

Tito ORLANDI

Informatica umanistica: realizzazioni e prospettive

Una definizione meditata del tema "informatica umanistica" pone non pochi problemi, (1) ma lasciandoli per il momento da parte, e intendendo provvisoriamente l'informatica umanistica come lo studio circa l'utilizzazione del "computer" per aiutare la ricerca e la diffusione dell'informazione nell'ambito delle discipline umanistiche, potremo tratteggiarne la storia tenendo presente che il computer, cioè il calcolatore elettronico digitale, può ragionevolmente considerarsi nato nel 1945. (2)

Fino al 1965 abbiamo a che fare con pochi pionieri. I dettagli non sono facili da trovare, (3) ma possiamo accennare ai lavori di Roberto Busa sul testo di S. Tommaso, Roy A. Wisbey sui testi antico-alto-tedeschi, Félicien de Tollenare sulla Bibbia gotica e poi sulla lessicografia olandese, Bernard Quemada sulla lessicografia francese, A. Ellegard sull'inglese del Settecento (problema dell'attribuzione delle c.d. Junius letters), Andrew Morton sul testo greco del Nuovo Testamento.

Dal 1966 (data della fondazione della rivista *Computers and the Humanities*, tuttora la più importante in questo campo) al 1980, si annovera un numero ancora ristretto, ma più consapevole e agguerrito, di ricercatori, che formano anche le prime strutture di contatto e di appoggio (soprattutto le associazioni *Computers and the Humanities*, *Association for Literary and Linguistic Computing* e *Association for Computational Linguistics*). Nel settore degli studi linguistici e letterari è questa la prima epoca della produzione di concordanze su vasta scala. Di poco precedenti sono le prime applicazioni in ambito archeologico, per le quali è da ricordare soprattutto l'équipe di Jean-Claude Gardin a Parigi.

Dal 1980, in concomitanza con la diffusione del Personal Computer, si assiste ad un moltiplicarsi di iniziative, sia singole sia di collaborazione o istituzionali, che ha portato le applicazioni umanistiche all'attenzione anche degli studiosi più tradizionalisti, oltre che di vaste parti di un'opinione pubblica, incuriosita ma anche sconcertata di fronte a questa novità.

Fino a ieri, in effetti, era consueto iniziare le manifestazioni dedicate all'informatica umanistica con la spiegazione di come uno strumento adatto a operazioni di tipo numerico potesse diventare un ausilio anche per discipline a carattere linguistico. Oggi è comune esperienza constatare come il computer possa immagazzinare e analizzare sequenze di parole e dunque quel genere di spiegazione è diventato superfluo. E' forse meno noto che fin dal 1949 William Weaver sollevò la

questione se i computer potessero essere usati per la traduzione meccanica da una lingua ad un'altra, (4) o quanto le applicazioni dell'informatica nelle discipline umanistiche siano ormai estese e numerose anche al livello propriamente della ricerca. Di questa situazione sembra opportuno dare un'idea concreta. (5)

Associazioni Vi sono almeno cinque associazioni internazionali di cultori di informatica umanistica o importanti parti di essa: Association for Computers and the Humanities (residenza non dichiarata); Association for Literary and Linguistic Computing (Oxford) e Association for Computational Linguistics (residenza non dichiarata); Association Internationale Bible et Informatique (Maredsous, Francia); History and Computing (Londra; ha branche nazionali, fra cui una italiana, Bologna). Queste stesse associazioni organizzano i **congressi** "stabili", periodici, sull'argomento di loro competenza; mentre sono ormai frequenti congressi occasionali su argomenti più ristretti.

Centri (6) Diamo un elenco dei maggiori centri di servizio e ricerca, con alcune specificazioni, dove questo non sia chiaro dal nome stesso del centro:

Literary and Linguistic Computing Center, Cambridge
 Oxford University Computing Services: Centre for Literature and Linguistic Studies
 Institut für deutsche Sprache, Mannheim (corpus di testi tedeschi)
 Zentrum für Datenverarbeitung, Tübingen (programma modulare di gestione di testi TUSTEP)
 CETEDOC Louvain-la-Neuve (corpus medievale latino)
 Istituto di Linguistica Computazionale, Pisa (corpus di testi italiani)
 Institut National de la Langue Française (Trésor de la Langue Française)
 Institut de Recherche et d'Histoire des Textes, Paris (banche dati su manoscritti e documenti)
 Global Jewish Database, Bar-Ilan University
 Centre Informatique et Bible, Maredsous
 Norwegian Computing Centre for the Humanities
 Centre de Recherche sur les Traitements Automatisés en Archéologie Classique, Nanterre
 National Center for Textual Data Base, Princeton
 Centre for Computing in the Humanities, Toronto
 Humanities Computing Facility Center, University of California at Santa Barbara

Humanities Computing Facility, University of California
 at Los Angeles
 Thesaurus Linguae Graecae Project, University of California
 at Irvine
 Lancaster-Oslo-Bergen Corpus of British English
 Brown University (USA) Corpus of American English
 Ricorderemo infine il neonato Centro per l'Automazione
 nelle Discipline Umanistiche (CISADU) presso la
 Facoltà di Lettere di Roma.

Ricerche e ricercatori (7) L'informatica si è introdotta ampiamente e profondamente nei procedimenti di realizzazione di vocabolari (eventualmente su CD-Rom), concordanze, analisi metriche, analisi di strutture linguistiche, edizioni di testi, pattern analysis di oggetti o luoghi archeologici, banche dati di oggetti archeologici e artistici, analisi statistica di fenomeni storici e archeologici, catalografia (museale, ma non solo).

Queste tecniche sono state utilizzate negli studi di: (8)

Lessicografia e lessicologia: Jacqueline Hamesse (filosofia medievale), CAAL (R. Busa, Gallarate: linguistica, soprattutto latina), Lessico Intellettuale Europeo (Roma; lessicografia filosofica), Ralph De Gorog (francese medievale), Félicien De Tollenaer (olandese mediev.), Saporetti (accadico)

Bibbia: E. Tov, H. van Dyke Parunak, E. Talstra, F. Postma (Antico Testamento); P. F. Johnson, Andrew Morton, Antony Kenny, M. Menken (Nuovo Testamento); Robert Kraft (Septuaginta).

Letterature classiche: CIPL (Liegi), J. Froger (ecdotica), Ghislaine Viré (ecdotica), Hubka (Sofocle, Commedia greca), David Packard (letteratura latina).

Letteratura: John Allen (filologia romanza), Joseph Duggan (filologia romanza), Patricia Galloway (filologia romanza), René Pellen (filologia romanza), Franklin Waltman (filologia romanza), D'Arco Silvio Avalle (italiano delle origini), Richard Bailey (stilometria) John Dawson (inglese), Gero Dolezalek (diritto medievale), Giuseppe Savoca (italiano del Novecento), Amedeo Quondam (poemi cavallereschi), Roy A. Wisbey (tedesco medievale), Barron Brainerd (Shakespeare), M.V.A. Smith (Shakespeare), J. F. Burrows (Jane Austen), Etienne Brunet (francese), Paul Fortier (francese), David Chisholm (tedesca), A.Q. Morton (stilometria), Anthony Kenny (Nuovo Testamento), Claude Cazalé Bernard (Boccaccio).

Archeologia: Jean-Claude Gardin, R. Ginouvès, Francois Djindjian, Anne-Marie Guimier-Sorbets, Mario Borillo, J. R. Clark, M. S. Lagrange, Paola Moscati, Clive Orton (temi e applicazioni generali); Amilcare Bietti, S. R. Binford, F. R. Hodson,

C. Renfrew (preistoria).

Storia: Domesday Project (Inghilterra contemporanea), C. Bourlet (medio evo francese), Jean-Philippe Genet (medio evo francese), Barbara Hanawalt (medio evo inglese), David Herlihy (Toscana rinascimentale), Christiane Klapisch-Zuber (catasto fiorentino), Marie-Thérèse Lorcin (Francia), Silio Scalfati (medio evo pisano), Karl Schmid (medio evo tedesco), D. van den Auweele (diritto medievale)
Storia dell'arte: Paola Barocchi (rinascimentale), Corrado Maltese (analisi stilistica), Maria Teresa Leoni Zanobini (analisi delle geometrie), Laura Corti (descrizione).

Riviste specializzate (9) Il solo elenco dei titoli dimostra l'estensione presa da questi studi: Computers and the Humanities, Le Médiéviste et l'Ordinateur, Literary & Linguistic Computing, ex ALLC Journal e Bulletin, Linguistica Computazionale, Bits & Bytes Review, Archeologia e Calcolatori, History and Computing, Histoire et mesure, Humanistke Data (Oslo), Sprache und Datenverarbeitung, Cahiers de Lexicologie, Archéologues et Ordinateurs.

Menzioneremo anche i **network**, reti di collegamento per la comunicazione in linea di informazioni e messaggi, gestiti da centri a cui viene attribuita la responsabilità di gestire razionalmente il traffico di dati. Oltre alle reti generali (Bitnet/Net-North/earn; tymnet; uninet; usenet), ve ne sono di specializzate: Humanet (a cura di North Carolina State University), Humanist (a cura del centro di Toronto, Willard McCarty), Humbul (University of Leicester), Archeological Information Exchange (S. Rahtz, University of Southampton). E finalmente le **Banche dati in linea**: Arts & Humanities Search, la bibliografia della Modern Language Association (USA), il Francis del CNRS (Francia), lo Humanities Index, i Dissertation Abstracts (sezione per le Humanities e Arts), il Philosophers Index.

Se si giudicasse dal solo punto di vista quantitativo, si concluderebbe facilmente che l'informatica umanistica è un settore di ricerca solidamente fondato, con connotati precisi e basi metodologiche riconosciute. Le cose purtroppo non stanno così. Una prima osservazione che si trae dalla lettura degli articoli e monografie che trattano delle ricerche sopra delinate è che troppe imprese annunciate sono poi fallite, talora per motivi di cambiamenti tecnologici, talora per errori di impostazione. Troppi dati (testuali, fattuali, grafici), posti con fatica e lungo lavoro su supporto magnetico, sono condannati a restare inutilizzati perché concepiti per un uso ristretto, che si è presto esaurito.

Più importante ci sembra una seconda osservazione. In quasi tutti i contributi manca una discussione sui metodi utilizzati, soprattutto in confronto a risultati ottenuti da altri, insomma un apparato di riferimenti bibliografici non generici, so-

prattutto su argomenti "militanti". E' sostanzialmente impossibile individuare correnti di metodo, caratteristiche di scuola, controversie tecniche di informatica, quasi che il delicato momento dell'interazione fra l'uso di programmi applicativi e il metodo di ricerca dello studioso non fosse un argomento discutibile, ma dipendesse esclusivamente da competenze in due settori del tutto distinti (informatica da una parte, disciplina umanistica dall'altra) che non si incontrano se non nella prassi.

Se si pensa che il repertorio bibliografico attualmente esistente, il *Humanities Computing Yearbook*, (10) nella sua seconda edizione consta di 700 pp. (di cui 100 di indici), per tutti i tipi di discipline umanistiche, e sia pure comprendendo la rassegna del software, si è portati a concludere che piuttosto che incoraggiare ad aumentare quella bibliografia, sarebbe opportuno cominciare a riflettere seriamente su quello che si sta facendo, prendendo posizione sul valore dei risultati ottenuti, da due punti di vista: la correttezza dei sistemi informatici utilizzati (sia per il hardware sia per il software), e la correttezza del metodo interno alle diverse discipline, che è stato applicato.

Vi sono in questo senso segnali positivi, che indicano come sia iniziata una riflessione su argomenti di metodo, che finalmente mette a fuoco i dati sostanziali del problema. Fino ad un tempo recente era usuale che opere collettive, che presentavano il panorama di tutti i campi applicativi, avessero introduzioni generiche che non giustificavano l'unità della pubblicazione; (11) ovvero poteva accadere che in un manuale come quello della Hockey, (12) che pure è rimasto per lungo tempo l'unico in questo campo, la parte metodologica non fosse neppure presa in considerazione; ovvero che si esaltassero soltanto le capacità tecniche del computer come macchina. (13) L'attenzione si sposta ora, in casi analoghi, a problemi di metodo, e per passarli in rassegna possiamo basarci su due saggi inglesi, pubblicati in una medesima raccolta, singolarmente in sintonia con il nostro punto di vista. (14)

David S. Miall (15) propone all'attenzione degli studiosi i seguenti temi:

Quali modelli mentali stiano alla base del crescente uso del computer nelle discipline umanistiche.

Quale ruolo i computer siano destinati a giocare nelle discipline umanistiche, e anche quale ruolo le discipline umanistiche possano giocare nello sviluppo del computing, inteso dunque non solo come settore tecnologico, ma come disciplina di carattere teorico.

Quali principi metodici vengano richiesti dall'uso dei computer nelle discipline umanistiche: come esse siano concettualizzate, quali modelli del pensiero e della comunicazione umana i computer incarnino.

Il fatto che l'introduzione dei metodi computerizzati richieda che definiamo, forse ridefiniamo, che cosa intendiamo per discipline umanistiche e come esse debbano

essere insegnate.

Se il "Humanities computing" sia identificabile come una disciplina. Vengono proposte due tesi: (16)

La tesi positiva nasce dalla semplice constatazione della molteplicità delle aree applicative, che dunque presuppone una unità diciamo così gestionale.

La tesi negativa (17) è basata sulla convinzione che il computer nelle mani di un umanista rimarrà uno strumento piuttosto che una "macchina logica", perché i mondi del pensiero umanistico sembrano particolarmente resistenti ad essere definiti in termini computazionali (cioè logico-formali!).

Se il modello di computer rimane quello derivato dalla capacità di analizzare elementi discreti, arbitrari, finiti e isolati (Bolter, cit., p. 75-8), le discipline umanistiche sono escluse. Ma questo cambierà con la possibilità di valersi su vasta scala di nuove tecniche di intelligenza artificiale e delle logiche alternative.

Nello stesso volume, Peter Denley intitola apertamente il suo saggio: *The Computer Revolution and 'Redefining the Humanities'*. (18) Ecco i temi che egli tocca:

La piena realizzazione del potenziale dello "humanities computing" è impedita dalla mancanza di comprensione di ciò che è in gioco.

La crisi (finanziaria) delle discipline umanistiche come parte dell'educazione dell'individuo, dall'impatto humboldtiano (scientificità) all'affermarsi dell'utilitarismo: la computerizzazione è potuta avvenire nonostante la crisi.

L'uso del computer ha portato un approfondimento e rinnovamento dei problemi metodologici.

Rapporti con gli scienziati.

Rapporti con gli "altri" umanisti: i cambiamenti di mentalità in senso pratico, ma anche la mancanza di una chiara direzione e chiari principi fondamentali.

Di nuovo il problema dell'informatica umanistica come disciplina. La domanda "che cos'è l'informatica umanistica" non ha ancora una risposta soddisfacente.

I problemi sono dunque sotto gli occhi di tutti coloro che vogliono riflettere su questa materia. Nè vale la pena di discutere dell'opportunità stessa dell'uso del computer nelle discipline umanistiche come vero e proprio strumento di ricerca, e non come una macchina da scrivere sofisticata. Un minimo di senso storico è sufficiente a comprendere come l'invenzione del computer abbia innescato una rivoluzione (prima tecnologica, e poi anche mentale) ineluttabile ed inarrestabile, anche nei confronti delle discipline umanistiche, dunque qualcosa che sovrasta la volontà dei singoli studiosi.

Quello che conta è dunque come utilizzare bene il computer, o meglio come

porre correttamente i rapporti fra informatica e discipline umanistiche. Questo della distinzione fra uso del computer e interazione con l'informatica è un punto delicato, che deve essere chiarito accuratamente. Infatti il computer non è uno strumento come altri, utilizzati da gran tempo, come per esempio la stampa. E' ben noto come la stampa abbia comunque portato una rivoluzione anche metodologica nelle discipline umanistiche, ma essa ha influito sulla quantità di informazione che si è potuta diffondere nella società, piuttosto che (e in via comunque secondaria o indotta) sulla loro qualità, perché la stampa è uno strumento che continua in maniera analogica quello precedente, cioè il manoscritto. Con il computer la questione è radicalmente diversa.

Gli informatici insegnano che, per poter valutare nei suoi aspetti generali la funzione del computer, occorre prendere in considerazione non le singole realizzazioni pratiche a cui esso contribuisce, ma i principi secondo i quali esso funziona. Infatti il computer è diverso dagli altri strumenti, in quanto è la realizzazione materiale di processi o procedure che hanno un fondamento prevalentemente teorico, in particolare logico. Ne consegue che, mentre una macchina "normale" si può definire dall'azione che compie una volta messa in moto; il computer è invece una macchina "general purpose", cioè capace di agire nelle maniere più diverse. Si aggiunge che esso per lo più elabora informazione, cioè un oggetto estremamente sfuggente, sul quale anche le elaborazioni possibili risultano in certo modo sfuggenti.

Il computer assume aspetti diversissimi e può essere costruito con materiali diversissimi. Ciò che sta alla base di questa diversità, è un meccanismo intimo, che guida le procedure operative. Sarà dunque questo "meccanismo" che dovrà fornire la chiave per intendere correttamente i rapporti fra il computer e le discipline umanistiche alle quali viene applicato. Il problema si sposta perciò dalla macchina alla disciplina "responsabile" della macchina.

Questo non semplifica la situazione, perché molte sono le discipline alle quali è possibile rivolgersi per avere un'idea esatta del funzionamento del computer; per esempio la matematica, la micro-elettronica, la fisica dei conduttori, e anche la cibernetica e la scienza delle comunicazioni. Nella storia del computo automatico tutte queste discipline hanno avuto un ruolo, e bene Paolo Ercoli (19) sottolinea la non ineluttabilità che l'informatica costituisca una disciplina a sè (contrariamente p.es. alla micro-elettronica o alla fisica dei semiconduttori, che sono facilmente individuabili).

Si constata tuttavia che ciascuna di quelle discipline riguarda un aspetto parziale del computer e di ciò che esso può fare. Il computer, come è stato argutamente osservato, (20) è costituito di un corpo e di un'anima, cioè di hardware e software. Sarà l'interesse per quest'ultimo, cioè per le istruzioni che mettono il computer in condizione di compiere le funzioni volute in un dato momento, a con-

durci sulla strada buona. L'informatica come disciplina è figlia del software, cioè dei programmi e dei linguaggi per esprimerli.

Gli specialisti hanno a lungo dibattuto (e in parte dibattono ancora) il problema se l'informatica sia una disciplina individuabile e con proprio fondamenti metodologici; e quali siano i rapporti con discipline che sono state proposte in tempi coevi, come la cibernetica, la teoria degli algoritmi, la teoria dell'informazione, etc. Per quanto è possibile dedurre dalla recente letteratura, si va imponendo un tipo abbastanza preciso di informatica (o per meglio dire di fondamenti di informatica, che è quanto ci interessa in questa sede), quella che viene chiamata Computer Science. Il termine esprimeva anni fa un contenuto anche generico, ma oggi viene inteso in modo preciso. Scrive Daniel Cohen (21) che la teoria dei Computer "è quella secondo la quale si formulano alcuni modelli matematici astratti, che descrivono con vari gradi di accuratezza parti di computer e tipi di computer e macchine simili. Essi sono usati per discutere le questioni più astratte dei limiti di capacità di questi strumenti meccanici". (22) Secondo questo punto di vista l'Informatica Teorica è la descrizione e lo studio di modelli simbolici e formali di macchine capaci di "computare"; e la computazione viene poi identificata con la "analisi di linguaggi".

Rileva Thomas Sudkamp: (23) "Lo studio dei linguaggi formali ha le sue radici nella linguistica, nella logica matematica, e nella computer science. Le grammatiche "phrase structured" sono state sviluppate per provvedere un meccanismo per descrivere linguaggi naturali. I matematici interessati alla computazione algoritmica hanno disegnato macchine astratte per studiare le possibilità e le limitazioni del processo computazionale. I principi della definizione del linguaggio dei computer e il disegno dei compilatori hanno i loro fondamenti in questi studi linguistici e matematici."

Anton Nijholt (24) pone l'accento sulla comunicazione fra il computer e l'uomo che lo utilizza. Essa presuppone l'uso di un linguaggio, che in teoria può essere un linguaggio di programmazione, un linguaggio formale, o un linguaggio naturale. Per ciascuno di essi la computer science ha sviluppato una appropriata teoria, ma tutte hanno in comune la possibilità del "parsing", cioè di un'analisi basata sull'identificazione delle componenti e delle regole per il loro inserimento in una frase significativa. Questa analisi deve poter determinare se una data sequenza di simboli fa parte di un linguaggio, e in tal caso renderne esplicita la struttura sintattica.

Tutto questo dovrebbe bastare a suscitare l'interesse degli umanisti, sia pure tenendo conto che il loro concetto di linguaggio è di solito differente da quello degli informatici. E' chiaro comunque che l'uso corretto del computer richiede la cosapevolezza dell'aspetto teorico che sta alla base dei vari tipi della macchina "fisica" che in pratica viene poi usata.

E' interessante vedere come gli studiosi stessi di "informatica teorica" ne descrivano l'utilità per gli informatici in generale: (25) "Gli specialisti in campo informatico, nella loro vita quotidiana, devono essere in grado di tradurre i problemi reali in astrazioni basate sull'impiego di modelli formali e di servirsi di tali descrizioni analizzandone le proprietà con rigore matematico. Questa attitudine, molto speciale, differenzia gli specialisti informatici dalla maggior parte dei professionisti di altri settori della tecnica." Questa capacità di astrarre e produrre modelli formali serve non a costruire macchine, ma ad applicarle. Si tratta di formalizzare non problemi di informatica, ma i problemi della realtà o comunque dell'oggetto che deve essere "trattato", sottoposto a procedimento, da parte del computer. Dunque questo compito non può essere lasciato ad un "informatico puro", nel caso delle discipline umanistiche, ma condiviso con lui da parte dello specialista di letteratura, di storia, di archeologia etc.

Proseguono Ghezzi e Mandrioli (p. 4): "...lo scopo principale di questo libro è di presentare l'informatica teorica come un fondamento unitario per tutti gli studi in campo informatico." E a p. 9: "...come l'ingegnere meccanico deve saper padroneggiare un bagaglio di conoscenze tecniche che include i principi della statica, della meccanica, della termodinamica, per poter progettare macchine e impianti efficaci... così l'informatico deve padroneggiare gli strumenti teorici della propria disciplina per progettare applicazioni che rispondano sempre meglio ai requisiti di qualità..."

A Cohen e a Ghezzi-Mandrioli è comune l'accento messo sull'opportunità che la Computer Science (o Theory) sia davvero messa a fondamento e non a coronamento della carriera di studio dello studente in informatica. La domanda allora si presenta: che cosa pensare nel caso di uno studente (o studioso) di informatica umanistica? Si deve ricordare che i legami fra l'informatica come disciplina e le discipline umanistiche sono per lo meno assai più forti di quelli con altre discipline "di applicazione", proprio a causa degli aspetti logici e linguistici; ed inoltre di quelli filosofici generali, per quanto attiene alla corrispondenza fra simboli e realtà, e al problema della formalizzazione del pensiero. Sorge a questo punto il problema se l'informatica o una sua parte costituisca "una" disciplina umanistica. Si sarebbe tentati di qualificare questo problema come futile, mentre è importante per le implicazioni anche operative, sia di critica sia di organizzazione.

Dal momento che le applicazioni del computer sono guidate dall'informatica teorica, la risposta a questa domanda non può essere se non che quello che conta non è il rapporto fra le discipline umanistiche e il computer, ma fra esse e l'informatica. Mario Borillo propone osservazioni molto giuste a questo proposito: "Dall'origine, si sono stabilite due concezioni sulla natura di questo rapporto. La prima, che è ancora la più corrente, lo considera soprattutto da un punto di vista operativo, ed in effetti il computer è un strumento capace di modificare profonda-

mente le condizioni materiali nelle quali si effettua la ricerca nelle scienze umane. L'altro punto di vista lega questa dimensione tecnica al suo fondamento concettuale (logico-matematico) e vede nel ricorso all'informatica un fattore suscettibile di far evolvere anche il quadro teorico e i referenti epistemologici delle scienze umane." (26) E più oltre (p. 6): "Le mie ricerche... si sono focalizzate su quelle che ritengo il nocciolo del problema, ... cioè le implicazioni metodologiche e teoriche delle tecniche del trattamento dell'informazione."

Il rapporto fra informatica e discipline umanistiche si può esprimere nella questione se vi sia un modo "informatico" di vedere (anche) le discipline umanistiche, che si differenzia a seconda delle discipline (che dunque in questo caso rappresentano l'oggetto di questa disciplina), ma che rimane unitario nel modo di considerarle. Il "modo informatico" prevede la formalizzazione dei dati (in senso molto lato), e la formalizzazione delle procedure per analizzarli e valutarli (algoritmizzazione); ma non consiste affatto nella quantificazione e tanto meno matematizzazione dei dati, nè nell'usare procedure, cioè metodi, quantitativi, in particolare statistici. Esso caso mai pone interrogativi sul valore di tali metodi (adottati in realtà da sempre! cf. sotto) all'interno della metodologia generale delle singole discipline, e in confronto con altri tipi di metodi non informatici. Esso inoltre studia il modo corretto di applicare in pratica le metodologie suggerite dall'informatica.

E' interessante constatare come l'esigenza di formalizzare sia i dati sia il metodo di studiarli, tipica dell'informatica, sia considerata come naturalmente antitetica al carattere delle discipline umanistiche da specialisti come Miall, che pure sono convinti della necessità dell'informatica, ma ritengono che la sua applicazione umanistica sia subordinata a progressi discutibili portati dall'intelligenza artificiale o dal piegare il computer alle cosiddette logiche alternative. In effetti è di moda attendersi soluzioni da tecnologie che assumono nomi accattivanti o misteriosi come: intelligenza artificiale, calcolo parallelo, calcolatori neuronali, e simili. Non mi sembra questa la strada corretta, perché tali tecnologie, utili quanto si voglia per l'informatica in senso stretto, non possono di per sé portare mutamenti o chiarimenti metodologici nei confronti delle applicazioni umanistiche.

Un'attenta considerazione storica delle discipline umanistiche rivela invece che le esigenze che l'informatica dovrebbe "imporre" alle discipline umanistiche sono state sentite abbastanza ampiamente ben prima che il computer fosse, non dico utilizzato, ma addirittura inventato. Infatti il desiderio da parte di alcuni studiosi, particolarmente attenti agli aspetti metodologici, di formalizzare le procedure relative alla propria disciplina, nasce almeno nel sec. XIX, e costituisce un movimento che coinvolge i migliori esponenti delle discipline a cui stiamo pensando.

Gli esempi che si possono fare sono parecchi. In primo luogo la linguistica, per la quale le teorie prodotte dal De Saussure (soprattutto con la distinzione fra studio diacronico e sincronico di una lingua, e quella fra lingue e parole) costituis-

cono un primo stadio di formalizzazione, che conduce direttamente alle teorie di Bar Hillel e Chomsky, che a loro volta si collegano all'uso del computer. Del resto anche nella glottologia la teorizzazione dell'indo-europeo prodotta da Franz Bopp si può considerare basata sulla formalizzazione del sanscrito, visto come griglia completa di fenomeni che, tenuto conto di mutamenti storici, si ritrovano parzialmente nelle altre lingue della famiglia indo-europea. (27)

In filologia, o meglio in ecdotica, le teorie di Lachmann, o comunque a lui attribuite, rappresentano la costruzione del sistema per formalizzare il rapporto fra i manoscritti in un albero genealogico (un "grafo" in termini informatici) per mezzo del quale ottenere un testo primitivo a partire dai dati offerti dai manoscritti attualmente disponibili. Un'elaborazione di queste teorie fu tentata da Froger in vista dell'automazione dei procedimenti ecdotici. (28)

In archeologia la scuola americana detta della New Archaeology, che propugnò un maggior rigore e una formalizzazione nei metodi euristici, in confronto alla scuola cosiddetta storicistica, basata maggiormente sull'apprezzamento di dati singoli, nacque del tutto indipendentemente dall'eventuale uso del computer, e comunque dalle ricerche speculative sull'informatica. (29)

Analogamente, la scuola della cosiddetta storia quantitativa, che pure ha avuto con l'uso del computer uno sviluppo che probabilmente i suoi stessi fondatori non si attendevano, è sorta per motivi ideologici interni alla problematica storiografica, e non certo sotto l'influsso dei progressi dell'informatica. (30)

E gli esempi potrebbero continuare. A noi importa sottolineare come l'informatica correttamente intesa risulti non come un'invasione di metodologie estranee nel campo umanistico, ma piuttosto come il forte potenziamento di metodologie già esistenti. La discussione deve allora spostarsi sul piano della relativa validità di tali metodologie, rispetto ad altre. Si tratta dunque di una discussione familiare agli studiosi, alla quale l'informatica può portare il bagaglio di esperienze e riflessioni fatte anche in altri settori. (31)

Aggiungeremo che l'informatica mette in forte risalto un altro aspetto, che pure era già parte del patrimonio metodologico delle discipline umanistiche, ma in maniera non del tutto consapevole. Alludo al fatto che ogni risultato dell'indagine umanistica è un momento di un processo che con Jean-Claude Gardin potremo definire "a spirale", (32) per il quale l'analisi e la "spiegazione" (storica, letteraria, archeologica, etc.) dei dati determina una diversa formalizzazione dei dati stessi (si pensi ad un testo memorizzato o ad un catalogo di oggetti o simili), che diventa a sua volta un diverso oggetto delle successive analisi. In fondo questo equivale a quanto è ovvio per ogni studioso, cioè che quanto è stato "pensato" fino a quel momento riguardo ad un dato argomento, diventa il punto di partenza per la continuazione del lavoro di ricerca.

Tuttavia il modo convenzionale è quello di vedere il lavoro umanistico come

la continua rielaborazione di un medesimo oggetto (un periodo storico, un'opera letteraria, etc.) da parte di diversi studiosi, che dialogano fra loro a distanza di tempo. Il punto di vista informatico porta invece a considerare anche l'oggetto, o meglio la sua rappresentazione, come un elemento in certo modo dinamico.

La "formalizzazione", che probabilmente è il punto focale dei procedimenti informatici, consiste nell'individuazione delle unità minime analizzabili di un dato gruppo di informazioni da studiare; della loro espressione mediante segni univoci; e nella descrizione precisa del loro ordinamento. Ma una volta sottoposto questo materiale all'analisi automatica, la valutazione dei risultati di questa analisi portano non solo e non tanto a sintesi critiche, ma soprattutto ad una nuova consapevolezza nei riguardi dell'oggetto della ricerca (cioè del gruppo di informazioni di cui si parla sopra) e dunque ad una diversa formalizzazione, cioè ad una diversa individuazione delle unità minime e conseguentemente ad una diversa codifica. Analoghi cambiamenti vengono portati per motivi analoghi anche nei programmi che consentono l'analisi automatica del materiale codificato.

Tutto questo diventa chiaro soprattutto quando si cerca di definire il criterio di codifica delle fonti, sia letterarie sia storiche sia archeologiche sia figurative, mediante la scelta degli elementi, dei punti d'interesse, etc. L'unità del "modo informatico" di operare, al di sopra delle particolarità suggerite dalle differenti applicazioni alle differenti discipline umanistiche, è difficile da teorizzare con precisione, ma viene riconosciuta abbastanza bene dagli esperti, e dunque rappresenta il fondamento della disciplina come disciplina a sè, individuata. Essa resta comunque una parte della disciplina informatica nel suo complesso, che in altre parti ha come materia (cf. sopra) discipline non umanistiche, cioè sia "scientifiche" sia "umane" (sociali, politiche, giuridiche, etc.).

Questa è la prospettiva, a nostro modo di vedere più corretta, che si apre con l'uso generalizzato del computer nelle discipline umanistiche. I problemi che esso pone devono essere affrontati necessariamente anche con il rinnovamento di alcune strutture, non solo tecnologiche, ma anche di organizzazione accademica, inserendo i necessari centri di ricerca e quindi anche di insegnamento nelle istituzioni umanistiche.

Si dovrebbero anzitutto istituire dei Centri Informatici, diciamo così, polifunzionali, allo scopo di facilitare riflessioni metodologiche che colleghino le diverse discipline e i diversi tipi di applicazione, in modo da poter fornire consulenza alle ricerche singole. In particolare sarebbe necessario che ciascuno dei momenti fondamentali del lavoro informatico nelle discipline umanistiche (penso alla codifica testuale; all'ideazione della struttura delle banche dati; alla codifica di oggetti; etc.) sia assistito da specialisti a tempo pieno, che non siano solo specialisti di informatica, ma abbiano esperienza nei settori delle discipline umanistiche. L'interazione fra metodo generale e dati specifici è, come abbiamo visto, essenziale.

E' anche necessario inserire sia i centri sia le singole ricerche in una "rete" aperta all'infinito, che riproduca per mezzo della telematica la rete informativa prodotta oggi convenzionalmente, in modo molto imperfetto, dalla stampa. Ricordo a questo proposito che il Centro Linceo Interdisciplinare sta iniziando una sperimentazione in tal senso.

Il lavoro di chi applica l'informatica alle discipline umanistiche è soprattutto di tipo sperimentale. La validità di un software viene provata mentre lo si usa, e si producono altre idee di ciò che si vuole dal software. Il software rappresenta così una parte del metodo stesso, ed esso viene cambiato anche tenendo presenti i risultati (sempre provvisori!) che fornisce. Per questo la formazione e l'aggiornamento del software deve essere sempre alla portata dello studioso. In pratica, poiché lo studioso non può essere un esperto in linguaggi di base, occorre che vi siano strumenti fondamentali e modulari, che si prestino facilmente in modo interattivo alle necessità dello studioso. Dunque egli dovrà avere a disposizione non pacchetti pre-costituiti, ma linguaggi molto potenti e semplici. Per linguaggio intendiamo tuttavia anche un modo qualsiasi di comandare alla macchina (aprire file, analizzarli, stamparli, organizzarli, etc.) o anche linguaggi di "mark-up" da inserire nei file, e interpretati in sede di analisi di essi da parte di altri moduli software.

Infine, da quanto si è detto circa la prospettiva metodologica, discende la soluzione del problema: chi deve essere responsabile dell'organizzazione informatica nel settore umanistico? Ci sembra inevitabile che siano persone con sperimentata esperienza nella ricerca umanistica, oltre che nell'informatica. Questo comporta anche la scelta del tipo di organizzazione. Tre raccomandazioni sembrano opportune: 1) Dare spazio alle macchine personali, ma considerate come stazioni di lavoro "in rete", cioè collegate con l'ambiente esterno. 2) Avere disponibilità di macchine di media potenza per applicazioni particolari, senza tuttavia favorire la moltiplicazione di macchine medie, che sono sempre sotto-utilizzate. 3) Creare raccordi fra macchine piccole e medie, e fra queste e i main-frame, e fra queste e le banche-dati in rete, vuoi testuali vuoi bibliografiche etc.

Questa sembra la strada più opportuna per favorire la crescita e la corretta applicazione dell'informatica nell'ambito delle discipline umanistiche.

NOTE:

1) Cf. David S. MIALL (cit. sotto, nota 15, p. 3; Tito ORLANDI, *Informatica umanistica. Riflessioni storiche e metodologiche, con due esempi*, in: G. GIGLIOZZI, ed. (cit. sotto, nota 14), p. 1-38.

2) Cf. Stan AUGARTEN, *Bit by Bit. An Illustrated History of Computers*, London, Unwin, 1985(2); Herman GOLDSTINE, *Il computer da Pascal a Von Neumann. Le radici americane dell'elaboratore moderno*, ETAS Libri, 1981.

3) Nella seguente rassegna non segnaleremo la relativa bibliografia, ma rimandiamo ai repertori dai quali si può dedurre: Susan HOCKEY, *A Guide to Computer Applications in the Humanities*, Reading, Addison-Wesley, 1977; *The Humanities Computing Yearbook 1989-90* (cit. sotto, nota 5); Bibliografia di Informatica umanistica (in corso di stampa, Roma, Bulzoni).

4) Cf. Anton NIJHOLT (cit. sotto, nota 24) p. 70.

5) I riferimenti bibliografici sarebbero troppo pesanti in questa sede. Rimandiamo a Ian LANCASHIRE e Willard McCARTY, *The Humanities Computing Yearbook 1988*, Oxford, Clarendon Press, 1988 (nuova edizione: Ian LANCASHIRE, *The Humanities Computing Yearbook 1989-90. A Comprehensive Guide to Software and Other Resources*, Oxford, 1991); C. BOURLET, Ch. DOUTRELEPONT, S. LUSIGNAN, *Ordinateur et études médiévales, Bibliographie I*, Montréal, 1982 (Institut d'Etudes Médiévales); e di nuovo alla bibliografia in corso di stampa (cf. nota 3).

6) Può essere difficile distinguere centri e ricerche (cf. sotto), perché spesso i centri tendono ad essere mono-indirizzati, sia nel personale sia nel tipo di attività. I centri saranno quelli prevalentemente di servizio generale, le ricerche quelle che hanno ottenuto risultati scientifici in un certo settore. Ma la distinzione che abbiamo fatto non pretende di essere più che un espediente espositivo.

7) Si tenga presente che la linguistica deve essere considerata un caso tutto a sè, perché sta a mezzo fra le "vere" discipline umanistiche (che trattano dei linguaggi come "già usati", dunque non dell'origine del linguaggio, della sua eventuale struttura universale etc.) e le "vere" discipline informatiche, come la strutturazione dei data base management systems. Per questo argomenti come la linguistica computazionale e la traduzione artificiale rientrano a fatica in un orizzonte "umanistico"; ma non così la cosiddetta textual analysis, che però è discutibile come disciplina a sè. Altro caso è lo studio di strutture narrative, che invece, se bene inteso, è del tutto storicistico. In verità sarebbe possibile vedere nello storicismo proprio il confine che separa quanto è umanistico da quanto è "altro".

8) Menzioniamo a questo punto ricerche di cui ci risulta siano stati pubblicati più volte dei risultati. Questo elenco è da considerare dunque puramente esemplificati-

vo, e l'ordine casuale. Occorre anche tener conto che sulle ricerche italiane è incentrato il presente congresso, e quindi esse sono qui entro certi limiti trascurate.

9) Innumerevoli sono anche i bollettini d'informazione, o newsletter, a circolazione limitata.

10) Cf. sopra, nota 5.

11) Cf. p.es. Peter C. PATTON, Renee A. HOLOIEN (eds.), *Computing in the Humanities*, Aldershot, 1981; Richard W. BAILEY, *Computing in the Humanities., Papers from the Fifth International Conference on Computing in the Humanities* (Ann Arbor, Michigan, May 1981), Amsterdam-New York-Oxford-Tokyo, 1982.

12) Cf. sopra, nota 3.

13) Cf. p.es. John J. HUGHES, *Bits, Bytes, and Biblical Studies. A Resource Guide for the Use of Computers in Biblical and Classical Studies*, Grand Rapids (MI), 1987.

14) Intendo le problematiche messe a fuoco dal gruppo di ricerca Informatica e Discipline Umanistiche della Sapienza, che si trovano in Tito ORLANDI, *Informatica Umanistica*, (Studi Superiori NIS/78), Roma, 1990 e in vari saggi pubblicati nelle miscellanee della collana del gruppo di ricerca: Giuseppe GIGLIOZZI, ed., *Studi di codifica e trattamento automatico dei testi*, Roma 1987; G. ADAMO (ed.), *Trattamento, edizione e stampa di testi con il calcolatore*, Roma, 1989; Paola MOSCATI, *Trattamento di dati negli studi archeologici e storici*, Roma, 1990.

15) David S. MIALL (ed.), *Humanities and the Computer*, Oxford, 1990, Introduction.

16) Cf. il dibattito svolto su questa questione al congresso di Vassar, 1986, pubblicato in un numero speciale di *Computers and the Humanities*, 21 (1986) fasc. 4 (p. 209-254).

17) Cf. anche J. David BOLTER, *Turing's Man: Western Culture in the Computer Age*, Chapel Hill, 1984.

18) In MIALL, ed. (cit. alla nota 15) p. 13-26.

19) *Formazione e ricerca*, in Antonio RUBERTI (ed.), *Tecnologia domani*, Roma-

Bari 1985, p. 161-223; cf. p. 165.

20) Angelo Raffaele Meo, al convegno di Siena sulla storia dell'informatica, 10 settembre 1991.

21) Daniel I.A. COHEN, *Introduction to Computer Theory*, New York, 1991 (revised printing), p. 3.

22) "...the Theory of Computers, which means that we form several abstract mathematical models that will describe with varying degrees of accuracy parts of computers and types of computers and similar machines. Our models will ... be used to discuss ... the more abstract questions of the frontiers of capability of these mechanical devices." In realtà mi sembra che si faccia molto di più, perché questo rende l'informatica indipendente dai "mechanical devices", cioè aiuta a progettarne di nuovi. Così anche Cohen p. 3-4.

23) Thomas A. SUDKAMP, *Languages and Machines: An Introduction to the Theory of Computer Science*, Reading etc., 1988, p. 3.

24) Anton NIJHOLT, *Computers and Languages, Theory and Practice*, Amsterdam - New York, 1988, p. 47 e V-VII.

25) Carlo GHEZZI, Dino MANDRIOLI, *Informatica teorica*, Milano, 1989, p. 3.

26) Mario BORILLO, *Informatique pour les sciences de l'homme. Limites de la formalisation du raisonnement*, Bruxelles, Mardaga, s.d. (ca. 1985), p. 5. E' interessante notare come Miall e Denley lo trascurino.

27) Cf. Tito ORLANDI, *La metodologia di Franz Bopp e la linguistica precedente*, Rendiconti dell'Istituto Lombardo di Scienze e Lettere, Classe di Lettere, 96 (1962) 529-549.

28) Cf. Mordenti, più oltre in questo volume...

29) Cf. Bruce G. TRIGGER, *A History of the Archaeological Thought*, Cambridge, 1989.

30) Cf. l'Editorial (anonimo) di *Histoire et Mesure* 1 (1986) 5: "Histoire, mesure... Accoupler ces deux termes n'est pas chose neuve. Dès le XIXème siècle, bien des esprits se sont posé le problème du recours à la mesure en histoire — entendons le

terme au sense le plus large... Pourquoi donc créer, aujourd'hui en 1985, une revue intitulée Histoire et Mesure? Les besoins, l'urgence sont là. Ils sont révélés en histoire comme dans les autres sciences humaines et sociale, par l'ordinateur, qui a profondément transformé la pratique de nombreux historiens."

31) L'informatica porta anche una verifica immediata dei risultati della propria metodologia. Ma questo tema non può essere sviluppato in questa sede.

32) Jean-Claude GARDIN, *Archaeological Constructs*, Oxford University Press, 1983, p. 135-146.