



Università  
di Catania

ASSOCIAZIONE per  
l'INFORMATICA UMANISTICA  
e la CULTURA DIGITALE

Consiglio Nazionale  
delle Ricerche

# ME.TE. DIGITALI

## MEDITERRANEO IN RETE TRA TESTI E CONTESTI

ATTI DEL XIII CONVEGNO ANNUALE  
AIUCD 2024



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

ISBN 978-88-942535-8-0



Copyright ©2024 AIUCD  
Associazione per l'Informatica Umanistica e la Cultura Digitale



Il presente volume e tutti i contributi sono rilasciati sotto licenza Creative Commons Attribution Share-Alike 4.0 International license (CC-BY-SA 4.0). Ogni altro diritto rimane in capo ai singoli autori.  
*This volume and all contributions are released under the Creative Commons Attribution Share-Alike 4.0 International license (CC-BY-SA 4.0). All other rights retained by the legal owners.*

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

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

Ultimo accesso agli URL in data 15 maggio 2024.

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

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

I contributi pubblicati nel presente volume hanno ottenuto il parere favorevole da parte di valutatori esperti della materia, attraverso un processo di revisione anonima mediante *double-blind peer review* sotto la responsabilità del Comitato di Programma di AIUCD 2024.

*All the papers published in this volume have received favourable reviews by experts in the field of DH, through an anonymous double-blind peer review process under the responsibility of the AIUCD 2024 Program Committee.*

## **Chair**

Antonio Di Silvestro (Università di Catania)

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

## **Comitato di programma / Program committee**

Emmanuela Carbé (Università di Siena)

Massimo Cultraro (CNR Istituto di Scienze del Patrimonio Culturale)

Christian D'Agata (Università di Catania)

Antonio Di Silvestro (Università di Catania)

Greta Franzini (Eurac Research)

Maurizio Lana (Università del Piemonte Orientale)

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

Marco Mazzone (Università di Catania)

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

Marianna Nicolosi-Asmundo (Università di Catania)

Marina Paino (Università di Catania)

Giuseppe Palazzolo (Università di Catania)

Jonathan Prag (University of Oxford Merton College)

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

Rachele Sprugnoli (Università di Parma)

Francesco Stella (Università di Siena)

## **Segreteria scientifica / Scientific Secretariat**

Liborio Barbarino (Università di Catania)

Denise Bruno (Università di Catania)

Giulia Cacciatore (Università di Catania)

Giuseppe Canzoneri (Università di Catania)

Elisa Conti (Università di Catania)

Milena Giuffrida (Università di Catania)

Miryam Grasso (Università di Catania)

Francesca Prado (Università di Catania)

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

Eliana Vitale (Università di Catania)

Alessandro Zammataro (Università di Catania)

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

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

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

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

## **Enti organizzatori / Organisers**

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

## **Supporter**

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

## Chair di area/ Track chair

### Le culture digitali nel Mediterraneo

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

Paola Moscati (CNR Istituto di Scienze del Patrimonio Culturale)

### Archivi ed edizioni digitali

Christian D'Agata (Università di Catania)

Greta Franzini (Eurac Research)

### Analisi computazionale dei testi

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

Simone Reborra (Università di Verona)

### Ontologie e Semantic Web

Marianna Nicolosi Asmundo (Università di Catania)

Francesca Tomasi (Università di Bologna)

### Preservazione della memoria e del patrimonio digitale

Fabio Ciraci (Università del Salento)

Anna Maria Marras (Università di Torino)

## Lista dei revisori /List of reviewers

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

# Il progetto digitale su Elisa Chimenti (LAI-ALEEF): la problematicità di un profilo mediterraneo tra reti e frontiere

Ada Desideri<sup>1</sup>, Bianca Vallarano<sup>2</sup>

<sup>1</sup> Sorbonne Université, Francia - ada.desideri@sorbonne-universite.fr

<sup>2</sup> Università di Napoli L'Orientale/Université de Lille, Italia/Francia - b.vallarano@unior.it

## ABSTRACT

Il paper presenta i lavori in corso sull'opera e sull'archivio di Elisa Chimenti (Napoli 1883-Tangeri 1969), per proporre delle riflessioni metodologiche sulle problematicità dei progetti digitali che implicino una rete trans-mediterranea. In primo luogo, viene presentata l'autrice e il suo corpus alla luce della sua pluri-appartenenza geografica e linguistica, che la inserisce a pieno titolo in una complessa rete mediterranea. In seguito, viene presentato il progetto Chimenti, un progetto di digitalizzazione del corpus che tenta di ovviare grazie al digitale agli impedimenti ed ai limiti geografici, metodologici e scientifici che un corpus tale presenta, dal plurilinguismo alla standardizzazione della codifica, dall'interdisciplinarietà alla transnazionalità. Infine, sono evidenziate le problematicità che rimangono aperte, tanto da un punto di vista teorico quanto pratico, riguardo agli strumenti e alle metodologie del digitale.

## PAROLE CHIAVE

Elisa Chimenti; plurilinguismo; spazio mediterraneo; archivi digitali; edizioni digitali.

## 1. INTRODUZIONE

Parlare di Mediterraneo non vuol dire solo parlare di reti, ma soprattutto di frontiere. Spazio transnazionale, fin dall'antichità luogo di incontro e scontro dei popoli che lo hanno abitato, il bacino Mediterraneo è rappresentato a seconda delle epoche e dei punti di vista come culla e come campo di battaglia, come ponte e come muro [18]. Ogni rappresentazione dello spazio mediterraneo dipende da ed è costruita su presupposti storico-culturali differenti, iscrivendosi in una specifica genealogia storico-culturale [14]. Attraversare il Mediterraneo, concretamente e metaforicamente, è un'operazione ancora oggi non scontata, ed asimmetrica [9]. Fare storia del Mediterraneo, fare letteratura del Mediterraneo, vuol dire ancora oggi trovarsi davanti continue frontiere [4]. Il caso di Elisa Chimenti si inserisce in tale panorama scientifico-disciplinare, e può essere considerato esemplificativo di una serie di tensioni che vi sono connaturate, tanto da un punto di vista geo-storico quanto metodologico. Cosa significa provare a fare rete nello spazio mediterraneo? Quali sono le difficoltà e le tensioni che si incontrano? Concretamente, come tali problematiche emergono e si ritrovano nei progetti digitali sul corpus? La riflessione si inserisce dunque nel secondo asse della *call for papers*, affrontando nel concreto temi e questioni inerenti alle edizioni scientifiche digitali nonché la rappresentazione e l'organizzazione delle tecniche e criticità nel campo della digitalizzazione, dell'organizzazione e dell'archiviazione del patrimonio, nel contesto della rete dei paesi gravitanti sul Mediterraneo.

## 2. ELISA CHIMENTI: UNA AUTRICE MEDITERRANEA

Elisa Chimenti è un chiaro esempio di un'autrice che a causa della sua pluri-appartenenza, nonché nel suo genere, non è entrata in nessun canone nazionale e ha avuto difficoltà, in vita come *post mortem*, a vedere riconosciuta la propria autorialità [6]. Di origine napoletana ma emigrata con la famiglia all'età di pochi mesi nel Maghreb, Chimenti ha vissuto e operato nel Marocco coloniale e postcoloniale, e specificamente nella città di Tangeri, zona internazionale fino all'indipendenza del paese ottenuta nel 1956. Italiana, Chimenti perde la cittadinanza ed assume quella tedesca in seguito al matrimonio, avvenuto a Tangeri nel 1912. Apolide dopo il divorzio, ha intrattenuto per tutta la vita relazioni problematiche con le istituzioni dei tre paesi. Autrice poliglotta, si è formata nello spazio internazionale e mediterraneo della regione di Tangeri, fin dall'antichità crocevia di lingue e popoli diversi [19, 2]. La sua opera è profondamente *métissée*, tanto da un punto di vista linguistico quanto di tradizioni ed immaginari [17]: nei suoi testi, fortemente ancorati nell'oralità, le diverse lingue del bacino mediterraneo si intersecano, dando voce ai popoli che lo hanno attraversato [28]. Così, sulla base francese risuona l'arabo e specificamente la *darija* marocchina, che si mescola ai dialetti italiani del Sud, all'ebraico sefardita, allo spagnolo vivo nello stretto di Gibilterra, all'amazigh della catena montuosa del Rif. Il risultato è

un ibridismo [5] che sfugge ad ogni definizione e categorizzazione univoca, presentandosi come un mosaico di storie letterarie, orali, culturali e religiose tanto distinte quanto strettamente interrelate [7].

Nonostante Chimenti abbia scritto moltissimo e collaborato con diversi giornali marocchini ed internazionali, ha pubblicato in vita solamente cinque opere [10]<sup>1</sup>. Migliaia di pagine rimangono inedite, conservate presso la *Fondation Méditerranéenne Elisa Chimenti*<sup>2</sup>. L'associazione, fondata nel 2010 con sede presso il *Palais des Institutions Italiennes* di Tangeri, è nata grazie all'impulso di intellettuali ed amici vicini all'autrice, con lo scopo di promuoverne l'opera. Purtroppo, nella mancanza di un ente o di una istituzione pubblica che garantisca la durabilità del progetto, le attività della fondazione sono da alcuni anni cessate, e la sistematizzazione dell'archivio non è stata terminata. La sede della Fondazione è oggi chiusa, e l'archivio non accessibile. Senza dubbio, il fatto che l'archivio interpellasse interlocutori nazionali diversi ha contribuito alle difficoltà nella sua sistematizzazione e durabilità: cittadina tedesca e italiana, Chimenti ha lavorato per tutta la sua vita in Marocco con gli enti italiani esteri (la scuola italiana, da lei fondata nel 1914), che, dopo la fine del periodo coloniale, sono stati definitivamente chiusi [27]. La condizione frammentata degli archivi del periodo coloniale è tutt'oggi una questione viva, e problematica [11].

Tanto la vita quanto l'opera e i lasciti di Elisa Chimenti si articolano dunque in una complessa rete mediterranea, che tenta di superare le frontiere nazionali per istituire un dialogo trasversale. È questa una delle cause che ha determinato l'esclusione dell'autrice da un canone che ancora oggi si identifica troppo spesso con un identitarismo eurocentrico nazionale e monolingue, ed è reticente ad accogliere identità plurali e composite [3, 26]. Il riconoscimento dell'autorialità andando di pari passo con l'istituzionalizzazione degli archivi – che non sono semplice luogo fisico ma luogo simbolico, che istituzionalizza una legge sull'altra [15] – il risultato è un circolo vizioso che permette ad un sistema chiuso di perpetuarsi, *ne varietur*. Il lavoro archivistico di preservazione, studio, organizzazione dei documenti è il primo passo per smuovere il canone e diffondere nuovi paradigmi di conoscenza. In questo senso, la riflessione sugli archivi non è una riflessione sul passato, ma sul presente – in quanto luogo di educazione e mobilitazione [13] – e sul futuro.

### 3. IL CORPUS CHIMENTI: I PROGETTI DIGITALI

Di fronte alle sfide poste da un archivio non sistematizzato, un corpus variegato ed una rete vasta e complessa che interpella interlocutori diversi, il digitale dovrebbe aprire maggiori possibilità rispetto ai metodi tradizionali. È per questo che i progetti di valorizzazione di tale corpus hanno intersecato fin da subito lo studio accademico e la preparazione delle edizioni a stampa degli inediti con progetti e approcci digitali. In un primo momento, questo aspetto è stato sviluppato nell'ambito del Laboratorio Associato Internazionale (LAI) “La scrittura dell'esilio al femminile. Il dialogo tra le lingue, le culture e le idee nello spazio europeo e mediterraneo XIX-XXI secolo” (2018-2022)<sup>3</sup>. In questo contesto, nato dalla collaborazione tra università diverse in Francia, Italia e Spagna, il lavoro informatico si è focalizzato sulla concezione e la creazione di edizioni critiche digitali, con lo scopo di diffondere i testi in ambito di ricerca, di salvaguardarli, ma anche e soprattutto di riuscire a rappresentare la loro complessità grazie alle visualizzazioni informatiche [8]<sup>4</sup>.

Un primo esperimento di edizione digitale [12] è stato effettuato su due testi molto diversi per forma e contenuto: un romanzo di fantascienza in bella copia (*Une étrange aventure*) e la bozza di un racconto ambientato durante la guerra del Rif (*Mennouch la Rifaine*, della raccolta *Contes Berbères*). In questi due progetti, l'annotazione si basa su un documento – concepito a partire da un'analisi generale del corpus e dei temi affrontati dall'autrice<sup>5</sup> – che aveva l'obiettivo di rappresentare un massimo di fenomeni ritenuti interessanti, senza distinzioni tra informazioni sulla struttura del testo ed elementi di analisi. Queste edizioni permettono di leggere due versioni del testo (semi-diplomatica e critica); di visualizzare parole e passaggi relativi a temi ricorrenti nel corpus; di creare delle liste a partire da queste parole; di far apparire la lingua utilizzata per le parole straniere; e di leggere le note legate all'apparato critico.

---

<sup>1</sup> Si tratta di: *Èves marocaines* (Tanger, 1935), *Chants de femmes arabes* (Paris, 1942), *Au cœur du harem* (Paris 1958), *Légendes marocaines* (Paris 1959), *Le sortilège et autres contes sephardites* (Tanger, 1964).

<sup>2</sup> Cfr. il sito internet della Fondazione: <https://www.elisachimenti.org/accueildef.html>.

<sup>3</sup> Progetto internazionale coordinato da Camilla Cederna (Università di Lille, lab. CECILLE) e Silvia Tatti (Università Sapienza di Roma), in collaborazione con l'Università di Pisa e l'Università di Siviglia. I due casi di studio focus del progetto erano Elisa Chimenti e Cristina Trivulzio di Belgiojoso (<https://cecille.univ-lille.fr/vie-scientifique>).

<sup>4</sup> Le edizioni sono depositate su GitLab, e possono essere condivise individualmente su richiesta a scopo di ricerca. Per motivi di diritto d'autore, al momento non possono essere rese pubbliche. Per questo motivo, non abbiamo potuto includere in questo articolo le immagini delle interfacce delle due edizioni.

<sup>5</sup> Il documento è stato creato da Camilla Cederna, Nathalie Gasiglia e dalla studentessa Camélia Ait-Mechedal.

Un secondo esperimento di edizione digitale è stato preparato su un racconto che presenta delle varianti genetiche (*Histoire du prince Moustapha et de la princesse Djamila*<sup>6</sup>), utilizzando il software di visualizzazione EVT. L'annotazione si è focalizzata – oltre che sull'arborescenza e su alcune annotazioni di base per esplorare il testo (persone, luoghi, lingue straniere) – sulla costruzione di un apparato critico, con l'obiettivo di rendere conto degli aspetti genetici. Si sono quindi combinati i due moduli TEI *basic apparatus* (per note, aggiunte, cancellature) e *critical apparatus* (per testimoni, lemmi, readings), entrambi codificati con il *parallel segmentation method*. Grazie a EVT, il testo è consultabile in diverse modalità di edizione (testo semplice, edizione fotografica, collazione). È possibile evidenziare persone, luoghi e parole in lingua straniera, creare delle liste a partire da questi elementi, e visualizzare i fenomeni legati allo sviluppo genetico del testo.

Parallelamente ai lavori di edizioni digitali, un nuovo progetto è iniziato nel 2023, finanziato dalla *Maison Européenne des Sciences de l'Homme et de la Société* (MESHS)<sup>7</sup>, dal titolo *Archives de l'écriture de l'exil au féminin* (ALEEF)<sup>8</sup>. Il progetto, che coinvolge il corpus di Elisa Chimenti, si pone come obiettivo di proseguire la ricerca iniziata in seno al LAI e portare avanti il lavoro sul digitale. In questa ottica, è nata la collaborazione con il laboratorio di ricerca Thalim<sup>9</sup> (CNRS, ENS, Sorbonne Nouvelle) per costruire una biblioteca digitale dell'autrice, gli *Archives Chimenti*, online sulla piattaforma EMAN<sup>10</sup>. Un altro obiettivo del progetto è stato quello di ampliare la rete di collaborazioni alla sponda meridionale del Mediterraneo: alle università in Europa si sono aggiunte università e enti in Marocco<sup>11</sup>, con i quali la collaborazione è in corso sia per lo studio linguistico e letterario del corpus, sia per la riflessione sullo stato dell'archivio.

#### 4. RIFLESSIONI SUL CAMPO OVVERO LE PROBLEMATICHE INCONTRATE NEI PROGETTI DIGITALI

Rispetto agli obiettivi e ai risultati attesi da tali progetti digitali, sono presto emerse le problematiche dovute al lavoro in un contesto ibrido, sia sul piano tecnico/materiale sia sul piano dei processi di formalizzazione dei testi.

Dal punto di vista tecnico, i problemi più evidenti sono legati all'OCR. In tutti gli esperimenti di edizione presentati, è stato fatto un primo tentativo di trascrizione usando dei software di OCR, prima di effettuare la trascrizione manuale. I risultati dei software usati (*Tesseract* e *Abby Fine Reader*) non erano soddisfacenti, anche per i testi in bella copia e ben conservati. I motivi sono due, ed entrambi dovuti alla complessità di una rete vasta, non organizzata e plurilingue. Da una parte, la bassa qualità delle immagini, diretta conseguenza della difficoltà di comunicazione tra paesi e istituzioni diverse – il che comporta la non sistematizzazione dell'archivio – e della conseguente insufficienza delle risorse. Le fotografie disponibili dei testi sono state scattate con i telefoni personali di due ricercatrici, che hanno potuto accedere al fondo una prima volta tra il 2018 e il 2019 ed una seconda volta nel 2022. Una scannerizzazione professionale era prevista, ma non ha potuto essere completata vista l'inaccessibilità del fondo. D'altra parte, seppure avessimo a disposizione immagini di buona qualità, ulteriore problema è l'ibridismo linguistico dei testi: i software di OCR sono ottimizzati per ricevere una sola lingua in input, e generano moltissimi errori in contesti plurilingui. A questo si aggiunge, ad un livello ulteriore, la problematicità delle lingue trattate. Tali software sono addestrati solo sulle lingue scritte e maggiormente sulle lingue occidentali politicamente maggioritarie<sup>12</sup>. Infine, neanche un OCR addestrato sulla *darja* del Marocco permetterebbe di risolvere questo problema, poiché i termini arabi sono traslitterati da Chimenti in alfabeto latino in maniera eterogenea. Abbiamo quindi la certezza che le parole in lingua straniera e soprattutto quelle in arabo e *darja* marocchina – che costituiscono una percentuale importante del lessico di Chimenti – non sarebbero trascritte correttamente.

Per quanto riguarda la rappresentazione formale dei testi, la pluralità che caratterizza il corpus ci pone di fronte a difficoltà metodologiche ed epistemologiche, legate allo specifico progetto di edizioni digitali. Per quanto riguarda le prime tipologie

<sup>6</sup> Il racconto è presente in due raccolte dattiloscritte di Chimenti, conservate entrambe nel suo archivio: *Contes d'autrefois* (con correzioni manoscritte) e *La veillée du harem* (in bella copia).

<sup>7</sup> CNRS, UAR 3185: [https://www.meshs.fr/page/Archives\\_Ecritures\\_Exil\\_Feminin.....titre-----asc](https://www.meshs.fr/page/Archives_Ecritures_Exil_Feminin.....titre-----asc).

<sup>8</sup> Archivi della scrittura dell'esilio al femminile, il progetto è coordinato da Camilla Cederna: <https://lai-ecriture-exil-au-feminin.univ-lille.fr/accueil>.

<sup>9</sup> *Théorie et histoire des arts et des littératures de la modernité XIXe-XXe siècle*, CNRS, UMR 7172: <https://www.thalim.cnrs.fr/>.

<sup>10</sup> EMAN (*Éditions de manuscrits et d'archives numériques*) è una piattaforma di pubblicazione digitale per la diffusione di manoscritti e di fondi d'archivio moderni. Si basa sul software Omeka, ma prevede numerose estensioni: <https://eman-archives.org/EMAN/>.

<sup>11</sup> L'Università Abdelmalek Essaâdi di Tetouan, il Centro culturale Beit Nedjma, gli Archives du Maroc.

<sup>12</sup> Il problema dei *bias* occidentali negli strumenti di DH è ben più ampio [29, 24]. D'altronde, i sistemi riflettono i *bias* delle società che li hanno creati. L'esperienza dimostra come gli strumenti di DH (dagli OCR ai database agli strumenti di *text analysis*) lavorano meglio con le lingue occidentali (e in particolare con l'inglese) che con le lingue definite "minori" (che spesso in realtà minori non sono, ma piuttosto marginali politicamente, cfr. [16]). Sono molte oggi le ricerche che si muovono per sopperire a queste mancanze. Cfr. non esaustivamente [22, 25, 1], negli atti della *7th IEEE Congress on Information Science and Technology (CiSt)*, Agadir - Essaouira, Morocco, 2023.

di edizioni, la volontà di rappresentare informazioni su aspetti diversi – strutturali ed analitici – ha fatto emergere una serie di problematiche. Oltre al tempo richiesto per trattare un grande numero di fenomeni, le annotazioni sono risultate talvolta non pertinenti e talvolta insufficienti rispetto ai nostri obiettivi. Da una parte, la codifica prestabilita risultava estremamente dettagliata riguardo le tipologie di correzione del dattiloscritto, il che non forniva, infine, dettagli interessanti e/o utili per l’analisi che ci si era preposte. D’altra parte, la codifica non prevedeva aspetti che sarebbero invece stati utili e pertinenti, come una annotazione che indicasse l’equivalenza tra le diverse forme del nome di uno stesso personaggio. Il caso di varianti grafiche è d’altronde estremamente rilevante nel corpus, vista la continua traslitterazione tra alfabeto arabo e latino, nonché la relativa instabilità della forma scritta della *darija* marocchina. Inoltre, a causa delle diverse forme dei testi, gli stessi elementi di codifica sono stati in alcuni casi usati in modi differenti (come le note, usate in un caso per le note d’autrice e nell’altro per quelle di editrice). Per quanto riguarda la seconda tipologia di edizione, anche qui il vocabolario delle annotazioni non è risultato completamente soddisfacente, e per i limiti di EVT e per i limiti della TEI. EVT essendo un software open source, lavorando sul codice probabilmente tali limiti si sarebbero potuti risolvere. Tuttavia, si tratta di un *workflow* pensato per essere portato avanti da persone senza una tale competenza informatica. Da una parte, si sono dovute fare delle scelte di tag obbligate per rispettare il vocabolario ottimizzato per EVT, mentre tag differenti sarebbero stati più adatti: ad esempio, si è dovuto marcare i termini in lingua straniera con il tag `<term>` seguito dall’attributo “ref” invece di `<foreign xml:lang>`, per poterne costruire una *list of entities*. Dall’altra parte, il vocabolario TEI in sé ha posto dei limiti non permettendo, ad esempio, di annidare i tag `<said who>` l’uno dentro l’altro, e quindi di marcare il meccanismo narrativo del racconto nel racconto, caratterizzato dalla presa di parola diretta di un personaggio all’interno del corpo del testo. L’espedito è nondimeno molto significativo nel corpus Chimenti, e si ispira non a caso al modello del racconto orientale come le *Mille e una notte*. In effetti, nei testi di Chimenti spesso troviamo raccolte che partono da una cornice narrativa per poi annidarvi all’interno ulteriori livelli di narrazione. Lo stesso racconto *Histoire du prince Moustapha et de la princesse Djamila*, oggetto della seconda edizione presentata, ne è un esempio. Il racconto è inserito ne *La veillée du harem*, raccolta che si apre con una cornice narrativa nella quale tutti i racconti che seguono sono iscritti. A sua volta, all’interno dei racconti sono presenti altri racconti narrati dai personaggi. Ancora, all’interno di questi racconti sono presenti dialoghi.

In maniera generale, è fondamentale decidere da subito un insieme di tag, che sia ragionevole e ben delimitato. Ciò ha due vantaggi diretti, e da un punto di vista del lavoro umano e dell’agilità dei passaggi del workflow. Il primo vantaggio evidente è che un insieme di tag delimitato ed agile facilita l’apprendimento e la padronanza del sistema di codifica. Dall’altra, un sistema complesso di codifica si ripercuote sulla produzione dei file XML, rischiando di generare documenti estremamente complessi, difficili da leggere e manipolare. Nel primo workflow, il documento XML/TEI dell’edizione può essere manipolato solo tramite trasformazioni XSLT, il che fa perdere al file XML le sue caratteristiche user-friendly – già relative. Nel caso del secondo workflow, è stato scelto un insieme di tag limitati, il che ha permesso che il lavoro fosse portato avanti da una sola persona in modo sì più ridotto ma ordinato e producendo infine un documento XML/TEI ben leggibile e manipolabile.

In conclusione, è risultato evidente che la formalizzazione dei dati permette di far emergere alcuni aspetti oscurandone necessariamente altri, oltre che facendo perdere alcune dimensioni interpretative [20, 23]. È estremamente complesso (se non impossibile) rappresentare tutta la varietà del corpus con un unico sistema di codifica. Il prossimo obiettivo sarebbe dunque di mettere a punto un nuovo *workflow* che permetta di differenziare l’annotazione in base alle tipologie di testo, mantenendo le diverse rappresentazioni coerenti ed omogenee tra loro. L’annotazione strutturale dovrebbe essere separata dall’annotazione critica, a sua volta variabile in base alle necessità. Infine, un’ulteriore sfida è quella della interdisciplinarietà. Trattandosi di un corpus ibrido, è importante che il lavoro presupponga un dialogo tra specialisti di ambiti diversi, anche geograficamente distanti. Un fitto scambio interdisciplinare è d’altronde pratica connaturata ai progetti di informatica umanistica, e in una misura inedita per il campo strettamente letterario [20, 21]. Tuttavia, nel pratico tale lavoro d’équipe non sempre è semplice da realizzare: i contesti di lavoro nei diversi lati del Mediterraneo non hanno accesso agli stessi strumenti, metodi ed infrastrutture, il che genera disparità a tutti i livelli del trattamento informatico dei dati.

## 5. CONCLUSIONI

Il plurilinguismo, l’interdisciplinarietà, la transnazionalità sono alcuni dei volti dello spazio mediterraneo, che se da una parte lo rendono un campo di studi fertile e proficuo, allo stesso tempo ne determinano i limiti e le difficoltà di azione. Fare rete in uno spazio simile vuol dire ancora oggi trovarsi davanti continue frontiere. Ad un livello metodologico, ancora le storie letterarie sono strettamente nazionali, così come rigidi sono i confini disciplinari, e i profili – di autori/autrici e di ricercatori/ricercatrici – che attraversano tali confini hanno ancora difficoltà ad essere integrati nel sistema. Ad un livello

scientifico, l'ibridismo e il *métissage* linguistici e culturali, se da un lato sono sempre più oggetto di studio nella ricerca, dall'altro pongono ancora diverse sfide, teoriche e pratiche, rispetto al modello di un monolinguisimo occidentale. Ad un livello tecnico, è complesso rappresentare la pluralità mediante un sistema di rappresentazione formale. Se da una parte è evidente che per definire necessitiamo di categorie determinate e delimitate, dall'altra è fondamentale riflettere, oltre che sulle possibilità che tali categorizzazioni aprono, anche sui limiti che portano con sé.

## BIBLIOGRAFIA

- [1] Aghzal, M., M. A. E. Bouni, S. Driouech, e A. Mourhir. «Compact Transformer-based Language Models for the Moroccan Darija». In *7th IEEE Congress on Information Science and Technology (CiSt)*, 299–304. Agadir - Essaouira, Morocco, 2023. <https://doi.org/10.1109/CiSt56084.2023.10409912>.
- [2] Bahri, Farid. *Tanger. Une histoire-monde du Maroc*. Paris: Les Éditions Bibliomonde, 2022.
- [3] Bhabha, Homi K. *The location of culture*. London/New York: Routledge, 1994.
- [4] Braudel, Fernand. *La Méditerranée et le monde méditerranéen à l'époque de Philippe II*. Paris: Armand Colin, 1949.
- [5] Budor, Dominique, e Walter Geerts, (a cura di). *Le texte hybride*. Paris: Presses Sorbonne Nouvelle, 2004.
- [6] Cederna, Camilla M. «Elisa Chimenti (Naples, 1883-Tanger, 1969): la traversée infinie de l'écriture entre les mondes, entre les langues». In *Plein feux sur les femmes (in)visibles*, a cura di Elsa Chaarani, Laurence Denooz, e Sylvie Thiéblemont-Dollet, 529–44. Nancy: Presses universitaires de Lorraine, 2021.
- [7] Cederna, Camilla M. «L'écriture d'exil d'Elisa Chimenti, une mosaïque des voix et de créations des femmes entre métissage et transgression». *Atlante. Revue d'Études romanes* 18 (2023a).
- [8] Cederna, Camilla M. «Un projet international de recherche et d'édition pour la valorisation et la sauvegarde de l'œuvre d'Elisa Chimenti». *Atlante. Revue d'Études romanes* 18 (2023b).
- [9] Chambers, Iain. *Mediterranean Crossings: The Politics of an Interrupted Modernity*. Durham, NC: Duke University Press, 2008.
- [10] Chimenti, Elisa. *Anthologie*. Mohammedia/Casablanca: Éditions du Sirocco/Senso Unico Éditions, 2009.
- [11] Deplano, Valeria. «Archivi d'Africa. Le carte dell'amministrazione coloniale in Italia e nei territori di nuova indipendenza». In *Archivi sul confine. Atti del convegno internazionale, Torino, Archivio di Stato, 6-7 dicembre 2017*, 41–52. Roma: Ministero per i beni e le attività culturali e per il turismo. Direzione generale archivi, 2019.
- [12] Desideri, Ada, e Pauline Modolo. «Traitements informatiques des données d'un traitement de texte à une structuration XML TEI pour la création d'interfaces de consultation en HTML». *Atlante. Revue d'Études romanes* 18 (2023).
- [13] Eichhorn, Kate. *The Archival Turn in Feminism: Outrage in Order*. Temple University Press, 2013.
- [14] Fabre, Thierry, e Robert Ilbert. *Regards croisés sur la Méditerranée. «Les» représentations de la Méditerranée*. Paris: Maisonneuve & Larose, 2000.
- [15] Foucault, Michel. *L'archéologie du savoir*. Paris: Gallimard, 1969.
- [16] Hogan, Christophe. «OCR for Minority Languages». In *1999 Symposium on Document Image Understanding Technology*. Annapolis: Maryland, 1999.
- [17] Laplantine, François, e Alexis Nouss. *Le Métissage*. Paris: Flammarion, 1997.
- [18] Matvejevic, Predrag. *Mediterraneo. Un nuovo breviario*. Milano: Garzanti, 1991.
- [19] Miège, Jean-Louis, Georges Bousquet, Jacques Denarnaud, e Florence Beaufre. *Tanger. Porte entre deux mondes*. Paris: ACR édition internationale, 1992.
- [20] Moretti, Franco. *Falso movimento. La svolta quantitativa nello studio della letteratura*. Milano: Nottetempo, 2022.
- [21] Moretti, Franco. *La letteratura vista da lontano*. Torino: Einaudi, 2005.
- [22] Moussa, H.N., e A. Mourhir. «Named Entity Recognition in the Moroccan Dialect». In *7th IEEE Congress on Information Science and Technology (CiSt)*, 282–86. Agadir - Essaouira, Morocco, 2023. <https://doi.org/10.1109/CiSt56084.2023.10409973>.
- [23] du Mouza, Cédric, Stéphane Lamassé, e Philippe Rygiel. «De l'histoire numérique à l'histoire données?» 42 (2023).
- [24] Murrieta Flores, Patricia. «Language as cosmovisión/Cēmānāhuac. Some reflexions about computational research on the “Conquest of America” and the cruciality of decolonial praxis in the Digital Humanities». In *DH Benelux*. Bruxelles, Belgio, 2023.
- [25] Nahli, O., E. Gugliotta, G. Benotti., e N. Khlif. «Challenges and Progress in Constructing Arabic Dialect Corpora and Linguistic tools: A Focus on Moroccan and Tunisian Dialects». In *7th IEEE Congress on Information Science and Technology (CiSt)*, 293–98. Agadir - Essaouira, Morocco, 2023. <https://doi.org/10.1109/CiSt56084.2023.10410009>.
- [26] Sapegno, Maria Serena. «Uno sguardo di genere su canone e tradizione». In *Dentro/fuori, sopra/sotto: critica femminista e canone letterario negli studi di italianistica*, a cura di Alessia Ronchetti e Maria Serena Sapegno. Ravenna: Longo Editore, 2007.
- [27] Tamburini, Francesco. «Le istituzioni italiane di Tangeri (1926-1956): “quattro noci in una scatola”, ovvero, mancati strumenti al servizio della diplomazia». *Africa: Rivista trimestrale di studi e documentazione dell'Istituto italiano per l'Africa e l'Oriente* 61, fasc. 3/4 (2006): 396–434.

- [28] Vallarano Bianca. «“Oltre la lingua franca”. Il plurilinguismo mediterraneo di Elisa Chimenti». *Annali. Sezione romanza* 65, fasc. 1 (2023): 139–56.
- [29] Windsor, Leah C. «Bias in Text Analysis for International Relations Research». *Global Studies Quarterly* 2, fasc. 3 (2022).