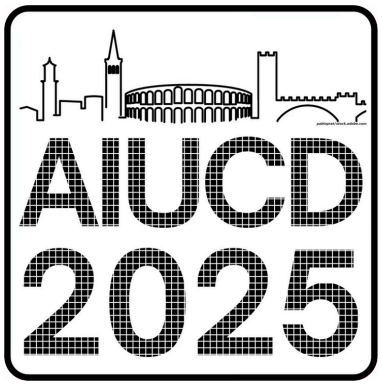

Atti del XIV Convegno Annuale

Diversità, Equità e Inclusione: Sfide e Opportunità per l'Informatica Umanistica nell'Era dell'Intelligenza Artificiale

Verona :: 11-13 giugno 2025

A cura di:
Simone Rebora • Marco Rospocher • Stefano Bazzaco

UNIVERSITÀ di VERONA
Dipartimento di LINGUE E LETTERATURE STRANIERE

daih

ASSOCIAZIONE per l'INFORMATICA UMANISTICA e la CULTURA DIGITALE

ISBN 978-88-942535-9-7

Copyright ©2025 AIUCD
Associazione per l'Informatica Umanistica e la Cultura Digitale

Il presente volume e tutti i contributi sono rilasciati sotto licenza Creative Commons Attribution ShareAlike 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: Simone Rebora; Marco Rospocher; Stefano Bazzaco (2025). Diversity, Equity, and Inclusion: Challenges and Opportunities for Digital Humanities in the Age of Artificial Intelligence, Proceedings del XIV Convegno Annuale AIUCD, Verona 11-13 giugno 2025, Università di Verona.

Ultimo accesso agli URL in data 8 maggio 2025.

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

Il programma della conferenza AIUCD 2025 è disponibile online
https://aiucd2025.dlls.univr.it/detailed-schedule/
The AIUCD 2025 Conference Program is available online
https://aiucd2025.dlls.univr.it/en-gb/detailed-schedule/

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, effettuata dai membri del Comitato di Programma sotto la supervisione del Comitato Scientifico di AIUCD 2025.
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, carried out by the members of the Programme Committee under the supervision of the Scientific Committee of AIUCD 2025.

Gli atti del convegno AIUCD 2025 sono pubblicati come raccolta di contributi in formato PDF forniti direttamente dagli autori e dalle autrici. I file sono stati raccolti e assemblati senza interventi redazionali da parte dei curatori.
The proceedings of the AIUCD 2025 conference are published as a collection of PDF contributions provided directly by the authors. The files have been collected and compiled without editorial intervention by the editors..

Il logo di AIUCD 2025 include l'immagine "Verona Dark Line Simple Minimalist Skyline With White Background" di @pabloprat/stock.adobe.com, ottenuta tramite la licenza Adobe Stock dell'Università di Verona.
The AIUCD 2025 logo includes the image "Verona Dark Line Simple Minimalist Skyline With White Background" by @pabloprat/stock.adobe.com, used under the Adobe Stock license of the University of Verona.

Il background della copertina è stato creato con tecniche di AI generativa con lo strumento "Magic Media" disponibile su Canva, usando un prompt con il tema del convegno.
The background of the cover was created using generative AI techniques with the "Magic Media" tool available on Canva, using a prompt based on the conference theme.

Comitato Organizzatore / *Organizing Committee*

General Chairs
Simone Rebora (Università degli Studi di Verona)
Marco Rospocher (Università degli Studi di Verona)

Local Chair
Anna Cappellotto (Università degli Studi di Verona)

Registration Chair
Giorgia Pomarolli (Università degli Studi di Verona)

Proceedings Chair
Stefano Bazzaco (Università degli Studi di Verona)

Sponsorship Chair
Matteo Lissandrini (Università degli Studi di Verona)

Publicity Chair
Sabrina Piccinin (Università degli Studi di Verona)

Comitato Scientifico / *Scientific Committee*

Program Chairs
Simone Rebora (Università degli Studi di Verona)
Marco Rospocher (Università degli Studi di Verona)

Digital Humanities e inclusione / *Inclusive DH*
Stefano Bazzaco (Università degli Studi di Verona)
Massimo Salgaro (Università degli Studi di Verona)

Archivi ed Edizioni Digitali / *Archives and Digital Editions*
Elisa Cugliana (Cologne Center for eHumanities)
Christian D'Agata (Università di Catania)

Metodi Computazionali / *Computational Methods*
Rachele Sprugnoli (Università degli Studi di Parma)
Sara Tonelli (Fondazione Bruno Kessler)

Rappresentazione di Dati e Conoscenza / *Data and Knowledge Representation*
Francesco Mambrini (Università Cattolica del Sacro Cuore)
Elena Spadini (Universität Bern)

Preservazione della Memoria e del Patrimonio Digitale / *Preservation of Memory and Digital Cultural Heritage*
Monica Berti (Universität Leipzig)
Daria Spampinato (Istituto di Scienze e Tecnologie della Cognizione-CNR)

Comitato di programma / *Program committee*

Stefano Allegrezza (Università di Macerata); Laura Antonietti (Université de Versailles Saint Quentin en Yvelines (Université Paris-Saclay); Luigi Bambaci (École pratique des hautes études, PSL); Liborio P. Barbarino (Università di Catania); Nicola Barbuti (Università degli Studi di Bari Aldo Moro (Dipartimento di Ricerca e Innovazione Umanistica); Sofia Baroncini (Leibniz Institute of European History); Andrea Bellandi (Institute for Computational Linguistics (CNR); Mario A. Bochicchio (University of Bari, Dep.t of Computer Science); Andrea Bolioli (Independent researcher); Marco Bombieri (University of Verona); Paolo Bonora (Università di Bologna); Flavia Bruni (Università di Chieti-Pescara); Marina Buzzoni (Ca' Foscari University of Venice); Alberto Campagnolo (KU Leuven); Anna Cappellotto (Università di Verona); Emmanuela Carbé (Università di Siena); Vittore Casarosa (ISTI-CNR); Raffaele Cioffi (Università di Napoli Federico II); Fabio Ciotti (Università di Roma Tor Vergata); Vincenzo Colaprice (University of Turin); Giuseppe Consolo (Università degli studi di Napoli, Federico II); Elisa Conti (Università di Catania); Salvatore Cristofaro (CNR ISTC); Giulia D'Agostino (TU Darmstadt); Elisa D'Argenio (HUN-REN Hungarian Research Centre for Linguistics); Enrico Daga (The Open University); Stefano Dall'Aglio (Università Ca' Foscari Venezia); Marilena Daquino (University of Bologna); Mauro De Bari (University of Bari Aldo Moro); Angelo M. Del Grosso (CNR-ILC); Matteo Di Franco (Università di Napoli Federico II); Giorgio Maria Di Nunzio (University of Padua); Stefano Ferilli (University of Bari); Lorenzo Ferroni (Università degli Studi di Verona); Franz Fischer (Ca' Foscari Università Ca' Foscari); Greta H. Franzini (Eurac Research); Francesca Frontini (CNR-ILC); Daniele Fusi (VeDPH, Stuttgart University); Mariangela Giglio (University of Bologna); Tiago Luis Gil (University of Brasilia); Luca Giovannini (University of Potsdam); Milena Giuffrida (Università di Catania); Edmondo Grassi (Università Telematica San Raffaele Roma); Miryam Grasso (Università di Catania); Piergiovanna Grossi (Università di Verona); Fahad Khan (CNR-ILC); Michele Lacriola (Università di Siena); Maurizio Lana (Univ. del Piemonte Orientale); Federica Lazzerini (Università degli Studi di Torino); Eleonora Litta (Università Cattolica del Sacro Cuore, Milano); Dominique Longrée (ULiège); Diego Mantoan (University of Palermo); Anna Maria Marras (University of Turin); Cristina Marras (CNR); Pietro Mazzarisi (University of Trieste); Barbara McGillivray (King's College London); Federico V. Meschini (Tuscia University); Alessio Miaschi (Istituto di Linguistica Computazionale "A. Zampolli" (CNR-ILC), Pisa); Andrea Micheletti (University of Padua); Giulia Miglietta (Università del Salento); Paolo Monella (Università Kore di Enna); Johanna Monti (Università degli Studi di Napoli "L'Orientale"); Rossana Morriello (Università degli Studi di Firenze); Gloria Mugelli (ILC CNR); Serge Noiret (AIPH (Associazione Italiana di Public History) (European University Institute); Giuseppe Palazzolo (Università di Catania); Mafalda Papini (CNR-ILC); Enrico Pasini (UniTO/CNR-ILIESI); Giulia Pedonese (CNR (Istituto di Linguistica Computazionale "Antonio Zampolli"); Paola Peratello (Università Ca' Foscari Venezia); Federico Pianzola (University of Groningen); Chantal Pivetta (Lund University (Sweden)); Igor Pizzirusso (AIPH); Giulia Re (ILC-CNR); Giulia Renda (University of Bologna); Dario Rodighiero (University of Groningen); Roberto Rosselli Del Turco (Università di Torino); Enrica Salvatori (Università di Pisa); Emilio M. Sanfilippo (CNR); Eva Sassolini (CNR-ILC); Andrea Schimmenti (University of Bologna); Flavia Sciolette (CNR-ILC); Pietro Sichera (ILIESI-CNR); Daniele Silvi (Università di Roma 'Tor Vergata'); Giulia Speranza (University of Naples L Orientale); Francesco V. Stella (UNISI); Timothy Tambassi (Ca' Foscari University of Venice); Mirko Tavosanis (Università di Pisa); Francesca Tomasi (Università di Bologna); Simona Turbanti (University of Milano); Marco Venuti (Università di Catania); Gennaro Vessio (University of Bari Aldo Moro); Gabriele Vezzani (Università di Verona / RWTH Aachen University); Fabio Vitali (University of Bologna).

Enti organizzatori / *Organisers*

AIUCD; Università di Verona: Dipartimento di Lingue e Letterature Straniere; Digital Arena for Inclusive Humanities (DAIH).

Sommario

Prefazione

Simone Rebora, Marco Rospocher, Stefano Bazzaco I-II

Digital Humanities and Inclusion

PrevNet. A FAIR and inclusive resource for the study of preverbs in historical languages 2
Andrea Farina, Barbara McGillivray

Potential bias in AI: cultural representation and the marginalization of African art 11
Francesca Bignotti

Exploring data-driven narratives in Digital Humanities web-based projects: features and impact 18
Tommaso Battisti, Marilena Daquino

Il Glossario delle Infrastrutture di Ricerca (GIR) 24
Lucia Francalanci, Alessia Scognamiglio, Irene Falini, Pietro Restaneo, Giulia Pedonese, Alessia Spadi

Educational Impact of Storytelling and Data Visualisation in the Interpretation of Humanities Data 29
Giulia Renda, Marilena Daquino

IncluInstIT: Un nuovo corpus per lo studio di linguaggio inclusivo su Instagram 35
Irene Caiazzo, Giovanna Maria Dimitri, Liana Tronci

Per un'analisi della rivista Umanistica Digitale in ottica DEI 40
Rossana Morriello, Lucia Sardo

Carpe bias, quam minimum credula queries 46
Sabato Danzilli

Evaluating bias within an epistemological framework for AI-based research in the humanities 52
Sarah Oberbichler, Cindarella Petz

Semplificare la lettura dei manoscritti utilizzando tecnologie WEB interattive e interazioni «hover» 60
Giacomo Marchioro, Andrea Brugnoli, Francesca Carnazzi, Paolo Pellegrini, Edoardo Ferrarini

Accessibilità e inclusione per la documentazione del restauro: gli archivi del Centro Conservazione e Restauro «La Venaria Reale» 66
Stefania De Blasi, Edi Guerzoni, Chiara Pipino

Il patrimonio culturale digitale delle minoranze etniche: il progetto DIGICHer tra le comunità Sámi, ladine ed ebraiche in Europa 72
Matteo Cova, Eleonora De Longis

Fulfilling GEN-der AIms: do image-generating tools discriminate? An on-field study 78
Francesco Meledandri

From Bias Paralysis to Bias as a Category of Analysis. Introducing the Bias-Aware Framework 86
Mrinalini Luthra, Amber Zijlma

Supporting Children with Linguistic Vulnerabilities Through Advanced, Theory-Driven Technological Solutions: The TELMI Approach for Italian children with DLD and Children with Italian as L2 94
Arianna Compostella, Giulia Valcamonica, Mattia Gianotti, Matteo Secco, Silvia Silleresi, Fabrizio Arosio, Franca Garzotto, Maria Teresa Guasti

Tecnologie AI per la didattica 104
Gabriele Prosperi, Giulia Miglietta, Eleonora Miccoli, Mario Bochicchio

DEA - An Innovative Technological Tool for Personalized Linguistic Training for Italian Children with Developmental Dyslexia 110
Marta Tagliani, Maria Vender, Giulia Valcamonica, Giovanni Caleffi, Franca Garzotto, Denis Delfitto

Il ruolo delle Infrastrutture nella costruzione di un ambiente di ricerca inclusivo. Un modello di buone pratiche 118
Marta Caradonna, Nicola Giampietro, Roberta Bianca Luzietti, Monica Monachini, Valeria Quochi, Emiliano Degl'Innocenti, Alessia Spadi, Alessandra Caravale, Antonio D'Eredità, Paola Moscati, Giacomo Mancuso

Predicting Grammatical Cases in Slovenian Varieties in Italy: A Use Case from the LORIS 1.1 Language Assistant 124
David Bordon

Verso un futuro senza barriere: l'accessibilità dei documenti elettronici nell 'European Accessibility Act 129
Stefano Allegrezza

La gestione del nuovo sapere digitale contemporaneo. Scenari, criticità, sfide, prospettive 134
Nicola Barbuti

Archives and Digital Editions

Learner Corpus of Creative Writing: An interdisciplinary challenge 143
Ioanna Tyrou, Katerina Florou

Retrieval-Augmented Generation systems for enhanced access to digital archives 149
Michele Ciletti

Preserving Clarity: The MAGIC project approach to ancient manuscripts 156
Yahya Momtaz, Stefania Conte, Guido Russo

Digital Explorations of Historical Multilingual Practices. The Challenges of the HyperAzpilcueta Project 160
Manuela Bragagnolo, Marcus Pöckelmann, Polina Solonets, Andreas Wagner

Digitalizzazione di un fondo archivistico per la creazione di un centro di documentazione digitale 165
Dario Baldini

Risorgimento Digitale: Un progetto di hyperedizione per i testi risorgimentali. Le Noterelle di Abba come caso di studio 171
Vincent Mobilia

SpaceLat: La geografia della letteratura latina tardoantica 177
Riccardo Consolini

Verso l'edizione digitale del carteggio Canneti-Fiacchi 184
Chiara Manca, Fiammetta Sabba, Bianca Sorbara, Silvia Tripodi

L'edizione critica digitale della 'Scienza nuova' di Giambattista Vico in Scholarly Digital Edition 190
Alessia Scognamiglio, Roberto Evangelista, Manuela Sanna, Salvatore Prinzi, Stefano Veneroni, Chiara Aiola, Luca de Santis

L'edizione digitale del papiro P.Tor.Choach. 12 in collaborazione con il Museo Egizio di Torino 197
Chiara Senatore

Dal palcoscenico al digitale: modelli di data visualization per la valorizzazione dell'Archivio Teatro delle Marionette di Gianni e Cosetta Colla 202
Elena Radaelli

Metodologie computazionali per l'organizzazione di archivi nati digitalmente 208
Mariangela Giglio

Archivi digitali per la conservazione e valorizzazione del patrimonio culturale: il caso del Santuario della Madonna di Carufo 215
Caterina Ciccotti

Il progetto ArPeR. Per un Archivio dei periodici romaneschi 222
Martina Ludovisi

Un modello integrato per il Roman d'Alexandre del codice Correr 1493: annotazione linguistica e edizione critica digitale 228
Giacomo Costa, Simone Zenzaro

Descrivere la catastrofe: documentare la diegesi per la catalogazione di opere distopiche e post-apocalittiche 232
Luca Paolo Bruno, Valeria Stabile, Juan Scassa, Carmelo Caruso, Ludovica Pannitto

Un'applicazione pratica per l'edizione digitale di testi agiografici e calendariali 238
Luca Avellis

Embracing flexibility: new EVT features for critical editing, accessibility and inclusivity 244
Roberto Rosselli Del Turco, Davide Cucurnia, Marina Buzzoni

Riscoperte poliane: edizione digitale di un manoscritto inedito de Il Milione 251
Giulia Fabbris, Samuela Simion, Fabio Soncin

«Proximior perfectioni»: criticità e future prospettive del progetto Dante Limina 257
Elisabetta Tonello

TEI Encoding as a Unified Structure for Multilingual Digital Editions: The LeggoManzoni Case Study 264
Mariia Levchenko, Beatrice Nava, Ersilia Russo

Verso l'implementazione di un sistema di riconoscimento di allusioni al lessico dantesco nelle testimonianze del Lager: il caso d'uso in Voci dall'Inferno 270
Carla Congiu, Angelo Mario Del Grosso, Marina Riccucci

Il progetto RETI (REndering Texts and Images): metodologia e primi risultati 276
Chiara Barbero, Matteo Di Franco, Federica Lazzerini, Annamaria Persia

Edizione digitale ed autorialità plurima: quali sfide? 283
Stefania Tesser

Il corpus di prosa letteraria del progetto RIND (1830-1930). Assunti teorici e vincoli pratici 289
Stefano Ondelli, Pietro Mazzarisi

ZoneRW: verso un'integrazione con Kraken ed eScriptorium per il riconoscimento e la gestione avanzata delle regioni di interesse 297
Pietro Sichera, Angelo Mario Del Grosso, Laura Mazzagufo, Daria Spampinato

L'edizione digitale dei primi 16 Taccuini di Paolo Orsi 303
Giuseppina Monterosso, Andrea Bolioli, Elisa Bonacini, Gianmario Cattaneo, Dario Gonella, Anna Maria Marras, Salvatore Spina, Paola Venuti

Artificial intelligence vs human handwriting: annotating damaged manuscripts 308
Dumitru Scutelnic, Laura Gazzani, Paolo Pellegrini, Claudia Daffara

Modellazione, interoperabilità e riuso in DiScEPT 314
Tiziana Mancinelli, Hansmichael Hohenegger, Federico Boschetti, Angelo Mario Del Grosso, Eleonora De Longis, Gloria Mugelli

Navigating the Digital Transition: Lessons from a Hybrid Critical Edition Project 319
Elisa Bastianello, Reto Baumgartner

Human–LLM Synergy in Higher Education Publishing: Two ChatGPT Use Cases within Editorial Pipelines 327
Gianluca Pavani

Taming the Hydra: A Model for Textual Dynamics and Constellations of Goethe's Venetian Epigrams 334
Daniele Fusi, Matteo Zupancic, Franz Fischer, Claus Zittel

Computational Methods

The Influence of AI Tools on University Students' Writing Style: A Stylometric Analysis of Narrative Texts 343
Dimitris Bilianos, Katerina Florou

Usare i Large Language Model per l'analisi del testo narrativo: strategie di prompt engineering per il riconoscimento del discorso indiretto libero nella narrativa italiana 1830-1930 349
Aurora Argenzio, Fabio Ciotti, Anna Chiara Corradino

Historical GIS e metodologie digitali per una storia della copertura boschiva 357
Vincenzo Colaprice

Experiments on the Use of LLMs for the Translation of the Babylonian Talmud 363
Mafalda Papini, Davide Albanesi, David Dattilo, Emiliano Giovannetti, Simone Marchi

Metodi di allineamento testuale bilingue per un'edizione genetica digitale dei Mémoires di Carlo Goldoni 368
Matteo Zibardi

Eastern Law in Western Words: Analyzing Roman Legal Terminology in Medieval Charters 375
Tamás Kovács, Angelos Nikolaou, Johannes Laroche, Georg Vogeler

Il corpus del Digesto: approcci e metodi computazionali per la creazione di risorse linguistiche 379
Alessandra Cinini, Paola Marongiu, Eva Sassolini

Preliminary Results for the Explanation of Neural Network-based Handwriting Identification in Historical Manuscripts 386
Riccardo De Cesaris, Valerio Caravani, Arianna Pastorini, Serena Ammirati, Paolo Merialdo

From Documents to Data: Digital Technologies in the Study of Notarial Charters 392
Franziska Decker, Sandy Aoun, Giuseppe Consolo

Verso la svolta computazionale della critica dantesca 397
Fara Autiero, Vittorio Celotto, Gennaro Ferrante, Chiara Fusco, Sandra Gorla, Giuseppe Andrea Liberti, Mariangela Palomba, Serena Picarelli, Stefano Angelo Rizzo, Silvia Tripodi

Phylo-1-preview. Un modello T5-Base per l'emendazione dei testi antichi 404
Giuseppe Ferrara

«Glottolab: A Linguistic Adventure»: Lo sviluppo di un'attività gamificata incentrata sulla linguistica 411
Cecilia Cattaneo, Claudia Roberta Combei, Chiara Zanchi

Concordanze e NLP: idee, metodi e regole per l'applicazione alla lingua italiana 419
Pietro Sichera, Christian D'Agata, Giuseppe Palazzolo

Reverse Engineering Critical Apparatuses for HTR Ground Truth Creation: The Case of Kennicott's Collation of the Hebrew Bible 426
Luigi Bambaci, Nachum Dershowitz, Daniel Stökl Ben Ezra

Data and Knowledge Representation

Prototyping an Atlas of Early Modern English Drama: An Experiment on DraCor Data 435
Luca Giovannini, Andreas Wagner

ATLAS: A data model for describing FAIR Digital Humanities research outcomes 440
Chiara Martignano, Giorgia Rubin, Sebastiano Giacomini, Alessia Bardi, Marina Buzzoni, Marilena Daquino, Riccardo Del Gratta, Angelo Mario Del Grosso, Franz Fischer, Roberto Rosselli Del Turco, Francesca Tomasi

Cantautorato e Digital Humanities. Per una valorizzazione dell'opera di Fabrizio De André, Lucio Dalla, Gianmaria Testa 448
Marcello Ranieri

OWL Ontology on the European Integration Process between 1949 and 1979 453
Lorenzo Galvagno

LiITA, una Knowledge Base di risorse interconnesse per l'italiano 460
Eleonora Litta, Marco Passarotti, Paolo Brasolin, Valerio Basile, Cristina Bosco, Andrea Di Fabio

Dai limina a LiMINA: un database per i marginalia alla Commedia 466
Serena Malatesta, Beatrice Mosca

Dai Materiali Didattici alle Piattaforme FAIR: Costruire un'Infrastruttura di Training in H2IOSC 473
Giulia Pedonese, Francesca Frontini, Roberta Ottaviani, Federico Boschetti, Alessia Spadi, Lucia Francalanci, Alessia Scognamiglio, Pietro Restaneo, Antonina Chaban, Jana Striova, Laura Benassi

IlluminAI: un sistema di navigazione interattivo per i manoscritti miniati rinascimentali 478
Valeria Minisini, Giorgio Gosti, Bruno Fanini

Making Germanic Cultural Heritage accessible to students: a proposal for a case study 485
Chiara De Bastiani, Giulia Fabbris

/DH.arc Vocabularies: Making semantic artefacts more visible and accessible using SKOS 492
Laurent Fintoni

Modeling an Ontology for Heritage Science: Challenges and Key Strategies 499
Erica Scarpa, Riccardo Valente, Irene Rossi

Linked Open Data and IIIF for connecting manuscripts images with their transcriptions: a case study from the Veneranda Biblioteca Ambrosiana 505
Lorenza Talarico

Describing Monastic Iconography Using Semantic Data: A Preliminary Investigation 511
Sofia Baroncini, Francesco Mele

A Linguistic Knowledge Graph of Word Borrowings from Portuguese 519
Anas Fahad Khan, Ana Salgado

E.T and Visual Culture Ontology (ETVCO): Perspectives on Extraterrestrial Influence in Visual Heritage 524
Kaosaier Wusiman, Simone Casazza

Automating XML-TEI Encoding of Unpublished Correspondence: A Comparative Analysis of two LLM Approaches 531
Marco De Cristofaro, Daniel Zilio

Modelli e tecnologie integrate e innovative per una cittadinanza digitale equa e sostenibile 538
Cristina Marras, Vittoria Fabiani, Enrico Pasini, Lisa Reggiani, Pietro Sichera, Paolo Ongaro, Martina Rossi

A Case Study in Cultural Heritage: A System Linking Three Open Data Tools – Digital Philology for Dummies (DPhD), Edition Visualization Technology (EVT), and a Relational Database 544
Renato Caenaro, Chantal Pivetta

Modeling Intermediality and Interpretations in Contemporary Combinatory Literature: Revealing Il Giuoco dell'Oca by Edoardo Sanguineti 551
Enrica Bruno, Maria Francesca Bocchi, Francesca Tomasi

From Metadata to Storytelling: A Framework For 3D Cultural Heritage Visualization on RDF Data 558
Sebastian Barzaghi, Simona Colitti, Arianna Moretti, Giulia Renda

Between Text and Icon: Towards A Representational Model for Ekphrastic Relations 566
Maria Francesca Bocchi, Carlo Teo Pedretti, Fabio Vitali

Preservation of Memory and Digital Cultural Heritage

Life and Death of DH Projects: A Preliminary Investigation of Their Lifecycles in Italy 575
Erica Andreose, Giorgia Crosilla, Remo Grillo, Gianmarco Spinaci

Research on Street Art in the Digital Space 581
Aleksandra Tselikova

Motion Visualisation of Dancers' Performances 587
Giacomo Alliata, Loïc Serafin, Samy Mannane, Sarah Kenderdine

Entità in relazione: policies, soluzioni tecnologiche e modelli lessicali per un (eco)sistema informativo integrato 593
Herbert Natta, Michela Tardella, Eleonora Lattanzi, Gianluca Rossi, Roberta Maggi

Preserving and enhancing cultural heritage: the Digest project 600
Alessandra Cinini, Paola Marongiu, Eva Sassolini, Monica Monachini

The Staccioli Digital Archive: Using Knowledge Graphs to power digital art history catalogues and art exhibitions 608
Klaus Werner, Pietro Liuzzo, Alessandro Adamou

Fantàsime: Interactive Drama per la valorizzazione del Patrimonio Culturale 613
Maria Chiara Provenzano, Eleonora Miccoli, Mario A. Bochicchio

Soluzioni phygital e mediazione culturale: riflessioni digiteconomiche nell'era dell'IA 619
Nicola Barbuti, Mauro De Bari

MeMo: Una mappa letteraria digitale per la memoria del Mezzogiorno 627
Laura Giurdanella, Giuseppe Palazzolo, Bernardo De Luca, Fara Autiero, Marco Gatto, Sabatino Peluso, Concetta Maria Pagliuca, Andrea Schembari

The relationship between art and sound: An experiment on the engagement of the cultural tourist 633
Sara Benetti, Nicola Orio

Analisi RTI delle iscrizioni runiche del Leone del Pireo (Arsenale di Venezia) 639
Paola Peratello, Elisa Corrò

Torino anni Ottanta. Digitalizzazione del patrimonio documentario e ricostruzione virtuale delle mostre negli spazi pubblici e privati 646
Filippo Yahia Masri

Linguistica dei corpora e informatica umanistica per la valorizzazione plurilingue del patrimonio culturale: implementazione del progetto UniVOCIttà 651
Rita Gramellini, Valeria Zotti

«Il mio sommario dunque è tutto qui?» Per Franco Fortini 658
Emmanuela Carbé, Mariangela Giglio, Pietro Orlandi, Jacopo Maria Romano, Giulio Quaresima

Per un'analisi della rivista *Umanistica Digitale* in ottica DEI

Rossana Morriello[1], Lucia Sardo[2],
[1] Università degli Studi di Firenze, Italia – rossana.morriello@unifi.it
[2] Alma Mater Studiorum Università di Bologna, Italia – lucia.sardo@unibo.it

ABSTRACT (ITALIANO)
L'intervento presenta il risultato di una analisi sugli articoli della rivista Umanistica Digitale dalla sua fondazione a oggi in un'ottica DEI, considerando che questo è il tema del convegno che sarà organizzato nel 2025. L'analisi ha rilevato una grande ricchezza di tematiche trattate ma una scarsa presenza di articoli dedicati alle tematiche DEI; inoltre, l'utilizzo di parole chiave non uniformate rende difficoltosa la ricerca di testi su argomenti simili e riduce la *discoverability* degli articoli. Il contributo propone l'uso di un vocabolario strutturato per la corretta disseminazione delle tematiche DEI.
Parole chiave: DEI, Umanistica Digitale, keywords, text mining

ABSTRACT (ENGLISH)
For an analysis of Umanistica Digitale journal from a DEI perspective
The paper presents the result of an analysis on the articles of the journal Umanistica Digitale from its foundation to the present day from a DEI perspective, considering that this is the theme of the conference to be organised in 2025. The analysis revealed a great richness of topics treated but a low presence of articles dedicated to DEI topics; furthermore, the use of non-uniform keywords makes it difficult to search for texts on similar topics and reduces the discoverability of the articles. The contribution proposes the use of a structured vocabulary for proper dissemination of DEI topics.
Keywords: DEI, Umanistica Digitale, keywords, text mining

1. INTRODUZIONE
Le keywords svolgono diverse funzioni importanti nell'era dei dati e delle applicazioni dell'intelligenza artificiale, tra cui innanzitutto permettere la *discoverability* e la *findability* delle risorse attraverso i motori di ricerca e nei sistemi di information retrieval delle banche dati. Sono quindi alla base delle strategie di Search Engine Optimization (SEO) per diverse tipologie di risorse. Nell'ambito della ricerca accademica e delle banche dati scientifiche le keywords favoriscono quindi il recupero, la disseminazione e la condivisione dei risultati che è il fine ultimo della ricerca scientifica. Consentono di mettere in contatto ricercatori che si occupano di tematiche affini poiché tramite le parole chiave si possono cercare articoli di argomento simile a ciò che è di interesse, ed eventualmente citarli, contribuendo alla determinazione degli indicatori bibliometrici. La presenza di keywords favorisce l'applicazione di strumenti di data analysis e information retrieval avanzati, sempre più basati sull'intelligenza artificiale.
Per queste ragioni gli editori nelle loro banche dati, siano esse citazionali, testuali, bibliografiche, attribuiscono grande importanza a questa forma di indicizzazione semantica che consente loro di organizzare efficacemente i contenuti e di raggrupparli in cluster per agevolare innumerevoli forme di data analysis, oltre alla *discoverability*, e di aggiornare costantemente l'organizzazione dei contenuti digitali sulla base dell'evoluzione della ricerca scientifica.[1] Gli articoli contenuti nelle principali banche dati scientifiche sono infatti dotati di varie keywords, o key phrases, come sarebbbe più opportuno chiamarle poiché spesso si tratta di frasi composte da più parole (Hartley & Kostoff, 2003), insieme al titolo e all'abstract dell'articolo. In misura crescente, questi elementi paratestuali, e in particolare le key phrases, vengono utilizzate per l'assegnazione della peer review a determinati revisori, tramite il matching tra le keywords di un articolo e le keywords con le quali il revisore ha sintetizzato le sue competenze, come spesso avviene sulle piattaforme editoriali. Per motivi analoghi le keywords sono fondamentali nei progetti di ricerca che si sottopongono per il finanziamento e che richiedono revisione paritaria.
Per quanto riguarda il nostro macrotema, ovvero diversità, equità e inclusione (DEI) le keywords sono tra i metadati che consentono di ricondurre le pubblicazioni agli ambiti DEI, poiché permettono di costruire, insieme agli altri elementi paratestuali, il contesto e il significato necessari al lettore per accedere consapevolmente all'informazione su questi temi. Nell'acronimo DEI, con "diversità" si intende una serie di

[1] Per esempio, è quanto dichiara esplicitamente Elsevier nella banca dati SciVal, che offre soluzioni di research intelligence e data visualization sui dati di Scopus, a cui faremo riferimento più avanti in questo contributo.

differenze socio-economiche e culturali potenzialmente svantaggiose quali disabilità, età, genere, orientamento sessuale, status socio-economico, lingua, pratiche di alfabetizzazione, background educativo, localizzazione geografica, razza, etnia, cultura e religione (Raju 2020; Morriello & Sardo 2023), e l''inclusione" insieme all'"equità" sono aspetti essenziali e integrali ai processi volti all'incremento della diversità nel mondo dell'informazione, della ricerca e dell'educazione (Raju 2020). Implicita nell'acronimo è l'azione svolta per superare le barriere e migliorare la diversità. Dunque, non si tratta solo di mappare il soddisfacimento dei criteri di diversità, equità inclusione in un settore quale può essere il settore o i settori disciplinari di copertura di una rivista scientifica, ma anche e soprattutto di agire per potenziare la diversità. Con tale obiettivo, costruttivo e non certamente critico, proponiamo il nostro contributo.

2. OBIETTIVI E METODOLOGIA

L'intervento presenta un'analisi degli articoli della rivista *Umanistica Digitale* dalla sua fondazione a oggi, per rilevare quanto e come le tematiche riconducibili a diversità, equità e inclusione siano presenti e in che modo vengano indicizzate tramite le keywords.

L'analisi è stata condotta utilizzando le informazioni e i dati presenti sul sito della rivista stessa. Nello specifico sono stati presi in considerazione gli autori e le parole chiave inserite negli articoli, al fine di rilevare le keywords legate ai temi DEI. Inoltre, sono stati rilevati i dati relativi al genere fra coloro che hanno presentato contributi alla rivista.

La raccolta ed elaborazione dei dati è stata effettuata utilizzando il software open source OpenRefine, con una successiva integrazione e verifica manuale nel caso in cui sia reso necessario. A questa fase è seguita l'analisi di quanto emerso dai dati raccolti, con una riflessione particolare sull'uso delle keywords. La sezione finale dell'intervento introduce una proposta per un ripensamento nell'uso delle parole chiave, al fine di facilitare il reperimento di documenti rilevanti e il posizionamento degli studiosi/e nei rispettivi ambiti disciplinari.

I dati analizzati con OpenRefine sono poi stati confrontati con i dati di Scopus e di SciVal. Da settembre 2021, infatti, la rivista *Umanistica Digitale* è indicizzata nella banca dati citazionale Scopus di Elsevier e di conseguenza nella banca dati SciVal, dello stesso editore, che attinge ai dati di Scopus per offrire soluzioni di research intelligence e data visualization.

L'obiettivo di questo lavoro, come sopra anticipato, non è certo una critica delle manchevolezze ma mira piuttosto ad agevolare il rafforzamento della *findability* degli articoli, in particolare in prospettiva DEI, e rientra in un più ampio lavoro di ricerca in corso su alcune riviste scientifiche italiane di ambito umanistico.

Sono stati analizzati tutti gli articoli presenti sulla rivista *Umanistica Digitale* dal primo numero del 2017 al numero 18 del 2024. In totale si tratta di 205 articoli, nei quali sono comprese le introduzioni e le recensioni per quanto riguarda gli autori,[2] e nel caso in cui fossero presenti, le parole chiave significative.[3]

3. RISULTATI

Un primo dato da cui partire riguarda gli autori in una prospettiva di genere. In totale sono stati individuati 327 autori. Di questi, 162 sono donne e 165 uomini. L'analisi degli autori vede perciò una sostanziale parità fra i generi se si considerano tutti i coautori dell'articolo, mentre se si considera solo il primo (o l'unico) autore di un paper, le percentuali vedono le donne con una maggior presenza (54% circa).

Come sovente accade, sia nelle riviste sia nelle banche dati di altro genere, si è riscontrata anche in questo caso una mancanza di uniformità nel modo in cui vengono indicate le istituzioni di affiliazione, per cui non è stato possibile elaborare in questo senso i dati. Si evidenzia la necessità di trovare degli identificatori anche per gli enti per favorire la ricerca e l'individuazione univoca dei medesimi con maggior facilità. Le banche dati commerciali lavorano a uniformare le affiliazioni e presentano dati più precisi a riguardo, tuttavia, il dato non risultava particolarmente rilevante per le finalità del lavoro e si è deciso che fosse sufficiente individuare le affiliazioni alle università e non ai dipartimenti. Anche in questo caso si sono uniformate le modalità citazionali, prima utilizzando OpenRefine e quindi con una ulteriore verifica di eventuali match non riscontrati dal software.

[2] Secondo i dati di Scopus gli articoli rappresentano comunque il 95,9% delle tipologie in *Umanistica Digitale*.
[3] Sono state ovviamente escluse le parole chiave "recensione" o "review."

La maggior parte degli autori ha una affiliazione accademica o legata a centri e fondazioni accademiche; molti lavorano per istituzioni legate alla ricerca (CNR, CNRS, ad esempio); solo pochi lavorano in biblioteche o archivi o in istituti di istruzione superiore; ancora minore è il numero di autori che lavorano per aziende private, amministrazioni statali o locali. Alcuni si qualificano come ricercatori indipendenti.
Gli articoli presenti hanno da 1 a 11 autori. La maggior parte degli autori è responsabile di un solo paper pubblicato in *Umanistica Digitale*, nel senso che ne è l'unico autore o uno dei co-autori.
Utilizzando OpenRefine e il servizio di riconciliazione con Wikidata, è emerso che, in percentuale, il 43% ha un match sicuro, mentre per il restante 57% si tratta di riconciliare manualmente o di creare l'item in Wikidata (in questi casi il sistema di riconciliazione non ha considerato certo il possibile match, qualora presente, o è appunto necessario creare l'item).
Questo dato a nostro avviso è significativo perché implica la necessità di prevedere l'inserimento degli autori (e degli articoli) in Wikidata, al fine di favorire la creazione di LOD (Linked Open Data) a maggior vantaggio di tutta la comunità scientifica.
Rispetto alle parole chiave riscontrate, si è operata una minima uniformazione utilizzate in modo da poterle analizzare e rappresentare graficamente (figura 1).

Figura 1 – Word cloud delle principali parole chiave

I risultati dell'analisi hanno riscontrato l'assenza di keywords in 18 articoli, ma si tratta principalmente di introduzioni, recensioni, etc.
Nei 205 articoli della rivista si contano oltre 750 keyword diverse. Ciascun articolo presenta da un minimo di tre a un massimo di undici parole chiave, con una media di circa quattro keywords per articolo.
L'elemento di attenzione non è costituito tanto dal numero di keywords, fatto questo che potrebbe semplicemente dimostrare l'ampiezza del campo di studio e la grande varietà di tematiche affrontate dalla rivista, quanto la ripetitività, ovvero che molte di esse fanno riferimento sostanzialmente allo stesso ambito semantico, e allo stesso argomento, il quale viene però definito e declinato in modi diversi dai diversi ricercatori, rendendo di fatto impossibile una ricerca efficace per parole chiave, qualora si volesse utilizzare questa modalità per trovare articoli simili.
A titolo d'esempio, si trovano keyword quali: "Topic Model, Topic modeling, Topic modelling", "archivi digitali, archivi nativi digitali, Archivio digitale", codifica TEI, codifica XML-TEI, TEI encoding, XML/TEI, TEI", "Digital Scholarly Edition, digital scholarly editing, Digital Scholarly Editions".[4]

[4] La presenza di keyword in italiano o in inglese è dovuta alle scelte degli autori.

Con pochissime eccezioni non sono presenti keywords riconducibili all'ambito DEI.
I dati analizzati con OpenRefine sostanzialmente coincidono con quanto si può rilevare in Scopus e SciVal, perlomeno dal punto di vista quantitativo. Ciò che differisce è l'organizzazione delle parole chiave che, nei due prodotti commerciali, sono standardizzate e fortemente strutturate. Per quanto riguarda *Umanistica Digitale,* la banca dati SciVal presenta 98 topics e 65 topic clusters, in aree disciplinari concentrate ovviamente in prevalenza tra le scienze umanistiche e le scienze sociali.[5]
SciVal utilizza il suo Fingerprint Engine per attribuire delle keyphrases tramite text mining e varie tecniche di Natural Language Processing applicate ad articoli, abstract e dati bibliografici e confrontate con alcuni thesauri di riferimento, e per l'indicizzazione in Topics, Topic Clusters e Research Areas.
Per fare un esempio, usiamo il Topic *Digital; Critical Edition; Semantic Web* che è il più rilevante in termini di preminenza globale (su base citazionale naturalmente) rispetto alla rivista *Umanistica Digitale* (include dodici articoli) e viene definito nella banca dati come relativo a "the development and impact of digital scholarly editions in humanities research. It encompasses the use of digital methods to enhance access, preservation, and analysis of historical texts and manuscripts". Il Topic presenta le seguenti keyphrases: Digital Edition; Digital; Archival Document; Scholarly edition; Online; Critical edition; Semistructured Data; Review; Scholars; Collaborative; Software; Development Tool; Online Edition; Development Version; Textual Culture; Search Engine.
In sintesi, in aggiunta alle keywords presenti sulla rivista che vengono attribuite dall'autore o dagli autori dell'articolo, e che sono sempre presenti su Scopus, SciVal presenta diversi metodi di indicizzazione propri, arricchendo significativamente il campo semantico di ciascun articolo tramite molteplici keywords o keyphrases.
L'attenzione con cui l'editore della banca dati lavora sulle keywords ne dimostra l'importanza per la possibilità di scoperta e condivisione dei risultati della ricerca scientifica pubblicati. L'uso di diversi metodi massimizza *discoverability* e *findability* dei contenuti e consente innumerevoli efficaci rielaborazioni dei dati, soprattutto alla luce delle tecniche di text and data mining e delle applicazioni crescenti dell'intelligenza artificiale.
Per tornare alla rivista di nostro interesse, l'analisi sui dati di *Umanistica Digitale* in Scopus e SciVal conferma la quasi nulla presenza di tematiche legate a diversità, equità e inclusione. Questo dato viene confermato anche dalla mappatura degli SDG (Sustainable Development Goals) applicata agli articoli della rivista sempre in SciVal. Solo ventidue articoli sono riconducibili agli Obiettivi di sviluppo sostenibile dell'Agenda 2030 (Nazioni Unite 2015). Di questi uno è ricondotto all'SDG 10 – Ridurre le diseguaglianze, quindi direttamente ed esplicitamente legato ai temi DEI, e quattro articoli sono ricondotti all'SDG 16 – Pace, giustizia e istituzioni sane, un goal non direttamente e necessariamente rilevante per DEI ma in realtà i quattro articoli lo sono, se li si va a consultare.

4. CONCLUSIONI

Gli editori sono attenti e quindi in molti casi offrono indicazioni precise agli autori e di norma utilizzano un soggettario o un thesaurus per l'indicizzazione dei contenuti bibliografici digitali. Il modo in cui i thesauri coprono le tematiche DEI sarebbe da indagare ma esula dagli obiettivi di questo contributo. Senz'altro vi sono ormai diversi studi che si soffermano sui temi DEI nella comunicazione scientifica, alcuni dei quali sono stati riportati nell'introduzione. Anche l'Agenda 2030 e l'attribuzione degli SDG ai contenuti scientifici aiuta a individuarli.
L'assenza di parole chiave relative a DEI non implica necessariamente la scarsa attenzione a queste tematiche dagli articoli nelle riviste. Nonostante la loro importanza di "connettori tra l'articolo e il lettore" spesso gli autori sottovalutano l'importanza delle parole chiave, come indicano diversi degli studi citati, e allo stesso tempo in molti casi non vi sono linee guida e indicazioni specifiche da parte degli editori e dei manager delle riviste. Questo uso poco organizzato può avere ricadute sull'attribuzione di significato e sulla ricerca di informazione in generale, e in ambito DEI in particolare. La poca attenzione attribuita dagli autori alle parole chiave non permette di evidenziare tali specificità. L'assenza di keywords legate a DEI potrebbe indicare l'assenza di tali argomenti oppure la non precisa attribuzione di parole chiave. L'uso di

[5] I *Topics,* come spiegato nella banca dati, sono raccolte di pubblicazioni con un interesse intellettuale comune, basato su citazioni dirette, che rappresentano campi di ricerca a livello globale, costantemente in evoluzione in base ai cambiamenti nelle prospettive di ricerca. Una pubblicazione può appartenere a un solo Topic, e un Topic può appartenere a un solo Topic Cluster. La copertura temporale della rivista in Scopus (e quindi in SciVal) va dal 2019 al 2023.

keywords non corrette può essere fuorviante oppure non permettere il recupero dell'articolo; l'uso di keywords che non evidenzino i temi DEI oscura questi aspetti anche se trattati all'interno di un articolo. Tra i punti di attenzione va di certo ricompreso il modo in cui viene scritta la keyword fa la differenza. Per esempio, "person with disability" ha una valenza diversa di "disabled person" in quanto nel primo caso si accentua l'enfasi sulla persona mentre nel secondo caso sulla disabilità, così come "persona con disabilità" ha un'implicazione diversa da "disabile", soprattutto se estrapolata dal contesto (Thompson, 2020).
Un tema cruciale è poi quello linguistico, al quale bisognerebbe senz'altro prestare maggiore attenzione in un contesto scientifico colonizzato dalla lingua inglese (Horvath, A., Wagner, C. Wrisley, D.J., et al. 2024), in quanto la lingua è uno strumento fortemente identitario per una comunità e la convergenza verso un'unica lingua nella comunicazione scientifica non favorisce la diversità.
Tutti questi aspetti sono essenziali nella gestione automatizzata tramite text mining dei contenuti di una rivista o di una banca dati. In generale queste sono questioni alle quali gli editori stanno attribuendo considerazione crescente (APA 2021, Cuellar 2021, Dewidar 2022, Hinton & Lambert 2022; Marabelli & Chan 2024) poiché la centralità di questi temi nella società è di fatto crescente, anche alla luce dell'Agenda 2030, degli effetti della pandemia e di episodi come la morte di George Floyd o il movimento #meetoo (Marabelli & Chan 2024). Tali eventi hanno accresciuto la sensibilità verso queste tematiche, e aumentato perfino gli studi che hanno dimostrato i benefici anche in termini economici di un approccio rivolto alla diversità, all'equità, all'inclusione (McKinsey & Co. 2022). L'evoluzione degli ultimi anni ha messo in luce il lavoro ancora da compiere ma allo stesso tempo ha rappresentato un punto di svolta nelle tematiche DEI, dal quale auspicabilmente non si può tornare indietro.
La proposta è quindi, per tutte le riviste che ancora non lo fanno, di fornire e delle indicazioni precise, adottare linee guida, offrire suggerimenti e uno schema di tematiche DEI in cui far rientrare le keywords, oltre ad incrementare la trattazione di questi argomenti nella rivista, magari attraverso specifiche call for papers così come in questo convegno.

BIBLIOGRAFIA

APA. (2021). Equity, Diversity, and Inclusion Toolkit for Journal Editors, https://www.apa.org/pubs/authors/equity-diversity-inclusion-toolkit.

Cuellar, NG. (2021). Diversity, Equity, and Inclusivity in Peer-Reviewed Journals. *Journal of Transcultural Nursing*, 32(5), 449-450. doi:10.1177/10436596211034902

Dewidar, O., Elmestekawy, N. & Welch, V. (2022). Improving equity, diversity, and inclusion in academia. *Research Integrity and Peer Review*, 7. https://doi.org/10.1186/s41073-022-00123-z

Ferreira Garcia, D.C., Chaves Gattaz, C., & Chaves Gattaz, N. (2019). The Relevance of Title, Abstract and Keywords for Scientific Paper Writing, *Revista de Administração Contemporânea*, 3. http://doi.org/10.1590/1982-7849rac2019190178.

Hartley, J., & Kostoff, Ronald D. (2003). How useful are 'key words' in scientific journals? *Journal of Information Science*, 29, 433-438.

Hinton, A. Jr., & Lambert, M.W. (2022). Moving diversity, equity, and inclusion from opinion to evidence. *Cell Reports Medicine*, 3(4), 100619. https://doi.org/10.1016/j.xcrm.2022.100619

Horvath, A., Wagner, C. Wrisley, D.J., et al. (2024). Multilinguality in Action: Towards Linguistic Diversity and Inclusion in Digital Humanities. *magazén*, 5(2). DOI: 10.30687/mag/2724-3923/2024/02/001/

Marabelli, M., Chan, Y.E. (2024). The strategic value of DEI in the information systems discipline. *The Journal of Strategic Information Systems*, 33(1) https://www.sciencedirect.com/science/article/pii/S0963868724000052

McKinsey & Company (2022). What is diversity, equity, and inclusion? https://www.mckinsey.com/featured-insights/mckinsey-explainers/what-is-diversity-equity-and-inclusion

Morriello, R., & Sardo, L. (2023). Diversity management e prospettive DEI in biblioteca. *Biblioteche oggi Trend*s, 9(2), 33-42. http://doi.org/10.3302/2421-3810-202302-033-1

Nazioni Unite (2015). Obiettivi per lo Sviluppo Sostenibile, https://unric.org/it/agenda-2030/

Raju, J. (2020). Diversity, Inclusion, and Social Justice in the Information Context: Global South Perspectives, *The International Journal of Information, Diversity, & Inclusion*, 4(3/4). DOI:10.33137/ijidi.v4i3/4.34974

Thompson, K.M. (2020). Socializing Engagement: From Words to Action, *The International Journal of Information, Diversity, & Inclusion*, 4(1). DOI: 10.33137/ijidi.v4i1.33251