

La ricerca “per soggetto” dell’informazione in Internet

di Carla Basili

1. L'INFORMAZIONE IN INTERNET

L'informazione disponibile attraverso Internet continua a proliferare ed a divenire sempre più importante e “legittimata” entro la comunità scientifica. Il ricercatore attinge informazione dalla rete con la stessa naturalezza con la quale attinge informazione dalla biblioteca: la rete va configurandosi quindi come una delle molteplici fonti di conoscenza utili alla comunità scientifica e, soprattutto, si appresta a diventare uno strumento di lavoro quotidiano, “invisibile” (come è accaduto, per esempio, per i programmi di elaborazione di testi), entrando così a far parte degli utensili da tenere sulla scrivania. Tutto ciò rende sempre più pressante l'esigenza di meccanismi sistematici e standardizzati per identificare, localizzare e descrivere le risorse informative di rete. La biblioteca, da un lato, vedrebbe utilmente l'estensione del catalogo locale per incorporare ed integrare anche l'informazione in rete. Il ricercatore, d'altro canto, laddove usa l'informazione in rete, vorrebbe disporre di regole per la citazione nei riferimenti bibliografici. A questo proposito, ricordiamo l'imbarazzo dell'Editrice Bibliografica quando, in occasione di una pubblicazione,^[5] ha evitato il titolo *Riferimenti bibliografici* a favore del più vago *Fonti citate*, proprio per la presenza di citazioni di informazione ottenuta da Internet.

Internet si pone dunque come uno strumento di lavoro capace di portare sulla scrivania un patrimonio informativo enorme, fin troppo esteso ed in continuo aumento. Chi si “affaccia” alla rete viene travolto da una miriade di informazioni, entro le quali a fatica riesce ad orientarsi. Questa situazione è descritta da espressioni coniate appositamente dalla comunità statunitense di Internet, quali *infoglut* (indigestione da informazione) o *drinking from a firehose* (bere da un estintore).

L'atteggiamento del neofita di Internet passa attraverso tre fasi fondamentali: l'entusiasmo iniziale che lo porta a lunghe sessioni di navigazione nell'informazione in rete, la perples-

sità, quando si accorge che non esistono supporti sistematici al recupero dell'informazione; la delusione — o quanto meno il ridimensionarsi dell'entusiasmo — quando si rende conto che solo alcune fonti o alcuni punti di partenza sono validi. La delusione è tanto più forte, poi, se abbinata alla aspettativa di trovare in rete una offerta di servizi informativi comparabile con quella delle basi di dati in linea, tecnologia del contenuto e del recupero e non solo dell'accesso.

2. INFORMAZIONE IN LINEA E INFORMAZIONE IN RETE

Uno degli scopi di questo articolo è — almeno nelle intenzioni — generare realistiche aspettative in chi si accinga per la prima volta ad utilizzare Internet e le sue risorse informative. A questo proposito, vogliamo subito fugare eventuali false aspettative, derivanti da un confronto tra l'informazione in rete e l'ormai maturo e consolidato ambito delle basi di dati in linea. Tale confronto sarebbe malposto, in quanto la natura di Internet è profondamente diversa da quella dell'offerta commerciale di basi di dati; si tratta, a nostro avviso, di due mondi complementari, ciascuno con le proprie caratteristiche e con la propria validità specifica. Una prima distinzione scaturisce banalmente dalla terminologia inglese utilizzata per indicare le due classi di informazione: *networked information* per l'informazione disponibile in rete e *online information* per l'informazione disponibile attraverso fornitori commerciali di basi di dati, quali Dialog. Passiamo ora ad esporre alcuni degli elementi di distinzione tra le due classi di informazione.

2.1 Collegamento vs comunicazione

Le basi di dati in linea sono fondate sull'architettura cosid-

CARLA BASILI, Istituto di studi sulla ricerca e la documentazione scientifica, Consiglio nazionale delle ricerche, Roma.

detta "a stella", dove un terminale "muto" — privo cioè della capacità di elaborazione locale che possiede, per esempio, un personal computer — si collega ad un elaboratore che eroga servizi (host). Il terminale è completamente asservito all'elaboratore centrale, sia per il linguaggio di interrogazione (che è quello che l'elaboratore centrale usa ed impone), sia per il tipo di servizio (il terminale può utilizzare solo i servizi e le risorse informative che l'elaboratore centrale mette a disposizione). Nella singola sessione di lavoro, l'utente può accedere ad un unico host ed usufruire, quindi, di un servizio specifico. Questo tipo di architettura è "sbilanciata" a favore dell'host, il quale ne costituisce il fulcro.

Nel caso di Internet, invece, la connessione è reticolare, in quanto Internet è una costellazione di nodi interconnessi, dove, nella medesima sessione di lavoro, è possibile "navigare" da un nodo all'altro della rete, "vagolando" tra le risorse informative di interesse "sparse" nella rete ed utilizzando alternativamente i diversi servizi di rete, passando dall'uno all'altro con grande facilità.

Qualunque tipo di calcolatore può essere connesso ad Internet (dal grande elaboratore al personal computer) e di qualunque fornitore (Digital, Macintosh, Sun, Ibm,...); Internet è dunque una rete di calcolatori eterogenei. Nella rete ogni calcolatore ha pari "dignità" dal punto di vista della comunicazione, ha cioè diritto di operare con linguaggio e modalità propri (*peer-to-peer connection*). Questa "democrazia tecnologica" è una conquista dovuta al protocollo Tcp/Ip, il quale funge da minimo comun denominatore tra i linguaggi di calcolatori eterogenei. Gli strumenti di navigazione in rete, quali Gopher e World wide web sono stati costruiti in base al protocollo Tcp/Ip; questo fatto spiega come da un singolo strumento di navigazione sia possibile accedere a più risorse informative residenti su calcolatori eterogenei.

2.2 Costo vs gratuità

È noto come l'offerta di informazione in linea sia vincolata ad un abbonamento al servizio. È pur vero che l'industria dell'informazione si sta adeguando alle pressioni del mercato e sta modificando i propri criteri di attribuzione dei costi, non più basati sul tempo di collegamento, bensì sulla quantità di informazione recuperata: l'accesso all'informazione resta comunque subordinato ad un costo. La maggior parte dell'informazione in Internet è invece gratuita, in virtù della matrice accademica di Internet, dove l'utente è anche produttore di informazione, in uno scambio tipico del ciclo di produzione della conoscenza in ambito scientifico.

2.3 Validazione e garanzia dell'informazione

Strettamente legato all'aspetto dei costi è il fattore qualità dell'informazione. L'informazione in Internet è messa a disposizione spontaneamente, come si è detto, e non esistono al momento procedure sistematiche di validazione, né organismi responsabili di valutare e filtrare l'informazione. Questi meccanismi potrebbero garantire la qualità dell'informazione, in termini di contenuto e di aggiornamento. Allo stesso modo, potrebbe essere risolto un ulteriore aspetto poco desiderabile

dell'informazione in rete: la sua volatilità. Non esiste infatti alcuna garanzia circa la disponibilità permanente dell'informazione oggi presente nella rete. È noto, al contrario, il livello di qualità e di aggiornamento dell'informazione in linea.

2.4 Formato dell'informazione

L'informazione in linea esiste prevalentemente sotto forma di base di dati, interrogabile per mezzo di un programma per il recupero dell'informazione, chiamato Information retrieval system (Irs). Forniamo di seguito una sommaria descrizione di Irs, limitata agli scopi di questo contributo. Questo tipo di software è capace di individuare nella base di dati l'insieme dei documenti che risultano pertinenti rispetto ad una specifica esigenza informativa. L'esigenza informativa (*query*) è esplicitata dall'utente attraverso un linguaggio di interrogazione (*query language*). Ogni information retrieval system ha associato il proprio linguaggio e programmi diversi utilizzano linguaggi di interrogazione diversi. La situazione è analoga a quella degli opac di biblioteca: opac diversi sono accessibili tramite linguaggi diversi.

La pertinenza del singolo documento rispetto alla esigenza informativa espressa nella *query* è valutata dal programma attraverso l'uso di un thesaurus; il thesaurus è uno strumento di classificazione e di rappresentazione semantica di un corpus di conoscenza, ed offre molteplici punti di accesso al singolo documento. Il thesaurus è un elemento fondamentale nella logica operativa dell'Information retrieval system, che si basa sul principio che nel documento siano presenti (apposti o derivati) uno o più dei descrittori presenti nel thesaurus. In altre parole, la base documentaria è classificata coerentemente con lo schema semantico espresso nel thesaurus. Un altro fattore fondamentale negli Irs è la strutturazione della base documentaria: ogni documento è costituito da almeno due parti: una parte strutturata ed una parte a testo libero; il formato dei documenti nella stessa base documentaria è omogeneo. Pertinenza, dunque, come risultato di due elementi fondamentali: thesaurus e strutturazione dei documenti.

Il formato base di dati è invece il meno diffuso in Internet, dove prevale informazione a testo completo, non strutturata e, soprattutto, non organizzata e classificata.

3. GOVERNARE L'INFORMAZIONE IN INTERNET

Internet è un gigante cresciuto troppo in fretta e in maniera selvaggia, che solo ora si sta cercando di "imbrigliare" in una struttura organizzata. Il problema di inventariare l'informazione in Internet, nonostante sia stato avvertito precocemente, resta tuttora aperto. I tentativi di soluzione sono numerosi, sia dal punto di vista degli strumenti di ricerca, sia da quello dell'organizzazione dell'informazione.

3.1 Gli strumenti di ricerca

Gli strumenti per l'accesso all'informazione in rete vengono indicati nella letteratura anglofona come *Nir tools* ➤

Fig. 1

- BUBL Beginners, Updates, Contacts, News, Sponsors, Star Items
- Library & Information Science on and via BUBL (A-Z Sections)
- Internet Resources by Subject; Reference Tools, Electronic Texts & Journals
- BUBL's Latest Additions and Amendments
- The BUBL Subject Tree Project and The CATRIONA Project
- Keyword Searching: BUBL Files, Internet: Archie, Veronica, Wais
- NISS, Mailbase, SALSER, HENSA, PORTICO, Irish National Archive Server...
- Services by Type: OPACs, USENET, Gophers, WWW, WAIS
- Networks and Networking: Tools, Userguides, Training and ...

(Networked information retrieval tools: strumenti per il recupero dell'informazione in rete); questa denominazione è, a nostro avviso, impropria in quanto sembra indicare una classe di strumenti capaci di realizzare le funzioni di recupero dell'informazione proprie degli Irs, applicandole, per di più, a risorse informative distribuite in rete. Si tratterebbe, pertanto, di strumenti ancora più potenti dei tradizionali Irs che operano su basi di dati locali. Allo stato attuale, invece, strumenti quali Gopher e World wide web ci sembrano lontani dal realizzare le complesse e sofisticate funzionalità degli Irs: i concetti di thesaurus e controllo di classificazione sono in fase di discussione o, al più, ad un livello embrionale di sviluppo.

Sistemi quali Gopher e World wide web assolvono invece egregiamente alla funzione di navigazione nell'informazione in rete, offrendo all'utente la possibilità di accedere in modo facilitato e unificato a risorse che, diversamente, sarebbero accessibili in modo più complesso (per ciascuna risorsa si dovrebbe conoscere, almeno, l'indirizzo di rete, nonché la procedura di collegamento). Per questa ragione per questa classe di sistemi preferiamo la denominazione "strumenti di navigazione in rete". La letteratura specializzata tratta ampiamente le funzionalità di questi strumenti,[2], [6] per questo motivo in questo articolo ci limiteremo ad una loro descrizione essenziale.

Il Gopher è un sistema di distribuzione (intesa sia come pubblicazione, sia come accesso) dell'informazione in rete. Nel Gopher l'informazione è organizzata secondo una gerarchia di menu del tipo in Fig. 1.

Chiameremo *voce di menu Gopher* la singola riga di menu. L'informazione è accessibile semplicemente selezionando la voce di menu di interesse.

World wide web (chiamato anche Www o W3) è anch'esso un sistema di distribuzione dell'informazione in rete, ma — a differenza di Gopher — è basato sulla tecnologia degli ipertesti. L'informazione è presentata sotto forma di testo (Fig. 2), le voci selezionabili per l'accesso sono quelle sottolineate o seguite da un numero. Basta specificare il numero della voce o "fare click" sulla voce sottolineata per ottenere la connessione con la risorsa di interesse.

Per ciascuna di queste funzionalità di base sono state svi-

luppate interfacce grafiche del tipo "punta e fai click" (*point and click*) quali TurboGopher per Gopher e Mosaic o Netscape per Web, che facilitano ulteriormente la navigazione in rete. Il ruolo dell'interfaccia non si limita alla presentazione — per così dire all'estetica — della schermata, ma include la facilitazione di servizi accessori, quali la stampa o la memorizzazione di schermate e/o documenti. Queste ed altre funzioni accessorie, grazie all'interfaccia, vengono integrate completamente nella operatività della postazione di lavoro, così che l'utente viene ad operare secondo modalità consuete: quelle della sua postazione di lavoro abituale.

Fin qui abbiamo citato solo Gopher e World wide web i quali, a nostro avviso, realizzano le due funzionalità di base per l'accesso facilitato all'informazione in rete. Esistono numerosi altri strumenti che, classificati come Nir tools, sono descritti nella guida prodotta da Earn (European academic research network) e reperibile all'indirizzo: <http://www.earn.net/gnrt>. Solo alcuni tra gli strumenti di navigazione sono orientati al recupero dell'informazione basato sulla ricerca per parola, seppure la realizzino ancora in forma rudimentale. Tra questi "strumenti di ricerca" elenchiamo i più diffusi, raggruppandoli in base allo spazio di ricerca, cioè all'insieme degli oggetti sul quale viene applicata la ricerca per "soggetto".

3.1.1 Ricerca di file

Archie: per individuare archivi elettronici di documenti o di dati, dove lo spazio di ricerca è un insieme di nomi di file; al sistema viene sottoposta una stringa di caratteri che costituisce la *query*; il risultato è un insieme di indirizzi di file che contengono quella stringa nel nome. È un sistema a due passi, nel senso che fornisce solo un elenco di indirizzi: l'accesso al singolo file deve avvenire in un secondo momento, mediante Ftp.¹

3.1.2 Ricerca di documenti a testo libero

Wais: per individuare documenti, dove lo spazio di ricerca è il testo di documenti presenti in centinaia di basi di dati, è capace di attivare la medesima interrogazione in parallelo su più basi di documenti; l'uso della tecnologia parallela (capace di impiegare contemporaneamente centinaia di unità di elaborazione), riducendo l'ordine di grandezza dei tempi di risposta, consente la scansione del testo completo dei documenti. Questo strumento impiega un protocollo² di recupero dell'informazione in rete chiamato Z39.50. Il protocollo Z39.50 è centrale nel problema della ricerca per soggetto in Internet. Nella discussione sulla biblioteca virtuale, infatti, un punto focale è la capacità di recuperare informazione da una gamma di fonti diverse, attraverso una interfaccia unica ed amichevole. Il primo passo in questa direzione è la connessione di risorse informative in una rete intercontinentale, obiettivo, questo, che è ormai una realtà. Il secondo passo è

¹ Ftp (File transfer protocol) è il servizio che consente il trasferimento entro il proprio calcolatore di archivi (di dati, di documenti, di programmi) residenti su altri nodi di Internet.

² La comunicazione tra calcolatori è regolata da un *protocollo* di comunicazione — una sorta di linguaggio tra calcolatori —, le cui funzioni generali sono: definizione delle regole che governano lo scambio di messaggi; definizione del formato dei messaggi; definizione di corretti meccanismi di instradamento del messaggio lungo la rete; definizione delle modalità di trasmissione dei singoli bit.

quello di accedere in modo uniforme a questo patrimonio informativo distribuito in rete. Pensiamo ai diversi passi logici che intervengono (per ogni fonte) in questo processo:

- a) conoscenza dell'indirizzo di rete delle risorse;
- b) conoscenza delle procedure di connessione e di login;
- c) conoscenza dell'uso dello specifico linguaggio di interrogazione della singola risorsa.

Il protocollo Z39.50 consente di superare questi problemi, eliminando la necessità di connessione esplicita con la risorsa informativa remota e dando luogo realmente al concetto completo di Networked information retrieval. Lo Z39.50, inoltre, elimina la necessità di conoscere il linguaggio di interrogazione della singola base di dati, offrendo un unico linguaggio — tarato al livello minimo di complessità — per tutte le fonti.

3.1.3 Ricerca di menu Gopher

Veronica: per individuare voci di menu Gopher, dove lo spazio di ricerca è l'insieme dei menu Gopher. Veronica (Very easy rodent-oriented net-wide index to computerized archives) è l'equivalente per le voci di Gopher di ciò che Archie è per i file; il risultato della ricerca è un insieme di voci di Gopher che vengono presentate sotto forma di un menu Gopher, contenente solo voci nelle quali è presente la stringa richiesta.

Abbiamo visto come con Gopher l'accesso alla singola risorsa o l'attivazione del singolo servizio si ottiene semplicemente selezionando la riga corrispondente; questa modalità rende molto semplice la navigazione (ricordiamo che

Gopher nasce per questo fine), ma se si cerca informazione relativa ad uno specifico argomento, scandire successivamente in questo modo i menu di tutti i Gopher in Internet richiederebbe un tempo troppo lungo. Di aiuto in questa direzione è Veronica, strumento capace di ricercare per chiave entro l'insieme dei menu dei Gopher in Internet (circa il 99 per cento dei Gopher esistenti); Veronica accetta gli operatori "and", "or", "not". È possibile quindi sottoporre al sistema una o più stringhe di caratteri; il risultato è un menu Gopher costituito da voci provenienti da svariati altri Gopher; ogni voce contiene la stringa specificata. A gennaio 1995 Veronica indicizzava 5.057 Gopher. Oltre all'informazione "originata in Gopher", Veronica indicizza (seppure non in maniera completa come per i Gopher) anche 3.905 World wide web e circa 1.000 servizi accessibili tramite Telnet.³ Veronica è attivabile dall'indirizzo: `gopher://gopher.tc.umn.edu/Other Gopher and Information Servers/veronica`. *Jughead*: (Jonzy's universal gopher hierarchy excavation and display) per individuare voci di menu Gopher, dove lo spazio di ricerca è un menu Gopher. Ha una funzionalità molto simile a quella di Veronica: opera la ricerca per parola su menu Gopher e restituisce come risultato un menu Gopher costruito dinamicamente in accordo con la query. Accetta, come Veronica, i connettori "and", "or", "not" ed anche un insieme di comandi speciali; la descrizione completa di Jughead è reperibile all'indirizzo: `gopher://gopher.musc.edu:70/00/computing/jughead/About.jughead`. Jughead è una realizzazione della Washington & Lee University (Virginia) e reperibile all'indirizzo: `gopher://liberty.uc.wlu.edu`. Le differenze fondamentali tra Veronica e Jughead si posso-

Fig. 2



Welcome to the BUBL WWW Server

Please Read

Does your site have a BUBL representative yet?

All About BUBL

Help, News, Statistics, Special Features

LIS

Library and Information Science on and via BUBL

Internet Resources

Sources for Arts, Science, Commerce, E-journals, Institutes

BUBL Updates

BUBL's Latest Additions and Amendments

BUBL Subject Tree

BUBL Subject Tree/CATRIONA Project

Keyword Searching

Searching BUBL and the Internet

Major Network Services

HENSA, Mailbase, NISS, SALSER, PORTICO ...

Services Arranged by Type

Gophers, OPACs, USENET, WAIS, WWW,

World-Wide-Web

Information about the Project, Software, Tools, Networking

Warning on Copyright!

This MUST be read by all users of the BUBL Information Service

BUBL CONTACTS

The BUBL Information Service has received support from: ISSC of JISC, UKOLN/SUN, OCLC Europe, LA, LINC, BLRDD, Mecklermedia, Academic Press, Dynix and Innovative Interfaces.

Fig. 3

Web Indexes

Here's a list of some of the current indexes available on the Web. They all work in slightly different ways, so try them all and see which one works best!

- The WebCrawler Index
- Jumpstation
- The Nomad
- RBSE's URL Database
- The NorthStar Database
- The ALIWEB Database
- The SIMON Index
- The WWW Worm
- Harvest
- Lycos

With many thanks to the author of this page, bp@cs.washington.edu

no così sintetizzare:

— Jughead opera generalmente su un singolo Gopher, ladove Veronica, come si è detto, è capace di indicizzare il 99 per cento dei Gopher esistenti in Internet;

— a causa delle diverse dimensioni dell'indice, Jughead è aggiornato quotidianamente, mentre Veronica con cadenza variabile dalla settimana al mese;

— entrambi effettuano la ricerca solo sui titoli delle voci di Gopher, non sul testo dei documenti. ➤

³ Telnet: è il servizio di rete che consente ad un utente situato presso un nodo A di connettersi ad un calcolatore situato in un nodo B; una volta stabilita la connessione, la stazione di lavoro nel nodo A opera come se fosse un terminale del calcolatore nel nodo B e può quindi usufruire delle risorse di calcolo di quel calcolatore.

3.1.4 Ricerca di server World wide web

La rapida moltiplicazione dei World wide web in Internet ha prodotto una situazione analoga al caso dei Gopher: la scansione manuale, alla ricerca di un particolare argomento, è impraticabile. Sono nati allora degli strumenti di ricerca nello spazio Web, citati variamente in letteratura come *spider*, *web wander*, *worm* o *robot*. Forniamo a p. 43 un elenco di questi strumenti (Fig. 3), reperibile in rete all'indirizzo: <http://www.bus.msu.edu/search> (http://web.nexor.co.uk/public/aliweb/search/doc/form.htm) ci sembra degno di nota in quanto più di altri consente di precisare il profilo della query. È infatti possibile specificare (Fig. 4): i connettori "and/or", la percezione delle maiuscole, la ricerca per sottostringa, il numero massimo di voci-risultato da visualizzare, il tipo di risorsa informativa che si vuole ottenere come risultato (documento, servizio, organizzazione, persona, altro).

3.2 L'organizzazione dell'informazione

Per quanto riguarda l'organizzazione dell'informazione, le iniziative sono state, e sono, numerose. I primi tentativi includevano l'offerta complessiva di informazione e di servizi in Internet e si sono concretizzati:

- in guide stampabili quali *Zen and the Art of The Internet* di Brendan P. Kehoe (1992), *Hitchhikers Guide to the Internet* (1989); queste guide costituiscono semplicemente la versione digitale di documenti a stampa;

- in guide interattive, quali HyperCard Internet Tour e Hytelnet di P. Scott;

- in "liste di connessione" di Yanoff, che si è sforzato di compilare un elenco di risorse informative e di servizi, specificando per ciascuna voce l'indirizzo di rete e le modalità di accesso. Attualmente questo documento è migrato in Web all'indirizzo: <http://info.cern.ch/hypertext/DataSources/Yanoff.html>.

Successivamente l'orientamento si è spostato a favore di raccolte specializzate in singoli settori disciplinari. L'organizzazione per argomento di risorse informative e servizi si è andata sempre più affermando e realizzando con il Gopher nei cosiddetti "subject trees" e — ancora più efficacemente — con il web nelle guide per soggetto attive.

3.2.1 Subject trees

Un subject tree è un Gopher che ha realizzato una sorta di classificazione per materia delle risorse informative disponibili in Internet. Una delle più famose realizzazioni è Gopher jewels, reperibile via Gopher all'indirizzo: gopher://cwis.usc.edu/Other_Gopher_and_Information_Resources/Gopher_by_Subject/Gopher_Jewels o via World wide web all'indirizzo: <http://galaxy.einet.net>.

Riportiamo estensivamente il contenuto di Gopher jewels in formato Web (Fig. 5); l'intero contenuto di Gopher jewels è indicizzato, e quindi interrogabile, via Jughead.

Nella classe "Internet and computer related resources", Gopher jewels contiene un elenco di Gopher a loro volta organizzati per argomento (A list of gopher with subject trees).

Fig. 4

ALIWEB Search Form

This form queries the ALIWEB database. You can provide multiple search terms separated by spaces, and the results will be displayed in a best-match order.

Search term(s):

There are several types of search: Case Sensitive

Which type records would you like to search:

Which fields would you like to search?
 Titles Descriptions Keywords URL's

What fields would you like displayed in addition to the title?
 Descriptions Keywords URL's Other fields

You can restrict the results to a domain (e.g. "uk"):

Stop after the first matches.

ALIWEB

Altre iniziative orientate alla organizzazione per soggetto di Gopher sono:

- 1) Riceinfo: <gopher://riceinfo.rice.edu> (corredato di ricerca via Jughead);
- 2) Adelaide: <gopher://jarrah.itd.adelaide.edu.au/Bysubject>.

In Europa non possiamo dimenticare Bubl (Bulletin board for libraries), una iniziativa nata per "incoraggiare, sviluppare, coordinare e supportare la emergente comunità Lis (Library and information science) in rete nel Regno Unito e di promuovere gli interessi di questa" (traduzione dal testo emanato in rete da Bubl). L'informazione presente in Bubl e la sua organizzazione costituiscono uno dei "fiori all'occhiello" tra le realizzazioni in Europa, non smentendo la tradizione bibliotecaria inglese. Bubl è presente in Internet nella versione Gopher (Fig. 1): <bubl.bath.ac.uk> e nella versione Web (Fig. 2): <http://www.bubl.bath.ac.uk/BUBL/home.html>. L'informazione in Bubl è inoltre interrogabile via Wais in entrambe le versioni (Gopher e Web). Un progetto fondato su solide basi teoriche è Catriona, dichiaratamente orientato a estendere l'opac locale con le risorse di rete mediante uno strumento di recupero dell'informazione basato sul protocollo Z39.50. L'analisi condotta nello sviluppo del progetto prevede anche la possibilità di ampliare il formato Marc per includere la descrizione dei dati di accesso delle risorse di rete (vedi punto 4). Per ulteriori informazioni sul progetto Catriona: <http://www.bubl.bath.ac.uk/BUBL/maincatriona.html> oppure via <gopher://bubl.bath.ac.uk:7070/11/Link/Catriona>. Infine ricordiamo il subject-tree della Library of Congress, all'indirizzo: [gopher://marvel.loc.gov/Global Electronic library \(BySubject\)](gopher://marvel.loc.gov/Global_Electronic_library_(BySubject)).

3.2.2 Guide attive

Una guida attiva è un documento elettronico formalmente simile ad una guida elettronica stampabile, con la differenza

che ogni risorsa o servizio elencato nella guida attiva è, rispettivamente, accessibile o attivabile entro la guida stessa. In effetti una guida attiva è un documento ipertestuale di World wide web.

Presso l'Università del Michigan la School for Information and Library Studies, insieme con la University Library, hanno sviluppato il progetto "Internet Resource Discovery". Uno dei risultati di questo sforzo è la Clearing house for subject-oriented Internet resource guides, all'indirizzo: <http://www.lib.umich.edu/chhome.html>, che offre un insieme di guide specializzate ciascuna ad uno specifico ambito disciplinare.

Sono guide veramente utili, ben strutturate e complete, sia dal punto di vista del contenuto, sia da quello degli strumenti di rete ove questa è reperibile (è inclusa tutta la gamma degli utensili di rete da Gopher e Web, alle conferenze elettroniche, ai nodi Ftp).

3.2.3 Biblioteche virtuali

Analoghe ai subject-trees per Gopher sono le *virtual libraries* per Web, organizzazioni per disciplina di risorse residenti su Web. Una delle più famose biblioteche virtuali è quella sviluppata presso il Cern di Ginevra, reperibile all'indirizzo: <http://info.cern.ch/hypertext/DataSource/bySubject/Overview.html>.

Nella classificazione è contenuta anche la biblioteca virtuale italiana, nella voce "Italian general subject tree", residente all'indirizzo: <http://www.mi.cnr.it/IGST> e contenente, organizzato per soggetto, l'elenco dei server Web italiani.

Altre virtual libraries sono raggiungibili dalla stessa virtual library del Cern, nella voce "Other virtual libraries".

4. LA CATALOGAZIONE DELL'INFORMAZIONE

Fin qui abbiamo descritto brevemente alcune delle iniziative per l'organizzazione delle risorse informative di rete e per lo sviluppo di strumenti di ricerca per argomento. La matrice di questi tentativi è prevalentemente pratica, dettata dalla necessità di migliorare le possibilità di orientamento entro l'informazione in rete.

Parallelamente a queste attività, si sono avviati da tempo — dalla fine del 1991, circa — dei gruppi di lavoro di matrice bibliotecaria, per l'analisi della possibilità di catalogare sistematicamente — in base a regole di catalogazione e formati di rappresentazione standard — le risorse di rete. L'American Library Association si è attivata precocemente, con il comitato Marbi (Machine-readable form of bibliographic information), che già nel maggio 1991 distribuiva in rete il discussion paper n. 49 (*Dictionary of data elements for online information resources*), seguito nel gennaio 1992 dal discussion paper n. 54 (*Providing access to online information resources*). L'obiettivo di Marbi, in collaborazione con Oclc e con la Library of Congress, è estendere le Anglo-American Cataloging Rules ed il formato Marc per permettere la descrizione delle risorse di rete.

La catalogazione delle risorse di rete, infatti, necessita di strumenti capaci non solo di descrivere la risorsa, ma anche di fornire dati sulle possibilità di accesso alla risorsa stessa (indirizzo di rete, procedura di connessione, ecc.). Non tutte

le risorse in rete sono bibliografiche o fattuali, come si è detto; gran parte dell'offerta di Internet è invece servizio (per esempio un nodo erogatore di servizio Ftp, un sistema informativo di ateneo, ecc.).

Il risultato dello studio condotto da Marbi (vedi discussion paper n. 69, *Accommodating online systems and services in Usmarc*) è la definizione di un nuovo campo Marc, il n. 856 (*Electronic location and access*) destinato appunto a contenere le informazioni per accedere alla risorsa remota. L'indirizzo del comitato Marbi, ove è possibile recuperare via Ftp gli articoli citati, è: <ftp://ftp.cni.org/pub/MARBI>, oppure, via gopher: <gopher://gopher.cni.org/cniftp/pub/M>.

In linea con la proposta Marbi, Oclc (Online computer library center) ha iniziato un progetto per la catalogazione delle risorse in Internet, il cui termine è previsto per il 31 marzo 1996, aperto a quanti vogliano collaborare; per ulteriori informazioni: <http://www.oclc.org/oclc/man/catproj/catcall.htm>.

5. CONCLUSIONI

Rispetto alle tradizionali fonti di informazione, le risorse informative di rete appaiono non organizzate, prive di struttura, non validate e, pertanto, utilizzabili con difficoltà. Del resto questo non deve stupire, se si valuta la impressionante velocità di sviluppo di Internet e la conseguente proliferazione di servizi e informazioni sulla rete.

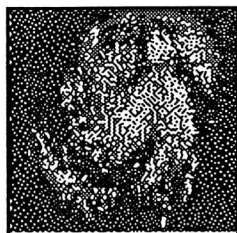
Dal versante tecnologico, per l'accesso a questo patrimonio informativo, si assiste ad un proliferare di strumenti tale che a stento si riesce ad essere aggiornati sulla loro evoluzione. Nessuno strumento, tuttavia, realizza ancora le sofisticate funzionalità dei sistemi di recupero dell'informazione in linea (per questo abbiamo scelto la più generica denominazione di "strumenti di ricerca", evitando la più precisa "strumenti di recupero dell'informazione"); tutti presentano una matrice pratico/tecnologica piuttosto che teorico/metodologica. In generale, infatti, nascono in risposta a pressanti esigenze di organizzare l'informazione; l'urgenza di fornire una rapida, seppur parziale, soluzione ha fatto sì che sia mancata una analisi teorica approfondita dei requisiti funzionali del problema della catalogazione dell'informazione in rete. Neppure Wais, che funzionalmente — operando su documenti a testo pieno — è il più vicino agli strumenti tradizionali di recupero dell'informazione, si avvale di supporti quali liste di autorità o thesauri. Parallelamente avanzano i progressi teorici per la catalogazione delle risorse di rete, nella direzione di sviluppare un censimento funzionale al recupero dell'informazione in Internet. Gli sforzi di sistematizzazione dell'informazione in Internet, dunque, sono focalizzati su tre aspetti: classificazione delle risorse; sviluppo di strumenti di ricerca; definizione di regole per la descrizione e la catalogazione delle risorse di rete.

Ciò che manca, a nostro avviso, è un meccanismo organizzativo, che:

— stabilisca dei referenti responsabili del mantenimento del "catalogo globale" (che potrebbe anche essere distribuito su più nodi);

— definisca delle procedure universali per la "nascita" e la "morte" delle risorse e dei servizi di rete, a governo della volatilità dell'informazione. ►

Fig. 5



EINet Galaxy

Up Home Help Search



We're outta here!

Check out new job opportunities at EINet!

The Galaxy is a guide to worldwide information and services and is provided as a public service by EINet and Galaxy guest editors. The information is organized by topic, and can be searched. You can also add your own information to Galaxy. For more help in moving about the Galaxy, click on the word "Help" in the bar at the top of this page.

See what Websurf Magazine says about EINet's World Wide Web clients - winWeb for Window; and MacWeb for the Macintosh.

Topics

Arts and Humanities

Architecture - Language and Literature - Performing Arts - Philosophy - Religion - Visual Arts

Business and Commerce

Business Administration - Business General Resources - Consortia and Research Centers - Consumer Products and Services - Electronic Commerce - General Products and Services - Investment Sources - Management - Marketing and Sales

Community

Charity and Community Service - Consumer Issues - Crime and Law Enforcement - Culture - Education - Environment - Family - Gender Issues - Health - Home - Immigration - Law - Liberties - Lifestyle - Net Citizens - Networking and Communication - News - Parascience - Politics - Religion - Safety - US States - Urban Life - Veteran Affairs - Workplace - World Communities

Engineering and Technology

Agriculture - Biomedical Engineering - Chemical Engineering - Civil and Construction Engineering - Computer Technology - Electrical Engineering -

Il coronamento di questi sforzi dovrebbe attuarsi nella convergenza verso un efficiente sistema di recupero dell'informazione e dei servizi di rete, affiancato da una robusta gestione, così da includere a pieno titolo l'informazione in rete nel novero delle fonti "canoniche" per la comunità scientifica.

6. RIFERIMENTI BIBLIOGRAFICI

[1] C. BASILI, *Strumenti informatici per biblioteche e centri di*

documentazione: interazione funzionale di software, Seminario Aib "Integrazione di sistemi bibliografici e documentari nel territorio geografico e culturale trentino", Trento, 25 ottobre 1991.

[2] C. BASILI, *Recupero dell'informazione in rete: la situazione della normativa*, "Aida Informazioni", n. 4/1992.

[3] C. BASILI - P. PICCOTTI - A.M. TAMMARO, *Reti telematiche e servizi bibliografici*, Firenze, Ifnia Didattica, 1993, p. 104.

[Human Factors and Human Ecology](#) - [Manufacturing and Processing](#) - [Materials Science](#) - [Mechanical Engineering](#) - [Technology Transfer](#) - [Transportation](#)

Government

[Government Agencies](#) - [Laws and Regulations](#) - [Military](#) - [Politics](#) - [Public Affairs](#)

Law

[Administrative](#) - [Commercial](#) - [Constitutional](#) - [Criminal](#) - [Environmental](#) - [Intellectual Property](#) - [Legal Profession](#) - [Military](#) - [Personal Finance](#) - [Research](#) - [Societal](#) - [Tax](#)

Leisure and Recreation

[Amateur Radio](#) - [Automotive](#) - [Aviation](#) - [Beverages](#) - [Boating](#) - [Collectibles](#) - [Dance](#) - [Fashion](#) - [Film and Video](#) - [Food](#) - [Games](#) - [Gardening](#) - [Hiking](#) - [Horses](#) - [Humor](#) - [Just For Kids](#) - [Metal Detecting](#) - [Motorcycles](#) - [Music](#) - [Personalities](#) - [Pets](#) - [Photography](#) - [Pictures](#) - [Radio](#) - [Reading](#) - [Recipes](#) - [Restaurants](#) - [Scuba Diving](#) - [Speleology](#) - [Sports](#) - [Television](#) - [Travel](#)

Medicine

[Chiropractic](#) - [Community Medicine](#) - [Dentistry](#) - [Exercise](#) - [History of Medicine](#) - [Holistic Medicine](#) - [Human Biology](#) - [Medical Applications and Practice](#) - [Medical Specialties](#) - [Medical Technologies](#) - [Nursing](#) - [Nutrition](#)

Reference and Interdisciplinary Information

[Census Data](#) - [Dictionaries etc](#) - [Directories](#) - [Events and Conferences](#) - [Grants](#) - [Internet and Networking](#) - [Library Information and Catalogs](#) - [Publications](#)

Science

[Astronomy](#) - [Biology](#) - [Chemistry](#) - [Geosciences](#) - [Mathematics](#) - [Physics](#)

Social Sciences

[Anthropology](#) - [Economics](#) - [Education](#) - [Geography](#) - [History](#) - [Library and Information Science](#) - [Psychology](#) - [Sociology](#)

[Top of Page](#)

EINet is a commercial provider of network communication and information services. The fab [Galaxy image](#) comes from the [Space Telescope Science Institute](#) (STScI).

Copyright 1993,1994,1995 EINet Corporation. All Rights Reserved.

Contact the [Galaxy Staff](#) for more information.

Search for: ▼
 [Galaxy Pages](#) [Galaxy Entries](#) [Gopher](#) [Hytelnet](#)
 [World Wide Web \(full text search\)](#) [World Wide Web Links \(link text only\)](#)

[Up](#) - [Home](#) - [Help](#) - [Search](#) - [Top](#) -- [EINet Galaxy](#)

[4] C. BASILI, *Accesso in rete ad informazione eterogenea*, atti del Convegno Aida "Documentazione e utenti: cultura del servizio, multimedialità e marketing", Feb. 1993.

[5] C. BASILI - C. PETTENATI, *La biblioteca virtuale: accesso a risorse informative in rete*, Milano, Editrice Bibliografica, 1994, p. 147.

[6] E. KROL, *The Whole Internet User's Guide and Catalog*, 2nd ed., Sebastopol, CA, O'Reilly and Associates, 1994.

[7] CLIFFORD A. LYNCH, *A Framework for Identifying, Locating, and Describing Networked Information Resources*, Draft for discussion at March-April 1993, Ietf Meeting.

[8] *Sulle ali di Internet: biblioteche al bivio*, a cura di R. Maini, "Biblioteche oggi", 13 (1995), 3, p. 6-11.

[9] Gran parte dell'informazione elaborata per questo articolo è stata tratta da Internet.