

Guido MILANESE, *Filologia, letteratura, computer. Idee e strumenti per l'informatica umanistica*, Vita e Pensiero, Milano 2020, 344 pp., ISBN 978-88-343-3751-6.

In questo manuale, diviso in 17 capitoli, Guido Milanese si propone di analizzare come l'impiego di strumenti informatici possa contribuire alla comprensione e all'interpretazione di testi letterari. Come Tito Orlandi in *Informatica testuale* (Roma-Bari 2010), l'obiettivo dell'A. è quello di «tenere sempre presente una considerazione teorica dei problemi, insieme a una guida pragmatica delle procedure» (9). In un paragrafo del primo capitolo, in cui vengono esposti i punti centrali che tratterà nel manuale, l'A. specifica esattamente «Che cosa non è questo libro»: l'intenzione non è quella di scrivere un testo di linguistica computazionale e nemmeno quella di analizzare il significato umanistico del *web* (12). L'A. passa poi a presentare in breve la storia del rapporto tra informatica e scienze umanistiche: grazie alla scoperta di vocabolari in ittita, sumerico, accadico e urrita, l'esigenza di analizzare e riordinare i testi nasce già nel secondo millennio a.C., esigenza che si ritrova anche nel mondo classico greco e latino. Attraverso l'analisi delle modalità di riordinare i testi e i loro contenuti nell'epoca ellenistica, romana e medievale, concentrandosi anche sulle schedature della Bibbia (18), l'A. spiega come grazie al «salto tecnologico», consistente prima nell'invenzione del libro e poi nel passaggio dal manoscritto alla stampa, sono cambiati i metodi di catalogazione. Tra l'Ottocento e il Novecento si ritrova nella fotografia uno strumento utile nel poter confrontare innumerevoli manoscritti, disponibili soprattutto in ambito biblico e liturgico (22-41). L'A. conclude la prima parte del suo volume esponendo come il computer sia entrato nel mondo della filologia e dell'umanistica con gli studi di Dom Jacques Froger e René-Jean Hesbert (42-45).

Nella seconda parte l'A. si sposta sulla letteratura digitale, esplorando implicazioni culturali e storiche dell'uso dei computer nella scrittura e nella pubblicazione di opere letterarie. In questo contesto, l'A. spiega come non vi sia alcuna certezza di permanenza nel futuro nell'affidare informazioni alla memoria dei computer, a causa dell'invecchiamento digitale, cioè di *hardware*, per esempio il *floppy disk* e la chiavetta usb, e di *software*, prendendo in esame due esempi: *WordPerfect* e *ChiWriter* (62-63). Una soluzione, definita dall'A. però insoddisfacente, potrebbe essere salvare il *file* in formato testo e visualizzarlo con un programma di edito-



re di testi, in cui scompaiono però tutte le caratteristiche metalinguistiche (67-70). La conclusione della seconda parte è dedicata all'idea del *Free-Open Software*, che risponde alla richiesta di mantenere integri e trasmettere nel tempo le informazioni digitali, idealizzato da Richard Stallman nei primi anni '80 con il progetto GNU.

Nella terza parte l'A. descrive la preparazione dei testi che devono essere analizzati sul computer. Spiega quindi che cosa si intenda per digitalizzazione, chiarendo come si sia passati dal copiare manualmente il testo dal libro cartaceo sul computer, rischiando ovviamente di commettere errori, alla scansione delle diverse pagine e all'invenzione di programmi di OCR, che permettono il riconoscimento ottico dei caratteri: ne è un esempio *Tesseract* (97-107). Di seguito, elenca i diversi archivi in cui è possibile consultare testi digitalizzati ovviamente senza diritto d'autore, che l'A. definisce «banche dati». L'A. dedica singoli paragrafi, analizzandoli brevemente, ma in modo efficace, a siti molto utili alla consultazione di libri, distinguendo opportunamente «banche dati generaliste», ovvero Gutenberg, Manuzio, Internet Archive, Google Books, HathiTrust (108-111), e «banche dati specializzate», ovvero siti dedicati principalmente alla letteratura classica greca e latina (*Perseus*, *Packard Humanities Institute*, *Musisque Deoque* e *digilibLT*, *Library of Latin Text*, *Thesaurus Linguae Graecae*), e siti che riguardano invece la letteratura moderna, come LION e *Bibliotheca Augustana* (111-118). L'A. passa poi alla codifica dei caratteri, spiegando a che cosa serve e come funziona la tabella ASCII (*American Standard Code for Information Interchange*), per poi definire lo sviluppo di *Unicode*, che «rappresenta lo standard comune che permette di unificare la comunicazione tra sistemi operativi diversi senza più i problemi che si verificavano prima dell'introduzione di questo standard» (121). Dopo aver descritto i metodi di acquisizione del testo, l'A. nel capitolo 7 si concentra sulla codifica del testo e su come sia possibile, attraverso il «tagging gerarchico», comunicare al computer caratteristiche metalinguistiche del testo. Ciò è possibile attraverso linguaggi di marcatura, ovvero i linguaggi XML (145-151) e HTML (164-171), che l'A. riesce a illustrare in modo molto esaustivo attraverso esempi semplici e calzanti (145- 151).

Nella quarta parte l'A. si propone di dimostrare come i risultati, che si ottengono attraverso l'utilizzo di strumenti informatici, siano utili per l'analisi del testo. È possibile, infatti, creare liste di parole, indici e concordanze attraverso programmi come *TextStat*, di cui l'A. esemplifi-

ca le capacità in modo semplice e chiaro. Il rapporto tra l'informatica e il filologo è oggetto del capitolo 9, in cui l'A. espone strumenti informatici utili alla collazione, come i programmi *TUSTEP* e *Juxta*, che è in grado, inoltre, di generare un apparato critico. Ormai esistono molti programmi che rendono il lavoro del filologo più veloce, ma è opportuno scegliere quelli più adatti alle proprie esigenze. Pertanto, l'A. spiega come «costruire i propri strumenti» nel capitolo 10, esaminando alcuni strumenti, come i linguaggi di programmazione per l'analisi dei testi, ovvero *Perl*, *Python*, *Prolog* e *Snobol* (237-244).

La quinta parte del manuale è centrata sull'analisi delle diverse modalità di scrittura di un testo sul computer, partendo da programmi di *word processing* fino a *TEX*, un sistema di tipografia digitale su cui l'A., nel capitolo 12, si concentra esaustivamente. Infine, l'A. affronta il problema della bibliografia in termini quantitativi, mostrando come funzionano in questo senso i cataloghi di biblioteche *online*.

«Tutto il resto»: in questo modo l'A. intitola la sesta e ultima parte del suo manuale, che riguarda la codifica, la scrittura e l'analisi della musica, la digitalizzazione dei monumenti e dei reperti e la glottodidattica (290-300). Qui l'A. elenca in modo utile e sintetico, indicando il numero di pagina del suo trattato, le persone che hanno contribuito significativamente allo sviluppo dell'informatica umanistica, il lessico tecnico, traducendo in italiano termini in inglese, i siti e le pagine *web* di cui ha trattato (309-313). Il testo si chiude con un indice delle figure, delle tabelle, degli esempi, che aiutano la comprensione dei vari argomenti affrontanti (315-325).

La lettura di questo manuale, che ben mantiene la promessa di coniugare teoria e prassi, risulta molto illuminante. Come con l'esempio dell'orologio o del cronometro, usato per spiegare il significato del verbo «digitalizzare» (97-99), a mio avviso l'A. riesce a spiegare al lettore in modo esaustivo gli argomenti affrontati attraverso l'utilizzo di esempi, che oserei definire inconsueti, ma che aiutano a comprendere concetti piuttosto complessi.

Eva MENGA