



Università  
di Catania

ASSOCIAZIONE per  
l'INFORMATICA UMANISTICA  
e la CULTURA DIGITALE

Associazione  
Italiana di  
Informatica  
Umanistica  
e Cultura  
Digitale  
AIUCD



Consiglio Nazionale  
delle Ricerche

# ME.TE. DIGITALI

MEDITERRANEO IN RETE TRA TESTI E CONTESTI

ATTI DEL XIII CONVEGNO ANNUALE

AIUCD 2024



28 - 30 MAGGIO

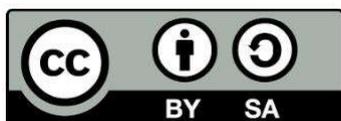
MONASTERO DEI BENEDETTINI

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[at]aiucd.org)  
*Please notify the publisher of any omissions or errors found: [aiucd.segreteria \[at\] aiucd.org](mailto:aiucd.segreteria[at]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 Tuscia), **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>	
<b>RELAZIONI DEI KEYNOTE SPEAKERS</b>	<b>VI</b>
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>	
<b>RELAZIONI DEGLI INVITED SPEAKERS</b>	<b>XXI</b>
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>	
<b>MEDITERRANEO TRA TESTI E CONTESTI</b>	<b>1</b>
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>	
<b>ARCHIVI E MUSEI DIGITALI PER IL PATRIMONIO CULTURALE</b>	<b>55</b>
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>	
<b>EDIZIONI SCIENTIFICHE DIGITALI</b>	<b>143</b>
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>	
<b>DIZIONARI E DIGITALIZZAZIONE DI BANCHE DATI</b>	<b>274</b>
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 Accacia, 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>	
<b>ANALISI COMPUTAZIONALE, INTELLIGENZA ARTIFICIALE E LINGUISTICA</b>	<b>318</b>
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>	
<b>ORGANIZZAZIONE DELLA CONOSCENZA CON SEMANTIC WEB</b>	<b>410</b>
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 DIOECESUM. 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>	

<b>PUBLIC HISTORY E ARCHEOLOGIA DIGITALE</b>	<b>473</b>
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>	
<b>INFRASTRUTTURE PER LA RICERCA UMANISTICA</b>	<b>555</b>
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

# Un futuro per la memoria. Strumenti, modelli e sinergie per l'integrazione dei dati nel Portale delle fonti per la storia della Repubblica italiana

Michela Tardella<sup>1</sup>, Roberta Maggi<sup>2</sup>, Giorgia Lodi<sup>3</sup>, Riccardo Albertoni<sup>4</sup>, Herbert Natta<sup>5</sup>, Gianluca Rossi<sup>6</sup>, Tiziana Pasciuto<sup>7</sup>, Paola Ciandrini<sup>8</sup>, Luca Sinopoli<sup>9</sup>, Maria Teresa Artese<sup>10</sup>, Isabella Gagliardi<sup>11</sup>, Eleonora Lattanzi<sup>12</sup>, Sara Ventroni<sup>13</sup>, Elisa Tizzoni<sup>14</sup>, Alessandro Russo<sup>15</sup>, Margherita Porena<sup>16</sup>

<sup>1</sup> CNR Istituto per il Lessico Intellettuale Europeo e Storia delle Idee, Italia – michela.tardella@cnr.it

<sup>2</sup> CNR Istituto di Matematica applicate e tecnologie informatiche, Italia – roberta.maggi@cnr.it

<sup>3</sup> CNR Istituto di scienze e tecnologie della cognizione, Italia – giorgia.lodi@istc.cnr.it

<sup>4</sup> CNR Istituto di Matematica applicate e tecnologie informatiche, Italia – riccardo.albertoni@ge.imati.cnr.it

<sup>5</sup> CNR Istituto di Matematica applicate e tecnologie informatiche, Italia – herbert.natta@ge.imati.cnr.it

<sup>6</sup> CNR Istituto di Matematica applicate e tecnologie informatiche, Italia – gianluca.rossi@ge.imati.cnr.it

<sup>7</sup> CNR Istituto di Matematica applicate e tecnologie informatiche, Italia – tiziana.pasciuto@ge.imati.cnr.it

<sup>8</sup> CNR Istituto di Matematica applicate e tecnologie informatiche, Italia – paola.ciandrini@ge.imati.cnr.it

<sup>9</sup> CNR Istituto di Matematica applicate e tecnologie informatiche, Italia – luca.sinopoli@ge.imati.cnr.it

<sup>10</sup> CNR Istituto di Matematica applicate e tecnologie informatiche, Italia – teresa@mi.imati.cnr.it

<sup>11</sup> CNR Istituto di Matematica applicate e tecnologie informatiche, Italia – gagliardi@mi.imati.cnr.it

<sup>12</sup> CNR Istituto per il Lessico Intellettuale Europeo e Storia delle Idee, Italia – eleonora.lattanzi@cnr.it

<sup>13</sup> CNR Istituto per il Lessico Intellettuale Europeo e Storia delle Idee, Italia – sara.ventroni@iliesi.cnr.it

<sup>14</sup> CNR Istituto per il Lessico Intellettuale Europeo e Storia delle Idee, Italia – elisa.tizzoni@gmail.com

<sup>15</sup> CNR Istituto di scienze e tecnologie della cognizione, Italia – alessandro.russo@cnr.it

<sup>16</sup> CNR Istituto di scienze e tecnologie della cognizione, Italia – margherita.porena@istc.cnr.it

## ABSTRACT

Il *Portale delle fonti per la storia della Repubblica italiana* nasce dalla volontà di offrire a un pubblico ampio e diversificato una raccolta di fonti per la storia politica e istituzionale nazionale relativa alla seconda metà del Novecento. L'intento è creare un'architettura software e un'infrastruttura dati in grado di favorire l'accesso effettivo al patrimonio culturale considerato nel progetto; un accesso il più possibile aperto, che agevoli la conoscenza dei documenti conservati nei complessi archivistici degli organi costituzionali e dagli apparati amministrativi dello Stato, integrati con quelli delle associazioni private che concorrono alla vita democratica del paese, in particolare le fonti prodotte dai partiti politici e dalle organizzazioni sindacali. Una sfida culturale e tecnologica importante, ricca di implicazioni civili, ma anche teoriche e metodologiche, che vede impegnati tre istituti del CNR in collaborazione con soggetti pubblici e fondazioni e istituti privati.

## PAROLE CHIAVE

Cultural heritage; Valorisation; Data Infrastructure; Linked Open Data.

## 1. INTRODUZIONE

Il *Portale delle fonti per la storia della Repubblica italiana* nasce dalla volontà di offrire a un pubblico ampio e diversificato una raccolta di fonti per la storia politica e istituzionale nazionale relativa alla seconda metà del Novecento. L'intento dei soggetti promotori dell'iniziativa è di creare un'architettura software e un'infrastruttura dati in grado di favorire l'accesso effettivo al patrimonio culturale oggetto del progetto; un accesso il più possibile aperto che agevoli la conoscenza dei documenti conservati nei complessi archivistici degli organi costituzionali e dagli apparati amministrativi dello Stato, integrati con quelli delle associazioni private che concorrono alla vita democratica del paese, in particolare le fonti prodotte dai partiti politici (art. 49 della Costituzione) e dalle organizzazioni sindacali (art. 39). Una sfida culturale e tecnologica importante, ricca di implicazioni civili, ma anche teoriche e metodologiche, che vede impegnati tre istituti del CNR in collaborazione con soggetti pubblici e fondazioni e istituti privati.

L'impegno teorico maggiore è stato dedicato all'elaborazione concettuale finalizzata a mettere a sistema risorse caratterizzate da una eterogeneità diffusa, per quanto riguarda sia le modalità e gli strumenti di descrizione delle risorse sia i formati e le tecnologie di trasmissione e condivisione dei dati. Sono ampiamente note, infatti, le specificità di formazione, gestione e conservazione di complessi archivistici prodotti dalle istituzioni pubbliche e private, nonché la vischiosità che caratterizza il rapporto esistente fra la vita istituzionale e la dimensione politico-sociale. Basta, ancora oggi, sfogliare la *Guida generale agli archivi di Stato*, curata da Claudio Pavone e Piero D'Angiolini o ancora il volume *Gli Archivi dei partiti politici* (che raccoglie gli atti dei seminari tenuti a Roma e Perugia del giugno-ottobre 1994) per avere un quadro

della complessità conservativa e descrittiva delle fonti a nostra disposizione [1, 8]. In questo orizzonte, il *Portale* ambisce a diventare uno spazio virtuale, un ecosistema nel quale arrivare ad una integrazione dei dati e dei sistemi, superando quel ‘particolarismo informativo’ che caratterizza l’eterogeneità e la frammentarietà delle informazioni, che si rispecchia anche nella proliferazione dei siti per la ricerca archivistica [3, 11]<sup>1</sup>. Per ottemperare a queste finalità, il nostro lavoro si è posizionato nella dimensione della sinergia, intendendo con questo termine un’azione coordinata e contemporanea di più elementi, suscettibile di numerose declinazioni: sinergia, appunto, tra i patrimoni conservati dai diversi soggetti coinvolti; sinergia di modelli concettuali per la descrizione e la rappresentazione del patrimonio culturale; sinergia di linguaggi e formati di rappresentazione e condivisione dei contenuti e dei contesti<sup>2</sup>. Data questa specifica connotazione del progetto, è chiaramente cruciale la contaminazione e l’integrazione tra esperti di vari ambiti disciplinari (archivisti, bibliotecari, informatici, ma anche filosofi, esperti di teoria dei linguaggi), che lavorando in sinergica armonia, mirano ad una integrazione di vari domini in cui nessuna delle competenze risulta essere ancella dell’altra. Per gli scopi descritti è stato dunque concepito un sistema che mira ad una piena interoperabilità, progettato per ingerire, gestire ed armonizzare grandi quantità di dati in formati e modelli eterogenei.

## 2. PER UN’INTEGRAZIONE MODULARE

La progettazione e sviluppo dell’infrastruttura digitale del Portale è stata orientata da scelte tecnologiche volte a ottimizzare sia l’interoperabilità del sistema con fonti dati esterne sia l’interazione di utenti e agenti software con il livello di presentazione, declinato in diverse soluzioni di interfaccia.

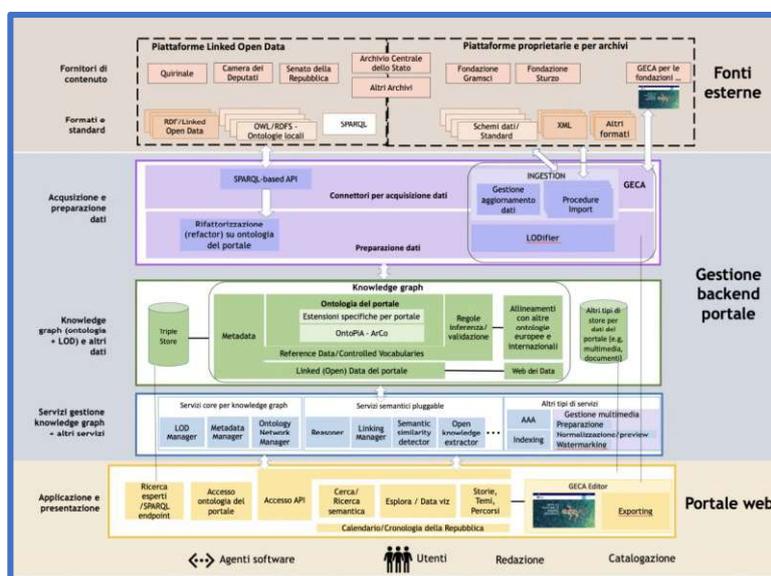


Figura 1. Schema architetturale del sistema con definizione dei livelli logici e delle componenti funzionali

Nella figura 1 è rappresentata l’architettura tecnica di alto livello del Portale, implementata sfruttando il modello architetturale a microservizi, che favorisce la possibilità di modificare e integrare modularmente le componenti open del sistema (la modularità, è utile precisare, caratterizza tutte le linee di attività del progetto e le metodologie utilizzate per il suo sviluppo, sia sul piano architetturale nel suo complesso, sia per la realizzazione delle ontologie). La parte centrale (in verde) costituisce la peculiarità di questa architettura che si struttura, oltre che intorno alle componenti software, anche attraverso il paradigma rappresentativo del knowledge graph. È opportuno soffermarsi sul doppio percorso di ingestione dei dati:

1. Per i portali istituzionali che dispongono di dati strutturati in formato Linked Open Data - ci riferiamo in particolare al portale dell’Archivio Storico della Presidenza della Repubblica e a quello dell’Archivio Storico della

<sup>1</sup> Sebbene particolare riguardo sia riservato al patrimonio archivistico, l’infrastruttura dedicata è aperta anche all’acquisizione e gestione di risorse bibliografiche e museali, garantendo la possibilità di trattare in modo trasversale i vari domini del patrimonio culturale.

<sup>2</sup> L’eterogeneità dei dati condivisi, in diversi formati, che variano da documenti testuali (.doc) e tabellari (.csv, .xls) e a dataset in XML e JSON strutturati in base a modelli definiti dagli standard di dominio e/o dalle specifiche dei sistemi software che li hanno originati, ha richiesto la definizione di specifiche *pipeline* di estrazione e trasformazione, finalizzate all’integrazione nel modello logico-semantic del Portale, basato su framework RDF e sul modello dei Linked Open Data.

Camera - si è proceduto con l'acquisizione dei dati esposti. Per quanto riguarda l'Archivio Storico del Senato, si è proceduto con la produzione di Linked Open Data (LOD) dei dati relativi alle persone, trasmessi al gruppo di lavoro in formato JSON e ricavati anche attraverso i LOD del Senato della Repubblica.

2. Per quanto riguarda invece i soggetti privati sono state predisposte le procedure di importazione di dati e di oggetti digitali, che hanno confermato la difformità delle situazioni che devono essere gestite: alcuni enti, infatti, hanno inviato i propri strumenti di ricerca in formato word, altri in excel accompagnato dal relativo pdf per fini di controllo, altri in XML (diversificati, in ragione della molteplicità dei in relazione alla diversità dei software di descrizione utilizzati o delle loro customizzazioni).

### 3. IL PERCORSO GECA

Le descrizioni e i documenti provenienti dalle istituzioni pubbliche e private che non dispongono di LOD (parte destra dello schema architetturale in Figura 1) vengono acquisiti attraverso il software GECA [6, 9]. Nato dall'evoluzione di GECA System (2004) e GECA RDC (2014) - strumenti per la catalogazione e la gestione di risorse bibliografiche - GECA è stato sviluppato come strumento destinato a differenti domini, quali il contesto museografico (tracciati ICCD, come schede F, OA, PST) e il contesto archivistico (complessi archivistici, unità archivistiche e documentarie, soggetti produttori e conservatori) proponendosi come collettore di descrizioni (cataloghi e strumenti di ricerca), nel rispetto dei relativi standard di settore (UNIMARC, MARC21, ISAD(G), EAD).

In GECA trovano piena attuazione tanto gli standard internazionali di dominio, quanto l'interoperabilità con i principali tracciati di interscambio, rappresentando quindi uno strumento utile sia alla granularità delle descrizioni previste dai diversi ambiti di applicazione, sia all'attivazione di relazioni cross dominio. Un esempio in questo senso è dato dalla possibilità fornita dal software di strutturare un albero archivistico (vd. Fig. 2) utilizzando non solo schede di descrizione per complessi archivistici e unità, ma anche schede dedicate a risorse bibliografiche o museali, integrate quindi con la descrizione archivistica [7].



Figura 2. L'albero archivistico di un fondo inserito in Geca, contenente sia risorse archivistiche sia bibliografiche

L'integrazione di GECA come componente modulare all'interno dell'architettura del sistema ha quindi l'obiettivo, da un lato, di fornire uno strumento utile all'inserimento nativo di nuove descrizioni, utilizzando tracciati specifici per i tre domini di riferimento, e, dall'altro, di importare dati garantendo sia la compatibilità con standard riconosciuti (quali ICAR-IMPORT, EAD, ecc.) sia la flessibilità necessaria a gestire l'eterogeneità delle fonti e dei formati di interscambio (XML, JSON, CSV, ecc.).

GECA offre inoltre un modulo di lodificazione che permette di esporre le risorse in catalogo attraverso il paradigma dei LOD, utilizzando ontologie di dominio di riferimento internazionale (quali Bibframe e RiC-O) e tenendo in considerazione le esperienze nazionali (quali SAN Lod, OAD ed ArCo). Il processo di lodificazione di GECA (modulo LODifier) si distingue per l'adozione di standard consolidati a livello internazionale e software a licenza libera, garantendo così l'assenza di dipendenze da software proprietari. Un esempio tangibile è l'utilizzo dello standard internazionale W3C R2RML come linguaggio per la definizione delle mappature dal modello relazionale, nativo in GECA 3, agli insiemi di dati RDF. Per l'archiviazione dell'RDF e l'interpretazione degli script R2RML, si fa affidamento sul software Open Link Virtuoso

Universal Server, il quale offre una soluzione per gestire questa trasformazione e la gestione dei dati in formato RDF. Il processo di lodificazione si raccorda con la/le ontologie realizzate/usate appositamente per il Progetto, fornendo sia le specificità dei singoli domini, sia la trasversalità delle diverse risorse del portale e permettendo in questo modo non solo l'esposizione in LOD dei dati, ma anche la loro integrazione con altre fonti già lodificate.

Nell'ottica di ottimizzare l'integrazione cross dominio, e di massimizzare le potenzialità offerte dalle tecnologie semantiche impiegate, è in corso di revisione il modello dati relativo alle voci di autorità, che rappresentano entità fondamentali per attivare relazioni all'interno delle risorse in catalogo. In particolare, la revisione riguarda principalmente la modellazione di un'entità agente, che rappresenta in modo univoco enti, persone, famiglie, tipizzata in funzione del tipo di relazione con le risorse (soggetto produttore, conservatore, autore, ecc.), e di altre entità (quali luoghi, soggetti) funzionali a esplicitare sia le connessioni interne al catalogo sia esterne (tramite rimandi a sistemi di rappresentazione quali VIAF, Geonames, Wikidata), massimizzando le possibilità offerte dai LOD in termini di arricchimento della conoscenza e recupero del contenuto informativo.

Tra le altre funzionalità offerte dal software, di particolare rilevanza per il Progetto è la possibilità di creare percorsi di esplorazione delle risorse su base tematica: una caratteristica che ha permesso di estendere l'ambito applicativo del modulo GECA dalla gestione del catalogo alla gestione dei contenuti previsti per il portale, permettendo di inserire e modificare prodotti redazionali e contenuti audio/video e di creare relazioni tra questi e le altre tipologie di risorse presenti nel sistema. La modularità e flessibilità di GECA ne permette l'utilizzo anche in contesti applicativi diversi dall'architettura del portale, rappresentando uno strumento utile sia per la creazione e gestione di basi dati di beni culturali, sia per la loro pubblicazione ed esplorazione.

#### 4. IL GRAFO DELLA CONOSCENZA DEL PORTALE

La parte sinistra dell'architettura di Figura 1 illustra gli strumenti e i servizi destinati all'acquisizione dei dati disponibili secondo il paradigma dei Linked Open Data, ovvero rappresentati attraverso standard del Web Semantico quali RDF. Da un punto di vista metodologico, si è avviata una prima analisi dei modelli dei dati (o ontologie) definite e usate dalle fonti nella pubblicazione dei loro LOD, con un particolare focus sulle fonti Camera dei Deputati, Senato della Repubblica, Presidenza della Repubblica e Archivio Centrale dello Stato. Dall'analisi è emerso un chiaro requisito di armonizzazione dei modelli dati esistenti, utile per il raggiungimento degli scopi specifici del portale. Sono poi stati raccolti requisiti mediante un'interazione diretta con esperti di dominio, facendo uso dello strumento delle domande di competenza (chiamate in inglese Competency question) ovvero domande che gli utenti potrebbero porre sui dati di loro interesse. Una buona modellazione infatti deve poter supportare interrogazioni, anche complesse, sui dati che siano capaci di rispondere a tali domande. Infine, si è svolta un'analisi dello stato dell'arte rispetto a ontologie esistenti a livello nazionale, più o meno vincolanti anche nel contesto dell'ecosistema di interoperabilità nazionale (nel caso degli archivi: SAN LOD<sup>3</sup> e OAD<sup>4</sup>, nel caso delle persone CPV<sup>5</sup> pubblicata nel catalogo nazionale della semantica schema.gov.it, Cultural-ON - Cultural Institute or Site ONtology [5]), rispetto a standard internazionali di riferimento (e.g., RiC-O per il settore archivistico, FRBR per il contesto bibliografico/lavori) e, infine, rispetto a cosiddetti ontology design pattern che potessero essere riutilizzati con successo rispetto alle esigenze di modellazione del portale (e.g., il pattern del ruolo che cambia nel tempo per agenti). Gli ontology design pattern (ODP) sono soluzioni riusabili usate per risolvere problemi ricorrenti nella modellazione ontologica. È stato provato che il riutilizzo di ODP migliora la qualità dell'ontologia, riduce l'arbitrarietà nel design dell'ontologia e favorisce così l'interoperabilità semantica dei dati [4, 10]. Tutte queste fasi rientrano nel contesto più ampio della metodologia eXtreme Design (XD) [9, 10] particolarmente efficace per lo sviluppo agile di ontologie o modelli dati. Dai requisiti raccolti è risultato tuttavia chiaro che quanto esistente poteva essere limitato per rispondere ad alcune esigenze del portale. Pertanto, si è proceduto con la progettazione di una vera e propria rete di moduli ontologici dove un approccio modulare è stato scelto in quanto particolarmente efficace in presenza di tipologie diverse di dati con, potenzialmente, esigenze di aggiornamento diverse. La modularizzazione infatti agevola la manutenzione complessiva in ambito software e di modellazione ontologica. La rete di ontologie sviluppate è mostrata in Figura 2. In particolare, sono stati creati otto moduli ontologici<sup>6</sup> che rappresentano le persone e le figure politiche, i loro incarichi, i loro mandati, le organizzazioni costituzionali, istituzionali e politiche, gli eventi pubblici di particolare rilevanza pubblica come l'elezione del presidente della repubblica, i bollettini dei vari rami del parlamento, i discorsi, gli atti e i diari storici nonché le principali informazioni sugli archivi (modulo in Figura 2), non più rappresentati come solo informazioni di catalogazione ma come

<sup>3</sup> <https://www.san.beniculturali.it/web/san/dati-san-lod>

<sup>4</sup> <https://labs.regesta.com/progettoReload/oad-ontology/>

<sup>5</sup> <https://w3id.org/italia/onto/CPV>

<sup>6</sup> Tutte le ontologie prodotte sono disponibili come risorse aperte su piattaforma github al seguente link <https://github.com/PortaleFontiRepubblica/assets/tree/main/ontologies>



Un ulteriore obiettivo del progetto consiste nella creazione di uno spazio dedicato alle scuole, con la realizzazione di percorsi didattici, strutturati a partire da documenti d'archivio, utilizzabili dagli insegnanti nel corso delle loro lezioni. La costruzione dei percorsi, realizzati attraverso le metodologie e le tecnologie riferibili alla didattica della storia, è supervisionata dal coordinamento dei consulenti storici.

## BIBLIOGRAFIA

- [1] Cacioli, Manuela. *Gli archivi dei partiti politici: atti dei seminari di Roma, 30 giugno 1994, e di Perugia, 25- 26 ottobre 1994*. Roma: Ministero per i beni culturali e ambientali, Ufficio centrale per i beni archivistici, 1996.
- [2] Carriero, Valentina A., Aldo Gangemi, Maria Letizia Mancinelli, Ludovica Marinucci, Andrea Giovanni Nuzzolese, Valentina Presutti, e Chiara Veninata. «ArCo: the Italian Cultural Heritage Knowledge Graph». In *The Semantic Web - {ISWC} 2019 - 18th International Semantic Web Conference, Auckland, New Zealand, October 26-30, 2019, Proceedings, Part {III}*, a cura di Chiara Ghidini, Olaf Hartig, Maria Maleshkova, Vojtěch Svátek, Isabel Cruz, Aidan Hogan, Jie Song, Maxime Lefrançois, e Fabien Gandon, 36–52. Lecture Notes in Computer Science. Auckland, New Zealand, 2019. [https://dx.doi.org/10.1007/978-3-030-30796-7\\_3](https://dx.doi.org/10.1007/978-3-030-30796-7_3).
- [3] Ciandrini, Paola, Eleonora Lattanzi, Roberta Maggi, e Michela Tardella. *Archivi e contaminazioni disciplinari: dai linguaggi ai modelli, dagli approcci alle tecniche*. Plurimi. Cnr-edizioni, in corso di stampa. <https://www.cnr.it/it/plurimi>.
- [4] Gangemi, Aldo, e Valentina Presutti. «Ontology design patterns». In *Handbook on ontologies*, a cura di Steffen Staab e Rudi Studer, 221–43. Berlin: Springer, 2009.
- [5] Lodi, Giorgia, Luigi Asprino, Andrea Giovanni Nuzzolese, Valentina Presutti, Aldo Gangemi, Diego Reforgiato Recupero, Chiara Veninata, e Annarita Orsini. «Semantic web for cultural heritage valorisation». In *Data Analytics in Digital Humanities*, 3–37. Springer, 2017.
- [6] Maggi, Roberta, Tiziana Pasciuto, Martina Mazzoleni, Maria Teresa Artese, Isabella Gagliardi, e Riccardo Albertoni. «GECA 3.0 - A new tool for cataloguing and enjoying cultural heritage». *La memoria digitale: forme del testo e organizzazione della conoscenza. XII Convegno Annuale AIUCD, Siena, 5-7/06/2023. Proceedings*, a cura di Emmanuela Carbé, Gabriele Lo Piccolo, Alessia Valenti, e Francesco Stella, 373-379, 2023. <https://doi.org/10.6092/unibo/amsacta/7721>.
- [7] Pasciuto, Tiziana, Riccardo Albertoni, Roberta Maggi, Maria Teresa Artese, Isabella Gagliardi, e Maurizio Gentilini. «Travelling Culture: Define, Implement, Enrich and Disseminate the Digital Cultural Heritage. The “DigitXL Project” Case Study». In *EDEN Research Workshop, Dubrovnik, 19-20/09/2022. Proceedings*, a cura di Josep M. Duart e Elena Trepule, 134–39. Dubrovnik, 2022.
- [8] Pavone, Claudio, e Piero D'Angiolini. *Guida generale degli archivi di Stato italiani*. Vol. 1–4. Roma: Ministero per i beni culturali e ambientali, Ufficio centrale per i beni archivistici, 1981.
- [9] Presutti, Valentina, Enrico Daga, Aldo Gangemi, e Eva Blomqvist. «eXtreme Design with Content Ontology Design Patterns». In *Proceedings of Workshop on Ontology Patterns*, a cura di Eva Blomqvist, Kurt Sandkuhl, Scharffe, Francois, e Svatek, Vojtek. CEUR Workshop Proceedings. CEUR-WS.org, 2009.
- [10] Presutti, Valentina, Giorgia Lodi, Andrea G. Nuzzolese, Aldo Gangemi, Silvio Peroni, e Luigi Aprino. «The role of ontology design patterns in linked data projects». In *Proceedings of Conceptual Modeling: 35th International Conference, ER 2016, Gifu, Japan, November 14-17, 2016*, a cura di Isabelle Comyn-Wattiau, Katsumi Tanaka, Il-Yeol Song, Shuichiro Yamamoto, e Motoshi Saeki, Vol. 9974. LNCS. Springer, 2016. [https://doi.org/10.1007/978-3-319-46397-1\\_9](https://doi.org/10.1007/978-3-319-46397-1_9).
- [11] Zanni Rosiello, Isabella. «Sul mestiere dell'archivista». In *L'archivistica sul confine. Scritti di Isabella Zanni Rosiello*, 371–88. Roma: Poligrafico Zecca dello Stato, 2000.