

Atti del IX Convegno Annuale
dell'Associazione per l'Informatica
Umanistica e la Cultura Digitale
(AIUCD)

LA SVOLTA INEVITABILE:
SFIDE E PROSPETTIVE PER
L'INFORMATICA UMANISTICA

15 – 17 gennaio 2020
Milano
Università Cattolica del Sacro Cuore

A CURA DI:
Cristina Marras
Marco Passarotti
Greta Franzini
Eleonora Litta

ISBN: 978-88-942535-4-2

ASSOCIAZIONE per
l'INFORMATICA UMANISTICA
e la CULTURA DIGITALE



Copyright © 2020

Associazione per l'Informatica Umanistica e la Cultura Digitale

Copyright of each individual chapter is maintained by the authors.

This work is licensed under a Creative Commons Attribution Share-Alike 4.0 International license (CC-BY-SA 4.0). This license allows you to share, copy, distribute and transmit the text; to adapt the text and to make commercial use of the text providing attribution is made to the authors (but not in any way that suggests that they endorse you or your use of the work). Attribution should include the following information:

Cristina Marras, Marco Passarotti, Greta Franzini, Eleonora Litta (a cura di),
Atti del IX Convegno Annuale AIUCD. La svolta inevitabile:
sfide e prospettive per l'Informatica Umanistica.

Available online as a supplement of Umanistica Digitale: <https://umanisticadigitale.unibo.it>

All links were visited on 29th December 2019, unless otherwise indicated.

Every effort has been made to identify and contact copyright holders and any omission or error will be corrected if notified to the editors.

Prefazione

La nona edizione del convegno annuale dell'Associazione per l'Informatica Umanistica e la Cultura Digitale (AIUCD 2020; Milano, 15-17 gennaio 2020) ha come tema “La svolta inevitabile: sfide e prospettive per l'Informatica Umanistica”, con lo specifico obiettivo di fornire un'occasione per riflettere sulle conseguenze della crescente diffusione dell'approccio computazionale al trattamento dei dati connessi all'ambito umanistico. Questo volume raccoglie gli articoli i cui contenuti sono stati presentati al convegno. A diversa stregua, essi affrontano il tema proposto da un punto di vista ora più teorico-metodologico, ora più empirico-pratico, presentando i risultati di lavori e progetti (conclusi o in corso) che considerino centrale il trattamento computazionale dei dati.

Dunque, la svolta inevitabile qui a tema va intesa innanzitutto come metodologica e, più nello specifico, computazionale. Ad essa la ricerca umanistica contemporanea assiste, con diversi gradi di accoglienza, critica addirittura rifiuto. La computabilità del dato empirico (anche) in area umanistica è, infatti, il tratto distintivo e il vero valore aggiunto che le innovazioni tecnologiche degli ultimi decenni hanno comportato in questo ambito. Nonostante negli anni il settore delle cosiddette Digital Humanities si sia voluto caratterizzare, anche a partire dalla propria denominazione, insistendo maggiormente sull'aspetto digitale che non su quello computazionale, i tempi sembrano ormai maturi perché il termine Computational Humanities, o il troppo precocemente accantonato Humanities Computing, (ri)prenda il posto oggi ancora occupato da Digital Humanities.¹ Digitale è, infatti, il formato dei dati con cui attualmente si ha in gran parte a che fare nel nostro settore: ma è computazionale l'uso che di questi dati si fa ed è un fatto che gran parte dei lavori prodotti nell'area delle Digital Humanities consista nel “fare conti” sui dati.²

Come tanti suoi predecessori, anche il formato digitale passerà; mentre il metodo, e la svolta che esso comporta, resterà, perché solidamente ancorato all'evidenza empirica del dato che è il punto di partenza e, quindi, il centro di analisi di molta ricerca umanistica. Per questa ragione, la svolta computazionale nelle scienze umanistiche è innanzitutto metodologica: a cambiare radicalmente non è tanto il formato dei dati, ma il modo con cui ad essi ci si approccia e l'uso che di essi si fa.

Non va negato un certo scetticismo reazionario che, ora esplicito, ora sottaciuto, parte del mondo della ricerca umanistica nutre nei confronti dei metodi e degli strumenti che la svolta computazionale ha messo a disposizione di noi ricercatori, che viviamo l'attuale scorcio di storia della scienza. Negli anni, tale scetticismo ha alimentato una irragionevole distinzione, e conseguente separazione, tra umanisti “tradizionali” e umanisti “digitali”, quasi che si debbano identificare due aree al fine di evitare che gli uni infastidiscano troppo gli altri con le proprie ricerche, trascurando che esse trattano i medesimi oggetti e hanno quale fine comune la produzione di nuova conoscenza.

Siffatta separazione è dovuta a errori imputabili all'una e all'altra parte. Da un lato, certi umanisti “digitali” tendono a produrre ricerca che rischia di scadere nella superficialità, assumendo che l'alta quantità dei dati trattati possa compensarne l'eventuale bassa qualità e dimenticando, così, che le ricerche di area umanistica molto raramente lavorano su Big Data e non possono (anzi, non vogliono) accontentarsi di tendenze percentuali fondate su dati imprecisi. Dall'altro lato, i “tradizionali” sono spesso afflitti da un conservatorismo protezionista incompatibile con la natura stessa del lavoro di ricerca, che è in sé progressivo e in costante evoluzione. Ne consegue un dialogo interrotto tra le due parti: i “digitali” sono considerati dei tecnici (inteso in senso riduttivo) che brutalizzano il delicato dato umanistico, mentre i “tradizionali” vengono derubricati a dinosauri incartapecoriti che ormai non hanno più niente di nuovo da dire.

Ma la svolta computazionale non è né “digitale”, né “tradizionale”. Semplicemente, essa è inevitabile. Chi ne fa cattivo uso, come certo mondo “digitale”, non sa valorizzarne la forza della portata; chi la rifiuta a priori, si pone fuori dalla realtà e, volutamente ignorando il nuovo, ferisce la ragione stessa del far ricerca.

¹ Una valida sintesi della questione relativa alla denominazione del settore, con una buona bibliografia a supporto, è riportata in un articolo di Leah Henrikson pubblicato su 3:AM Magazine (24 Ottobre 2019) e disponibile presso <https://www.3amagazine.com/3am/humanities-computing-digital-humanities-and-computational-humanities-whats-in-a-name/>

² Da, Nan Z. “The computational case against computational literary studies.” *Critical Inquiry* 45.3 (2019): 601-639.

Ma resta che la svolta è inevitabile: non si comprende perché sul tavolo dell'umanista del 2020 non possano trovarsi al contempo un'edizione critica cartacea e i risultati di un analizzatore morfologico automatico proiettati sullo schermo di un computer. Entrambi sono strumenti che diversamente trattano il comune oggetto d'interesse di tanta ricerca, ovvero i dati.

Ma di una svolta non solo metodologica questa edizione 2020 del convegno AIUCD vuole trattare e farsi carico, aspirando anzi a mettere in atto anche una piccola, ma sostanziale svolta organizzativa. Per la prima volta, la call for papers di un convegno dell'Associazione, ha richiesto l'invio non di abstract, ma di articoli completi della lunghezza di un massimo di 4 pagine (bibliografia esclusa). Di concerto con il Comitato Direttivo dell'Associazione, abbiamo deciso di orientarci in tal senso per due ragioni principali. Primo, crediamo che, giunto alla propria nona edizione, il convegno annuale della AIUCD sia ormai sufficientemente maturo per passare a una fase il cui obiettivo sia quello di accogliere nel programma del convegno proposte che nel formato dell'articolo completo consentissero ai revisori una valutazione piena e più accurata. Ciò si lega anche alla seconda ragione. Il nostro settore come è noto è molto veloce: i dati (e i risultati su di essi basati) tendono a cambiare nel giro di poco tempo. Ricevere articoli completi ci ha consentito di mettere i contenuti del presente volume nelle mani dei partecipanti (e più in generale della comunità tutta) il primo giorno del convegno, fornendo così una realistica fotografia dello stato dei lavori al gennaio 2020.

Tutti gli articoli selezionati per essere presentati al convegno hanno cittadinanza in questo volume. Anche questa è una svolta: diversamente dall'uso fino ad oggi adottato, gli articoli pubblicati non sono più il risultato di una selezione a posteriori rispetto al convegno, ma tutti quelli effettivamente apparsi nel programma di AIUCD 2020. In tal senso, una certa esclusività promossa a livello di selezione scientifica si fa inclusività in termini di pubblicazione e, dunque di visibilità dei lavori presentati. Ogni proposta è stata valutata da tre revisori; si è dovuto ricorrere a una quarta valutazione solo nel caso di due proposte su cui i tre revisori avevano espresso opinioni che rendevano difficile prendere una decisione in merito alla loro accettazione, o meno. Al proposito delle differenze tra i revisori, abbiamo constatato divergenze piuttosto frequenti e, in alcuni casi, nette tra coloro che provengono dall'area linguistico-computazionale e quanti, invece, sono a vario titolo legati ai diversi settori dell' "umanistica digitale". Mentre i linguisti computazionali sono tradizionalmente usi a valutare articoli completi e tendono a richiedere che i contenuti di essi descrivano motivazioni, metodi e risultati (preferibilmente replicabili) di lavori di ricerca in corso, o completati, i revisori di area umanistico-digitale sono disposti a valutare positivamente anche idee e proposte che ancora non si siano incarnate in una reale applicazione ai dati. La constatazione di tale diversità è il risultato della composizione volutamente inter- e trans-disciplinare del comitato dei revisori, a rappresentare la natura trasversale di AIUCD e, di riflesso, del suo convegno annuale. Nel prendere le decisioni in merito alle proposte, abbiamo cercato un equilibrio tra gli atteggiamenti delle due parti, favoriti dall'aver a disposizione un livello di dettaglio sul lavoro descritto. La richiesta di articoli completi ha avuto un impatto non molto rilevante sul numero delle proposte inviate, che sono state 71, di cui 67 sottoposte al processo di revisione, mentre 4 sono state escluse perché non confacenti ai criteri richiesti dalla call for papers (tra cui anonimato e originalità). Alla precedente edizione del convegno AIUCD (Udine, 23-25 gennaio 2019) erano state inviate 82 proposte, di cui 75 sottoposte a revisione. Conseguenze più sostanziali si sono, invece, riscontrate sulla percentuale delle proposte accettate e rifiutate. Delle 67 proposte valutate, 45 sono state accettate per apparire nel programma del convegno e, quindi, in questo volume, mentre 22 sono state rifiutate, risultando così in una percentuale di accettazione pari al 67.16%. All'edizione udinese, la percentuale si era attestata intorno all'84%. La contrazione del numero di proposte accettate è strettamente connessa alla richiesta di articoli completi invece che di abstract.

Il programma del convegno ha incluso due sessioni poster. Dei 45 contributi accettati, 21 sono stati giudicati adatti alla presentazione in modalità poster. Rispetto alle consuetudini del settore, che tende a relegare le proposte meno interessanti o più problematiche nelle sessioni poster, abbiamo deciso di assegnare la modalità di comunicazione in forma di poster non secondo la qualità, ma piuttosto in base alla tipologia della proposta. Dunque, tendenzialmente le proposte che presentano lavori che hanno portato a risultati pratici (come strumenti, risorse, o interfacce) sono state giudicate più adatte a una presentazione in formato poster, mentre le discussioni teoriche, disciplinari, o metodologiche hanno occupato le sessioni di comunicazioni orali. Resta che non sussiste differenza alcuna in termini di selezione qualitativa tra un articolo i cui contenuti sono stati

presentati al convegno in forma orale, o in forma di poster, come dimostra l'aver riservato il medesimo numero di pagine a tutti gli articoli presenti in questo volume.

I contenuti dei testi qui raccolti in ordine alfabetico testimoniano la varietà dei temi che usualmente sono trattati nei convegni della AIUCD. Essi spaziano da riflessioni generali sul settore di ricerca alla realizzazione di nuove risorse linguistiche e strumenti di analisi dei dati, da lavori di filologia ed editoria digitale a temi connessi alla digitalizzazione delle fonti in ambito bibliotecario. Oltre alla presentazione dei contenuti degli articoli di questo volume, il programma del convegno ha previsto tre relazioni su invito (una per ciascuno dei tre giorni della sua durata), che sono state rispettivamente presentate da Roberto Navigli (Sapienza, Università di Roma), Julianne Nyhan (University College London) e Steven Jones (University of South Florida).

Il contributo di Roberto Navigli, intitolato *Every time I hire a linguist my performance goes up (or: the quest for multilingual lexical knowledge in a deep (learning) world)*, è un esempio di ricerca che dice della ineludibilità del legame e, auspicabilmente, della collaborazione tra mondo scientifico e mondo umanistico e, nello specifico, tra la comunità che si riconosce nella AIUCD e quella della linguistica computazionale. Gli interventi di Julianne Nyhan (*Where does the history of the Digital Humanities fit in the longer history of the Humanities? Reflections on the historiography of the 'old' in the work of Fr Roberto Busa S.J.*) e Steven Jones (*Digging into CAAL: Father Roberto Busa's Center and the Prehistory of the Digital Humanities*) si posizionano nell'alveo della storia della disciplina, particolarmente riferendo in merito ai loro studi sulle attività di padre Roberto Busa. La figura di Busa è strettamente legata all'Università Cattolica del Sacro Cuore di Milano, dove a partire dalla fine degli anni settanta il gesuita tenne un corso di Linguistica Computazionale e Matematica e fondò un gruppo di ricerca che, nel 2009, fu trasformato in un Centro di Ricerca; quel CIRCSE che con l'AIUCD ha organizzato il convegno annuale dell'associazione di cui questo volume raccoglie gli Atti. Nel 2010, un anno prima di lasciarci, padre Busa volle donare alla Biblioteca della Cattolica il proprio archivio personale. Una ricchissima documentazione del lavoro di Busa e della sua diffusione, oltre che delle sue relazioni personali e professionali (ricostruibili attraverso il vasto epistolario), l'Archivio Busa è attualmente in fase di catalogazione e digitalizzazione da parte della Biblioteca d'Ateneo. Una selezione di materiale tratto dall'Archivio è stata resa direttamente accessibile ai partecipanti dell'edizione milanese del convegno AIUCD in una piccola mostra allestita nell'atrio dell'aula dei lavori congressuali. Le teche della mostra raccolgono fogli di lavoro, lettere, schede perforate, nastri e articoli di quotidiani che trattano del lavoro di padre Busa: una forma di ringraziamento che l'Università Cattolica, il CIRCSE e la comunità scientifica tutta vuole riservare a uno dei pionieri dell'analisi linguistica automatica.

I nostri ringraziamenti vanno innanzitutto alla Presidente di AIUCD Francesca Tomasi e a Fabio Ciotti, che in quel ruolo l'ha preceduta, per aver scelto Milano quale sede dell'edizione 2020 del convegno. Da loro è venuto il primo, fondamentale, sostegno alla “svolta organizzativa” di cui abbiamo voluto farci portatori. Ringraziamo altresì il Consiglio Direttivo dell'Associazione, il Comitato di Programma e tutti i revisori, che hanno lavorato alacremente per metterci nelle condizioni di definire il miglior programma possibile. La sede milanese dell'Università Cattolica del Sacro Cuore ci ha supportato a livello amministrativo e logistico; teniamo particolarmente a ringraziare l'Ufficio Formazione Permanente, nello specifico di Elisa Ballerini, la Biblioteca d'Ateneo, e specificatamente Paolo Senna, che ci ha messo a disposizione i materiali dell'Archivio Busa, l'Ufficio Eventi e la Direzione di Sede, che hanno fornito gli spazi per il convegno. Grazie soprattutto a chi ha inviato proposte, ai relatori e ai partecipanti tutti, perché sono loro i protagonisti essenziali dell'evento.

La nostra speranza è che il lavoro fatto sia utile ancora prima che apprezzato. E che i suoi risultati si mantengano nelle edizioni a venire, con l'obiettivo di migliorare sempre, guardando avanti; perché saper vedere le svolte e affrontarle è la ragione stessa della ricerca.

Cristina Marras
Marco Passarotti
Greta Franzini
Eleonora Litta

Chair e Comitati

General Chair

- Cristina Marras

Chair del comitato scientifico e di programma

- Marco Passarotti

Comitato scientifico e di programma

- Maristella Agosti
- Stefano Allegrezza
- Federica Bressan
- Cristiano Chesi
- Fabio Ciraci
- Greta Franzini
- Angelo Mario Del Grosso
- Eleonora Litta
- Pietro Maria Liuzzo
- Federico Meschini
- Johanna Monti
- Federico Nanni
- Marianna Nicolosi
- Dario Rodighiero
- Marco Rospocher
- Chiara Zuanni

Comitato Organizzatore

- Greta Franzini
- Eleonora Litta

Indice dei Contenuti

EcoDigit-Ecosistema Digitale per la fruizione e la valorizzazione dei beni e delle attività culturali del Lazio	1
Luigi Asprino, Antonio Budano, Marco Canciani, Luisa Carbone, Miguel Ceriani, Ludovica Marinucci, Massimo Mecella, Federico Meschini, Marialuisa Mongelli, Andrea Giovanni Nuzzolese, Valentina Presutti, Marco Puccini, Mauro Saccone	
Encoding the Critical Apparatus by Domain Specific Languages: The Case of the Hebrew Book of Qohelet	7
Luigi Bambaci, Federico Boschetti	
600 maestri raccontano la loro vita professionale in video: un progetto di (fully searchable) open data	14
Gianfranco Bandini, Andrea Mangiatordi	
Ripensare i dati come risorse digitali: un processo difficile?	19
Nicola Barbuti	
Verso il riconoscimento delle Digital Humanities come Area Scientifica: il catalogo online condiviso delle pubblicazioni dell'AIUCD	24
Nicola Barbuti, Maurizio Lana, Vittore Casarosa	
Il trattamento automatico del linguaggio applicato all'italiano volgare. La redazione di un <i>formario</i> tratto dalle prime dieci <i>Lettere</i> di Alessandra M. Strozzi	28
Ottavia Bersano, Nadezda Okinina	
Annotazione semantica e visualizzazione di un corpus di corrispondenze di guerra	34
Beatrice Dal Bo, Francesca Frontini, Giancarlo Luxardo	
The Use of Parallel Corpora for a Contrastive (Russian-Italian) Description of Resource Markers: New Instruments Compared to Traditional Lexicography	39
Anna Bonola, Valentina Nosedà	
PhiloEditor: Simplified HTML Markup for Interpretative Pathways over Literary Collections	47
Claudia Bonsi, Angelo Di Iorio, Paola Italia, Francesca Tomasi, Fabio Vitali, Ersilia Russo	
An Empirical Study of Versioning in Digital Scholarly Editions	55
Martina Bürgermeister	

ELA: fasi del progetto, bilanci e prospettive Emmanuela Carbé, Nicola Giannelli	61
Digitized and Digitalized Humanities: Words and Identity Claire Clivaz	67
La geolinguistica digitale e le sfide lessicografiche nell'era delle digital humanities: l'esempio di VerbaAlpina Beatrice Colcuc	74
Una proposta di ontologia basata su RDA per il patrimonio culturale di Vincenzo Bellini Salvatore Cristofaro, Daria Spampinato	82
Biblioteche di conservazione e libera fruizione dei manoscritti digitalizzati: la Veneranda Biblioteca Ambrosiana e la svolta inevitabile grazie a IIF Fabio Cusimano	89
Repertori terminologici plurilingui fra normatività e uso nella comunicazione digitale istituzionale e professionale Klara Dankova, Silvia Calvi	98
The Digital Lexicon Translaticium Latinum: Theoretical and Methodological Issues Chiara Fedriani, Irene De Felice, William Michael Short	106
Selling Autograph Manuscripts in 19th c. Paris: Digitising the Revue des Autographes Simon Gabay, Lucie Rondeau du Noyer, Mohamed Khemakhem	113
Enriching a Multilingual Terminology Exploiting Parallel Texts: an Experiment on the Italian Translation of the Babylonian Talmud Angelo Mario Del Grosso, Emiliano Giovannetti, Simone Marchi	119
Towards a Lexical Standard for the Representation of Etymological Data Fahad Khan, Jack Bowers	125
Workflows, Digital Data Management and Curation in the RETOPEA Project Ilenia Eleonor Laudito	130

Il confronto con Wikipedia come occasione di valorizzazione professionale: il case study di Biblioteca digitale BEIC Lisa Longhi	136
Making a Digital Edition: The Petrarchive Project Isabella Magni	142
Extending the DSE: LOD Support and TEI/IIIF Integration in EVT Paolo Monella, Roberto Rosselli Del Turco	148
Mapping as a Contemporary Instrument for Orientation in Conferences Chloe Ye-Eun Moon, Dario Rodighiero	156
Argumentation Mapping for the History of Philosophical and Scientific Ideas: The TheSu Annotation Scheme and its Application to Plutarch's Aquane an ignis Daniele Morrone	163
Leitwort Detection, Quantification and Discernment Racheli Moskowitz, Moriyah Schick, Joshua Waxman	171
From Copies to an Original: The Contribution of Statistical Methods Amanda Murphy, Raffaella Zardoni, Felicita Mornata	178
FORMAL. Mapping Fountains over Time and Place. Mappare il movimento delle fontane monumentali nel tempo e nello spazio attraverso la geovisualizzazione Pamela Palomba, Emanuele Garzia, Roberto Montanari	185
Paul is Dead? Differences and Similarities before and after Paul McCartney's Supposed Death. Stylometric Analysis of Transcribed Interviews Antonio Pascucci, Raffaele Manna, Vincenzo Masucci, Johanna Monti	191
Digital Projects for Music Research and Education from the Center for Music Research and Documentation (CIDoM), Associated Unit of the Spanish National Research Council Juan José Pastor Comín, Francisco Manuel López Gómez	198
Prospects for Computational Hermeneutics Michael Piotrowski, Markus Neuwirth	204
EModSar: A Corpus of Early Modern Sardinian Texts Nicoletta Puddu, Luigi Talamo	210
Shared Emotions in Reading Pirandello. An Experiment with Sentiment Analysis Simone Reborà	216
DH as an Ideal Educational Environment: The Ethnographic Museum of La Spezia Letizia Ricci, Francesco Melighetti, Federico Boschetti, Angelo Mario Del Grosso, Enrica Salvatori	222

A Digital Review of Critical Editions: A Case Study on Sophocles, Ajax 1-332 Camilla Rossini	227
Strategie e metodi per il recupero di dizionari storici Eva Sassolini, Marco Biffi	235
Encoding Byzantine Seals: SigiDoc Alessio Sopracasa, Martina Filosa	240
Preliminary Results on Mapping Digital Humanities Research Gianmarco Spinaci, Giovanni Colavizza, Silvio Peroni	246
Epistolario De Gasperi: National Edition of De Gasperi's Letters in Digital Format Sara Tonelli, Rachele Sprugnoli, Giovanni Moretti, Stefano Malfatti, Marco Odorizzi	253
Visualizing Romanesco; or, Old Data, New Insights Gianluca Valenti	260
What is a Last Letter? A Linguistics/Preventive Analysis of Prisoner Letters from the Two World Wars Giovanni Pietro Vitali	265
L'organizzazione e la descrizione di un fondo nativo digitale: PAD e l'Archivio Franco Buffoni Paul Gabriele Weston, Primo Baldini, Laura Pusterla	273

Verso il riconoscimento delle Digital Humanities come Area Scientifica: il Catalogo online condiviso delle pubblicazioni dell'AIUCD.

Nicola Barbuti
Università degli Studi di Bari
Aldo Moro
nicola.barbuti@uniba.it

Maurizio Lana
Università degli Studi del
Piemonte Orientale
maurizio.lana@uniupo.it

Vittore Casarosa
ISTI-CNR
casarosa@isti.cnr.it

Abstract

English. The recognition of Digital Humanities (DH) as Scientific Sector is a first-level issue in the debate on innovation of scientific research in Italy. This issue is among the most recurrent even within the AIUCD. In this regard, during the 2019 Conference of the Association, the need emerged to launch activities aimed at keeping the attention of political decision makers to this claim. The Assembly opted for a "bottom-up" action: to produce an online digital catalogue of national sector publications for providing concrete elements to support the awareness that DH are no longer a nebula dotted with indistinct and doubled edged scientific entities, but are today a very well established and widespread reality in the panorama of Italian research.

Italiano. Il riconoscimento delle Digital Humanities (DH) quale Area Scientifica autonoma è uno dei temi di maggiore attualità nel dibattito sui nuovi orientamenti della ricerca scientifica in Italia. Anche in seno all'AIUCD, la questione di riconoscere alle DH la dignità di settore specifico di ricerca è tra i più ricorrenti. A riguardo, nel corso del Convegno 2019 dell'Associazione, è emersa la necessità di avviare attività orientate ad attirare l'attenzione dei decisori politici sul tema. Si è dunque optato per un'azione "dal basso": produrre un catalogo digitale on line delle pubblicazioni nazionali di settore, di modo da fornire elementi concreti a sostegno della consapevolezza che le DH non sono più una nebulosa costellata di entità scientifiche indistinte e ancipiti, ma sono ormai una realtà molto ben consolidata e diffusa nel panorama della ricerca italiana.

1 Introduzione

Nell'ambito della Conferenza AIUCD 2019, uno dei temi di maggior confronto, anche in seno all'Assemblea annuale dei soci, è stata la necessità ormai non più rinviabile di attivare iniziative finalizzate a riconoscere alle Digital Humanities – o Humanities Computing che si vogliono ridefinire – la dignità di Area Scientifica, emancipandole definitivamente dalla dimensione di come nebulosa in cui fluttuano in modo frammentario e caotico ricercatori e studiosi rinvenienti da una pletera di SSD tradizionali delle Humanities, in cui sono considerati delle entità scientifiche ancipiti.

L'analisi dello stato dell'arte della disciplina ha evidenziato la consistenza decisamente ampia di contributi scientifici che, pur riferendosi ad ambiti umanistici riconducibili a singoli SSD tradizionali, sono confluenti nella comune prospettiva di ricerca su digitale e computazionale applicati alle humanities e, perciò, assolutamente associabili in un'unica area di contenimento.

Pertanto, al fine di sostanziare l'istanza di riconoscimento delle DH come Area con una sua dignità scientifica, si è scelto un indirizzo operativo "dal basso": produrre un catalogo digitale delle pubblicazioni nazionali di settore che possa essere riconosciuto dall'AIUCD quale proprio riferimento ufficiale e, in prospettiva, possa diventare nodo di un più ampio e condiviso catalogo internazionale di pubblicazioni sulle DH.

Il Gruppo di Lavoro individuato per occuparsi di progettare e strutturare il catalogo è composto dai soci AIUCD Maurizio Lana, Vittore Casarosa e Nicola Barbuti.

Il GdL ha intrapreso le attività immediatamente dopo la chiusura del Convegno 2019 e attualmente sta provvedendo agli ultimi passaggi per la realizzazione esecutiva di quanto progettato e proposto all'Associazione agli inizi dell'estate 2019.

2 Il Catalogo delle DH AIUCD

Primo tema di riflessione è stato definire i limiti geografici e cronologici delle pubblicazioni da inserire in prima istanza. Si è deciso di limitare inizialmente l'inserimento agli autori italiani con priorità per i soci AIUCD, con un orizzonte temporale non superiore agli ultimi 10 anni. È stata ipotizzata la creazione dello spazio online con credenziali di accesso da condividere con tutti i soci in modo che, per quanto possibile, ognuno possa inserire da sé i record bibliografici relativi alle proprie pubblicazioni, e di consentire anche l'associazione dei pdf ai record, ove legalmente disponibili. Tuttavia, per conferire al catalogo veste ufficiale, pur consentendo a ciascuno di inserire i dati direttamente, sembra opportuno stabilire regole definite e condivise, nominando un organismo deputato a eseguire un controllo annuo sulla coerenza delle nuove risorse bibliografiche caricate per evitare l'insorgere di situazioni caotiche.

Dal momento che una delle principali criticità connesse con la multiforme produzione scientifica delle DH è proprio l'elevata varietà di fonti e quindi di formati citazionali bibliografici, si è passati ad analizzare software open source per la gestione e la fruizione di record bibliografici digitali che avessero le caratteristiche necessarie a favorire un import di risorse di diversa tipologia e struttura.

Primi a essere presi in considerazione sono stati Zotero, in considerazione sia del fatto che l'AIUCD utilizzando questo software aveva già attivato un repository per l'allocazione di risorse digitali relative alle DH sia dell'ampio uso che se ne sta facendo per la creazione di bibliografie online con pubblicazioni relative ad altri ambiti di ricerca scientifica, e Zenodo per il repository delle fonti non altrimenti online.

Zotero nacque proprio come strumento per la gestione di bibliografie di area storica (Cohen, 2008) ed è attualmente usato, per esempio, dall'associazione tedesca di Digital Humanities¹ per uno scopo simile a quello cui stiamo pensando anche per AIUCD o dalla *American School of Classical Studies at Athens* (ASCSA) per catalogare e gestire i metadati di tutte le pubblicazioni della Scuola stessa (libri e articoli)². Tuttavia, quest'ultimo è risultato un open repository per "prodotti della ricerca" generici ed ha evidenziato il limite non secondario che ogni oggetto caricato deve essere linkato a mano al record corrispondente. Zotero invece è mirato a collezioni di articoli e bibliografie di varia natura e i dati caricati in uno suo spazio online non sono soggetti al problema rilevato per Zenodo (O'Donnell, Manola, Manghi, Porter, Esau, Viejou, Rosselli Del Turco, and Singh, 2018; Peters, Kraeker, Lex, Gumpenberger and Gorriaz, 2017). È sufficiente, infatti, che il responsabile del caricamento tagghi la pubblicazione con l'indicazione del/dei SSD in cui si colloca e con le keywords che egli ritiene ne identifichino correttamente il contenuto per rendere il record interrogabile anche in questo modo oltre che con i consueti criteri di autore, titolo, etc. Inoltre, altre caratteristiche interessanti di Zotero sono la separazione dei (meta)dati dalla loro presentazione secondo uno stile citazionale piuttosto che un altro, la possibilità di esportare in RDF e altri formati open il database della bibliografia.

Altro punto a favore di Zotero è la possibilità di definire modalità di recupero delle informazioni bibliografiche incrociando i seguenti dati:

- dati dei titoli;
- dati delle keywords inserite ufficialmente nei lavori o, in mancanza, indicate espressamente dagli autori rilevando parole poi riscontrabili nel testo;
- dati che possono essere estratti dagli abstract o dai full text dei contributi inseriti;
- SSD degli autori.

Queste rilevazioni possono essere utilizzate per produrre report annui sullo stato dell'arte della ricerca scientifica sulle DH da ufficializzare per rimarcare la fertilità produttiva del settore.

Per iniziare a popolare il catalogo, si è concordato di proporre ai soci (e non, purché italiani) vari modi per inserire i dati nella bibliografia (Vahdati, Arndt, Auer and Lange, 2016):

¹ https://www.zotero.org/groups/372575/dhd_ag_publicationen ² https://www.zotero.org/groups/80651/american_school_of_classical_studies_at_athens

- coloro che usano un Bibliographic Reference Software (BRS: Zotero, Mendeley, Endnote, Bibref, Refworks, etc.), possono esportare i (meta)dati citazionali delle loro pubblicazioni completi di URL a ciascuna pubblicazione, quindi inserirli in Zotero e dare l'accesso pubblico per consentire la fruizione diretta delle risorse caricate;
- coloro che possono utilizzare l'ISBN per le monografie o il DOI per gli articoli possono inserirli direttamente nella bibliografia online, utilizzando Zotero, i dati delle loro pubblicazioni, completandoli con i tag che indicano il SSD e quant'altro può essere necessario al recupero del record.

Per tutte le forme di pubblicazione grigia (presentazioni, abstract, raccolte di dati, etc.) Zotero non può gestire in modo ottimale i dati perché gestisce per lo più risorse bibliografiche. Giunge utile a questo punto Zenodo, poiché assegna automaticamente un DOI alle pubblicazioni o alle fonti in genere che non lo hanno già, e quindi permette di salvare nel suo repository aperto anche altri prodotti della ricerca, come dati e software (Potter and Smith, 2015), oltre alle classiche pubblicazioni.

Si è dunque concluso di creare il catalogo sfruttando al massimo le diverse opportunità offerte dai due software, creando una soluzione che, non comportando attività di sviluppo software ad hoc, dia ragionevoli prospettive di sostenibilità e permetta agevoli importazione, esportazione e migrazione dei dati.

Di seguito riportiamo l'articolazione dell'ipotesi progettuale, che si articola nelle seguenti fasi.

Base di partenza sarà il catalogo AIUCD già esistente in Zotero, sebbene esso presenti alcuni fisiologici punti di debolezza (duplicazioni, item incompleti, assenza di rimando con link alle fonti dove disponibili in OA, etc.).

I cataloghi Zotero sono accessibili in lettura-scrittura per gli editors, in sola lettura per tutti quelli che hanno il link.

Per facilitare l'acquisizione e l'inserimento dei dati si partirà dalle pubblicazioni che hanno già il DOI, quindi seguiranno, in ordine progressivo, quelle che hanno ISBN, poi quelle con ISSN, infine le pubblicazioni i cui dati devono essere raccolti e inseriti manualmente.

Le pubblicazioni cosiddette "grigie" prive di DOI (a es.: presentazioni) saranno invece preliminarmente caricate in Zenodo, di modo che il sistema le renda accessibili assegnandovi un DOI.

Scegliere di gestire le risorse digitali associando Zotero e Zenodo, oltre a presentare il vantaggio di rendere le pubblicazioni inserite facilmente ricercabili utilizzando chiavi di accesso uniformi, consente di collocare il catalogo in un contesto di ricerca aperta.

Ciascun socio/autore, quindi, potrà provvedere all'inserimento dei propri dati nel modo che segue:

- chi ha pubblicazioni in OA ma senza DOI le carica in Zenodo per ottenere il DOI;
- chi ha pubblicazioni già provviste di DOI comunica la lista dei DOI, uno per riga, mandando un messaggio all'indirizzo email dedicato pubblicazioni@aiucd.it; se la pubblicazione non ha keywords internamente, si possono indicare fino a un massimo di 5 keyword (parole o espressioni) accanto al DOI, separate da virgole; a riguardo, una buona chiave identificativa può essere il (o i) SSD in cui l'autore ritiene si collochi la sua pubblicazione nello spazio cloud di zotero;
- chi ha libri manda gli ISBN;
- chi ha articoli senza DOI manda l'URL.

Le pubblicazioni provviste di DOI saranno caricate in Zotero utilizzando il codice. Il DOI dovrà essere presente e ben visibile nel catalogo pubblico, in quanto è la chiave di accesso principale alla pubblicazione inserita.

Relativamente alle necessità di gestione del catalogo, si prospettano le seguenti soluzioni.

Sarà necessario ridefinire chi avrà accesso in scrittura al catalogo sia per inserire i DOI in Zotero e costruire l'elenco, sia per intervenire a correggere eventuali errori nei dati inseriti. Un accesso indiscriminato, infatti, creerebbe rischi di rumore notevole e soluzioni caotiche e difformi nell'organizzazione dei dati.

Ogni inizio d'anno, i soci saranno invitati a inviare all'indirizzo sopra indicato i DOI e gli ISBN delle nuove pubblicazioni dell'anno precedente.

Dal momento che esiste un problema concreto di autodefinizione sulla pertinenza delle pubblicazioni al campo DH, al fine di evitare l'afflusso nel catalogo di pubblicazioni che poco hanno a che fare con il settore sarà necessario che il Direttivo AIUCD definisca delle linee guida sui temi di ricerca coerenti con esso.

Una volta analizzato l'impatto del catalogo nell'ambito della ricerca scientifica (Sample, 2011), si potranno prendere in considerazione altri aspetti che renderanno necessarie opportune implementazioni si pensi, nello specifico, alle pubblicazioni senza codici, che sono in linea di massima quelle più datate e richiedono notevole lavoro per essere inserite nel catalogo.

Resta da definire come agire operativamente, cioè chi si occuperà di inserire i DOI in Zotero per costruire l'elenco: potrebbe essere un'attività laboratoriale per studenti di biblioteconomia/scienze bibliotecomiche e dell'informazione o per i futuri digital librarians?

3 Conclusioni

Tirando le conclusioni sul catalogo progettato, siamo del parere che un'integrazione Zotero-Zenodo in un'unica soluzione on line, sfruttando al massimo l'assegnazione DOI e, in prospettiva, la funzione di repository del secondo, prende il massimo da due mondi open source. Ciò apre opportunità di analisi della ricerca del settore prima impossibili da attuare (Winslow, Rains, Skripsky and Kelly, 2016), e si configura come un elemento qualificante il riconoscimento delle DH come settore di primo livello nella ricerca italiana.

Bibliografia

https://www.zotero.org/groups/372575/dhd_ag_publicationen

https://www.zotero.org/groups/80651/american_school_of_classical_studies_at_athens

Daniel J. Cohen. 2008. *Creating Scholarly Tools and Resources for the Digital Ecosystem: Building Connections in the Zotero Project*, «First Monday 13», n. 8. <https://doi.org/10.5210/fm.v13i8.2233>

Daniel Paul O'Donnell, Natalia Manola, Paolo Manghi, Dot Porter, Paul Esau, Carey Viejou, Roberto Rosselli Del Turco and Gurpreet Singh. 2018. Using Zenodo as a Discovery and Publishing Platform. <https://doi.org/10.5281/zenodo.1297110>

Isabella Peters, Peter Kraker, Elisabeth Lex, Christian Gumpenberger and Juan Ignacio Gorraiz. 2017. *Zenodo in the spotlight of traditional and new metrics*, «Frontiers in Research Metrics and Analytics» 2: 13. <https://doi.org/10.3389/frma.2017.00013>

Megan Potter and Tim Smith. 2015. *Making code citable with Zenodo and GitHub*, Software Sustainability Institute. <https://www.software.ac.uk/blog/2016-09-26-making-code-citable-zenodo-and-github>

Rachel Rains Winslow, S. Skripsky and Savannah L. Kelly. 2016. *Not just for citations: Assessing Zotero while reassessing research*. In: *Information literacy: Research and collaboration across disciplines*, 299–316. Fort Collins, CO: WAC Clearinghouse and University Press of Colorado.

Mark Sample. 2011. *Sharing Research and Building Knowledge Through Zotero*. In: *Learning Through Digital Media Experiments in Technology and Pedagogy*, ed. by Trebor R. Scholz. New York: New School. <http://mcpres.media-commons.org/artoflearning/sharing-research-and-building-knowledge-through-zotero/>

Sahar Vahdati, Natanael Arndt, Sören Auer and Christoph Lange. 2016. *OpenResearch: Collaborative Management of Scholarly Communication Metadata*. In: *Knowledge Engineering and Knowledge Management*, ed. by Eva Blomqvist, Paolo Ciancarini, Francesco Poggi and Fabio Vitali, 10024:778–93. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-49004-5_50