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 Mazzagufo; 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
Please notify the publisher of any omissions or errors found: 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 Rebora (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 Ciracì (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** (Universitè 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 Ciracì** (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 Rebora** (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
Antonio Di Silvestro, Daria Spampinato

RELAZIONI DEI KEYNOTE SPEAKERS VI

INFORMATICA UMANISTICA, INFOCRAZIA, E INTELLIGENZE ARTIFICALI VII
Giuseppe Savoca

THE MEDIEVAL MEDITERRANEAN IN…DATA? INTERPRETATION, CONJECTURE AND DIGITAL METHODS XIII
Tara L. Andrews

RELAZIONI DEGLI INVITED SPEAKERS XXI

RICOSTRUZIONE DEL TESTO E BANCHE DATI. LA FILOLOGIA DIGITALE ALLA PROVA DELL'ESEGESI ANTICA DELLA *COMMEDIA* XXII
Vittorio Celotto, Andrea Mazzucchi

IL MEDITERRANEO: UN MARE DI OPPORTUNITÀ E SFIDE XXX
Salvatore Capasso

MEDITERRANEO TRA TESTI E CONTESTI 1

COMBINING GENERATIVE AI AND ARCHAEOLOGY TO BUILD DATA-DRIVEN STORIES 2
Francesca Buscemi, Angelica Lo Duca

DIGITAL PRESERVATION E SOSTENIBILITÀ AMBIENTALE 9
Adele Gorini

IL PROGETTO DIGITALE SU ELISA CHIMENTI (LAI-ALEEF): LA PROBLEMATICITÀ DI UN PROFILO MEDITERRANEO TRA RETI E FRONTIERE 15
Ada Desideri, Bianca Vallarano

MEDITERRANEO VERDE FANTASTICO: UN REPERTORIO DIGITALE ATTRAVERSO LA COLLEZIONE BOTANICA DI ULISSE ALDROVANDI (1522-1605) 21
Sara Obbiso

PER LA DIGITALIZZAZIONE DEL PATRIMONIO LINGUISTICO E CULTURALE ITALIANO IN EGITTO: I PERIODICI ITALIANI (1892- 1940) 26
Wafaa El Beih

PIATTAFORME WIKI PER L'INSEGNAMENTO UMANISTICO: SPERIMENTAZIONI IN CORSO NEL LICEO DE COSMI DI PALERMO
Antonino Fiorino, Paolo Monella, Francesca Saieva, Antonella Sorci 32

THE ORGANIZATION AND MANAGEMENT OF THE MAGIC PROJECT FOR ANCIENT MANUSCRIPTS DIGITIZATION: CONNECTIONS BETWEEN MEDITERRANEAN CULTURES 38
Stefania Conte, Andrea Mazzucchi, Guido Russo, Augusto Tortora, Giorgia Tortora

TOWARDS A RESEMANTISATION OF THE CONCEPT OF MODELLING IN DIGITAL HUMANITIES 43
Cristina Marras, Arianna Ciula, Øyvind Eide, Patrick Sahle

VALORIZZARE UN ARCHIVIO 'MEDITERRANEO': STUDI PER UN'EDIZIONE CRITICA DIGITALE DELLE OPERE DI GIOVANNI COMISSO 48
Marco Borrelli

ARCHIVI E MUSEI DIGITALI PER IL PATRIMONIO CULTURALE 55

A WORKFLOW FOR GLAM METADATA CROSSWALK 56
Arianna Moretti, Ivan Heibi, Silvio Peroni

DIGITALIZZAZIONE E MODELLAZIONE DELLA *DRAMMATURGIA* DI LEONE ALLACCI 63
Luca Giovannini, Giorgia Gallucci

From Data Complexity to User Simplicity: A Framework for Linked Open Data Reconciliation and Serendipitous Discovery 67
Marco Grasso, Giulia Renda, Marilena Daquino
Giuseppe Chiarini: un'opera inedita 73
Elena Almangano, Mirko Castaldi, Eleonora De Longis, Daniele Pasqualetti
HERITRACE: Tracing Evolution and Bridging Data for Streamlined Curatorial Work in the GLAM Domain 78
Arcangelo Massari, Silvio Peroni
Libri e biblioteche tra museabilità e musealizzazione digitale: sogno o realtà? 84
Nicola Barbuti, Mauro De Bari
Listening2Painting: an Audio Augmented Reality approach for Arts 89
Nicola Orio, Daniel Zilio, Andrea Micheletti
Luoghi comuni: metodi e strategie di sviluppo software in ambito GLAM, dalle voci di autorità all'esplorazione cartografica 92
Herbert Natta, Gianluca Rossi, Roberta Maggi
Messaggistica Istantanea e Archivi Digitali. Quali soluzioni? Best practices e considerazioni dal contesto internazionale 98
Alessia Del Bianco
Nuove interazioni con collezioni digitali: l'"Archivio digitale del Capitolo di Laterza" 104
Stefania Riso, Nicola Barbuti
Preservare e valorizzare la memoria di archivi storici di ex-ospedali psichiatrici 110
Grazia Serratore
Preservazione del patrimonio culturale e cloud computing: caratteristiche e criticità 117
Manuela Grillo
Preserving Culinary Traditions. A Crowdsourced Digital Collection of Cookbooks 122
Giulia Renda, Giulia Manganelli, Mila Fumini, Marilena Daquino
Serious games e gamification: a che punto sono le istituzioni culturali italiane? 128
Vincenzo Colaprice
The Tree of Philosophers: design and implementation of a digital resource for the history of academic philosophy 133
Guido Bonino, Nicola Ruschena
Thinking Outside the Black Box: Insights from a Digital Exhibition in the Humanities 138
Sebastian Barzaghi, Alice Bordignon, Bianca Gualandi, Silvio Peroni

EDIZIONI SCIENTIFICHE DIGITALI 143

Il progetto Corr<si>Ca: edizione digitale della corrispondenza Canioni 144
Anna Giaufret, Beatrice Dal Bo, Elena Margherita Vercelli, Laura Bonanno
Il progetto "MaximHum": Italia umanistica e Moscovia cinquecentesca dialogano in digitale 152
Francesca Romoli, Letizia Ricci, Angelo Mario Del Grosso
L'Archivio di Giuseppe Fava: conservazione e valorizzazione attraverso il digitale 159
Giuseppe Davide Di Mauro, Marzia D'Amico
L'archivio digitale di una casa editrice: l'esempio del Saggiatore e della sua prima pubblicazione 165
Giada Di Pino
Marcare la poesia del Novecento: uno studio per *Ossi di seppia* 171
Chiara Cauzzi, Martina Corti, Anna Guadagnoli, Maria Grazia Schiaroli
MetaScript: a framework proposal for screenplay encoding 175
Erica Andreose, Giorgia Crosilla, Leonardo Zilli
OpenData: OpenGadda 181
Eleonora Pasquale, Martina Pensalfini
Paul Klee, *Tunisreise* e *Bildnerische Formlehre*: un caso studio di DiScEPT (Digital Scholarly Editions Platform and Aligned Translations) 186
Hansmichael Hohenegger, Tiziana Mancinelli, Fabio Ciotti, Federico Boschetti, Angelo Mario Del Grosso, Eleonora De Longis

PAVES-E: PER UNA HYPEREDIZIONE DELL'OPERA DI CESARE PAVESE 191
Christian D'Agata, Angelo Mario Del Grosso, Laura Nay, Giuseppe Palazzolo, Antonio Sichera, Daria Spampinato
PER UN'EDIZIONE DIGITALE DI *SE QUESTO È UN UOMO* 197
David Tagliacozzo
PER UN'EDIZIONE SCIENTIFICA DIGITALE DELLO *SPECULUM GUY OF WARWICK* 204
Omar Khalaf, Sibilla Siano
PER UNA LETTURA ANTROPOLOGICA DI VERGA: TRA CODIFICA E GEOREFERENZIAZIONE 210
Giovanna Zisa
PROGETTO DI EDIZIONE GENETICA DIGITALE DEL *CANZONIERE* MANOSCRITTO DI U. SABA (1919-20) 215
Marina Buzzoni, Davide Cucurnia, Cristina Fenu, Roberto Rosselli Del Turco, Giulia Tancredi
RAPPRESENTARE LA STORIA SACRA: UN'IMPRESA IERI, UNA SFIDA OGGI. PROPOSTA DI EDIZIONE SCIENTIFICA DIGITALE DEL "COMPENDIUM" DI PIETRO DI POITIERS 221
Franz Fischer, Agnese Macchiarelli
TOWARDS AN INTEGRATED DIGITAL EDITION OF THE *LEGES LANGOBARDORUM* 226
Marina Buzzoni, Roberto Rosselli Del Turco
UN *CORPUS* ONLINE DELLA LETTERATURA SECONDARIA (1872- 1890) DEL VERISMO ITALIANO 232
Denise Bruno, Giuseppe Canzoneri, Antonio Di Silvestro, Daria Spampinato, Alessandro Zammataro
UN PROGETTO DI EDIZIONE DIGITALE *IMAGE-BASED* DELLE *MERAVIGLIE D'ORIENTE* NEL MS COTTON VITELLIUS A.XV 240
Andreea Mihaela Toma
UNA PROPOSTA DI CODIFICA IN XML/MEI PER TESTI MUSICALI AUTOGRAFI DI VINCENZO BELLINI 246
Laura Mazzagufo
VERISMO DIGITALE. PER UN'EDIZIONE DIGITALE COMMENTATA DELLE OPERE DI VERGA, CAPUANA, DE ROBERTO 252
Liborio Pietro Barbarino, Elisa Conti, Christian D'Agata, Miryam Grasso, Ninna Maria Lucia Martines, Eliana Vitale
VERSO L'HYPEREDIZIONE. LO SVILUPPO DI PIRANDELLO NAZIONALE TRA DIDATTICA E RICERCA 260
Milena Giuffrida, Christian D'Agata, Giulia Cacciatore, Fabrizio Lo Presti
VOCI DALL'INFERNO: DANTE PER DIRE IL LAGER - DIGITALIZZARE E STUDIARE LE TESTIMONIANZE 267
Angelo Mario Del Grosso, Marina Riccucci, Elvira Mercatanti

DIZIONARI E DIGITALIZZAZIONE DI BANCHE DATI 274

IL VIVER (VOCABOLARIO DELL'ITALIANO VERISTA) 275
Gabriella Alfieri, Marco Biffi, Stephanie Cerruto, Giovanni Salucci
L'INFORMATIZZAZIONE DEL GDLI: RISULTATI, PROSPETTIVE, SFIDE FUTURE 281
Eva Sassolini, Sebastiana Cucurullo, Marco Biffi
LA DIGITALIZZAZIONE DEL DIZIONARIO LATINO LANA 1978 287
Francesca Michelone
LEXICAD: PIATTAFORMA LESSICOGRAFICA DIGITALE PER L'ITALIANO DELLE ORIGINI 293
Salvatore Arcidiacono, Antonella Zammataro
STRATEGIE PER LA CREAZIONE E LA CONDIVISIONE DI UNA COLLEZIONE DIGITALE DI TESTI GRECO-LATINI 298
Vincenzo Ortoleva, Maria Rosaria Petringa, Salvatore Cammisuli, Mariarosaria Villareale
THE CORR<SI>CA PROJECT: ENHANCING AND "QUERYING" THE CANIONI FAMILY CORRESPONDENCE 303
Tiziana Pasciuto, Selenia Anastasi, Daniele Zolezzi, Simonetta Acacia, Giada D'Ippolito, Chiara Storace, Maria Tolaini
XML/TEI E DIZIONARI *BORN-DIGITAL*: UNA PROPOSTA PER LE RISORSE LESSICOGRAFICHE DELLA RETE LEXICAD/PLUTO
Giuseppe Leonardo Zappalà 310

ANALISI COMPUTAZIONALE, INTELLIGENZA ARTIFICIALE E LINGUISTICA 318

ANALISI COMPUTAZIONALE DEI REPORT DI SOSTENIBILITÀ: LA VAGHEZZA COME STRATEGIA DI GREENWASHING 319
Erica Cutuli
ANALISI STILOMETRICA APPLICATA ALLE CAPACITÀ EMULATIVE DI GPT-4 325
Marco De Cristofaro, Mariangela Giglio
C'È UN TESTO IN QUESTA CHAT? INTELLIGENZA ARTIFICIALE E COOPERAZIONE INTERPRETATIVA 332
Daniele Silvi

GENERE E GEOPOLITICA NELLE DISCIPLINE UMANISTICHE DIGITALI IN ITALIA. LE CONFERENZE AIUCD (2012-2023)
Selenia Anastasi 336
GLI LLM COME LETTORI MODELLO ARTIFICIALI 342
Fabio Ciotti
I SIMILI SI ATTRAGGONO. LA VALUTAZIONE LETTERARIA SULLE PIATTAFORME DI DIGITAL SOCIAL READING 348
Gabriele Vezzani, Simone Rebora, Massimo Salgaro
IL *DISTANT READING* È L'ORNITORINCO 354
Pietro Mazzarisi
L'IMPIEGO DELL'INTELLIGENZA ARTIFICIALE PER LA RICOSTITUZIONE DELLE AGGREGAZIONI ARCHIVISTICHE E L'ARRICCHIMENTO DEI METADATI NEGLI ARCHIVI DIGITALI 361
Stefano Allegrezza
MACCHINE PER LEGGERE: LA TEXT ANALYSIS COME STRUMENTO PER IMPARARE A LEGGERE I CLASSICI DELLA NARRATIVA… E AD AMARLI 367
Fabio Ciotti
PRESERVARE LA DIVERSITÀ NELL'ERA DELL'INTELLIGENZA ARTIFICIALE: IL DILEMMA ETICO DI BIAS E DISCRIMINAZIONI NEGLI ALGORITMI 371
Gianluca Pavani
QUI PRO QUO? DATI TESTUALI E STRUMENTI PER LA RISOLUZIONE DI COREFERENZE IN LATINO 377
Roberta Grazia Leotta, Eleonora Delfino
STRUMENTI DIGITALI PER LA TRASCRIZIONE E LA LEMMATIZZAZIONE DI TESTI IN ITALIANO ANTICO 382
Emiliano Degl'Innocenti, Alessia Spadi, Federica Spinelli, Lucia Francalanci, Michela Perino, Irene Falini, Francesco Coradeschi, Francesco Pinna
TESTI ALLOGRAFICI: CONTATTI TRA LINGUE E SCRITTURE DEL MEDITERRANEO 388
Antonio Pagliara, Federico Boschetti, Daniele Baglioni
THE DARK MIRROR OF ARTIFICIAL INTELLIGENCE: HOW AI AFFECTS CLIMATE CHANGE 393
Mauro De Bari
UN SISTEMA DI CLASSIFICAZIONE AUTOMATICA DI IMMAGINI RELATIVE A MATERIALI LIBRARI ANTICHI E MODERNI 398
Nicola Barbuti, Tommaso Caldarola
UNCOVERING THE SPREAD OF LEXICAL INNOVATION IN ITALIAN TWEETS 405
Greta Franzini, Paolo Brasolin, Stefania Spina

ORGANIZZAZIONE DELLA CONOSCENZA CON SEMANTIC WEB 410

AFFINARE IL CONTESTO: ESTRAZIONE DI INFORMAZIONI STRUTTURATE PER L'ARRICCHIMENTO DEI CONTESTI ARCHIVISTICI 411
Lucia Giagnolini, Paolo Bonora, Francesca Tomasi
CLEF 2.0. SOLUZIONI PER LA CATALOGAZIONE NATIVA LINKED DATA DEL PATRIMONIO DIGITALE CULTURALE ITALIANO
Marilena Daquino, Laurent Fintoni, Sebastiano Giacomini, Francesca Tomasi 417
L'ONTOLOGIA BIGRAFO: VERSO UN MODELLO SEMANTICO PER L'OPERA DI FRANCO FORTINI 423
Laura Antonietti, Emilio M. Sanfilippo, Emmanuela Carbé
LOST IN DATIFICATION? THE JOURNEY OF DATA FROM THE PRIMARY SOURCE TO THE FINAL INTERPRETATION 429
Enrica Bruno, Sofia Baroncini, Francesca Tomasi
ORBIS DIOECESIUM. CREATING AUTHORITY DATA ON THE LEGAL-HISTORICAL CHANGES OF CATHOLIC DIOCESES 435
Benedetta Albani, Rowan Dorin, Yohan Park
PER L'INTEROPERABILITÀ E LA SOSTENIBILITÀ DELLE RISORSE DIGITALI DANTESCHE: IL PROGETTO LIDA 441
Cesare Concordia, Gaia Tomazzoli, Nicola Aloia, Carlo Meghini, Luca Trupiano
PER UN'ANALISI DEI PERSONAGGI TRA LETTERATURA, FILOSOFIA E ONTOLOGIA APPLICATA 448
Emilio M. Sanfilippo, Gaia Tomazzoli, Michele Paolini Paoletti, Jansan Favazzo, Roberta Ferrario
REPRESENTING TEXTS AS LOD: A SYSTEMATIC LITERATURE REVIEW 455
Michela Bandini, Valeria Quochi
STRUCTURING AUTHENTICITY ASSESSMENTS ON HISTORICAL DOCUMENTS USING LLMS 463
Andrea Schimmenti, Valentina Pasqual, Francesca Tomasi, Fabio Vitali, Marieke van Erp
VALORIZZAZIONE DI REGISTRAZIONI STORICHE DI CANTO LIRICO NEL WEB SEMANTICO 469
Marcello Ranieri, Angelo Pompilio

PUBLIC HISTORY E ARCHEOLOGIA DIGITALE **473**

ARCHEOLOGIA E RILIEVO 3D: UNA RIFLESSIONE SULLE METODOLOGIE. DUE CASI STUDIO DI AREA MEDITERRANEA 474
Graziana D'Agostino, Pietro Maria Militello

DAL CATASTO BORBONICO ALLA GENOMICA. PIATTAFORME DIGITALI E INTERDISCIPLINARITÀ: IL PROGETTO "WE ARE WHAT THEY WERE" DI RIPOSTO 481
Salvatore Spina

DRAWING HISTORY FROM THE CODICE PELAVICINO DOCUMENTS: GRAPH VISUALIZATION FOR HUMAN RESEARCHERS
Natthida Wiwatwicha 489

GESTIONE INFORMATICA DELLA DOCUMENTAZIONE ARCHEOLOGICA "MINORE". METODOLOGIE E APPLICAZIONI NELL'AMBITO DEL PROGETTO STORAGE 494
Marianna Figuera, Erica Platania

GESTIRE L'IMMATERIALE. CONTESTI SENSORIALI A SERVIZIO DEL PATRIMONIO ARCHEOLOGICO 501
Serena D'Amico

GODSCAPES: MODELING SECOND MILLENNIUM BCE POLYTHEISMS IN THE EASTERN MEDITERRANEAN THANKS TO THE STORAGE PROJECT 507
Nicola Laneri, Marianna Nicolosi-Asmundo, Daniele Francesco Santamaria, Chiara Pappalardo

IL PROGETTO STORAGE: DAI DATI AL WEB 512
Simone Faro, Pietro Maria Militello, Marianna Nicolosi-Asmundo

LA MODELLAZIONE 3D AL SERVIZIO DELL'ARCHEOLOGIA: NUOVE PROSPETTIVE PER L'APPLICAZIONE AD EDIFICI MULTIPIANO DI ETÀ PROTOSTORICA 517
Dario Puglisi, Marco Chiricallo

MLS CON SENSORE LIDAR *APPLE* PER LO SCAVO ARCHEOLOGICO: APPLICAZIONI PRATICHE 523
Luigi M. Caliò, Antonello Fino, Gian Michele Gerogiannis

ODONIMI D'ITALIA E DIGITAL PUBLIC HISTORY: LE PROBLEMATICHE DI UNA SCHEDATURA PARTECIPATA 528
Enrica Salvatori, Vittore Casarosa, Riccardo Chiari

OPENSTREETMAP: UNO STRUMENTO E UNO SPAZIO PER LA DIGITAL PUBLIC HISTORY? 535
Camilla Zucchi

PRESERVARE LA MEMORIA: IL PROGETTO STORAGE E L'ARCHIVIO DELL'EX ISTITUTO DI ARCHEOLOGIA 539
Giovanni Fragalà, Pietro Maria Militello

UN ATLANTE DIGITALE PER LA STORIA MARITTIMA DEL REGNO DI SARDEGNA 545
Giampaolo Salice

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
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

INFRASTRUTTURE PER LA RICERCA UMANISTICA **555**

COPHI EDITOR: FROM GREEKSCHOOLS TO THE PROJECTS MULTIVERSE 556
Simone Zenzaro

DIGITAL HUMANITIES AND HERITAGE SCIENCE: MOVING FROM LANDSCAPING TO A DYNAMIC RESEARCH OBSERVATORY IN AN OPEN SCIENCE CLOUD 559
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

FUNZIONI E SOSTENIBILITÀ DI UNA PIATTAFORMA DIGITALE PER LE LINGUE ARCAICHE 566
Michele Mallia, Riccardo Del Gratta, Valeria Quochi

INFRASTRUTTURE DI RICERCA COME STRUMENTI DI "INTERCULTURALITÀ DIGITALE" 572
Salvatore Cristofaro, Vittoria Fabiani, Cristina Marras, Enrico Pasini, Pietro Sichera, Mingyang Yu

MATERIALI DIDATTICI COME OGGETTI DIGITALI FAIR: UNA METODOLOGIA CONDIVISA PER LA FORMAZIONE IN H2IOSC
Giulia Pedonese, Francesca Frontini, Roberta Ottaviani, Federico Boschetti, Alessia Spadi, Lucia Francalanci, Alessia Scognamiglio, Pietro Restaneo, Antonina Chaban, Jana Striova, Laura Benassi 577

RETHINKING SCHOLARLY DIGITAL OBJECTS AS CULTURAL HERITAGE: THE KNOT PROJECT 582
Laurent Fintoni

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

Il *Distant reading* è l'ornitorinco

Pietro Mazzarisi
Università degli studi di Trieste, Italia - pietro.mazzarisi@gmail.com

ABSTRACT

Negli studi di *Distant reading*, l'oggetto base su cui vengono condotte le ricerche è composto dai testi letterari e gli aspetti, caratteristiche, costruzioni e categorie linguistiche al loro interno. L'invalso che si è sviluppato è di non operare alcuna distinzione all'interno dei *corpora* su cui vengono condotte le indagini. Ai fini di validità statistica l'invalso è corretto, pena sarebbe, diversamente, la non rappresentabilità del campione. Ma ai fini dell'attività critico-interpretativa l'invalso non riconosce almeno due aspetti fondamentali: (i) tutte le classificazioni di linguistica testuale pongono il testo letterario a sé, riconoscendo a un tale oggetto base di ricerca caratteristiche peculiari e strutturali; (ii) aspetti, costruzioni, categorie e caratteristiche linguistiche oggetto base di indagine acquisiscono salienze diverse quando si presentano nella diegesi e quando si presentano nella mimesi di una stessa opera, poiché le porzioni testuali diegetiche e le porzioni testuali mimetiche di uno stesso testo letterario hanno salienze diverse. L'intervento propone una riflessione e proposta teorico-metodologica per ovviare a tali criticità, seguita da una sua applicazione concreta su un *corpus* di letteratura italiana premiata.

PAROLE CHIAVE

Distant reading; linguistica testuale; diegesi/mimesi; letteratura premiata.

1. INTRODUZIONE

Una delle monografie più recenti sul *Distant reading* lo rinomina *Literary mathematics* e lamenta la scarsa attenzione teorica che la *Cultural analytics*, le *Digital humanities* e questo specifico sottocampo hanno indirizzato verso l'oggetto base d'indagine, richiamando l'attenzione di ricercatori e ricercatrici sullo studio del *corpus* in sé: «what literary scholars bring to the interdisciplinary table is their robust and sustained attention to textual forms. We have elaborate theories for poetic and novelistic structures, yet cultural analytics and the digital humanities have proceeded to date without a working theory of the corpus as an object of inquiry» [9: 6]. Dopo aver riaffermato le persistenti necessità e mancanze di quadri teorici per il *Distant reading* [5], lo studio individua in questi termini le domande di ricerca atte a colmare queste mancanze: «We need a theory of quantitative literary analysis. Such a discourse would ask: What kinds of things are the objects of textual computation? […] A much better question to begin with is this one: What relations exist between the corpus and its source texts?» [9: 165-166]. Nel rispondere a queste domande di ricerca, lo studio di Gavin ha il pregio di portare la necessaria e dovuta attenzione verso l'intersecazione, all'interno delle pratiche di *Distant reading*, tra le analisi computazionali dei testi letterari e dei metadati; questi ultimi intesi nell'accezione più ampia che superi le sole autorialità, anno, luogo e casa editrice di pubblicazione e si soffermi anche sulle vicende inter/intrapersonali di tali autorialità non dimenticando le eventuali e ulteriori differenze etniche, di genere e di altra natura. In questa direzione, lo studio ricostruisce la nascita dell'ipotesi distribuzionale di Zellig Harris [12] dagli studi linguistici firthiani [8] per estenderla alle pratiche di *Distant reading*, dimostrando infine l'assunto secondo il quale «*Similar words tend to appear in documents with similar metadata* […]. Similar words will be used by similar authors, at similar times, when describing similar things in similar places» [9: 175].

Come la maggior parte degli studi di *Distant reading*, anche il sopraccitato studio manca però di un passaggio preliminare e fondamentale: ovvero la necessità di fare un passo indietro e non trattare il testo letterario alla stregua di un qualsiasi altro testo. Da decenni, la linguistica testuale ha infatti proposto classificazioni ampiamente note e condivise e che, pur basandosi su differenti prospettive tipologiche, tutte distinguono il testo letterario da altre forme testuali: nelle prospettive funzionali più note è definito "testo narrativo" [22] e "testo letterario e poetico" [6], "testo a libertà interpretativa alta" [19] nella prospettiva interpretativa.

Il testo letterario è un mammifero, condivide caratteristiche comuni a tutte le altre tipologie testuali. Il testo letterario è un ornitorinco, come le echidne e a differenze di tutti gli altri mammiferi possiede caratteristiche peculiari. Ipotizzando uno studio sulla riproduzione nei mammiferi, si dovrebbero trattare diversamente i monotremi da tutti gli altri mammiferi sulla base della oviparità dei primi rispetto alla viviparità dei restanti. A discapito di tali peculiarità, fino ad oggi l'invalso nelle ricerche di *Distant reading* è stato studiare il testo letterario come un testo comune. Malgrado la specificità sottolineata dalla linguistica testuale e la forte attenzione alle forme testuali della ricerca letteraria rilevata nello studio di Gavin, nelle

proprie pratiche il *Distant reading* non ha riconosciuto che i *corpora* su cui concentra gli studi sono composti di testi strutturalmente diversi dai testi dei *corpora* impiegati negli studi politici, giuridici, psicologici ecc.
Gli studi più entusiasti che hanno seguito la nascita del concetto di *Distant reading* sviluppandolo e mettendolo in pratica si sono spesso proposti come approcci coincidenti con la condizione descritta da Umberto Eco a cui echeggia il titolo di questo intervento [7]: scoperte che metterebbero in discussione categorizzazioni ritenute ormai storicamente organiche. Ma la maggior parte di essi sembra essersi comportata come l'ornitorinco, almeno secondo una visione antropocentrica sulla coscienza: ignaro della propria specificità rispetto a tutti gli altri mammiferi. Un esempio recente può essere lo stesso studio sopraccitato: oltre al proporre la definizione di *Literary mathematics* in sostituzione di *Distant reading* e a non concentrare la necessaria attenzione sulle specificità del testo letterario, ha svolto le proprie indagini su un *corpus* costituito non solo di testi letterari, ma anche da testi di critica letteraria in aggiunta ai testi letterari [9].

2. DIEGESI E MIMESI

Nella visione del *Distant reading* qui proposta, la disciplina dovrebbe anzitutto riconoscere e rispettare l'oggetto base della ricerca e le sue peculiarità, ovvero le peculiarità del testo letterario. Di conseguenza, la creazione di un quadro teorico base dovrebbe porsi come principale domanda la seguente: come si studia il testo letterario computazionalmente entro un quadro teorico che (i) alla fonte ne riconosca la specificità testuale e (ii) sia anche operazionalizzabile [14]?
Per rispondere a questa domanda, l'approccio proposto propone di ripartire dalla millenaria opposizione tra diegesi (διήγησις) e mimesi (μίμησις). La tradizione occidentale fa risalire a Platone le prime categorizzazioni e distinzioni tra ciò che è riconducibile al narratore e ciò che, invece, è riconducibile ai personaggi. Ne *La Repubblica*, Platone segue la concezione greca per cui l'arte e la poesia sono forme di imitazione, ma ne offre un approfondimento attraverso il confronto con Adimanto quando gli chiede se, in definitiva, tutti i racconti fatti da mitologi e poeti non si riducano a una narrazione diretta, imitativa o delle due forme intrecciate [18: III, 392d]. Per esemplificare la distinzione, Platone parafrasa lo scambio dialogico tra il sacerdote apollineo Crise e i carcerieri della figlia all'inizio dell'*Iliade*, offrendo così una riformulazione degli scambi dialogici in diegesi e sottolineando la distinzione tra discorso diretto e discorso indiretto [18: III, 393d-394b]. Infine, àncora questa distinzione principalmente a quattro generi letterari: «Nell'invenzione poetica c'è un genere completamente imitativo, come tu dici, ed è rappresentato dalla tragedia e dalla commedia. Poi ce n'è un altro in cui è il poeta stesso a narrare, come accade particolarmente nei ditirambi. Infine c'è un terzo tipo, misto di narrazione e di imitazione, che si trova nella poesia epica e in molti altri componimenti. Mi capisci, non è vero?» [18: III, 394c].
La riflessione teorica sull'opposizione diegesi/mimesi è millenaria, nella tradizione occidentale essa parte già da Aristotele [1: I-II, 1447a-1448a] e continua a svilupparsi fino alla narratologia classica [10: 52-67, 11: 71-72] e oltre. Infatti, per l'attività critico-interpretativa letteraria risulta fondamentale poter continuare a contare sulle categorie della diegesi e della mimesi, poiché su queste ha costruito una buona fetta dei suoi strumenti critici, dallo *skaz* di Ejchenbaum, per l'analessi di Genette alla polifonia di Bachtin, solo per citare pochi esempi. Il *Distant reading*, invece, non ha ereditato l'opposizione diegesi/mimesi, privando la disciplina di una basilare riflessione teorica, benché risulti fondamentale per l'attività critico-interpretativa non solo, per esempio, ottenere un *dataset* con il tipo di lessico che caratterizza un *corpus* letterario, ma poter anche distinguere quale delle due dimensioni tra diegesi e mimesi è caratterizzata da quale tipo di lessico e incrociare, raffrontare, interpretare il differenziale tra i dati.

3. APPROCCIO TEORICO-METODOLOGICO: ASPETTI E LIMITAZIONI

Il nuovo approccio teorico-metodologico, denominato "analisi dell'asse diegetico-mimetico" (*Diegetic-Mimetic Axis Analysis*), ha per obiettivi (i) riconoscere la specificità del testo letterario operazionalizzando alla sorgente l'opposizione teorica diegesi/mimesi, (ii) preservare la validità statistica dei campioni di indagine e (iii) creare le condizioni per poter ottenere una miglioria nella raccolta dati, arricchendoli qualitativamente in granularità e salienza. Per raggiungere questi obiettivi la procedura prevede una previa partizione automatica dei testi letterari in porzioni testuali diegetiche e in porzioni testuali mimetiche con tecnica di *text mining*. La partizione automatizzata esclude l'intervento e la discrezione umana, basandosi su elementi interni dei testi e predeterminati da chi ha creato i testi: i segni di interpunzione e la convenzioni formali a cui la maggior parte dei testi letterari in prosa ricorre per segnare l'avvicendamento tra diegesi e mimesi. Pertanto, la validità statistica del campione non viene compromessa, poiché esso viene segmentato e non modificato.
La procedura prevede poi la conduzione di analisi computazionali sulle distinte porzioni testuali così ottenute. Infine, prevede la conduzione di analisi differenziali tra le distinte entità testuali ottenute, ovvero tra quella diegetica e quella mimetica. L'analisi differenziale è un tipo di indagine preso in prestito da altri campi del sapere con la finalità di migliorare gli aspetti interpretativi di fronte a un problema. Nei campi della diagnostica medica e psicologica include la presa in considerazione di fenomeni diversi per giungere a una determinazione quanto più corretta possibile. In economia compara

analisi di voci di bilancio diverse per determinare la sostenibilità di una scelta. Nella critica letteraria computazionale mira a resocontare come e in che misura aspetti, caratteristiche, costruzioni e categorie linguistiche oggetto base di ricerca si manifestino nelle due dimensioni della diegesi e della mimesi e quali aspetti è possibile cogliere dalle diverse manifestazioni. In sintesi, l'analisi differenziale è un procedimento utile nel formare addizionali interpretativi per l'espressione dei giudizi critici.

Il binomio teorico diegesi/mimesi è impiegato con accezioni operative in prospettiva di operazionalizzazione e connota la diegesi come "la porzione testuale riconducibile a chi narra", la mimesi come "la porzione testuale riconducibile ai personaggi". L'approccio ingloba le riflessioni teoriche che si sono sviluppate in narratologia da tale dicotomia [10: 52-67], ma semplifica le distinzioni teoriche alle due sole dimensioni diegetica e mimetica per le necessità pratiche che si presentano quando si deve tradurre una teoria letteraria in una sequenza di istruzioni formali da comunicare con il linguaggio di programmazione; ovvero quando si deve operazionalizzare una teoria letteraria affinché le misurazioni sui testi avvengano alla luce di quel concetto letterario [14].

Di conseguenza, la metodologia ha delle limitazioni circa la propria applicabilità. La prima e più ovvia limitazione riguarda il fatto che l'approccio delimita da sé il proprio campo di applicazione, perimetrandolo alla testualità letteraria in prosa, così come la linguistica testuale suddivide il testo letterario in prosa e poesia [6]. La seconda limitazione riguarda il fatto che se la maggior parte dei testi letterari sono suddivisi in mimesi/diegesi e di norma queste due dimensioni sono graficamente indicate con segni interpuntivi, esistono eccezioni che a volte costringono un intervento umano quantomeno di controllo su come effettivamente è stato reso l'avvicendamento tra mimesi e diegesi (nel *corpus* preso in analisi nel seguente paragrafo un caso su centotré). Per ovviare a tali problematicità, la metodologia nasce come un approccio di natura *mixed method*. In questa prospettiva (e)segue il concetto di *scalable reading* proposto da Martin Mueller - che consiste in una alternanza di fasi di letture ravvicinate a fasi di letture a distanza [15]. La metodologia va così a coincidere con il fine indicato da Andrew Piper sul versante di un ricomponimento con la critica classica. Piper, infatti, ha proposto il carattere circolare e dunque ermeneutico dei passaggi da micro-analisi a macro-analisi e di nuovo a micro come base di un *mixed method* che consiste in una ermeneutica computazionale fatta di letture a fasi alterne tra vicino e lontano e altrettante alterne analisi tra i dati qualitativi e quantitativi [17]. Un approccio *mixed method* oltre a risanare la frattura con la critica classica può, inoltre, far progredire la disciplina del *Distant reading*, passando appunto dai punti intermedi e inevitabili quando una disciplina giovane si trova nelle sue fasi iniziali, dimostrando che se dapprima non si è in grado di trattare con efficienza teorica le scale delle centinaia e delle migliaia, difficilmente si potranno trattare teoricamente ed efficacemente le scale di milioni di testi letterari (e, in ogni caso, l'esperienza maturata nelle scale inferiori giova alle superiori).

4. APPLICAZIONE SU *CORPUS* DI LETTERATURA PREMIATA

L'approccio viene applicato su un *corpus* di letteratura italiana premiata composto da tutte le opere narrative scritte in lingua italiana vincitrici dei premi Strega, Campiello e Bancarella dal 1980 al 2020 per mostrare come possa migliorare qualitativamente la raccolta e interpretazione dei dati aumentandone le caratteristiche in granularità e salienza. L'intervallo storico inserisce il *corpus* nell'epoca del postmoderno. I premi letterari dovrebbero, ipoteticamente, intercettare parte delle opere che in futuro verrà inclusa nel canone letterario; questo a sua volta dovrebbe riflettere il periodo storico-letterario in cui tali opere sono state concepite. Ma in che modo, effettivamente, i premi letterari si pongono in relazione agli stili che riflettono il postmoderno? Una caratteristica del postmoderno sarebbe l'impiego di un linguaggio alleggerito da norme e restrizioni, non controllato da una mente unificatrice tramite un processo di centralizzazione, ma che rifletta una maggiore spontaneità. I premi hanno apprezzato lo stile postmoderno alleggerito da norme e restrizioni o si sono volti indirizzando la eco da loro prodotta verso gli stili "alti, controllati, sublimi" dove, proporzionalmente, è stata vigilata l'insorgenza di attenuazioni da norme e restrizioni?

Per rispondere alle domande di ricerca e mostrare la miglioria apportata dall'approccio, nel *corpus* viene preso in analisi l'impiego del *presente pro-futuro*. Una marcatezza che nella mimesi rifletterebbe la scelta di rendere un linguaggio più autentico e spontaneo, nella diegesi uno stile narrativo alleggerito da norme e restrizioni. Infatti, per Berruto, ormai «*verrò domani* risulta quasi funzionare da forma enfatica rispetto a *vengo domani*» [2: 70]. Ricercare su testi letterari scritti in lingua italiana ha portato a escludere dal *corpus* alcune opere vincitrici del premio Bancarella: alcune perché scritte in lingua straniera, dunque traduzioni non esattamente cumulabili con una ricerca sull'uso di una forma marcata; altre perché di *non-fiction*. Strega e Campiello hanno sempre e solo premiato opere di autori e autrici italiane e nell'intervallo storico non figurano testi poetici tra i vincitori. Nel 2005, il Campiello registra una vincita *ex aequo* tra Pino Roveredo (*Mandami a dire e altri racconti*) e Antonio Scurati (*Il sopravvissuto*). È stato preferito Roveredo secondo il criterio di maggiore inclusione e diversità all'interno del *corpus*. Infatti, Scurati è già presente con un'altra opera (*M. Il figlio del secolo*, Strega 2019); Roveredo no e, conseguentemente, sarebbe stato del tutto escluso. La raccolta di racconti di Roveredo è stata

misurata come libro, non suddividendola in sotto-testi e lo stesso criterio è stato applicato alle altre raccolte di racconti vincitrici (ovvero 1982S, 1997S, 1999S). In questo modo, la dimensione finale del *corpus* è di 103 opere e 1.762.929 occorrenze totali di forme verbali (41 opere e 837.639 occorrenze Strega, 41/606.712 Campiello, 21/318.578 Bancarella; lo Strega include *La scuola cattolica* di Edoardo Albinati che da solo apporta 149.228 occorrenze. Vd. Fig. 1?).
Passando al presente in funzione futura, sembra essere diffuso e trasversale nelle lingue naturali, indipendentemente dalla disponibilità di tempi preposti all'espressione della futurità [3: 97-98]. Attestato già nell'italiano antico [13: 75], in epoca contemporanea è frequentemente coadiuvato da avverbi ed espressioni di tempo [16: 99] e solitamente circoscritto alla lingua colloquiale [20: 467]. Per escludere occorrenze di presente con funzioni indesiderate - come, per esempio quella storica o iterativa - il *data cleaning* ha limitato l'output alle *key words in context* (KWIC) circoscrivendolo alle co-occorrenze di presente semplice in concomitanza degli avverbi di tempo *oggi*, *stamattina*, *stasera*, *stanotte*, *domani*, *indomani*, *domattina*, *dopodomani* nelle precedenti e seguenti 17 parole.
Passando ai premi letterari, un recente studio di Simonetti prende in analisi soprattutto vittorie e cinquine finaliste degli ultimi sei anni, notando punti comuni nelle vittorie dello Strega. Sarebbero opere inclusive che non impegnano eccessivamente chi legge e propongono sempre la conformità alle coeve idee politiche. Si porrebbero dunque lontano dallo sperimentalismo, scevre di fratture allo standard, sperimentando al massimo nell'ibridazione dei generi. Sempre a supporto dei deboli, tradirebbero, una falsa coscienza poiché questa *pietas* non si traduce in solidarietà stilistica e autentica adesione linguistica al mondo popolare cui esprimono compassione [21: 163]: «La lingua di questo romanzo non deve essere troppo complicata o resistente alla decifrazione, soprattutto nella struttura sintattica […]. Lo sperimentalismo formale nel complesso continua a essere bandito, come pure […] gli stereotipi della narrativa di consumo; mentre giocare a mescolare generi diversi […] viene senz'altro incoraggiato, incluso il ricorso a un po' di blasone regionale» [21: 178].

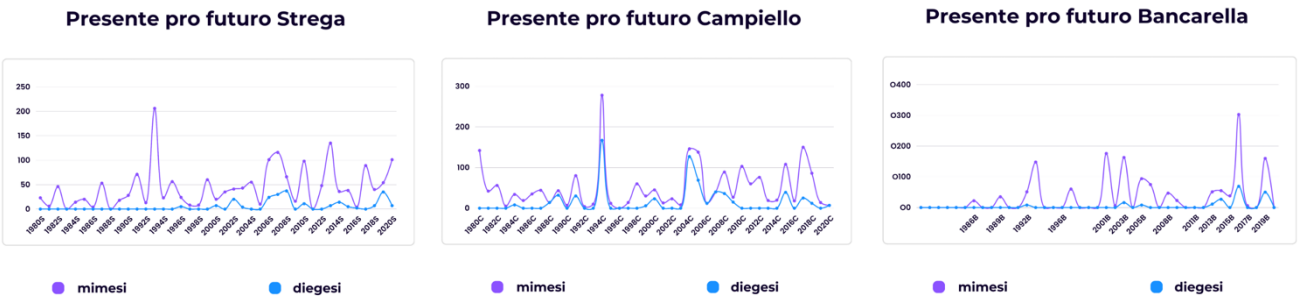

Figura 1. Differenziali delle occorrenze nei premi Strega, Campiello e Bancarella

Lo Strega registra nei suoi testi premiati 332 occorrenze di presente pro-futuro registrate, di cui 289 in mimesi e 43 in diegesi. In questo premio si può notare un lento aumento di forme del presente pro-futuro e quella che potrebbe essere una maggiore presenza di stili meno alti e sublimi a partire dal 2000 circa, mentre in precedenza sembrerebbe confermato il bando a un linguaggio più spontaneo. L'eccezione è il testo vincitore nel 1993, ovvero *Ninfa plebea* di Domenico Rea e questo sembrerebbe confermare un carattere dello Strega, peraltro condiviso con la maggior parte dei più famosi premi letterari stranieri: l'essere particolarmente sensibile alle pressioni commerciali esercitate dalle maggiori case editrici [21: 17]. Infatti, l'unicità dell'opera (pubblicata nel 1992) combacia con l'identità della casa editrice che l'ha portata sul mercato editoriale: è l'unica opera, delle 41 nel relativo premio, a non essere stata data alla stampa da un editore blasonato. L'opera di Rea è stata pubblicata dalla Leonardo editore, un marchio indipendente fino al 1993 e con una linea editoriale che non ha seguito i filoni tradizionali, ma ha preferito piuttosto la promozione della creatività all'interno delle pratiche di scrittura[1], fino a quando non ha incrociato le vicissitudini del Lodo Mondadori. Le altre opere che sembrano segnare delle controtendenze nel nuovo millennio sono quelle vincitrici delle edizioni 2006, 2007, 2010, 2013 e 2017; ovvero *Caos calmo* di Sandro Veronesi (2006S), *Come Dio comanda* di Nicolò Ammaniti (2007S), *Canale Mussolini* di Antonio Pennacchi (2010S), *Resistere non serve a niente* di Walter Siti (2013S) e *Le otto montagne* di Paolo Cognetti (2017S). Pennacchi e Siti hanno in comune il materiale su cui hanno costruito le loro narrazioni, cioè realtà sociolinguistiche a cui solitamente viene associata una minore vigilanza sulle forme. In Pennacchi è protagonista una famiglia di contadini della bassa padana, i Peruzzi, che muove verso il Lazio per la bonifica dell'omonimo canale. In Siti è protagonista una differente prospettiva generazionale all'interno della criminalità organizzata, dove un bagaglio di nuove competenze permette sì ai più giovani di allargare gli orizzonti criminali, ma non cancellerebbe l'oralità tipica dei contesti di socializzazione primaria

[1] «Leonardo editore srl (1986 -)», https://www.lombardiabeniculturali.it/archivi/soggetti-produttori/ente/MIDB001358/.

di provenienza. Il carattere linguistico che i dati evincono in Veronesi, Ammaniti e Cognetti sembra coincidere con i risultati del *close reading* di Simonetti, per cui Veronesi e Ammaniti sarebbero accomunati da una scrittura "scenografica" che pare *ab origine* confezionata per il salto transmediale e audiovisivo, come poi avvenuto [21: 138]. Anche Cognetti dimostrerebbe sintassi e ritmiche analoghe, ma oltre a questi aspetti, la sua scrittura risulterebbe essere anche la più distante dai modelli sublimi e più conservatori: «*Le otto montagne* […] risulta tra i cinque il romanzo più consapevole e forse, in fondo, il più orgoglioso della propria distanza culturale e linguistica da una certa tradizione italiana» [21: 140].
La doppia giuria del Campiello, per Simonetti, è nata per smarcarsi dalle logiche più meramente editoriali (Strega) e più meramente commerciali (Bancarella) e puntare sul valore della narrazione in sé, non sulle possibili auree di chi ha pubblicato e scritto. Dal 2000, la concorrenza avrebbe spinto il Campiello verso la prosa insieme commerciale e pseudo-impegnata che connoterebbe lo Strega [21: 105-106]. Delle 250 occorrenze registrate nel Campiello, 181 sono in mimesi e 69 in diegesi. Si possono notare alcune analogie lo Strega. Entrambi i grafici mostrano un picco nella prima metà degli anni Novanta (presagio degli imminenti cannibali?) e un lento aumento di forme del presente e di quella che potrebbe essere una maggiore presenza di stili meno alti e sublimi a partire dal 2000 circa. Nel Campiello il picco negli anni Novanta è *Sostiene Pereira* di Antonio Tabucchi (1994C). Qui è significativo notare come la granularità nei dati riesca a rilevare un una qualità diegetica riflessa nella complessità narratologica di un testo che sovrappone la voce di Pereira filtrata dal narratore a quella del narratore stesso. È anche significativa la granularità data dall'approccio nel differenziare i maggiori tre picchi degli anni 2000. Distingue in qualità diegetica *La barca nel bosco* di Paola Mastrocola (2004C) da *L'ultimo arrivato* di Marco Balzano (2015C) e *L'arminuta* di Donatella Di Pietrantonio (2017C), ma rileva anche una minima differenza diegetica tra 2015C e 2017C. I tre testi hanno in comune l'essere tutti autodiegesi condotte da pre-adolescenti e adolescenti. Da qui il primo, secondo e terzo posto per quel che riguarda gli anni 2000 (viene escluso 2005C di Roveredo perché l'averlo misurato non dividendolo in sotto-testi non permette un discorso organico e comparato). Ma i dati differenziano ulteriormente 2004C. Perché è una narrazione simultanea che riporta i fatti adolescenziali mentre avvengono per la maggior parte del libro. Infatti, 2004C presenta come narratore autodiegetico un adolescente (dai 13 ai 25 anni circa) e offre il lavoro autoriale nel filtrare una storia attraverso occhi e lingua di un ragazzino che da una piccola isola del Sud si trasferisce a Torino per studiare al liceo. 2015C ripresenta la figura del giovane emigrato da una più specificata Sicilia verso Milano, ma alla più tenera età di 9 anni. A differenza di 2004C, in 2015C il narratore autodiegetico a volte racconta in narrazione simultanea, altre invece compie un salto temporale narrando già da adulto e anziano, con *amarcord* e distanza che ne consegue. Infine, l'io narrante di 2017C è una ragazza che, nelle prime due parti del libro, cerca le risposte a quello che le è successo fra i 13 e 14, quando è stata rispedita indietro nella sua famiglia biologica, dopo che fin dai sei mesi era cresciuta in una famiglia adottiva di parenti. 2004C si differenzia per l'autodiegesi in narrazione simultanea che in 2015C e 2017C è più limitata. Inoltre, a dividere le tre opere nel frattempo c'è stata anche una certa Elena Ferrante che pur non vincendo nessuno di questi premi dal 2011riscuote successo con un io narrante che a distanza di tempo ripercorre le fasi adolescenziali della sua formazione (e in ambito di attribuzione autoriale se ne potrebbero ipotizzare due distinte usando questo approccio per darle la caccia).
Anche nel Bancarella le 198 occorrenze registrate, di cui 178 sono in mimesi e 20 in diegesi, suggeriscono analogie. Il più alto picco, *La ragazza di fronte* di Margherita Oggero (2016B), è un romanzo che segue due personaggi principali, lei archivista torinese e lui un meridionale che dal Sud ha fatto base a Torino, impiegando due eterodiegesi che si avvicendano in narrazioni simultanee. I dati sembrerebbero collegare queste caratteristiche di 2016B a quelle analoghe già viste in 2004C e 2015C. I dati, infine, segnano altri tre picchi: *La gita a Tindari* di Andrea Camilleri (2001B), *Amiche di salvataggio* di Alessandra Appiano (2003B) e *Il ladro gentiluomo* di Alessia Gazzola (2019B). Anche qui l'approccio permette di inferire e notare analogie: 2001B è un giallo; 2003B è un *chick lit*; 2019B un *chick lit* ibridato di giallo.

5. CONCLUSIONI E IPOTESI DI SVILUPPI FUTURI

Se notiamo i diversi rapporti tra il numero totale delle occorrenze verbali di ogni premio, le percentuali delle occorrenze del presente pro-futuro nei tre premi e il differenziale dei dati, il Bancarella segna allo stesso tempo le più alte percentuali di uso del presente pro-futuro (0,062%) e di impieghi mimetici (89,90% in mimesi e 10,10% in diegesi). Il Campiello e lo Strega hanno valori simili sull'uso generale del presente pro-futuro (Campiello 0,041% e Strega 0,039%), ma differenziali che li distinguono nettamente sull'asse diegetico-mimetico. Il Campiello è al primo posto per gli impieghi diegetici (27,6% in diegesi e 72,4% in mimesi) e doppia l'impiego diegetico del presente pro-futuro nello Strega (12,95% in diegesi e 87,05% in mimesi).
Se continuiamo a ipotizzare che una maggiore percentuale generale di impieghi del presente coincida con una lingua più spontanea e meno vigilata mentre una maggiore percentuale di impieghi diegetici coincida con una complessità delle strutture narrative, troviamo il premio popolare Bancarella aver apprezzato di più la spontaneità linguistica, il Campiello

aver apprezzato di più la complessità narrativa e lo Strega equidistante: essersi tenuto a debita distanza da complessità narrativa e spontaneità linguistica, come sostenuto nello studio di Simonetti. È un risultato significativo che già con pochi dati l'approccio sia in grado di raccogliere al contempo indicazioni diegetiche sulle complessità narrative e mimetiche su tipo di lingua e milieu finzionali. Infine, l'approccio proposto appare promettente perché già con un basso output di dati riesce a far inferire ed emergere analogie tra realtà linguistico-narrative effettivamente presenti nei testi letterari, ma diversamente perse limitando lo studio nell'attuale invalso del *Distant reading*.

L'approccio, ovviamente, necessita di ulteriori studi e conferme con altri *corpora*, ma è già possibile avanzare linee di sviluppi futuri a cui dà base e strumenti epistemici. Uno degli sviluppi futuri della ricerca riguarda la valutazione dei dati sotto una cornice teorica arricchita dalla *Game Theory*. In questo caso si ipotizza che la scrittura diegetica e la scrittura mimetica siano due *player* per come intesi nella *Game Theory* e che il genere letterario sia la risultante dei differenti rapporti che tra essi si possono instaurare. Tali rapporti sono misurabili con analisi dei sentimenti, modellazione degli argomenti, liste di vocaboli e analisi differenziali. Infatti, una applicazione teorica della *Game Theory* al rapporto tra le scritture diegetiche e mimetiche risulta estremamente interessante per tre motivi.

Il primo motivo è che negli studi letterari la *Game Theory* è stata applicata solo in ambito di scrittura mimetica, nel senso che è stata fatta la cosa più ovvia: analizzare il comportamento dei personaggi alla luce delle categorie della teoria dei giochi [4: 1-26]. Ma non è stato ipotizzato un profilo per le scritture diegetica e mimetica in quanto due *player* né un rapporto tra esse in quanto *player*.

Il secondo motivo è che le varie categorie teoriche previste dalla teoria dei giochi sembrano coincidere con i rapporti che sussistono tra le scritture diegetiche e mimetiche, realizzati, di solito, per differenziare generi letterari e specifici espedienti letterari. Così, la categoria del gioco cooperativo sembra riflettere il rapporto tra le scritture diegetiche e mimetiche nei generi letterari del romanzo storico, della fiaba, dell'epica, dove la comunanza di interessi (la cooperazione) risiede in narrazioni nelle quali, tendenzialmente, non si avrà una delle due scritture che contraddice l'altra, piuttosto appunto appaiono coordinate a un comune fine narrativo. Dall'altra parte, la categoria del gioco non cooperativo sembra riflettere il rapporto tra le scritture diegetiche e mimetiche nei generi letterari del thriller e dell'horror, dove il non coordinamento, lo scarto tra le due scritture è atto alla creazione della necessaria suspense. La categoria dei giochi ripetuti nel tempo e i loro continui *payoff* sembrano riflettere il rapporto tra le scritture diegetiche e mimetiche che connotano le serialità letterarie e i godimenti cognitivi dati dall'isomorfismo. La categoria dei giochi a informazione perfetta sembrano riflettere i tipici rapporti tra le scritture diegetiche e mimetiche delle narrazioni con narratore onnisciente; dall'altra parte, la categoria dei giochi a informazione imperfetta sembrano riflettere i tipici rapporti delle scritture diegetiche e mimetiche nelle narrazioni con narratore inattendibile. La categoria dei giochi finiti sembrano riflettere i tipici rapporti tra le scritture diegetiche e mimetiche nei generi letterari giallo e rosa. Infine, la categoria dei giochi a somma zero coincide con i tipici rapporti della scrittura mimetica impiegata nella tragedia, il genere letterario in cui maggiormente è stata applicata la *Game Theory*. È questa una ipotesi che chiede approfondimenti, ma che già allo stadio iniziale sembra promettente, anche quando si considera il fatto che la rivoluzione storico-letteraria del metateatro introdotta da Pirandello, in questa prospettiva, risulta appunto dall'aggiunta del nuovo *player* autoriale, laddove per millenni c'era stata la sola dimensione mimetica. E anche nella narrativa le infrazioni tra realtà e finzione tipiche dei testi metaletterari scaturiscono dall'aggiunta del nuovo *player* autoriale, laddove per millenni erano prevalse soprattutto le sole contrapposizioni tra le scritture diegetiche e mimetiche. In questa ipotesi i diversi rapporti (e infrazioni) tra le scritture diegetiche e mimetiche che distinguono i generi letterari assumono l'immagine di concerti ad archi; un'immagine rafforzata dall'elemento *game* della teoria che quando messo in azione esprime in diverse lingue il significato di "suonare" oltre che di "giocare": inglese *to play*, francese *jouer*, tedesco *spielen*, russo играть (*igrat'*).

Il terzo motivo è che la *Game Theory* è nata in ambito matematico, ma ha avuto le sue migliori applicazioni nelle scienze sociali prima di tutto e poi nell'evoluzionismo e nel cognitivismo: tre ambiti teorico-applicativi verso cui il collegamento del *Distant reading* viene sempre più indirizzato e considerato necessario.

BIBLIOGRAFIA

[1] Aristotele. *Poetica*. (a cura di) M. Valgimigli. Bari: Laterza, 1964.

[2] Berruto, Gaetano. *Sociolinguistica dell'italiano contemporaneo*. Roma: La Nuova Italia Scientifica, 1987.

[3] Bertinetto, Pier Marco. «Metafore tempo-aspettuali». *Linguistica* 32, fasc. 2 (1992): 89-106.

[4] Brams, J. Steven. *Game Theory and The Humanities. Bridging two Worlds*. 2011a ed. Cambridge-London: The MIT Press, 2011.

[5] Ciotti, Fabio. «Distant reading in literary studies: a methodology in quest of theory». *Testo e Senso* 23 (2021): 195-213. https://testoesenso.it/index.php/testoesenso/article/view/509.

[6] De Beaugrande, Robert-Alain, e Dressler Wolfgang. *Introduction to Text Linguistics*. New York: Routledge, 1981.

[7] Eco, Umberto. *Kant e l'ornitorinco*. Milano: Bompiani, 1997.

[8] Firth, John Rupert. «The technique of semantics». *Transactions of the Philological Society* 34, fasc. 1 (1935): 36-72.

[9] Gavin, Micheal. *Literary Mathematics. Quantitative Theory for Textual Studies*. Stanford: Stanford University Press, 2022.

[10] Genette, Gérard. *Figures* II. Paris: Seuil, 1969.

[11] Genette, Gérard. *Figures* III. Paris: Seuil, 1972.

[12] Harris, S. Zellig. «Distributional Structure». *Word* 10, fasc. 2-3 (1954): 146-162.

[13] Lorenzetti, Luca. *L'italiano contemporaneo*. Roma: Carocci, 2002.

[14] Moretti, Franco. «'Operationalizing'». *New Left Review* 84 (novembre 2013): 103-119.

[15] Mueller, Martin. «Shakespeare His Contemporaries: collaborative curation and exploration of Early Modern drama in a digital environment». *Digital Humanities Quarterly* 8, fasc. 3 (2014). http://www.digitalhumanities.org/dhq/vol/8/3/000183/000183.html.

[16] Patota, Giuseppe. *Grammatica di riferimento dell'italiano contemporaneo*. Milano: Garzanti, 2010.

[17] Piper, Andrew. «Novel Devotions: Conversional Reading, Computational Modeling, and the Modern Novel». *New Literary History* 46, fasc. 1 (2015): 63-98.

[18] Platone. *La Repubblica*. (a cura di) G. Lozza. Milano: Mondadori, 1990.

[19] Sabatini, Francesco. «"Rigidità-esplicitezza" vs "elasticità-implicitezza": possibili parametri massimi per una tipologia dei testi». (a cura di) G. Skytte e F. Sabatini, 141-172. Copenhagen: Museum Tusculanum Press, 1999.

[20] Serianni, Luca. *Grammatica italiana. Italiano comune e lingua letteraria*. Torino: UTET Università, 2006.

[21] Simonetti, Gianluigi. *Caccia allo Strega. Anatomia di un premio letterario*. Milano: Nottetempo, 2023.

[22] Werlich, Egon. *A text grammar of English*. Heidelberg: Quelle und Meyer, 1976.