



ASSOCIAZIONE ITALIANA SAN ROCCO DI MONTPELLIER
CENTRO STUDI ROCCHIANO

FRANCESCA RIZZI

« MEDIOEVO IN RETE. L'UTILIZZO DI INTERNET PER LE RICERCHE STORICHE »



FRANCESCA RIZZI

« MEDIOEVO IN RETE. L'UTILIZZO DI INTERNET PER LE RICERCHE STORICHE »

L'argomento di questo saggio della dott.ssa Francesca Rizzi è fondamentale per un ente come il nostro, che inizialmente – e significativamente – si è dotato di uno strumento informatico, e solo in un secondo momento ha predisposto anche i più tradizionali supporti cartacei.

Fermo restando che entrambe le cose hanno per noi lo stesso interesse e la stessa rilevanza, dobbiamo comunque ammettere che l'incidenza sulla nostra attività del portale Internet è ad un livello neppure lontanamente paragonabile ai risultati, peraltro assai positivi, delle pubblicazioni a stampa, che pure hanno riscosso un unanime apprezzamento sia negli ambienti culturali, sia tra le persone meno addentro agli studi specialistici.

L'andamento statistico dei contatti al nostro portale si è infatti attestato su livelli francamente inimmaginabili all'inizio del nostro percorso, già significativamente superiori all'andamento medio dei siti riconducibili al nostro settore specifico; peraltro, ancor prima di esordire con le versioni in lingua straniera, abbiamo rilevato un'ampia quota di visitatori esteri.

E' evidente, allora, che la *questione Internet* assume anche per noi un forte interesse, in modo particolare per quanto attiene al delicato rapporto fra le attività culturali, le ricerche storiche ed il supporto informatico. Il tema è tutt'altro che definito, nel senso che il dibattito in corso è ancora segnato da posizioni contrastanti ed in alcuni casi inconciliabili.

Da parte nostra, è inutile dire che il «Centro Studi» è ovviamente convinto del valore positivo e delle grandi potenzialità del mezzo, ed anzi riteniamo che l'utilizzo di Internet dischiuda prospettive nuove e possibilità, fino ad oggi precluse, per un netto miglioramento della diffusione della cultura e delle ricerche storiche e scientifiche.

D'altro canto siamo pienamente consapevoli anche dei limiti e dei rischi, altrettanto forti, di un uso non corretto di Internet, e non pensiamo affatto che esso debba *sostituire* i tradizionali mezzi a stampa: semmai li deve *affiancare*. E' del tutto ovvio, poi, che il lavoro dello storico si debba sempre basare sui documenti scritti, sui reperti materiali, sulle tracce *fisiche* del lungo cammino dell'uomo verso una sempre più compiuta civiltà.

In tal senso, il seguente saggio ci darà la possibilità di approfondire tali tematiche, a partire, peraltro, da una opportuna ricapitolazione delle principali caratteristiche (e sigle!) di Internet; i già esperti ci perdoneranno, ma l'intento principale del nostro «Centro Studi» è sempre quello di rendere il più possibile accessibile al pubblico dei non addetti ai lavori le materie generalmente riservate ad una ristretta cerchia di studiosi.



FRANCESCA RIZZI

« MOYEN-ÂGE EN WEB. L'UTILISATION D'INTERNET POUR LES RECHERCHES HISTORIQUES »

Le sujet de cet essai est fondamental pour un organisme comme le nôtre, qui a démarré son expérience sur Internet en premier lieu, et s'est doté ensuite d'un support papier.

Nous constatons par ailleurs que les résultats obtenus par le site se rapprochent des résultats très positifs de la revue. Nous apportons donc une attention toute particulière au support Internet afin de créer un juste équilibre entre les activités culturelles, les recherches historiques et les systèmes informatiques. Nous sommes bien sûr convaincus d'avoir en main un excellent instrument de travail, s'il est utilisé correctement; il ne peut toutefois se substituer au support papier.

Nous allons approfondir ce sujet avec l'essai qui suit, en commençant par rappeler les principales caractéristiques et les symboles utilisés sur Internet, pour faciliter la navigation aux lecteurs moins experts.



FRANCESCA RIZZI

« MIDDLE AGES IN THE WEB. USE OF INTERNET FOR HISTORICAL RESEARCHES »

Our «Centro Studi» first got himself a portal and only later a review, thus it is patent that the *Internet question* assumes a great relevance for us as well, especially in the area of the delicate relationship between cultural activities, historical research and informatics; still, even if we are sure of the positive effects and of the great chances offered by this mean, we are also conscious of its limits and risks. Francesca Rizzi, responsible of the Historical Section of our «Centro Studi», is going to tell us about it.



FRANCESCA RIZZI

« EDAD MEDIA EN WEB. UTILIZAR INTERNET PARA LAS INVESTIGACIONES HISTÓRICAS »

Nuestro «Centro Studi» se dotó inicialmente de una página web y solamente en un segundo momento de una revista, y por lo tanto es evidente que el problema Internet adquiere para nosotros una importancia considerable, sobretodo a propósito de la delicada relación entre las actividades culturales, las investigaciones históricas y la informática; pero estamos convencidos del valor positivo y de las grandes posibilidades de este instrumento, también somos conscientes de sus límites y de sus riesgos. Este ensayo de Francesca Rizzi nos hablará de esas temáticas, empezando por recordar las principales características y los símbolos utilizados en Internet para facilitar la navegación a los lectores menos expertos.



FRANCESCA RIZZI

« MEDIOEVO IN RETE. L'UTILIZZO DI INTERNET PER LE RICERCHE STORICHE »

PARTE PRIMA

*"Il mondo appartiene a coloro
che amano le cose nuove"*

Marc Bloch

1. www.sanroccodimontpellier.it

Quando circa un anno fa ho incontrato il dott. Ascagni e per la prima volta ho sentito parlare del progetto del sito su san Rocco, sono stata immediatamente travolta dall'entusiasmo nei confronti di un lavoro che mi sembrava senz'altro ambizioso ma altrettanto stimolante. Ad accendere il mio interesse, a parte la mia sincera passione per la storia, hanno certamente contribuito la dedizione e l'entusiasmo che il dott. Ascagni ha dedicato allo studio di questo santo ed alla progettazione di questo portale; ma, in parte, è stata proprio l'idea che fosse *un sito*, e che quindi tutto il lavoro che avremmo fatto sarebbe andato in rete, su Internet.

In effetti per la mia generazione convivere con il computer un po' in ogni campo, dagli studi al lavoro come a casa, è un fatto di normalità acquisita, così come utilizzare Internet per reperire ogni genere di informazione, dalle ricerche scolastiche all'acquisto di biglietti aerei.

Già all'Università ero entrata in contatto con diversi progetti nati in alcuni dipartimenti della Facoltà di Lettere che in qualche modo tentavano di sfidare le tradizionali vie di comunicazione, così radicate specie in un ambiente tradizionalmente conservatore come quello universitario; e già da allora mi affascinava l'enorme potere divulgativo di Internet, che ho sempre percepito come sconfinato.

Nel progetto su san Rocco pubblicato nel portale, sono la responsabile della Sezione Storica, che raccoglie documenti, saggi ed articoli sull'argomento ma che dedica spazio anche a qualche spaccato di storia medievale e che quindi sarà sempre oggetto di aggiornamenti, come richiede la natura stessa di un sito. Il mio rinnovato stupore per la potenzialità di questo strumento è stata confermata, se non accresciuta, dai primi dati che abbiamo raccolto sulla consultazione del sito da quando, nel gennaio del 2006, esso è stato pubblicato in rete¹.

Certo è che, come ogni strumento estremamente potente, va anche maneggiato con cautela, perché di fatto Internet è un enorme contenitore nel quale entra ogni genere di informazione riguardante qualsiasi settore culturale e non; ed è per questo che per poterlo usare al meglio occorre capirlo e sapere come funziona.

2. Internet nelle discipline umanistiche

Come già ho già accennato più sopra, la telematica ha ormai raggiunto tutti i settori della nostra vita, compresi gli ambienti accademici più refrattari come quelli umanistici, anche grazie ad un utilizzo dell'informatica più duttile e più compatibile alle varie esigenze espresse da chi si occupa di scienze umane. La rete Internet, in particolare, riesce a mettere a disposizione di studiosi e scienziati risorse inaccessibili al di fuori dei centri informatici delle grandi università. In rete si sono poi sviluppate velocemente altre applicazioni con finalità prettamente comunicative, permettendo la condivisione di dati ed informazioni attraverso la comunicazione interpersonale, come la posta elettronica e le liste di discussione, allargando notevolmente il bacino degli utenti e diventando così uno dei principali canali della comunicazione scientifica.

Tra i mezzi di reperimento e di accesso alle risorse primarie ci sono diversi strumenti *on-line* che consentono di individuare, reperire e consultare le fonti, come i cataloghi bibliotecari, le bibliografie

¹ Si veda, a tal proposito, la sezione del sito dedicata alle statistiche.

on-line, gli archivi di testi e immagini in formato digitale². Una grande novità nel mondo delle biblioteche è stato lo sviluppo degli OPAC, *On-line Public Acces Catalog*, i primi ad introdurre il concetto di indipendenza delle risorse dalla localizzazione territoriale, nel settore umanistico. Accanto agli OPAC, si trovano su Internet molte risorse bibliografiche settoriali prodotte da singoli studiosi o da progetti di ricerca istituzionali. Il testo elettronico ha raggiunto una funzione autonoma, tanto che alcune grandi istituzioni hanno avviato progetti di vere e proprie biblioteche digitali³.

La novità principale nel mondo della comunicazione scientifica è rappresentata dall'introduzione di sistemi di comunicazione interpersonali:

- la posta elettronica, che si basa sull'uso di un indirizzo *e-mail* individuale.
- le conferenze e i gruppi di discussione *on-line*, che possono essere di due tipi: *asincrone*, cioè con un intervallo di tempo tra l'invio del messaggio, la ricezione e la risposta, oppure *in tempo reale*. Queste ultime si chiamano CHAT quando tutti i partecipanti sono collegati a un ripetitore in uno spazio virtuale; oppure MUD, *Multi User Dungeons*, una evoluzione della *chat*, con regole e scopi da perseguire sempre in un ambiente virtuale. Non sono mai stati molto usati dalla comunità scientifica perché considerati troppo artificiali, fatta eccezione per qualche caso come il famoso PMC2, nato nel 1994 – e non più attivo – dalla redazione di *Postmodern Culture*, nel quale si svolgevano vere discussioni seminariali⁴.

Navigando in rete attraverso un qualsiasi *motore di ricerca* e impostando vari argomenti legati alla storia medievale, si ha la sensazione che Internet stia entrando piano piano in questo ambito di studi, anche se spesso la qualità delle informazioni che si ottiene non è soddisfacente. Molti dipartimenti ed istituti di ricerca, infatti, sono presenti su Internet a vari livelli, ma spesso non possiedono una pagina *web* propria oppure restituiscono solo informazioni essenziali come l'indirizzo, lo staff e la posta elettronica⁵.

La medievistica italiana è senz'altro in ritardo nell'utilizzo dell'informatica, nonostante qualche eccezione che però non basta ad eliminare la sensazione di forte scetticismo che circonda ancora il settore.

L'editoria *on-line* potrebbe essere d'aiuto per tutte quelle pubblicazioni scientifiche molto specializzate che non trovano spazio nel mercato editoriale e che spesso vengono pubblicate solo a fronte di finanziamenti. Non credo comunque che il mondo digitale potrà mai sostituire quello culturale tradizionale, ma è altrettanto vero che esso offre un vasto ventaglio di risorse che, se ben utilizzate, possono integrare le pratiche della tradizione. Questa integrazione si è già attuata per gli archivi e le biblioteche ai quali si è affiancato un terzo luogo dallo spazio illimitato, costituito appunto da Internet.

La digitalizzazione delle informazioni è in assoluto la principale novità in questo processo di integrazione che ha investito ogni strumento di lavoro: fonti, testi, documenti, studi, riviste, bibliografie, banche dati, provocando un mutamento che non è solo tecnologico ma anche di modi di ricerca e di linguaggi. Il libro ad oggi resta comunque il canale di comunicazione preferito da chi fa ricerca, mentre la pubblicazione sul *web* è ancora spesso priva di autorevolezza ed è considerata una sede secondaria. Infatti in Italia nulla pare ancora sostanzialmente cambiato né nei modi di fare ricerca, né nei metodi e nemmeno nei linguaggi⁶.

Il problema principale che si deve affrontare navigando in rete resta comunque quello della valutazione e della selezione qualitativa delle risorse digitali; questa selezione critica è tutt'altro che semplice e immediata poiché richiede grande dimestichezza con Internet, costituendo così il principale ostacolo per tutti coloro che non utilizzano spesso questo tipo di strumento. La prima distinzione da applicare a un qualsiasi risultato di ricerca in rete è quella tra i siti divulgativi e quelli

² Cfr. T. NUMERICO e A. VESPIGNANI (curatori), *Informatica per le scienze umanistiche*, Bologna 2003.

³ Cfr. F. CIOTTI, *La comunicazione telematica e le discipline umanistiche*, in *Medioevo in rete tra ricerca e didattica*, a cura di R. GRECI, Bologna 2002, pagg. 1-21.

⁴ Cfr. F. CIOTTI, *ibidem*, pagg. 1-21.

⁵ Cfr. E. SALVATORI, *Tra didattica e ricerca: l'esperienza dei Dipartimenti e Istituti di Storia*, in *Medioevo in rete...*, op. cit., 2002, pagg. 75-81.

⁶ Cfr. A. ZORZI, *Comunicazione del sapere ed editoria digitale: problemi e prospettive per gli studi medievali*, in *Medioevo in rete...*, op. cit., 2002, pagg. 183-235.

amatoriali che sono, di fatto, la maggioranza e che non si possono riconoscere senza la capacità critica nell'uso di risorse digitali a fini di studio.

Per accedere ad un sito senza conoscerne l'U.R.L. (cioè l'indirizzo preciso del sito stesso), si può utilizzare un motore di ricerca (*Altavista, Virgilio, Yahoo, Lycos, Arianna, Google*, ecc.) oppure un *link* da un altro sito; per ottenere qualche risultato apprezzabile con un motore di ricerca, è però opportuno sforzarsi ad impostare la ricerca digitando parole il meno generiche possibili, lemmi mirati e specifici.

Dal motore di ricerca si passa al sito, dal sito all'indice, dall'indice ad un elenco di voci e da questo all'informazione desiderata e non ci sono altre vie per accedere direttamente al risultato; questo è senz'altro un limite funzionale che rende difficile consultare il *World Wide Web* (più comunemente detto *www*) che infatti non può essere considerato una vera e propria biblioteca digitale. I testi sono riprodotti in formati compatibili con la rete Internet; il *file* HTML, acronimo di *Hyper Text Markup Language* (linguaggio per la marcatura di ipertesti), in informatica è il linguaggio di programmazione con il quale sono scritti tutti i documenti ipertestuali destinati al *World Wide Web*. I linguaggi di marcatura (*Markup Languages*), o metalinguaggi, servono ad indicare ai programmi di decodifica il modo in cui la parte di un testo dev'essere trattata; danno, quindi, informazioni di tipo *procedurale* (spaziature, interlinee, formattazione, ecc.) o di tipo *descrittivo* (l'articolazione logica e strutturale di un testo, cioè capitoli, paragrafi ma anche titoli, citazioni, ecc.). Il *Document Type Definition* (DTD) indica lo specifico linguaggio di marcatura valido per una determinata tipologia di documenti. I più noti linguaggi dichiarativi sono SGLM (*Standard Generalised Markup Language*) e XML (*EXstensible Markup Language*)⁷.

Di testi elettronici ne esistono di vari tipi, tutti genericamente inseribili nel più vasto concetto di *biblioteca digitale*, anche se spesso le realizzazioni hanno caratteristiche molto diverse. Ma lo stesso *World Wide Web* in realtà, può essere visto come un immenso archivio globale della conoscenza umana all'interno del quale però non tutto ha lo stesso fine.

Seguendo l'esempio del *Progetto Gutenberg* nato in America negli anni settanta, l'italiano *Progetto Manuzio* nato nei primi anni novanta (www.liberliber.it) si propone lo scopo di "*portare la cultura a disposizione di tutti*": dal sito si possono scaricare copie digitali di opere, specialmente di classici, della letteratura italiana, così come traduzioni italiane di opere straniere, in edizioni per le quali sono scaduti i termini del diritto d'autore.

Diverse sono le biblioteche digitali nate in ambito accademico, che pongono maggiore attenzione all'autorevolezza e all'affidabilità delle trascrizioni e che offrono la possibilità oltre che di leggere e scaricare i testi, anche di effettuare ricerche più complesse, ad esempio linguistiche o lessicali. È questo il caso della *Biblioteca telematica italiana del Cibit* (www.bibliotecaitaliana.it, Centro interuniversitario biblioteca italiana telematica) con i suoi oltre mille testi letterari, giuridici, filosofici, storici e scientifici. In questo caso lo scopo non è solo quello di utilizzare il *Web* per archiviare e comunicare ma anche di mettere a disposizione percorsi di ricerca e di esplorazione di testi non attuabili con i media tradizionali⁸.

Ma la maggioranza delle raccolte di documenti presenti nel *Web* sono però semplici trascrizioni realizzate nei formati che ne consentono la lettura e la stampa (.txt, .html, .pdf, .doc.; anche i saggi e gli articoli contenuti nel nostro sito sono tutti in formato .pdf e si possono leggere e stampare facilmente da tutti i computer). Lo scopo primario di queste raccolte è spesso quello di arricchire la quantità di testi disponibili in rete; come spesso accade, la quantità va a discapito della qualità e non sempre chi naviga su Internet tiene presente che un testo appartiene sempre a una pluralità di contesti e che alcuni di questi hanno una propria manifestazione materiale non più visibile nel testo digitale.

Se a intendimenti diversi possono corrispondere biblioteche digitali di diverso tipo, anche la pubblicazione in linea di un singolo testo può rispondere a finalità diverse (un conto è ad esempio lo scopo didattico, un conto è la ricerca storica); dovrebbe sempre esistere una certa coerenza tra le soluzioni che si scelgono di adottare rispetto alla finalità prevista dalla pubblicazione di un testo⁹.

⁷ Cfr. S. VITALE, *Passato digitale. Le fonti dello storico nell'era del computer*, Milano 2004, pagg. 48-50.

⁸ Cfr. S. VITALE, *ibidem*, pagg. 111-116.

⁹ Cfr. D. BUZZETTI, *Risorse digitali e strutture di dati*, in *Medioevo in rete...*, op. cit., 2002, pagg. 47-62.

3. Le fonti digitali

Il rapporto con le fonti costituisce un aspetto fondamentale del lavoro dello storico. La storia si distingue dalle altre narrazioni del passato perché deve sempre fondarsi su delle prove¹⁰ che non sono proprio le fonti stesse, ma ciò che lo storico riesce ad estrapolare dai documenti attraverso l'elaborazione critica degli stessi.

Nell'epoca dell'era digitale, sono ancora adeguati i metodi di analisi e di critica delle fonti storiche usati finora? Queste procedure comprendono la *critica esterna*, tesa a stabilire se un documento è un originale o una copia e, in questo caso, qual è la sua fonte, e la *critica interna* che porta a stabilire la credibilità del documento riguardo al suo contenuto; consolidate nel XIX secolo, avevano e hanno il compito di ricercare l'autenticità e la credibilità dei documenti¹¹.

L'applicazione dell'informatica al lavoro dello storico ha posto nuove problematiche nel suo rapporto con le fonti; queste esprimono, al di là dei contenuti, anche tanti altri elementi che "parlano", in maniera non esplicita, di chi ha composto la fonte e rivelano, quindi, particolari appartenenti a una pluralità di contesti; si tratta spesso di elementi di difficile valutazione già nei formati tradizionali delle fonti, che diventano pressoché invisibili nell'edizione digitale con il rischio di essere perduti.

Inoltre, il rapporto tra una fonte e una realtà storica è reso ancor più complesso dai modi in cui questa viene scelta e trasmessa nel tempo. Le future generazioni avranno una visione della nostra società che dipenderà in buona parte dalle testimonianze conservate e tramandate da noi stessi, necessità che fa nascere spesso in chi riveste questo compito (istituzioni pubbliche e private ma anche individui, comunità locali, associazioni) la necessità di ridurre il più possibile le masse documentarie, molto diversificate, ma spesso ripetitive¹².

Le nuove fonti digitali hanno, tra i problemi più rilevanti legati alla loro gestione, quello della durata e permanenza delle informazioni che contengono; si pensi, ad esempio, alle fotografie, alle registrazioni audiovisive su supporto magnetico, alle pellicole cinematografiche. A differenza delle fonti tradizionali, le digitali hanno come limite principale quello che per potervi accedere è necessaria la mediazione di componenti *hardware* e *software* dei computer.

L'*hardware* ha come difetto assoluto quello di "invecchiare" molto velocemente: questo genera incompatibilità tra i supporti di archiviazione ormai obsoleti e quelli di nuova generazione. Lo stesso invecchiamento precoce lo soffre il *software*, sia per il sistema operativo che per gli applicativi specifici. A questi due grandi limiti, bisogna aggiungere e tenere ben presente la fragilità dei supporti utilizzati, spesso vulnerabili alle variazioni di temperatura e umidità e ai campi magnetici, tanto da rischiare la perdita di leggibilità in maniera irreversibile¹³.

Altro problema in cui ci si imbatte spesso per le fonti presenti in Rete è quello della volatilità: la permanenza in linea delle pagine Web è stimata in media tra sessanta e cento giorni; inoltre le pagine Web sono molto dinamiche e soggette a trasformazioni continue e rapide, nella grafica, nelle modalità di presentazione ma spesso anche nel contenuto. Così capita a volte di cercare una pagina visitata e di non trovarla più ma al suo posto vedere comparire una scritta, familiare per chi naviga di frequente in Internet, che è "error 404 page not found" dell'*Hypertext Transfer Protocol*. Dal punto di vista tecnico, l'informazione digitale non è altro che una sequenza di *bit*, cioè di 1 e di 0 che servono a veicolare istruzioni verso i programmi di visualizzazione, che li decodificano rendendoci comprensibili. Così come si fa con le immagini digitali, un documento anch'esso digitale si può copiare, correggere, tagliare e incollare, se ne può cambiare la forma di presentazione o la dimensione, ingrandendo o riducendo le pagine, e così via.

Ecco perché nell'utilizzarli e nel riferirsi a documenti digitali il rigore deve essere equivalente, se non maggiore, a quello che gli storici utilizzano verso i supporti tradizionali. Tutti dovrebbero avere, attraverso la Rete, la possibilità di prendere visione delle fonti citate in un testo di storia. Biblioteche, archivi, musei hanno cercato finora di provvedere alla conservazione e trasmissione di testi, manufatti e opere d'arte, provvedendo nel contempo anche a una loro prima autenticazione.

¹⁰ Cfr. J. TOPOLSKI, *Narrare la storia. Nuovi principi di metodologia storica*, Milano 1996, pag. 19

¹¹ Cfr. C.V. LANGLOIS e C. SEIGNOBOS, *Introduction aux études historiques*, Parigi 1898.

¹² Cfr. S. VITALE, *Passato digitale...*, op. cit., 2004, pagg. 129-149.

¹³ Cfr. C. DOLLAR, *La memoria elettronica e la ridefinizione della preservazione*, in T. GREGORY e M. MORELLI (curatori), *L'eclisse delle memorie*, Roma-Bari, pag. 175.

I documenti elettronici e, più in generale, i *files* digitali, pongono, come abbiamo visto, molti nuovi problemi di archiviazione e trasmissione. Tra i vari tentativi di soluzione, le due più adottate finora non sono certo soddisfacenti né prive di controindicazioni:

- la prima propone di ricopiare periodicamente i *files* digitali su nuovi supporti (in gergo tecnico si chiama *refresh* delle memorie digitali), cercando in questo modo di mantenere gli oggetti digitali all'interno di un contesto tecnologico il più possibile vicino a quello originario; questo sistema mira a conservare le piattaforme *hardware* e *software* in una sorta di «museo del computer» che consenta l'accesso ai *files* originari;
- la seconda propone di rendere compatibili gli oggetti digitali trasformandoli, affinché divengano adattabili ai nuovi *hardware* e *software*; in questo caso si tratta di migrare i dati su nuovi formati compatibili con le piattaforme *hardware* e con gli specifici *software*, in modo che si possa sempre accedervi con continuità.

Recentemente, si preferisce convertire i dati in formati standard (per esempio, i metalinguaggi di marcatura SGML o XML) che permettono di svincolare i processi di conservazione dal rapido evolversi della tecnologia¹⁴. Queste soluzioni, che possono sembrare di difficile applicazione a proposito di un testo, diventano di facile comprensione se le pensiamo rivolte a uno dei fenomeni culturali più diffusi della nostra epoca: i *videogiochi*. Questi programmi sono destinati a girare su piattaforme *hardware* che vengono superate con una rapidità incredibile da altre, con prestazioni ed effetti grafici sempre più evoluti; in questo caso, è inutile cercare di conservare in perfetto stato il supporto originario (ad esempio un CD-ROM), se non esiste più un supporto *hardware* in grado di far funzionare il programma.

Internet sembra un infinito archivio da cui si può attingere materiale su qualsiasi argomento ma, ad oggi, ci si è preoccupati poco della sua stabilità e permanenza verso il futuro, e l'esperienza quotidiana di navigazione in Rete attesta questa carenza più che qualsiasi statistica. Inoltre, la dimensione del *World Wide Web* continua a aumentare molto rapidamente, nonostante le sue dimensioni già incredibili: a gennaio 2004 *Google*, uno dei più popolari motori di ricerca, indicizzava circa 3 miliardi e mezzo di pagine *Web*¹⁵. L'azione prioritaria resta quella di individuare ciò che va conservato e cosa no; ma, premesso questo, speriamo soprattutto che resti una questione irrinunciabile e di prossima soluzione la ricerca e l'individuazione di politiche di archiviazione del *Web*, dalla quale non si può prescindere ma che per il momento non trova una soluzione univoca e soddisfacente.

PARTE SECONDA

*"I computer sono incredibilmente veloci, accurati e stupidi.
Gli uomini sono incredibilmente lenti, inaccurati e intelligenti.
Insieme sono una potenza che supera l'immaginazione"*
Albert Einstein

1. Consigli... e precauzioni

Nella prima parte scrivevo, quasi in forma di consigli pratici, su come avvicinarsi alla rete, esprimendo comunque le mie perplessità sui risultati che si possono ottenere effettuando una qualsiasi ricerca in Internet, specie se molto generica, tramite un qualunque motore di ricerca. Qualsiasi risultato si ottenga, io credo che vada considerato, analizzato e soprattutto utilizzato con molta attenzione, tenendo sempre presente che proprio perché Internet è di tutti, l'informazione che contiene è totalmente senza limiti, nel bene e nel male. Ma al di là di questo, si deve sempre consigliare, per quanto ovvio, di non effettuare mai, in nessun caso, alcuna ricerca sulla rete se il *computer* che si utilizza non è ben protetto da un «antivirus»; infatti, insieme alle informazioni circolano – quasi altrettanto numerosi e vari per tipologia – i cosiddetti *virus*, che attaccano il sistema del vostro *computer*, nella migliore delle ipotesi bloccandolo o limitandone la potenza e la velocità, ma spesso distruggendo ed attaccando il sistema stesso che lo fa funzionare, escludendo

¹⁴ Cfr. S. VITALE, *Passato digitale...*, op. cit., 2004, pagg. 182-186.

¹⁵ Cfr. S. VITALE, *Passato digitale...*, op. cit., 2004, pag. 210.

ogni possibilità di recupero dei vostri dati.... ed in quest'ultimo caso, nemmeno il nostro caro san Rocco, che pure proteggeva e guariva dalla peste, potrà in qualche modo esservi d'aiuto!

2. La storia di Internet

I sistemi tecnologici sono una produzione sociale, modellata sulla cultura, ed Internet non fa eccezione. La cultura dei produttori di Internet ha plasmato il mezzo ed essi stessi ne sono stati i primi utilizzatori, cioè produttori-utenti che hanno provocato un effetto diretto sul sistema. Vediamo ora, brevemente, la storia di questo grande sistema.

Internet nasce da ARPANET¹⁶, un network di computer creato nel 1969 dalla *Advanced Research Project Agency* (ARPA), facente parte del dipartimento della Difesa degli Stati Uniti, con lo scopo di migliorare e raggiungere la superiorità tecnologica e militare rispetto all'Unione Sovietica; per questo tipo di ricerca, il mondo a cui attingeva energie e intelligenze era principalmente quello universitario. Nello specifico, ARPANET era un programma minore elaborato da un dipartimento interno dell'ARPA, l'*Information Processing Techniques Office* (IPTO) e il suo scopo era quello di condividere nello stesso momento l'utilizzo del *computer* tra diversi centri di elaborazione-dati e gruppi di ricerca dell'agenzia.

ARPANET fu presentato ufficialmente nel 1972. In seguito, la nuova ricerca puntò sul modo di far comunicare non più solo i *computers* di un *network* (cioè, una rete di *computers*), ma diversi *networks* tra di loro, cioè diverse reti di *computers*, ognuna delle quali costituita da un gruppo di *computers* che dialoga solo al suo interno, a cominciare da quelli già presenti in ARPA (come PRNET e SATNET). Per fare questo, era necessario che queste reti di *computers* utilizzassero dei protocolli standardizzati, cioè validi per tutti. Nel 1973 il risultato venne in parte raggiunto grazie all'utilizzo del TCP, il protocollo di controllo trasmissione¹⁷.

L'evoluzione del sistema portò nel 1978 a distinguere questo protocollo, il TCP, in due parti, inserendo un protocollo tra rete e rete (IP) ed un secondo protocollo, il TCP/IP, lo standard con cui ancora oggi lavora Internet. Nel 1990 ARPANET, ormai considerata obsoleta, venne smantellata. Nel frattempo Internet si era liberata da quel contesto militare nel quale era nata, grazie alla formazione di MILNET, rete separata da ARPANET e dedicata, dal Dipartimento della Difesa statunitense, solo ad impieghi militari, mentre ARPANET diventava appunto ARPA-INTERNET, finalizzata solo alla ricerca.

Una volta smantellata, il governo degli Stati Uniti la affidò alla «National Science Foundation», ma per breve tempo. Ormai, infatti, la tecnologia informatica di Internet era di dominio pubblico ed il Dipartimento della Difesa aveva deciso di commercializzare questa tecnologia, finanziando i produttori di *computers* affinché introducessero nei loro protocolli il TCP/IP. Ciò avvenne già attorno al 1990; da allora Internet è cresciuta rapidamente come rete globale delle reti di *computers*.

ARPANET però non è stata l'unica fonte di Internet; la sua forma attuale è anche il prodotto di una tradizione di base di comunicazione tra *computers*. Nel 1977 due studenti di Chicago scrissero un programma, battezzato MODEM, che permetteva di trasferire i *files* tra i loro *personal computers* (i cosiddetti «pc»); di seguito idearono un programma sulla base dei già noti sistemi di bacheca elettronica (chiamati BBS, *Bulletin Board Systems*), che consentiva ai *pc* di archiviare e trasmettere messaggi. Entrambi i programmi furono messi a disposizione del pubblico dai loro stessi ideatori.

Nel 1983 Tom Jennings ideò un programma, FIDO, per le BBS, dando vita a una rete di BBS denominata FIDONET; questa è ancora oggi la rete di comunicazione tra *computers* più accessibile al mondo, perché necessita solo di *pc* e di linee telefoniche standard.

E ancora, una tendenza decisiva nella comunicazione tra *computers* è emersa dalla comunità degli utenti UNIX, un sistema che nel 1974 venne rilasciato alle università e che divenne la lingua franca della maggior parte dei dipartimenti di Informatica: ben presto gli studenti impararono a manipolarlo. Nel 1980 il dipartimento di Informatica dell'università di Berkeley, in California – che era un nodo di ARPANET – sviluppò un programma che facesse da ponte tra le due reti UNIX e ARPANET, che prese il nome di USENET.

¹⁶ Per la storia di Internet, cfr. M. CASTELLS, *Galassia Internet*, Milano 2006, pp. 21- 44; K. HAFNER e M. LYON, *La storia del futuro. Le origini di Internet*, Milano 1998.

¹⁷ Il protocollo, lo ricordiamo ai meno esperti, sostanzialmente è come un linguaggio.

Solo alla fine di questo complesso processo di unione, di ricerche e di programmi, nacque Internet¹⁸. Ma va citato anche un altro importante sviluppo scaturito dagli utenti di UNIX: quello denominato *Open Source Movement*, cioè un tentativo di tenere aperto l'accesso a tutte le informazioni relative ai *softwares*¹⁹. Nel 1984 venne lanciata la «Free Software Foundation», con l'intento di sostituire *copyright* con *copyleft*²⁰, termine che introduce questo concetto: chiunque utilizzi un *software* messo liberamente a disposizione, in cambio deve distribuire sulla rete il codice migliorato.

Mettendo in pratica questo principio, nel 1991 Linus Torvalds, studente di 22 anni, creò un nuovo sistema basato su UNIX, chiamato LINUX e distribuito gratuitamente su Internet. Oggi LINUX è largamente riconosciuto come uno dei sistemi operativi più avanzati del mondo. Dagli utenti UNIX sono nati altri gruppi di *software* cooperativo che utilizzavano l'*open source* e ciò ha permesso lo sviluppo del *World Wide Web*, applicazione nata nel 1990 proprio dalla condivisione delle informazioni – e dalla mente – di Tim Berners-Lee. Fu lui ad ideare il programma ENQUIRE nel 1980, programma che realizzò quella che fino ad allora era comunque rimasta un'utopia: il recupero e la distribuzione di informazioni *da* e *a* qualunque *computer* collegato via Internet. Nel 1990, insieme a Robert Cailliau²¹, creò un programma *browser/editor* ipertestuale²², appunto il *World Wide Web*, il notissimo WWW.

Microsoft, la celebre azienda di Bill Gates, scoprì Internet nel 1995 ed insieme al proprio *software* «Windows 95» introdusse anche il relativo *browser*, «Internet Explorer»; così, entro la metà degli anni novanta, Internet è privatizzata. L'applicazione più utilizzata della rete, già dal 1970, era ed è la posta elettronica.

E siamo ai nostri giorni. Avvicinandosi a Internet, credo valga la pena di conoscerne la storia per fare qualche considerazione. Ad esempio, è forse utile tenere presente, per capirne lo spirito, che Internet non è nata nel mondo dell'impresa, bensì all'interno di enti governativi, grandi università e centri di ricerca; si trattava di una tecnologia troppo coraggiosa, di un progetto troppo costoso e rischioso per interessare le grandi aziende, ovviamente orientate principalmente al profitto. Internet si è quindi sviluppata in un ambiente sicuro, garantito da risorse pubbliche, orientato alla ricerca, contesto che ha permesso che non venisse soffocata la libertà di pensiero e di innovazione. L'architettura aperta di Internet è stata la sua più grande forza, apertura che le ha permesso di svilupparsi ed evolversi autonomamente, nella quale gli utenti diventano produttori di tecnologia, la adattano ai loro usi ed infine la trasformano. Nessuno ha commissionato a Tim Berners-Lee la creazione del *World Wide Web*, ma lui ha sempre continuato la ricerca, sostenuto dalla comunità di Internet che lo aiutava e stimolava.

E in più, Internet ha una caratteristica speciale: il suo progresso, le sue modifiche, vengono comunicati di volta in volta al mondo intero, in tempo reale. E' per questo che è cresciuta e cresce anche oggi ad una velocità senza precedenti, non solo per il numero delle sue reti, ma anche per la portata delle sue applicazioni.

Internet è una creazione culturale. La cultura di Internet è la cultura degli utilizzatori stessi di Internet.

3. Il *computer* nella professione dello storico

Computer e ricerca storica si sono incontrati per la prima volta in America attorno al 1960, con l'avvento della «storia quantitativa»²³, e più precisamente della storia politica, attraverso l'analisi di tipo statistico (ad esempio dei dati elettorali), ma anche nell'ambito della storia sociale, per svilupparsi poi in Europa, prima fra tutti in Francia, attraverso le «Annales»²⁴. Diversamente da

¹⁸ Cfr. M. CASTELLS, *La nascita della società in Rete*, Milano 2002.

¹⁹ Il *software* è l'insieme dei programmi di un *computer*, mentre l'*hardware* è l'insieme delle parti fisiche e meccaniche.

²⁰ L'inglese *copyright* indica il principio giuridico del diritto d'autore, tutelato dalle leggi; il neologismo *copyleft*, volendo esprimere un concetto di libero utilizzo, si basa sul gioco di parole, intraducibile, tra *right* (diritto, ma anche destra) e *left* (sinistra).

²¹ Cfr. J. GILLIES e R. CAILLIAU, *Com'è nato il Web*, Milano 2000.

²² Il *browser* è un programma grafico che permette di operare e di cercare informazioni in Internet.

²³ Cfr. S. VITALI, *Passato digitale*, Milano 2004, p.7

²⁴ Cfr. P. BURKE, *Una rivoluzione storiografica. La scuola delle «Annales»*, in "Annales E.S.C.", XXXIV (1979), n. 6, pp. 1360-1376.

quanto accadeva in America, gli storici delle «Annales» fecero dell'analisi seriale e quantitativa uno strumento per andare oltre l'economia e la società, come campo applicativo di indagine, ed ampliarlo a più largo respiro ai *sistemi di civiltà* nel loro complesso.

Le diverse «storie quantitative» avevano comunque delle caratteristiche comuni. In primo luogo, condividevano l'aspirazione ad introdurre nella ricerca storica metodi, tecniche e problematiche tipiche delle scienze sociali – l'economia e la demografia, per esempio – che sembrava potessero offrire agli studi storici nuovi quadri teorici di riferimento. Inoltre, le accomunava l'utilizzo del *computer*, divenuto indispensabile avendo sempre più a che fare con statistiche, numeri, tabelle; la «storia quantitativa» richiedeva infatti la raccolta e quindi l'elaborazione di notevoli quantità di dati, non analizzabili con i mezzi artigianali e rudimentali dello storico tradizionale. Il largo prevalere di questo tipo di utilizzo, prettamente quantitativistico, impedì lo sviluppo di altri usi del *computers* che tentavano di farsi strada all'interno di altre discipline, come quelle letterarie e umanistiche.

Un esempio di una fonte medievale in edizione elettronica, è quella del «Catasto fiorentino del 1427»²⁵, che risale agli anni settanta, e che costituisce una sorta di censimento a fini fiscali di Firenze, del suo contado e del suo distretto. Il passaggio da una fonte alla sua rappresentazione elettronica necessita di alcune opzioni metodologiche di fondo. Nel caso del Catasto fiorentino, furono scelti alcuni elementi da privilegiare, decidendo di trascurarne altri. L'ipotesi di ricerca, diretta all'analisi demografica e alla distribuzione della ricchezza, indirizzò la raccolta dei dati verso alcuni indicatori, ignorando necessariamente altre informazioni come, ad esempio, quelle relative agli assetti fondiari o al paesaggio agrario²⁶. Le informazioni raccolte erano codificate in schede perforate di ottanta colonne ed ogni informazione era tramutata in un codice numerico. Queste scelte, che privilegiavano solo alcune informazioni, comportavano chiaramente degli interventi sulla fonte tutt'altro che neutrali, ed il sostituire le espressioni linguistiche con codici numerici significava perdere informazioni preziose, oltre che impoverire la fonte stessa dei molteplici e complessi significati in essa contenuti. Inoltre, tradurre dei testi in simboli numerici predefiniti lasciava ampio margine all'interpretazione personale e quindi all'errore. Un altro aspetto critico di questo processo di codificazione era rappresentato dalla rigidità imposta dal formato di rilevazione dei dati, vale a dire le ottanta colonne, che non permetteva di modificare il tracciato in corso d'opera, apportando aggiunte o mutando codici.

Il problema ancora irrisolto, quindi, era ed è quello di riuscire ad elaborare modelli di rappresentazione delle fonti che permettano di trattare le informazioni in modo efficiente, senza impoverirle o senza alterare i molteplici significati in esse contenute.

Nella «storia seriale», il concetto stesso di fatto storico subiva una trasformazione radicale: se nella storiografia tradizionale il fatto storico era qualcosa di dato, che bisognava solo ritrovare *nelle* e *attraverso* le fonti, in questa nuova storia era invece il risultato di un processo di astrazione attraverso il quale lo storico definiva il suo oggetto di studio. Lo sviluppo della «storia quantitativa» nel corso degli anni ottanta aprì serrati dibattiti e proprio gli studiosi francesi, i primi ad averla abbracciata, cominciarono a prenderne le distanze. L'attenuarsi alla fine degli anni settanta, anche a causa del generale clima culturale, degli entusiasmi e delle aspettative verso la «storia quantitativa», coinvolse anche la modalità di utilizzo del *computer* nella ricerca storica. Per di più, le risorse informatiche si concentravano nei centri di calcolo e comportavano un alto investimento in tempo e denaro, oltre che difficoltà organizzative e tecnologiche, per non parlare poi di quelle metodologiche. Alla fine degli anni settanta, l'applicazione dell'informatica alla ricerca storica era ormai in fase di stallo²⁷.

Il rapporto tra informatica e storiografia si riconfigurò agli inizi degli anni ottanta con l'avvento e la diffusione del *personal computer*, che ribaltava la concezione fortemente accentrata dei processi di elaborazione. Il *computer* iniziò ad essere concepito come uno strumento di liberazione individuale contro i grandi sistemi centralizzati, come uno strumento flessibile, adatto a molti usi ed alla portata di tutti, non solo di una *élite* di tecnici specializzati. Negli anni successivi, la rapida evoluzione dei

²⁵ Cfr. D.HERLIHY, C.KLAPISCH-ZUBER, *I toscani e le loro famiglie. Uno studio sul Catasto fiorentino del 1427*, il Mulino, Bologna 1988 (ed. orig. Parigi 1978)

²⁶ Cfr. S. VITALI, *Passato digitale...*, op. cit., 2004, p. 13

²⁷ Cfr. S. VITALI, *Passato digitale...*, op. cit., 2004, p. 23-26

personal computers – a proposito della dimensione della memoria, della potenza di elaborazione, dei dispositivi quali stampanti, lettori e *masterizzatori* di CD-ROM e DVD, dei sistemi operativi, delle tecnologie di Rete, dello sviluppo di *softwares* sempre più semplici da utilizzare – hanno fatto sì che la flessibilità di questo strumento divenisse sempre maggiore.

Esauritosi il rapporto privilegiato con l'economia, la sociologia e la demografia, anche per la storiografia si sono aperte nuove possibilità di ricerca e di utilizzo; e la ricerca storica, in particolare, si è per lo più identificata con una specifica applicazione, il *word processor*, cioè i programmi di scrittura per scrivere ed elaborare bibliografie²⁸.

La velocità di esecuzione, la facilità di correzione, la rapidità di confezionare prodotti *puliti*, hanno incrementato la produttività, almeno dal punto di vista quantitativo: così come i correttori ortografici, grammaticali e sintattici e la grande innovazione dell'*ipertestualità*²⁹. Ma il *personal computer*, specie nella sua versione portatile, e i suoi programmi di scrittura hanno trasformato altri aspetti del lavoro dello storico. E' diventata realizzabile, infatti, l'idea di mettere un archivio dentro al *computer*, sfruttandone le capacità di archiviazione e stoccaggio dei dati; il *computer* consente quindi allo storico di crearsi il proprio archivio, di aggiornarlo, di organizzarlo, di prevedere percorsi di ricerca al suo interno. Si è aggiunta poi la possibilità di utilizzare anche fogli elettronici o di calcolo (come *Excel*, per esempio), che consentono di strutturare le informazioni in tabelle suddivise in celle, disposte in righe e colonne, con la possibilità di ottenere calcoli sulla base di formule predefinite. Infine, si sono diffusi i programmi di gestione di *database* – DBMS (*Database Management System*) – come il popolare *Access* della Microsoft o i più professionali *Oracle* o *MySQL*.

Il ricorso a questo genere di programmi ha riportato l'attenzione dello storico su delicate questioni metodologiche, questioni che sono a tutt'oggi aperte e oggetto di dibattito, specialmente in merito alla valutazione degli esiti dell'applicazione delle nuove tecnologie informatiche al trattamento dei documenti storici.

FRANCESCA RIZZI

Francesca Rizzi, nata a Cremona nel 1970, ha frequentato per tre anni l'Istituto di Archeologia dell'Università di Bologna, partecipando a due importanti campagne di scavo; trasferitasi per ragioni di lavoro all'Università di Parma, ha conseguito nel 2003 la laurea in Lettere presso l'Istituto di Storia, con una tesi in storia medievale. Impiegata di banca e dirigente sindacale, si occupa principalmente della formazione; ha pubblicato testi a contenuto storico, sia sulla storia del movimento sindacale, sia sulla figura di san Rocco.

²⁸ Cfr. R. JENSEN, *Scrivere col personal computer*, in "Passato e Presente", VI (1988), n. 16, pp. 165-185.

²⁹ Cfr. F. CARLINI, *Lo stile del Web. Parole e immagini nella comunicazione di Rete*, Torino 1999.