

Cronache

EDIZIONI DIGITALI: RAPPRESENTAZIONE, INTEROPERABILITÀ, ANALISI DEL TESTO E INFRASTRUTTURE (Venezia, 7-9 settembre 2016)

FEDERICO BOSCHETTI E MARINA BUZZONI

Il quinto convegno annuale dell'Associazione di Informatica Umanistica e Cultura Digitale (AIUCD) si è svolto dal 7 al 9 settembre 2016 presso l'Aula Magna di Ca' Dolfin dell'Università Ca' Foscari di Venezia e ha avuto per tema: "Edizioni digitali: rappresentazione, interoperabilità, analisi del testo e infrastrutture".

Il convegno è stato pensato per far dialogare studiosi provenienti da discipline diverse ma con interessi convergenti sul testo (digitale): filologi, storici, insegnanti e studenti di materie umanistiche, umanisti digitali, linguisti computazionali, logici, informatici e ingegneri informatici. Si è mirato infatti a far interagire i protagonisti di questi settori per ragionare insieme sulla necessità di coordinare i mezzi per raggiungere obiettivi comuni, che coinvolgono l'acquisizione, la codifica, l'annotazione e l'elaborazione dei dati per lo studio del testo contenuto nell'edizione digitale.

Sono stati identificati cinque aspetti, o assi portanti, attorno ai quali si è deciso di strutturare il convegno: i modelli formali per la rappresentazione del testo; le risorse digitali per la costituzione di corpora testuali e linguistici; gli strumenti computazionali per l'analisi dei dati; le infrastrutture di ricerca per gli studi del testo; le comunità di pratica per le Digital Humanities. Ciascuno di questi aspetti coinvolge attori con professionalità (e quindi competenze) diverse, che spesso hanno difficoltà a dialogare, in parte perché non parlano un linguaggio comune e in parte perché non riconoscono fino in fondo il ruolo degli altri. Ad esempio, un *digital humanist* che codifica in TEI un'edizione critica può avere difficoltà a far comprendere al linguista computazionale l'importanza di gestire la variantistica; parimenti, un ingegnere del software che implementa l'in-

frastruttura di ricerca può avere difficoltà a soddisfare la comunità dei filologi digitali che hanno la necessità di collegare su una piattaforma integrata fonti primarie (ad esempio immagini di manoscritti), secondarie (articoli specialistici) e terziarie (dizionari, encyclopedie, etc.).

Il convegno è stato animato da sei relazioni su invito riguardanti i temi chiave di ciascuna sessione, ventitré presentazioni orali (di cui un panel), ventidue poster, un workshop organizzato dal Centre for Digital Humanities dell'università di Göttingen sul riuso testuale, cinque note informative su invito relative a progetti in corso nell'ambito delle Digital Humanities e una tavola rotonda conclusiva sul tema della *cross-fertilization* fra cultura umanistica e cultura scientifica.

Pur avendo scelto una politica di inclusività verso tutti i contributi del convegno, la soglia di valutazione da parte dei *reviewers* necessaria ad ammettere le presentazioni orali nella forma originaria proposta dagli autori stessi è stata posta a sette decimi. Sotto tale soglia, gli autori sono stati invitati a convertire il loro contributo in poster.

Per questo motivo quindi, anche se sono stati ammessi al convegno tutti i 45 contributi sottoposti a *review* (100%), soltanto 23 contributi orali su 35 (66%) sono stati accettati nella forma originaria, mentre 12 su 35 (34%) sono stati accettati dopo la conversione in poster. I 10 interventi presentati dagli autori stessi come poster sono stati ammessi indipendentemente dal superamento o meno del livello di soglia, anche se vale la pena notare che due di essi hanno ricevuto valutazione eccellente.

Pur essendo concepito come il convegno annuale di un'associazione nazionale, AIUCD 2016 ha visto la presenza non solo di autori italiani (70%) ma anche di altri Paesi europei (Svizzera: 10%; Germania: 7%; Francia: 5%; Spagna: 2%; Olanda: 1%) e nordamericani (Canada: 3%; Stati Uniti: 1%), in linea del resto con la composizione del comitato scientifico, composto da studiosi che operano in Italia (86%), Germania (7%) e Stati Uniti (7%).

Tutti i materiali del convegno (*abstracts, slides, posters* e *videos*) sono accessibili online dalla *homepage* del convegno: <http://www.aiucd2016.unive.it> (ultima visita: 1 febbraio 2017), cui si rimanda per le affiliazioni degli autori, gli estremi dei progetti e i riferimenti bibliografici.

La prima sessione, dedicata ad infrastrutture, modelli e metodologie, si è aperta con l'intervento di Monica Monachini, che ha illustrato l'infrastruttura di ricerca per le risorse linguistiche CLARIN-it. Monachini ha messo in evidenza l'importanza delle infrastrutture per la reperibi-

lità, la citabilità e l'usabilità delle risorse depositate e documentate dagli utenti stessi. CLARIN-it garantisce inoltre la *long term preservation* degli oggetti digitali registrati e corredati di metadati, e ne favorisce la visibilità internazionale.

Odd Einar Haugen ha proseguito con il secondo intervento su invito, illustrando il flusso di lavoro per la costituzione del Medieval Nordic Text Archive (MeNoTA Project), accessibile tramite l'infrastruttura norvegese di CLARIN, denominata CLARINO. L'intervento ha insistito sulla necessità di codificare in parallelo diversi livelli di rappresentazione del testo, in modo da tenere allineate fra di loro l'edizione facsimile, l'edizione diplomatica e l'edizione normalizzata, usate dallo studioso per scopi differenti. Questo aspetto, declinato su altri casi di studio, è stato ripreso anche negli interventi di Raul Mordenti e di Paolo Monella. L'intervento di Elena Pierazzo ha spostato l'attenzione dai modelli teorici ai modelli economici e ai processi produttivi delle edizioni digitali, auspicando il passaggio dalle edizioni digitali elitarie (*haute couture*), di alta qualità ma molto costose rispetto ai mezzi messi a disposizione per la ricerca a causa di soluzioni *ad hoc*, alle edizioni digitali standardizzate (*prêt-à-porter*), di qualità forse inferiore ma comunque soddisfacente, con il vantaggio di essere molto meno costose grazie al riuso di soluzioni condivise.

La seconda sessione, dedicata alle edizioni digitali, ha ospitato interventi su alcune soluzioni adottate in diverse tipologie di documenti da codificare, ad esempio manoscritti o lettere (Roberto Vetrugno e Cecilia Russo), e su aspetti salienti del processo editoriale, come la trascrizione collaborativa di manoscritti (Troy Griffitts) e la collazione semi-automatica (Elena Spadini). Nella maggior parte degli interventi è emersa l'esigenza di fornire all'editore, in un ambiente integrato, strumenti di analisi e di comparazione dei contenuti trattati, come supporto all'attività editoriale stessa.

Forte rilievo è stato dato all'applicazione EVT (acronimo di *Edition Visualization Technology*), sviluppata presso l'Università di Pisa e impiegata in molteplici progetti, illustrati durante la presentazione dei poster.

Francesco Stella e Neven Jovanović hanno concluso con alcune riflessioni sui protocolli di digitalizzazione di grandi collezioni di testi come l'*Archivio della Latinità Italiana del Medioevo* (ALIM) e *Croatiae Auctores Latini* (CroALa).

La terza sessione, dedicata alla didattica e alla disseminazione, ha visto protagonisti ricercatori, insegnanti, dottorandi e studenti coinvolti in iniziative di alternanza scuola-lavoro (ASL), stage curriculare e progetti-

pilota universitari. Il gruppo di liceali coordinato da Paola Tomè e Toti Rizzetto ha illustrato un progetto di trascrizione collaborativa di un *incunabulum*. Serena Agodi e Daria Spampinato hanno mostrato i risultati del lavoro svolto con gli studenti in ASL per l'edizione digitale del *corpus* epigrafico del Museo Civico Castello Ursino di Catania. Antonio Stanzone e Giulia Re, studenti dell'Università di Pisa, hanno parlato della loro esperienza di stage per la validazione di alcune sezioni tematiche di Homeric Greek WordNet, sviluppata presso il CNR-ILC. Barbara Balbi infine ha illustrato un sistema di visualizzazione 3D dei dati acquisiti con l'*eye tracker* per studiare la fruizione di opere d'arte da parte del pubblico tramite l'osservazione del movimento degli occhi sulla tela.

La sessione è stata chiusa da un panel proposto dai membri dell'*Illustrated Dante Project*, basato sulla digitalizzazione del patrimonio iconografico dantesco.

La quarta sessione, dedicata all'analisi testuale, è stata aperta dall'intervento del presidente dell'AIUCD Fabio Ciotti, che ha messo in luce alcuni limiti dei metodi statistici applicati a *corpora* diacronici di testi letterari, soprattutto perché i testi, in quanto oggetti intenzionali, richiedono un atto interpretativo che sfugge alla regolarità misurabile tramite metodi quantitativi. I successivi tre interventi hanno affrontato aspetti diversi di *Memorata Poetis*, progetto collaborativo per l'annotazione di temi e motivi in testi poetici brevi multilingui e multiculturali. Nel primo, Silvia Arrigoni ha fornito alcuni dati statistici per misurare il modo di operare degli annotatori. Nel secondo, Ouafae Nahli ha illustrato le caratteristiche del *subcorpus* di testi arabi contenuti in *Memorata Poetis* e le difficoltà di adattare la lista di temi e motivi pensata per testi della tradizione occidentale a testi di altre tradizioni. Nel terzo, Daniele Silvi ha parlato della trasposizione della tassonomia di temi e motivi in una ontologia di dominio, iniziata presso il CNR-ILC, e dell'applicazione di una ontologia per la georeferenziazione dei luoghi nominati nei testi poetici. Infine, Andrea Bolioli ha discusso sulla necessità di creare strumenti per l'annotazione semantica delle edizioni scientifiche digitali, per rendere autonomo l'umanista digitale da figure intermedie, come lo sviluppatore web.

La sessione è stata chiusa da due interventi su invito. Susanna de Beer ha parlato del progetto *Digital Roman Heritage*, dove i testi sono semanticamente annotati e collegati a informazioni di tipo geografico, intertestuale e iconografico. Rodolfo Delmonte ha discusso i risultati di uno studio approfondito sulle strutture sintattiche non canoniche nella

poesia italiana, mostrando come gli strumenti della linguistica computazionale possono intervenire molto efficacemente a sostegno della critica letteraria.

La quinta sessione, dedicata a strumenti ed architetture, si è aperta con una riflessione presentata da Angelo Mario Del Grosso sui vantaggi dell'astrazione per il *Digital Scholarly Editing*, al fine di implementare moduli software manutenibili, estendibili, sostituibili, riusabili e dinamici. Gli *Abstract Data Types* (ADT) permettono infatti di spostare l'attenzione dal valore e dalla rappresentazione del dato al comportamento del componente, definito dalla sua *Application Programming Interface* (API).

Anche l'intervento di Renzo Orsini ha insistito sull'importanza di modellare non soltanto i fenomeni testuali, come già fanno i linguaggi di mark-up, ma anche le azioni, i comportamenti sul testo. Per questo ha proposto un linguaggio di *scripting* dedicato al *Digital Scholarly Editing*.

L'intervento di Nicola Barbuti si è focalizzato su tecniche di *graphic matching* per identificare regioni graficamente simili all'interno di basi di dati contenenti immagini di parole estratte da manoscritti.

Eleonora Litta e Marco Carlo Passarotti hanno presentato infine una risorsa lessicale per lo studio della formazione delle parole (derivazione e composizione) nella lingua latina. La sessione è stata chiusa da brevi presentazioni di progetti istituzionali di Digital Humanities.

La tavola rotonda, moderata da Marina Buzzoni, ha avuto per titolo *Intersezioni*, in quanto è stata dedicata al rapporto fecondo (*cross-fertilization*) fra cultura umanistica e cultura scientifica.

Il dibattito si è aperto con le osservazioni di Federico Meschini sul *digital thinking*, nella sua componente qualitativa che va verso la capacità di narrare, di cucire fra di loro frammenti di informazione, e la sua componente quantitativa, che va verso la tracciatura dei contatti e dei consensi con la conseguente promozione della visibilità, nel contesto del *social web*.

Federico Boschetti ha notato come le Digital Humanities non siano il frutto di una sola intersezione (fra cultura umanistica da un lato e cultura scientifica dall'altro), ma di grappoli di intersezioni (fra singole discipline umanistiche e singole discipline scientifiche, come ad esempio la filologia e la bioinformatica oppure la storia dell'arte e il *visual computing*). Ciò comporta la difficoltà a riconoscere in una terminologia condivisa, in una bibliografia fondativa. Anche per questo è necessario costruire infrastrutture che favoriscano l'interazione fra gli attori

delle Digital Humanities. Boschetti ha poi suggerito un triplice ruolo dell'umanista digitale: in testa ad un percorso di ricerca, come ideatore del tema di ricerca e coordinatore di un gruppo di lavoro multidisciplinare; oppure al centro del percorso, come mediatore fra l'umanista tradizionale, che si fa promotore del progetto di ricerca e gli informatici, che assumono il ruolo di esecutori; oppure in coda al percorso, per disseminare i risultati di un progetto multidisciplinare presso le diverse comunità d'interesse.

Monica Monachini ha ribadito l'importanza della *cross-fertilization* fra le due culture e il valore delle infrastrutture per promuoverla.

Maristella Agosti ha denunciato il fatto che in Italia la cultura scientifica è sentita ancora secondaria rispetto alla cultura umanistica e ha evidenziato il rischio corso da molti di confondere la formazione scientifica con la formazione tecnico-ingegneristica. Agosti ha poi spostato l'attenzione sulla didattica, asserendo la necessità da parte dei docenti universitari di fornire agli studenti gli strumenti per creare competenze nuove, capaci di dar loro la giusta collocazione professionale senza costringerli a dover emigrare all'estero, tenendo conto anche dell'enorme patrimonio culturale dell'Italia.

Fabio Ciotti, in parziale disaccordo con Agosti, ha precisato che oggi è la cultura umanistica ad essere profondamente in crisi. Per gli studi umanistici, è il momento di ripensare una giustificazione storica e sociale all'altezza dei tempi. Per questo le Digital Humanities ricoprono un ruolo cruciale.

Marina Buzzoni ha osservato che chi si occupa di Digital Humanities spesso considera gli informatici come dei tecnici, e questo non deve avvenire, perché sia gli umanisti che gli informatici devono avere un approccio scientifico al testo, devono discutere insieme e collaborare alla modellazione del loro oggetto di studio.

Renzo Orsini ha dichiarato che l'umanista non deve autolimitarsi ma deve avere obiettivi ambiziosi, utili per la società, e deve pensare che l'informatico possa realizzare tali sogni.

Federico Boschetti, riprendendo la parola, ha sottolineato ancora una volta l'importanza di un intermediario, un interprete dei sogni, fra l'umanista e l'informatico, rappresentato appunto dal *digital humanist*, che traduce le richieste del primo in un modo operativamente comprensibile al secondo.