio, Daniel Zilio, Andrea Micheletti</i>	
LUOGHI COMUNI: METODI E STRATEGIE DI SVILUPPO SOFTWARE IN AMBITO GLAM, DALLE VOCI DI AUTORITÀ ALL'ESPLORAZIONE CARTOGRAFICA	92
<i>Herbert Natta, Gianluca Rossi, Roberta Maggi</i>	
MESSAGGISTICA ISTANTANEA E ARCHIVI DIGITALI. QUALI SOLUZIONI? BEST PRACTICES E CONSIDERAZIONI DAL CONTESTO INTERNAZIONALE	98
<i>Alessia Del Bianco</i>	
NUOVE INTERAZIONI CON COLLEZIONI DIGITALI: L'“ARCHIVIO DIGITALE DEL CAPITOLO DI LATERZA”	104
<i>Stefania Riso, Nicola Barbuti</i>	
PRESERVARE E VALORIZZARE LA MEMORIA DI ARCHIVI STORICI DI EX-OSPEDALI PSICHIATRICI	110
<i>Grazia Serratore</i>	
PRESERVAZIONE DEL PATRIMONIO CULTURALE E CLOUD COMPUTING: CARATTERISTICHE E CRITICITÀ	117
<i>Manuela Grillo</i>	
PRESERVING CULINARY TRADITIONS. A CROWDSOURCED DIGITAL COLLECTION OF COOKBOOKS	122
<i>Giulia Renda, Giulia Manganelli, Mila Fumini, Marilena Daquino</i>	
SERIOUS GAMES E GAMIFICATION: A CHE PUNTO SONO LE ISTITUZIONI CULTURALI ITALIANE?	128
<i>Vincenzo Colaprice</i>	
THE TREE OF PHILOSOPHERS: DESIGN AND IMPLEMENTATION OF A DIGITAL RESOURCE FOR THE HISTORY OF ACADEMIC PHILOSOPHY	133
<i>Guido Bonino, Nicola Ruschena</i>	
THINKING OUTSIDE THE BLACK BOX: INSIGHTS FROM A DIGITAL EXHIBITION IN THE HUMANITIES	138
<i>Sebastian Barzaghi, Alice Bordignon, Bianca Gualandi, Silvio Peroni</i>	
EDIZIONI SCIENTIFICHE DIGITALI	143
IL PROGETTO CORR<SI>CA: EDIZIONE DIGITALE DELLA CORRISPONDENZA CANIONI	144
<i>Anna Giaufret, Beatrice Dal Bo, Elena Margherita Vercelli, Laura Bonanno</i>	
IL PROGETTO “MAXIMHUM”: ITALIA UMANISTICA E MOSCOVIA CINQUECENTESCA DIALOGANO IN DIGITALE	152
<i>Francesca Romoli, Letizia Ricci, Angelo Mario Del Grosso</i>	
L'ARCHIVIO DI GIUSEPPE FAVA: CONSERVAZIONE E VALORIZZAZIONE ATTRAVERSO IL DIGITALE	159
<i>Giuseppe Davide Di Mauro, Marzia D'Amico</i>	
L'ARCHIVIO DIGITALE DI UNA CASA EDITRICE: L'ESEMPIO DEL SAGGIATORE E DELLA SUA PRIMA PUBBLICAZIONE	165
<i>Giada Di Pino</i>	
MARCARE LA POESIA DEL NOVECENTO: UNO STUDIO PER <i>OSSI DI SEPPIA</i>	171
<i>Chiara Cauzzi, Martina Corti, Anna Guadagnoli, Maria Grazia Schiaroli</i>	
METASCRIP: A FRAMEWORK PROPOSAL FOR SCREENPLAY ENCODING	175
<i>Erica Andreose, Giorgia Crosilla, Leonardo Zilli</i>	
OPENDATA: OPENGADDA	181
<i>Eleonora Pasquale, Martina Pensalfini</i>	
PAUL KLEE, <i>TUNISREISE</i> E <i>BILDNERISCHE FORMLEHRE</i> : UN CASO STUDIO DI DISCEPT (DIGITAL SCHOLARLY EDITIONS PLATFORM AND ALIGNED TRANSLATIONS)	186
<i>Hansmichael Hohenegger, Tiziana Mancinelli, Fabio Ciotti, Federico Boschetti, Angelo Mario Del Grosso, Eleonora De Longis</i>	

PAVES-E: PER UNA HYPEREDIZIONE DELL'OPERA DI CESARE PAVESE	191
<i>Christian D'Agata, Angelo Mario Del Grosso, Laura Nay, Giuseppe Palazzolo, Antonio Sichera, Daria Spampinato</i>	
PER UN'EDIZIONE DIGITALE DI <i>SE QUESTO È UN UOMO</i>	197
<i>David Tagliacozzo</i>	
PER UN'EDIZIONE SCIENTIFICA DIGITALE DELLO <i>SPECULUM GUY OF WARWICK</i>	204
<i>Omar Khalaf, Sibilla Siano</i>	
PER UNA LETTURA ANTROPOLOGICA DI VERGA: TRA CODIFICA E GEOREFERENZIAZIONE	210
<i>Giovanna Zisa</i>	
PROGETTO DI EDIZIONE GENETICA DIGITALE DEL <i>CANZONIERE</i> MANOSCRITTO DI U. SABA (1919-20)	215
<i>Marina Buzzoni, Davide Cucurnia, Cristina Fenu, Roberto Rosselli Del Turco, Giulia Tancredi</i>	
RAPPRESENTARE LA STORIA SACRA: UN'IMPRESA IERI, UNA SFIDA OGGI. PROPOSTA DI EDIZIONE SCIENTIFICA DIGITALE DEL "COMPENDIUM" DI PIETRO DI POITIERS	221
<i>Franz Fischer, Agnese Macchiarelli</i>	
TOWARDS AN INTEGRATED DIGITAL EDITION OF THE <i>LEGES LANGOBARDORUM</i>	226
<i>Marina Buzzoni, Roberto Rosselli Del Turco</i>	
UN <i>CORPUS</i> ONLINE DELLA LETTERATURA SECONDARIA (1872- 1890) DEL VERISMO ITALIANO	232
<i>Denise Bruno, Giuseppe Canzoneri, Antonio Di Silvestro, Daria Spampinato, Alessandro Zammataro</i>	
UN PROGETTO DI EDIZIONE DIGITALE <i>IMAGE-BASED</i> DELLE <i>MERAVIGLIE D'ORIENTE</i> NEL MS COTTON VITELLIUS A.XV	240
<i>Andreea Mihaela Toma</i>	
UNA PROPOSTA DI CODIFICA IN XML/MEI PER TESTI MUSICALI AUTOGRAFI DI VINCENZO BELLINI	246
<i>Laura Mazzagufò</i>	
VERISMO DIGITALE. PER UN'EDIZIONE DIGITALE COMMENTATA DELLE OPERE DI VERGA, CAPUANA, DE ROBERTO	252
<i>Liborio Pietro Barbarino, Elisa Conti, Christian D'Agata, Miryam Grasso, Ninna Maria Lucia Martines, Eliana Vitale</i>	
VERSO L'HYPEREDIZIONE. LO SVILUPPO DI PIRANDELLO NAZIONALE TRA DIDATTICA E RICERCA	260
<i>Milena Giuffrida, Christian D'Agata, Giulia Cacciatore, Fabrizio Lo Presti</i>	
VOCI DALL'INFERNO: DANTE PER DIRE IL LAGER - DIGITALIZZARE E STUDIARE LE TESTIMONIANZE	267
<i>Angelo Mario Del Grosso, Marina Riccucci, Elvira Mercatanti</i>	
DIZIONARI E DIGITALIZZAZIONE DI BANCHE DATI	274
IL VIVER (VOCABOLARIO DELL'ITALIANO VERISTA)	275
<i>Gabriella Alfieri, Marco Biffi, Stephanie Cerruto, Giovanni Salucci</i>	
L'INFORMATIZZAZIONE DEL GDLI: RISULTATI, PROSPETTIVE, SFIDE FUTURE	281
<i>Eva Sassolini, Sebastiana Cucurullo, Marco Biffi</i>	
LA DIGITALIZZAZIONE DEL DIZIONARIO LATINO LANA 1978	287
<i>Francesca Michelone</i>	
LEXICAD: PIATTAFORMA LESSICOGRAFICA DIGITALE PER L'ITALIANO DELLE ORIGINI	293
<i>Salvatore Arcidiacono, Antonella Zammataro</i>	
STRATEGIE PER LA CREAZIONE E LA CONDIVISIONE DI UNA COLLEZIONE DIGITALE DI TESTI GRECO-LATINI	298
<i>Vincenzo Ortoleva, Maria Rosaria Petringa, Salvatore Cammisuli, Mariarosaria Villareale</i>	
THE CORR<SI>CA PROJECT: ENHANCING AND "QUERYING" THE CANIONI FAMILY CORRESPONDENCE	303
<i>Tiziana Pasciuto, Selenia Anastasi, Daniele Zolezzi, Simonetta Acacia, Giada D'Ippolito, Chiara Storace, Maria Tolaini</i>	
XML/TEI E DIZIONARI <i>BORN-DIGITAL</i> : UNA PROPOSTA PER LE RISORSE LESSICOGRAFICHE DELLA RETE LEXICAD/PLUTO	310
<i>Giuseppe Leonardo Zappalà</i>	
ANALISI COMPUTAZIONALE, INTELLIGENZA ARTIFICIALE E LINGUISTICA	318
ANALISI COMPUTAZIONALE DEI REPORT DI SOSTENIBILITÀ: LA VAGHEZZA COME STRATEGIA DI GREENWASHING	319
<i>Erica Cutuli</i>	
ANALISI STILOMETRICA APPLICATA ALLE CAPACITÀ EMULATIVE DI GPT-4	325
<i>Marco De Cristofaro, Mariangela Giglio</i>	
C'È UN TESTO IN QUESTA CHAT? INTELLIGENZA ARTIFICIALE E COOPERAZIONE INTERPRETATIVA	332
<i>Daniele Silvi</i>	

GENERE E GEOPOLITICA NELLE DISCIPLINE UMANISTICHE DIGITALI IN ITALIA. LE CONFERENZE AIUCD (2012-2023)	
<i>Selenia Anastasi</i>	336
GLI LLM COME LETTORI MODELLO ARTIFICIALI	342
<i>Fabio Ciotti</i>	
I SIMILI SI ATTRAGGONO. LA VALUTAZIONE LETTERARIA SULLE PIATTAFORME DI DIGITAL SOCIAL READING	348
<i>Gabriele Vezzani, Simone Rebora, Massimo Salgaro</i>	
IL <i>DISTANT READING</i> È L'ORNITORINCO	354
<i>Pietro Mazzarisi</i>	
L'IMPIEGO DELL'INTELLIGENZA ARTIFICIALE PER LA RICOSTITUZIONE DELLE AGGREGAZIONI ARCHIVISTICHE E L'ARRICCHIMENTO DEI METADATI NEGLI ARCHIVI DIGITALI	361
<i>Stefano Allegrezza</i>	
MACCHINE PER LEGGERE: LA TEXT ANALYSIS COME STRUMENTO PER IMPARARE A LEGGERE I CLASSICI DELLA NARRATIVA... E AD AMARLI	367
<i>Fabio Ciotti</i>	
PRESERVARE LA DIVERSITÀ NELL'ERA DELL'INTELLIGENZA ARTIFICIALE: IL DILEMMA ETICO DI BIAS E DISCRIMINAZIONI NEGLI ALGORITMI	371
<i>Gianluca Pavani</i>	
<i>QUI PRO QUO?</i> DATI TESTUALI E STRUMENTI PER LA RISOLUZIONE DI COREFERENZE IN LATINO	377
<i>Roberta Grazia Leotta, Eleonora Delfino</i>	
STRUMENTI DIGITALI PER LA TRASCRIZIONE E LA LEMMATIZZAZIONE DI TESTI IN ITALIANO ANTICO	382
<i>Emiliano Degl'Innocenti, Alessia Spadi, Federica Spinelli, Lucia Francalanci, Michela Perino, Irene Falini, Francesco Coradeschi, Francesco Pinna</i>	
TESTI ALLOGRAFICI: CONTATTI TRA LINGUE E SCRITTURE DEL MEDITERRANEO	388
<i>Antonio Pagliara, Federico Boschetti, Daniele Baglioni</i>	
THE DARK MIRROR OF ARTIFICIAL INTELLIGENCE: HOW AI AFFECTS CLIMATE CHANGE	393
<i>Mauro De Bari</i>	
UN SISTEMA DI CLASSIFICAZIONE AUTOMATICA DI IMMAGINI RELATIVE A MATERIALI LIBRARI ANTICHI E MODERNI	398
<i>Nicola Barbuti, Tommaso Caldarola</i>	
UNCOVERING THE SPREAD OF LEXICAL INNOVATION IN ITALIAN TWEETS	405
<i>Greta Franzini, Paolo Brasolin, Stefania Spina</i>	
ORGANIZZAZIONE DELLA CONOSCENZA CON SEMANTIC WEB	410
AFFINARE IL CONTESTO: ESTRAZIONE DI INFORMAZIONI STRUTTURATE PER L'ARRICCHIMENTO DEI CONTESTI ARCHIVISTICI	411
<i>Lucia Giagnolini, Paolo Bonora, Francesca Tomasi</i>	
CLEF 2.0. SOLUZIONI PER LA CATALOGAZIONE NATIVA LINKED DATA DEL PATRIMONIO DIGITALE CULTURALE ITALIANO	417
<i>Marilena Daquino, Laurent Fintoni, Sebastiano Giacomini, Francesca Tomasi</i>	
L'ONTOLOGIA BIGRAFO: VERSO UN MODELLO SEMANTICO PER L'OPERA DI FRANCO FORTINI	423
<i>Laura Antonietti, Emilio M. Sanfilippo, Emmanuela Carbé</i>	
LOST IN DATIFICATION? THE JOURNEY OF DATA FROM THE PRIMARY SOURCE TO THE FINAL INTERPRETATION	429
<i>Enrica Bruno, Sofia Baroncini, Francesca Tomasi</i>	
ORBIS DIOECESIVM. CREATING AUTHORITY DATA ON THE LEGAL-HISTORICAL CHANGES OF CATHOLIC DIOCESES	435
<i>Benedetta Albani, Rowan Dorin, Yohan Park</i>	
PER L'INTEROPERABILITÀ E LA SOSTENIBILITÀ DELLE RISORSE DIGITALI DANTESCHE: IL PROGETTO LIDA	441
<i>Cesare Concordia, Gaia Tomazzoli, Nicola Aloia, Carlo Meghini, Luca Trupiano</i>	
PER UN'ANALISI DEI PERSONAGGI TRA LETTERATURA, FILOSOFIA E ONTOLOGIA APPLICATA	448
<i>Emilio M. Sanfilippo, Gaia Tomazzoli, Michele Paolini Paoletti, Jansan Favazzo, Roberta Ferrario</i>	
REPRESENTING TEXTS AS LOD: A SYSTEMATIC LITERATURE REVIEW	455
<i>Michela Bandini, Valeria Quochi</i>	
STRUCTURING AUTHENTICITY ASSESSMENTS ON HISTORICAL DOCUMENTS USING LLMs	463
<i>Andrea Schimmenti, Valentina Pasqual, Francesca Tomasi, Fabio Vitali, Marieke van Erp</i>	
VALORIZZAZIONE DI REGISTRAZIONI STORICHE DI CANTO LIRICO NEL WEB SEMANTICO	469
<i>Marcello Ranieri, Angelo Pompilio</i>	

PUBLIC HISTORY E ARCHEOLOGIA DIGITALE	473
ARCHEOLOGIA E RILIEVO 3D: UNA RIFLESSIONE SULLE METODOLOGIE. DUE CASI STUDIO DI AREA MEDITERRANEA	474
<i>Graziana D'Agostino, Pietro Maria Militello</i>	
DAL CATASTO BORBONICO ALLA GENOMICA. PIATTAFORME DIGITALI E INTERDISCIPLINARITÀ: IL PROGETTO “WE ARE WHAT THEY WERE” DI RIPOSTO	481
<i>Salvatore Spina</i>	
DRAWING HISTORY FROM THE CODICE PELAVICINO DOCUMENTS: GRAPH VISUALIZATION FOR HUMAN RESEARCHERS	489
<i>Natthida Wiwatwicha</i>	
GESTIONE INFORMATICA DELLA DOCUMENTAZIONE ARCHEOLOGICA "MINORE". METODOLOGIE E APPLICAZIONI NELL'AMBITO DEL PROGETTO STORAGE	494
<i>Marianna Figuera, Erica Platania</i>	
GESTIRE L'IMMATERIALE. CONTESTI SENSORIALI A SERVIZIO DEL PATRIMONIO ARCHEOLOGICO	501
<i>Serena D'Amico</i>	
GODSCAPES: MODELING SECOND MILLENNIUM BCE POLYTHEISMS IN THE EASTERN MEDITERRANEAN THANKS TO THE STORAGE PROJECT	507
<i>Nicola Laneri, Marianna Nicolosi-Asmundo, Daniele Francesco Santamaria, Chiara Pappalardo</i>	
IL PROGETTO STORAGE: DAI DATI AL WEB	512
<i>Simone Faro, Pietro Maria Militello, Marianna Nicolosi-Asmundo</i>	
LA MODELLAZIONE 3D AL SERVIZIO DELL'ARCHEOLOGIA: NUOVE PROSPETTIVE PER L'APPLICAZIONE AD EDIFICI MULTIPIANO DI ETÀ PROTOSTORICA	517
<i>Dario Puglisi, Marco Chiricallo</i>	
MLS CON SENSORE LIDAR APPLE PER LO SCAVO ARCHEOLOGICO: APPLICAZIONI PRATICHE	523
<i>Luigi M. Calì, Antonello Fino, Gian Michele Gerogiannis</i>	
ODONIMI D'ITALIA E DIGITAL PUBLIC HISTORY: LE PROBLEMATICHE DI UNA SCHEDATURA PARTECIPATA	528
<i>Enrica Salvatori, Vittore Casarosa, Riccardo Chiari</i>	
OPENSTREETMAP: UNO STRUMENTO E UNO SPAZIO PER LA DIGITAL PUBLIC HISTORY?	535
<i>Camilla Zucchi</i>	
PRESERVARE LA MEMORIA: IL PROGETTO STORAGE E L'ARCHIVIO DELL'EX ISTITUTO DI ARCHEOLOGIA	539
<i>Giovanni Fragalà, Pietro Maria Militello</i>	
UN ATLANTE DIGITALE PER LA STORIA MARITTIMA DEL REGNO DI SARDEGNA	545
<i>Giampaolo Salice</i>	
UN FUTURO PER LA MEMORIA. STRUMENTI, MODELLI E SINERGIE PER L'INTEGRAZIONE DEI DATI NEL PORTALE DELLE FONTI PER LA STORIA DELLA REPUBBLICA ITALIANA	549
<i>Michela Tardella, Roberta Maggi, Giorgia Lodi, Riccardo Albertoni, Herbert Natta, Gianluca Rossi, Tiziana Pasciuto, Paola Ciandrini, Luca Sinopoli, Maria Teresa Artese, Isabella Gagliardi, Eleonora Lattanzi, Sara Ventroni, Elisa Tizzoni, Alessandro Russo, Margherita Porena</i>	
INFRASTRUTTURE PER LA RICERCA UMANISTICA	555
COPHI EDITOR: FROM GREEKSCHOOLS TO THE PROJECTS MULTIVERSE	556
<i>Simone Zenzaro</i>	
DIGITAL HUMANITIES AND HERITAGE SCIENCE: MOVING FROM LANDSCAPING TO A DYNAMIC RESEARCH OBSERVATORY IN AN OPEN SCIENCE CLOUD	559
<i>Roberta Bianca Luzietti, Alessia Spadi, Nicola Giampietro, Giacomo Mancuso, Alessandra Caravale, Antonio D'Eredità, Marta Caradonna, Paola Moscati, Valeria Quochi, Monica Monachini, Emiliano Degl'Innocenti</i>	
FUNZIONI E SOSTENIBILITÀ DI UNA PIATTAFORMA DIGITALE PER LE LINGUE ARCAICHE	566
<i>Michele Mallia, Riccardo Del Gratta, Valeria Quochi</i>	
INFRASTRUTTURE DI RICERCA COME STRUMENTI DI “INTERCULTURALITÀ DIGITALE”	572
<i>Salvatore Cristofaro, Vittoria Fabiani, Cristina Marras, Enrico Pasini, Pietro Sichera, Mingyang Yu</i>	
MATERIALI DIDATTICI COME OGGETTI DIGITALI FAIR: UNA METODOLOGIA CONDIVISA PER LA FORMAZIONE IN H2IOSC	577
<i>Giulia Pedonese, Francesca Frontini, Roberta Ottaviani, Federico Boschetti, Alessia Spadi, Lucia Francalanci, Alessia Scognamiglio, Pietro Restaneo, Antonina Chaban, Jana Striova, Laura Benassi</i>	
RETHINKING SCHOLARLY DIGITAL OBJECTS AS CULTURAL HERITAGE: THE KNOT PROJECT	582
<i>Laurent Fintoni</i>	

THE ATLAS: A KNOWLEDGE GRAPH OF DIGITAL SCHOLARLY RESEARCH ON ITALIAN CULTURAL HERITAGE	588
<i>Marilena Daquino, Alessia Bardi, Marina Buzzoni, Riccardo Del Gratta, Angelo Mario Del Grosso, Franz Fischer, Francesca Tomasi, Roberto Rosselli Del Turco</i>	
INDICE DEGLI AUTORI	593

The Corr<si>Ca Project: enhancing and “querying” the Canioni family correspondence

Tiziana Pasciuto¹, Selenia Anastasi², Daniele Zolezzi³, Simonetta Acacia⁴,
Giada D’Ippolito⁵, Chiara Storace⁶, Maria Tolaini⁷

¹ CNR IMATI Genoa & Department of Modern Languages and Cultures, University of Genoa, Italy -
tiziana.pasciuto@edu.unige.it

² Department of Modern Languages and Cultures, University of Genoa, Italy &
Language Technology Group, Hamburg University, Germany - selenia.anastasi@edu.unige.it

³ Department of Modern Languages and Cultures, University of Genoa, Italy - daniele.zolezzi@edu.unige.it

⁴ Department of Modern Languages and Cultures, University of Genoa, Italy - simonetta.acacia@edu.unige.it

⁵ Department of Modern Languages and Cultures, University of Genoa, Italy - giadadippolito30@gmail.com

⁶ Department of Modern Languages and Cultures, University of Genoa, Italy - chiara.storace@edu.unige.it

⁷ Department of Modern Languages and Cultures, University of Genoa, Italy - maria.tolaini@edu.unige.it

ABSTRACT¹

The Corr<si>Ca project, undertaken by professors and PhD students from the University of Genoa's PhD program in Digital Humanities, involves the digitization and the enhancement of a family correspondence comprising 270 letters. This correspondence belongs to the Canioni family, originally from Olmi-Cappella, an inland village in the northern part of Corsica. The letters span the years 1882-1918, featuring contributions from both male and female writers with varying levels of literacy in the two languages of correspondence, namely Italian and French.

This paper outlines the ongoing development of (i) a blog designed to enrich the document corpus, engaging not only general users but also secondary school students, and (ii) an ontology for the future querying of this archival material. The ontology aims to extract valuable references and explore the correlations between language, gender, and the authors' origins. The objective is to delve into archival, philological, and sociolinguistic aspects related to the reference population.

KEYWORDS

Dissemination; Correspondence; Corsica; Blog; Ontology.

1. INTRODUCTION

The Canioni family correspondence, accessed with the permission of the family descendants, presents an interesting case of familial correspondence spanning the late 19th to early 20th centuries, specifically between 1882 and 1918. The 270 letters exchanged among Canioni family members and their entourage, primarily based in the Haute Corse village of Olmi-Cappella, reveal a linguistic and cultural transition from Italian to French. Written by three generations of Canioni individuals, including both semi-literate men and women, the letters provide linguistic insights and historical perspectives. Beyond linguistic shifts, the correspondence offers valuable details on daily life, material culture, local economy, politics, and trade between Corsica and the mainland, making it a significant source for linguistic and historical studies.

With the aim of enhancing this interesting corpus, valuable both from a linguistic and cultural point of view (since the chronological range of correspondence coincides with the end of the First World War), ongoing efforts include the creation of an accessible blog and the development of an ontology. The details will be discussed in the following paper's sections. Simultaneously, another sub-working group within the project is dedicated to the digitization of letters and transcription in XML/TEI format².

2. ENHANCING THE PROJECT: DISSEMINATION. Corr<si>Ca blog

In order to disseminate the project to as wide an audience as possible, have created a blog, providing a platform to share all the gathered material related to the Canioni family. Although opinions on the matter may vary, the term ‘weblog’ is widely credited to Jorn Barger [7]. ‘Blog’ is employed as an English onomatopoeic term, evoking the act of vomiting, and

¹ Author's statement - the author of the present paper contributed to the writing of this article as follows: T. Pasciuto (1. Introduction, 2. Corr<si>Ca blog, 3. Corr<si>Ca Ontology, 4. Conclusions and future works), S. Anastasi (3. Corr<si>Ca Ontology), D. Zolezzi (2. Corr<si>Ca blog), S. Acacia (2. Corr<si>Ca blog), G. D’Ippolito (3. Corr<si>Ca Ontology), C. Storace (2. Educational section), M. Tolaini (2. Educational section).

² For further information on this aspect, please refer to the contribution of Giaufret *et al.*, «Il progetto Corr<si>Ca: edizione digitale della corrispondenza Canioni», in the present volume.

infusing an additional meaning to the term, depicting it as a platform where one can freely regurgitate and share their content [6]. In recent years, several digital platforms emerged with the goal of enhancing cultural heritage and narrating historical and cultural events, both on a local and global scale, to raise awareness within the worldwide community. These initiatives aim to disseminate knowledge and foster appreciation for cultural richness through digital tools accessible to a broad audience.

For the Corr<si>Ca project, various platforms served as inspiration, for example the WarSampo Portal³ [8-9], which reassembles life stories of soldiers fallen in World War II, the Italian Bellini Digital Correspondence⁴ [2-4, 12], the digital edition of the letters of the Sicilian composer Vincenzo Bellini, and the platform MythLod⁵ [10, 11], born for the formal representation of cultural metadata.

Corr<si>Ca blog is powered by WordPress, a highly customizable Content Management System (CMS) widely used for creating and managing websites. This decision was motivated by our commitment to make the project's website as inclusive and open as possible, encouraging interaction from all visitors through the ability to comment on content. The flexibility of WordPress enables us to tailor the appearance and functionality of our site to effectively meet the needs of our community and visitors [13]. The user-friendly interface facilitates navigation and interaction, fostering the development of meaningful connections, promoting the sharing of knowledge, and nurturing a constructive dialogue to cultivate an active community.

The structure and the contents of the Corr<si>Ca blog are accessible in two languages (French and Italian), while the letters are available only in the language in which they are written, in order to respect the original language and avoid providing translations that could be artificial, as they would be unable to reproduce many features of the original such as words hypo- or hypersegmentation.

A minimalist and straightforward style has been upheld to highlight the authentic content of the platform, avoiding excessive visual distractions. The selected theme is 'Why Minimalist Blogger X', crafted to impart an elegant look to the blog. It ensures full responsiveness and optimization for search engines, making the content more attractive to search engines and easily locatable⁶.

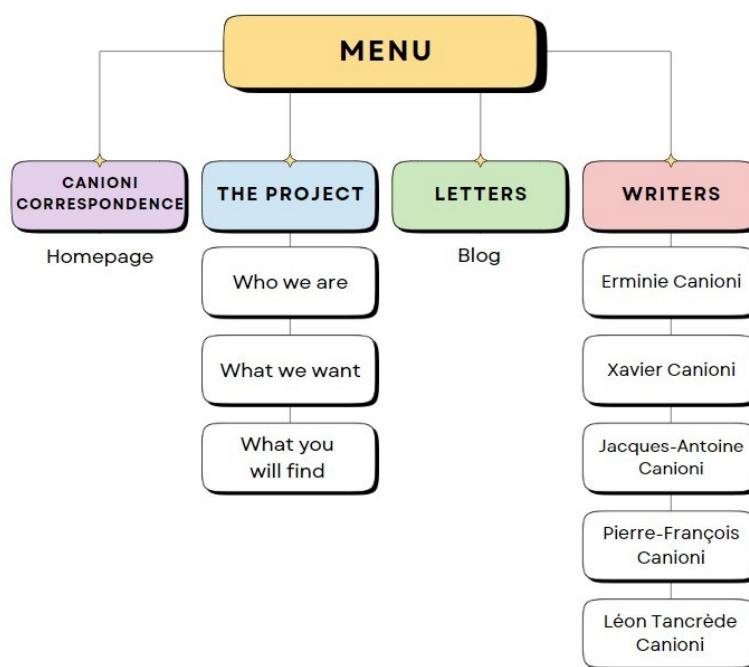


Figure 1. Graphic scheme of the main menu items of the Corr<si>Ca blog, at the present time

The goal of the project, aimed at bringing the epistolary heritage of the Canioni family back to life within the blog's virtual context, led us to structure the site to explain not only the objectives of the project and their creators, but also to provide

³ <https://www.sotasampo.fi/en/>

⁴ <http://belliniconrespondence.cnr.it/>; <http://licodemo.ilc.cnr.it/demo/bellini/build/>

⁵ <https://dharc-org.github.io/mythlod/>

⁶ Minimalist-blogger_link

access to all documents in the corpus (see Fig. 1). The site supplies two different access modes for users to explore the letters' transcriptions, which are explained below.

The first option is represented by the “Letters” section (marked with the term *Lettere* in Italian, and *Lettres* in French), where users can peruse the transcribed letters (at the moment, only a sample, see Fig. 2a). The letters are labeled with the names of the sender and recipient, for example, ‘*Da Xavier a Jacques Antoine*’ (*From Xavier to Jacques Antoine*), to clearly indicate to readers which family members the letter pertains to. The letters have been published on the blog by adjusting the post's publication date to the actual date on which the letter was sent (e.g., February 9, 1884), establishing a link with the present, where each letter is treated as a post on social networks. The letters feature a **diplomatic transcription of the text**, where only word segmentation, punctuation and use of capital letters have been modified to make the letters easier to read, a **scanned version** of the original document available for consultation as you progress through the reading, and a **geographical map** showing where the letter was written and received, and other places mentioned in the text. The points of interest (POI) offer additional information about the protagonists, not only pinpointing the location but also detailing their activities at that moment and specifying their degree of relationship.

Regarding the visualization of the places mentioned in the letters, in addition to the map for each document, a general map for the entire corpus is embedded in the introductory section on the Canioni correspondence. Here users can filter the places based on data deduced from the transcription, in particular fields such as author, recipient, date, type of place (origin, destination, cited), etc.

The second method for accessing information is through a search bar that can retrieve letters containing keywords entered by visitors, enabling them to quickly find content based on their interests.

The blog also includes a section featuring profiles of the main writers and recipients of the letters (see Fig. 2b). Each profile provides a brief biography and, where available, a picture of the person being discussed, allowing users to put a face to the stories and empathize with them. Users can comment on both the letters and the profiles of the authors, creating an opportunity for interaction and discussion. Furthermore, it is planned to add an educational section, which will be presented in the next paragraph.

The main objective is to build a community of practice to generate new, organized, and high-quality knowledge [16] around the themes discussed in the letters and the historical period they refer to.



Figure 2. a) A letter of the Canioni family, featured in the “Letters” section of the site. Besides the digitalized version and the text transcription, valuable information (sender and recipient, date, and place) are provided.

b) Short biography of Erminie, one of the writers of the Canioni corpus.

Educational section

Besides offering an inclusive and engaging website interface that allows all kinds of users to enjoy the contents, it is intended to create a specific educational section for students. Thanks to the variety of topics covered, the Canioni correspondence constitutes an interesting case of epistolary exchange that can be used for linguistic, socio-cultural, economic, and historical studies. For these reasons, we believe that not only experts and enthusiasts can benefit from the dissemination of these letters but also secondary school students. The information reported by the numerous letters offers insights on different school subjects, such as Foreign Language Learning, History, Literature, and Active Citizenship. We present a proposal for the educational section to be included on the website.

To formulate it, we carried out a short analysis of existing sites dedicated to digitizing and disseminating correspondence⁷. For this purpose, we defined a set of observation criteria, starting with the main question: is there a section dedicated to education? If it was present, we also observed the focus of this section, the target audience chosen, the interactivity of the web page, what resources were made available both in terms of materials provided, exercises or teaching activities proposed and, finally, the strengths and weaknesses of the section.

This brief analysis showed that only two of these websites had a dedicated educational section, thus highlighting the originality and the need for more such contributions. In particular, the Pirandello Nazionale Website, a digital project aimed to publish the integral corpus of Pirandello's works on-line, presents a captivating educational section. It provides digital resources that include educational paths characterized by mixed media materials whose exploration is facilitated by a well-designed legend. Given the brief analysis results, and considering the highlighted valuable examples, we decided to create a rich and multimedia educational section illustrated by an index page that shows the main focuses of the Canioni Project. Each focus contains two subsections: one with in-depth materials and one with suggested educational activities.

Regarding the in-depth materials, our proposal includes open-source links for further exploration of the themes, as well as videos and audio materials. These materials allow students to explore the topics independently and enable teachers to enhance and prepare their lessons. Multimedia elements are designed to meet the diverse educational needs of students.

Concerning the teaching activities, Canioni correspondence offers various points of historical and sociological interest that could be addressed by students through the analysis of the letters. For example, a design for the teaching unit could be to guide the exploration and the in-depth study of topics discussed in various letters across different historical periods, such as weights and measures, agricultural practices, epidemics, and medical treatments. The teaching unit will consist of two parts. Initially, introductory materials will be provided to contextualize the specific themes. Teachers will use these materials to design classroom activities or to guide students in their autonomous use of the blog. The second segment of the teaching unit involves learners using keywords in the blog's search bar to retrieve information on the chosen topics. Next, students will create PowerPoint presentations to expose their understanding of the evolution of these topics.

The value of the Canioni correspondence does not end in its being historical evidence of the cultural and linguistic transition but also encompasses its being an interdisciplinary starting point for reflections at the scholastic level, at the academic one.

3. QUERYING THE PROJECT: THE CORR<SI>CA ONTOLOGY

A crucial part of the Corr<si>Ca project involves the deep encoding of the texts according to the XML/TEI standard, followed by the representation and organization of data and metadata within an ontological model, named **OntoCorr<si>Ca** (see Fig. 3). This formalization aims to enable complex queries on the database from the perspective of interactions among the actors involved in the exchanges, also extracting useful references to the places where these exchanges occurred, and the relationship between language, gender, and authors' origin, for those who are interested in investigating the archival, philological, and sociolinguistic aspects of the population. This section aims to describe, at a high level, the ontological schema created specifically for the management of the information related to the corpus. Future works dedicated to modeling sociolinguistics characteristics such as the development of written proficiency first in Italian, then in French, will also be discussed within a multilingual repertoire context where the Corsican language was the only spoken one utilized. This section of the project represents an effort in line with the European enterprise of digitizing the rich cultural heritage of the UE⁸, and the subsequent creation of interoperable web resources that may provide in the future an in-depth understanding of the relationships between places, cultures, and languages whose history and tradition are deeply connected.

An initial phase of the project was dedicated to the study and selection of foundational ontological resources existing for the modeling of the archival domain and, in particular, for the representation of epistolary correspondence. In recent years, proposals concerning the semantic organization of cultural heritage have proliferated, many of which are based on the CIDOC-CRM model (Conceptual Reference Model)⁹ [5], which has emerged as an international standard since 2006 for the controlled exchange of information on cultural heritage. Based on CIDOC-CRM, many ontological models have been developed to improve the semantic expressiveness of the typical features of epistolary correspondence and address specific issues previously overlooked. Nevertheless, we believe that this standard is not without problems concerning the terminology used and concepts, such as the conceptual opacity of some classes as discussed in [14-15] and the indiscriminate use of the over-extended masculine for classes such as E-22 – Man-Made Object, E24 – Physical Man-Made

⁷ MythLOD; Bellini Digital Correspondence; WarSampo; Visual Correspondence; Claviuson on the web; Digitising experiences of migration; Early Modern Letters Online; Mapping the Republic of Letters; Storia digitale UniCA. Il portale della Storia Digitale dell'Università degli Studi di Cagliari; Transcribe Bentham; Edizione Nazionale dell'Opera Omnia di Luigi Pirandello.

⁸ <https://www.europarl.europa.eu/factsheets/en/sheet/64/1-agenda-digitale-europea/>

⁹ <https://www.cidoc-crm.org/>

Thing, E25 – Man-Made Feature, E71 – Man-Made Thing¹⁰. For this reason, it has been deliberately chosen to operate, at least in this initial phase, on a parallel path, without neglecting the purposes of alignment with other exemplary Italian and international models such as OntoBelliniLetters¹¹, MythLod, and WarSampo.

In the early stages of OntoCorr<si>Ca modeling, we explored other ontologies describing archival and bibliographic domains: this analysis will lead to the alignment of classes and properties between ontologies later on. Specifically, we examined bibliographic standards such as Functional Requirements for Bibliographic Records (FRBR)¹² and Bibliographic Framework (BIBFRAME)¹³, focusing on distinguishing between paper and digital documents, copies, and drafts. Additionally, since the heritage in question consists of a corpus of letters related to archival heritage, the ontology was modeled after the Records in Contexts Ontology (RiC-O)¹⁴, which has been available in a stable first version since December 2023.

XML-TEI elements	OntoCorr<si>Ca classes and properties
<body>	[C] LetterBody
<opener>	[subC] Opener
<closer>	[subC] Closer
<salute>	[subC] Salute
<postscript>	[subC] PostScript
<signed>	[subC] Signature
<date>	[C] ChronologicalDate
<placeName>	[C] Place
<persName>	[C] foaf:Person - [DP] hasName

Table 1. Example of mapping between XML/TEI elements (selected for labelling the transcriptions of Canioni letters corpus) and some of the classes [C], subclasses [subC] and datatype properties [DP] of the OntoCorr<si>Ca ontology.

To begin outlining the ontology classes and properties, XML/TEI elements, used to label transcriptions, have been considered (Table 1). For the modeling of data related to the corpus, the material has been represented according to three different conceptual levels mirroring possible planes of analysis.

The **first level** is the analysis of the **agents** (*foaf:Agent*, *foaf:Person*) and **places** (*Place*) involved in the interactions, with a particular focus on one hand on *name*, *gender*, and *birth* and *death dates* of the authors, and on the other hand the *source* and *destination places* of the letters.

A **second level** of analysis is aimed at capturing the internal organizational **structures of texts** in relation to the archival domain. Every *Document* is written in one or more languages (*Language*), has a unique *Identifier*, is characterized by one or more types of medium (*Matter*) and by a specific state of conservation (*StateOfConservation*). On this second level, an important ontological distinction has been made to distinguish the *Form* of the document, that can be *Paper* or *Digital*, and the *Status* of realization, such as a *Draft*, a *Copy* or a *Final* document, handwritten or not; the last three classes are in a *disjunctive* relationship. Each letter (*Letter*), which is a *type* of document (*hasType*) may have one or more *authors* (that often – but not always – in letters coincide with the *sender*) and *recipients*, who in turn may contain references to other entities such as people, organizations, places, works, events, and temporal ranges. A subdivision has been made considering the internal organization of the text (*isStructuredIn*): in fact, each letter is annotated according to the XML/TEI standard, taking into account general structural elements of the body of the text (*LetterBody*), such as opening (*Opener*) and closing phrases (*Closer*), postscripts (*PostScript*), author’s signatures (*Signature*), and greetings (*Salute*). Other elements of the body of the text that belong to the original copy and that are also intended to be captured through the digital version are the presence of deletions, corrections, ambiguous or illegible graphic elements, marginal notes, and the subdivision of the text into paragraphs.

The final but important **third level** of modeling concerns the **textual and linguistic information**, with particular reference to the terminological, morphological, syntactic, and grammatical characteristics.

¹⁰ For the latest version of the classes mentioned in CIDOC-CRM: <https://www.cidoc-crm.org/Entity/e24-physical-man-made-thing/version-6.2.1>.

¹¹ <http://bellininrete.istc.cnr.it/OntoBelliniLetters.html/>

¹² https://www.ifla.org/files/assets/cataloguing/frbr/frbr_2008.pdf

¹³ <https://www.loc.gov/bibframe/>

¹⁴ https://github.com/ICA-EGAD/RiC-O/tree/master/ontology/current-version/HTML_view

Regarding the ontological editor, the first part of the work was developed on Protégé, while for the linguistic domain the possibility of integrating the Tedi software¹⁵, designed for the construction of multilingual ontoterminologies, will be evaluated. Tedi allows exporting ontoterminologies in RDF/OWL format, permitting the importation of the ontoterminology into Protégé.

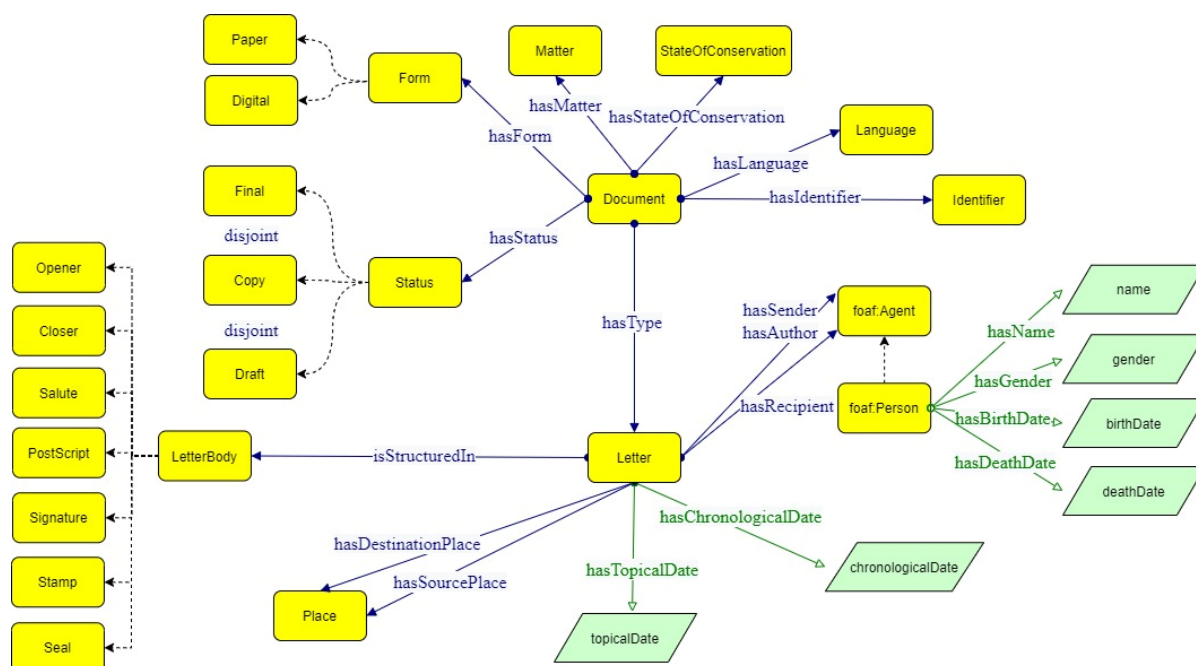


Figure 3. Graphical representation of the ontological model *OntoCorr<si>Ca*. The yellow boxes and green parallelograms indicate respectively the classes and the datatypes of the ontology, while the blue and green lines connecting classes and datatypes represent respectively object and datatype properties.

4. CONCLUSIONS AND FUTURE WORKS

In this paper, we have outlined the choices and methodologies employed in creating a blog dedicated to the study and enhancement of the archival material of the Canioni family’s correspondence. Additionally, we have illustrated the preliminary formalization of an ontological model designed for the querying of data.

The project is currently ongoing, and every initiative illustrated in the paper is at a prototype level. Letters and transcriptions, customizing the enjoyment of the heritage for both general users, secondary school students and possibly university students, will further enrich the blog. In addition, the ontological module will be extended with more detailed definitions related to the physical and digital versions of the letters. Finally, we plan to validate the model by instantiating the data provided by the project and enabling semantic querying of the corpus.

For the representation of linguistic features, Ontolex-Lemon will be considered. Ontolex-Lemon is a standard model for the multilingual, lexical, syntactic, and semantic modeling of terms and allows, for example, the decomposition of compound nouns, highlighting variations considering contextual meanings, and the use of metadata that provide descriptive elements on how, why, by whom, etc. [1]. Note that utilizing this ontology requires a manual intervention in transcribing letters' content into the XML/TEI format: specifically, it involves the targeted labelling of lemmas using the <choice> TEI element. Consequently, this process will probably be executed on a restricted set of letters, chosen for their linguistic significance.

5. ACKNOWLEDGEMENTS

We express our gratitude to the descendants of the Canioni family, who graciously provided us with access to the original correspondence that served as the focal point of this study, and to Santu Massiani, precious guardian of local history.

REFERENCES

[1] Cimiano, Philipp, Christian Chiarcos, John P. McCrae, and Jorge Gracia. *Linguistic Linked Data: Representation, Generation and Applications*. New York Inc: Springer, 2020. <https://doi.org/10.1007/978-3-030-30225-2>.

¹⁵ <http://ontoterminology.com/tedi/>

- [2] Del Grosso, Angelo Mario, Erica Capizzi, Salvatore Cristofaro, Maria R. De Luca, Emiliano Giovannetti, Simone Marchi, Graziella Seminara, and Daria Spampinato. 'Bellini's Correspondence: A Digital Scholarly Edition for a Multimedia Museum'. *Umanistica Digitale* 3, no. 7 (2019): 23-47. <https://doi.org/10.6092/issn.2532-8816/9162>.
- [3] Del Grosso, Angelo Mario, and Daria Spampinato, eds. *Bellini Digital Correspondence*. CNR Edizioni, 2023. <http://bellinicorrespondence.cnr.it>.
- [4] Del Grosso, Angelo Mario, Daria Spampinato, Erica Capizzi, Salvatore Cristofaro, and Graziella Seminara. 'Promoting Bellini's Legacy and the Italian Opera by Scholarly Digital Editing His Own Correspondence'. In *What Is Text, Really? TEI and Beyond*. Graz, Austria, 2019.
- [5] Doerr, Martin, Christian-Emil Ore, and Stephen Stead. 'The CIDOC Conceptual Reference Model - A New Standard for Knowledge Sharing'. *Proceedings of the Tutorials, Posters, Panels and Industrial Contributions at the 26th International Conference on Conceptual Modelling, Auckland, New Zealand* 83 (5 November 2007): 51-56. <https://doi.org/10.13140/2.1.1420.6400>.
- [6] Garden, Mary. 'Defining Blog: A Fool's Errand or a Necessary Undertaking'. *Journalism* 13, no. 4 (2012): 483-499. <https://doi.org/10.1177/1464884911421700>.
- [7] Hookway, Nicholas. 'Entering the Blogosphere': Some Strategies for Using Blogs in Social Research'. *Qualitative Research* 8, no. 1 (2008): 91-113. <https://doi.org/10.1177/1468794107085298>.
- [8] Hyvönen, Eero. 'Digital Humanities on the Semantic Web: Sampo Model and Portal Series'. Edited by Christoph Schlieder. *Semantic Web* 14, no. 4 (April 2023): 729-744. <https://doi.org/10.3233/SW-223034>.
- [9] Koho, Mikko, Esko Ikkala, Petri Leskinen, Minna Tampe, Jouni Tuominen, and Eero Hyvönen. 'WarSampo Knowledge Graph: Finland in the Second World War as Linked Open Data'. Edited by Christoph Schlieder. *Semantic Web* 12, no. 2 (January 2021): 265-278. <https://doi.org/10.3233/SW-200392>.
- [10] Pasqual, Valentina, and Francesca Tomasi. 'Data Narratives with Linked Open Data, the Case of MythLOD Storytelling'. In *Digital Humanities 2023: Book of Abstracts*, edited by Walter Scholger, Georg Vogeler, Toma Tasovac, Anne Baillet, and Paul Helling, 283-284. Graz, Austria, 2023.
- [11] Pasqual, Valentina, and Francesca Tomasi. 'Linked Open Data per La Valorizzazione Di Collezioni Culturali: Il Dataset MythLOD.' *AIB Studi* 62, no. 1 (May 2022): 149-168. <https://doi.org/10.2426/aibstudi-13301>.
- [12] Pellino, Santa, Pietro Sichera, Angelo Mario Del Grosso, and Daria Spampinato. 'Dalla Codifica Alla Fruizione: L'edizione Digitale Bellini Digital. AIUCD 2022, Lecce 1-3 Giugno 2022 Correspondence'. In *Culture Digitali. Intersezioni: Filosofia, Arti, Media. AIUCD2022*, edited by Fabio Ciraci, Giulia Miglietta, and Carola Gatto, 163-168. Lecce, 2022.
- [13] Sabin-Wilson, Lisa, and Matt Mullenweg. *WordPress for Dummies*. John Wiley & Sons, 2011.
- [14] Sanfilippo, Emilio M. 'Ontologies for Information Entities: State of the Art and Open Challenges'. *Applied Ontology* 16, no. 2 (2021): 111-135.
- [15] Sanfilippo, Emilio M., Béatrice Markhoff, and Pittet Perrine. 'Ontological Analysis and Modularization of CIDOC-CRM'. In *Information Systems: Proceedings of the 11th International Conference (FOIS 2020)*, edited by Boyan Brodaric and Fabian Neuhaus, 107-121. IOS Press, 2020.
- [16] Wenger, Etienne. *Comunità Di Pratica. Apprendimento, Significato e Identità*. Milano: Raffaello Cortina Editore, 2007.