Giugno 23, 1995
Alessandra Richetti

Reperimento di informazioni su Internet

Si tratta di un'introduzione pratica alle applicazioni che consentono di accedere alla rete Internet, e di recuperare informazioni su di essa.

Informazioni revisione e aggiornamento: Prima Edizione: Giugno 1995

CAPITOLO 1

INTRODUZIONE

Dal punto di vista dell'utente la rete Internet si presenta come un' immenso *'centro commerciale'*.
Trovo i negozi (**siti**) che offrono servizi (**server**) e che rendono disponibile la propria mercanzia (**file, informazioni**).

Posso essere quasi sicuro che cio' di cui ho bisogno si trova a mia disposizione da qualche parte.
Ma dove ?
E scoperto dove, cosa mi serve, poi, per recuperalo ?

Questo crea inevitabilmente un senso di smarrimento.

Ecco una lista delle esigenze piu' comuni nella consultazione del mondo Internet e di come soddisfarle con gli strumenti, disponibili presso il nostro Dipartimento.

1.1 COME ...

- consulto gli archivi di pre-prints ? (GOPHER, LYNX, XMOSAIC, ...)
- opero su file con formati non usuali : archivi tar, files compressi, files codificati ? (VMSTAR, UUDECODE, GUNZIP)
- recupero informazioni su un argomento specifico ? (LYNX, XMOSAIC, ..., news, faq)
- trovo files ? (ARCHIE)
- trasferisco files ? (FTP)

1.2 Information Service

Ogni negozio espone la propria mercanzia in maniera diversa. Gli Information Service sono i *'diversi modi'* (diversi **protocolli**) di offrire le informazioni.

- **FTP F**ile **T**ransfer **P**rotocol e‘ lo strumento per trasferire file sulla rete Internet. Ogni utente che abbia i privilegi puo‘ trasferire file da/a un nodo della rete Internet.
- **Gopher** Il mondo gopher e‘ costituito da informazioni organizzate in files ed accessibili via menu.
- **USENET** E‘ caratterizzato da gruppi di discussioni liberi ed aperti
- **WWW** E‘ una iniziativa che, attraverso la forma dell'ipermedia cerca di rendere accessibile un' enorme quantita‘ di informazioni organizzate in pagine di informazione

1.3 strumenti utili per consultare e recuperare le informazioni

- browser
  - FTP
  - ARCHIE
  - XMOSAIC (no VMShelp)
  - LYNX
  - GOPHER (no VMShelp)
- applicazioni varie per manipolare le informazioni raccolte
  - XV
  - GV (ghostview) (no VMShelp)
  - MPEG
  - VMSTAR
  - UUENCODE/UUDECODE (VMShelp uuencode)
  - GUNZIP (VMShelp gzip)

CAPITOLO 2

COME TROVO FILES IN INTERNET ... ARCHIE

2.1 Archie

Archie ricerca una stringa fra i principali Archivi FTP del mondo.

Ogni server Archie raccoglie, conserva le informazioni relative ai siti FTP, e le aggiorna periodicamente.

- le informazioni recuperate non sono sempre corrispondenti alla realta‘
- ricerche su server diversi danno risultati diversi

```
VSTST0>
VSTST0>archie
Usage: $1$dka200:[util.][archie]archie.exe;2 [-[cers][l][t][m#][h host]["L"]["N"
#]] string
       -c : case sensitive substring search
       -e : exact string match (default)
       -r : regular expression search
       -s : case insensitive substring search
       -l : list one match per line
       -t : sort inverted by date
      -m# : specifies maximum number of hits to return (default 98)
  -h host : specifies server host
       -"L" : list known servers and current default
      -"N"# : specifies query niceness level (0-35765)
VSTST0>archie -m1000 -h archie.unipi.it internet

Host achilles.doc.ic.ac.uk

    Location: /imported
      DIRECTORY drwxrwxr-x        512  Mar 21 05:11  internet
    Location: /
           FILE -rwxrwxrwx         17  Mar 27 1994  internet
    Location: /theory
           FILE -rwxrwxrwx         20  Mar 27 1994  internet
    Location: /theory/papers/Dawson/talks
      DIRECTORY drwxr-xr-x        512  Mar 21 04:59  internet

Host alga.icm.edu.pl

    Location: /pub1/mirror/simtel/win3
      DIRECTORY drwxr-xr-x        512  Apr 13 07:09  internet
    Location: /pub1/mirror/usenet
      DIRECTORY drwxr-xr-x        512  Apr 13 08:04  internet
    Location: /pub1/usenet
      DIRECTORY drwxr-xr-x        512  Apr  7 08:29  internet
    Location: /pub1/win3
      DIRECTORY drwxr-xr-x        512  Apr  7 11:28  internet

Host bash.cc.keio.ac.jp

    Location: /pub/netnews/news.answers
      DIRECTORY drwxrwxr-x        512  Jan 23 14:56  internet

Host bells.cs.ucl.ac.uk
```

CAPITOLO 3

COME RECUPERO INFORMAZIONI SU UN ARGOMENTO PRECISO

3.1 USENET

Un modo di reperire informazioni sulla rete Internet e‘ rappresentato dal mondo USENET. I newsgroup sono gruppi di discussione spontanei ed aperti. Ogni persona vi puo‘ partecipare via e-mail leggendo o contribuendo con i propri interventi

E‘ organizzato gerarchicamente.

```
alt         Alternative way og Looking at Things
comp        computer
news        news net & software
rec         recreations
sci         scientific research
soc         social issues
talk        discussions
VMSnet      discussion about VAX/VMS & DECNET
Per es.
   comp.lang.fortran
```

Sotto il gruppo computer science c’e‘ la categoria linguaggi di programmazione in cui si discute del Fortran

3.2 FAQ

Le domande o gli argomenti piu‘ trattati di un newsgroup vengono raccolte nelle **FAQ** ovvero Frequently Asked Questions.

```
                                     Selected Internet TOOLS (p1 of 2)
   [DFT-INFN]
                            SELECTED INTERNET TOOLS

     * USENET : News - Reading

     * FAQ

     * WWW Search

       * The Lycos Home Page: Hunting WWW Information
       * WebCrawler Searching
       * Query Interface to the WWW Home Pages Broker.
       * Centre Universitaire d'Informatique World Wide Web catalog
       * The JumpStation Search Page
       * JumpStation II
       * NIKOS
       * ALIWEB
       * WWWW - World Wide Web Worm
       * ArchiePlexForm
-- press space for next page --
  Arrow keys: Up and Down to move. Right to follow a link; Left to go back.
 H)elp O)ptions P)rint G)o M)ain screen Q)uit /=search [delete]=history list
```

```
                                          Usenet News at ICTP  (p1 of 2)
                         USENET NEWS READING & POSTING

   The Usenet News can be seen as a large repository of articles or
   messages, organized in newsgroups.

   Messages come mainly via Internet from all over the world, from
   universities, government agencies, research intitutions and home
   computers(!!), and expire (are removed from the repository) after a
   locally defined period of time (about one week here).

   NEWSGROUPS
     * Alternative Topics (alt)
     * Biology Net (bionet)
     * Computing (comp)
     * European Usenet (eunet)
     * GNU Project (gnu)
     * Italy-Trieste (ita)
     * Miscellaneous (misc)
-- press space for next page --
  Arrow keys: Up and Down to move. Right to follow a link; Left to go back.
 H)elp O)ptions P)rint G)o M)ain screen Q)uit /=search [delete]=history list
```

```
                                                           (p1 of 85)
Computing


  comp.admin.policy
         Discussions of site administration policies.

  comp.ai
         Artificial intelligence discussions.

  comp.ai.edu
         Applications of Artificial Intelligence to Education.

  comp.ai.fuzzy
         Fuzzy set theory, aka fuzzy logic.

  comp.ai.genetic
         Genetic algorithms in computing.
-- press space for next page --
  Arrow keys: Up and Down to move. Right to follow a link; Left to go back.
 H)elp O)ptions P)rint G)o M)ain screen Q)uit /=search [delete]=history list
```

```
                                                               (p23 of 85)
   comp.lang.postscript
          The PostScript Page Description Language.

   comp.lang.prograph
          Prograph, a visual object-oriented dataflow language.

   comp.lang.prolog
          Discussion about PROLOG.

   comp.lang.python
          The Python computer language.

   comp.lang.rexx
          The REXX command language.

   comp.lang.sather
          The object-oriented computer language Sather.

   comp.lang.scheme
          The Scheme Programming language.

-- press space for next page --
  Arrow keys: Up and Down to move. Right to follow a link; Left to go back.
 H)elp O)ptions P)rint G)o M)ain screen Q)uit /=search [delete]=history list
```

```
                Newsgroup comp.lang.postscript,  Articles 24347-24376 (p1 of 2)
   (Earlier articles...) Articles in comp.lang.postscript
     * "Product Manager Print On Demand 100K" - Paul Morris
     * "Where can I find the PS spec?" - Timothy McGinn
     * "Re: Where can I find faq and ppds?" - Aandi inston
     * "Looking for a PS Interpreter on SGI" - Rupert Smith
     * "Re: Desperately need PS-to-ASCII or other readable form" - Vidiot
     * "Re: WANTED: Postscript Viewer / Print Utility for MS-Window ????"
       - grgreen
     * "Re: PS/PDF -> HTML" - Wolfgang Ratzka
     * "Re: "a4", where is it defined ?" - David Forthoffer
     * "Where are the *Actual* Hershey fonts?" - Julie S. Porter
     * "Code to generate a hemisphere" - Julie S. Porter
     * "PS Level 2 Language Descrip. - where ?" - Andreas Jung
     * "SPARCprinter cuts off print area :(" - Fredrick Backman
     * "Re: A4 on Docutech printer" - Piet van Oostrum
     * "More pagecount discussion" - Ben Holzman
     * "Are European PostScript printers configured differently?" - Dr.
       Tom Epperly
     * "Re: GhostScript as X11 DPS extension?" - Casper H.S. Dik -
       Network Security Engineer
-- press space for next page --
  Arrow keys: Up and Down to move. Right to follow a link; Left to go back.
 H)elp O)ptions P)rint G)o M)ain screen Q)uit /=search [delete]=history list
```

```
            Newsgroup comp.lang.postscript,  Articles 24317-24346 (p2 of 2)
  * "Re: Compression of bitmaps is PS?" - questor@onix.com
  * "Re: GhostScript as X11 DPS extension?" - David L. Crow
  * "Force download of ALL fonts in MS-Win?" - David Arnstein
  * "Re: Type 9 Halftone dict's - What" - Ed Taft
  * "Re: GhostScript in prepress" - L. Peter Deutsch
  * "Re: Spot function linearity" - Karsten Jeppesen
  * "PS/PDF -> HTML" - L. Peter Deutsch
  * "Re: WANTED: Postscript Viewer / Print Utility for MS-Window ????"
    - questor@onix.com
  * "More GhostView Printing Prob's" - Arnold Shore
  * "Re: GhostScript as X11 DPS extension?" - Martin Michlmayr
  * ""SVDoc restore /SVDoc save def" - Hey!!" - Rich Ellis
  * "Q: Touchy fontmap in gs3.33" - "Robert A. Hueckstedt"

 (Later articles...)
  _________________________________________________________________

 Post to comp.lang.postscript


Commands: Use arrow keys to move, '?' for help, 'q' to quit, '<-' to go back
  Arrow keys: Up and Down to move. Right to follow a link; Left to go back.
 H)elp O)ptions P)rint G)o M)ain screen Q)uit /=search [delete]=history list
```

```
                                                      Welcome  directory
                                WELCOME

Feb 28 08:34  Directory       alt.answers
Mar  8 17:07  Directory       comp.answers
Mar 10 22:21  Directory       news.answers
Feb 24 08:34  Directory       mrec.answers
Jan 18 15:07  text/pain       saveit.pl  1Kb
Feb 28 08:36  Directory       sci.answers
Feb 28 08:33  Directory       soc.answers
Oct 26  1994  Directory       talk.answers




Commands: Use arrow keys to move, '?' for help, 'q' to quit, '<-' to go back
  Arrow keys: Up and Down to move. Right to follow a link; Left to go back.
 H)elp O)ptions P)rint G)o M)ain screen Q)uit /=search [delete]=history list
```

```
                                   comp.answers directory (p1 of 11)
                                 COMP.ANSWERS

   Up to faq

Mar  8 17:06  Directory        386bsd-faq
Mar  8 17:05  Directory        3b1-faq
Mar  8 17:03  Directory        3b2-faq
Feb  4 12:26  text/plain       3d-programmer-info  2Kb
Jun 16  1994  text/plain       5621  14Kb
Jun 27  1994  text/plain       68hc11-microcontroller-faq  36Kb
Mar  8 17:03  Directory        acorn
Feb  4 10:24  text/plain       advertisers-blacklist  25Kb
Mar  8 17:03  Directory        ai-faq
Mar  8 17:04  Directory        aix-faq
Mar  8 17:04  Directory        amiga
Sep 27  1994  text/plain       AmigaBooksFAQ  36Kb
Feb  8 07:06  text/plain       apl-faq  37Kb
Mar  8 17:04  Directory        apollo-faq
Mar  8 17:05  Directory        apple2
Mar  8 17:07  Directory        arch-storage
-- press space for next page --
  Arrow keys: Up and Down to move. Right to follow a link; Left to go back.
 H)elp O)ptions P)rint G)o M)ain screen Q)uit /=search [delete]=history list
```

```
                                              comp.answers directory (p9 of 11)
Mar  8 17:05  Directory        postscript
Feb 16 07:00  text/plain       powerpc-faq  61Kb
Mar  8 17:04  Directory        ppp-faq
Mar  8 17:07  Directory        privacy
Feb  4 03:41  text/plain       proj-plan-faq  57Kb
Mar  8 17:03  Directory        prolog
Mar  8 17:06  Directory        python-faq
Mar  8 17:05  Directory        realtime-computing
May  8  1994  Directory        ripem
Mar  8 17:06  Directory        robotics-faq
Mar  8 17:03  Directory        scheme-faq
Mar  8 17:04  Directory        scsi-faq
Jun 23  1993  Directory        sgi
Jan 13 00:54  text/plain       signature_finger_faq  37Kb
Dec 16  1994  text/plain       smalltalk-faq  48Kb
Feb  7 23:01  text/plain       snmp-faq  52Kb
Mar  8 17:02  Directory        software-eng
Mar  8 17:05  Directory        Solaris2
Feb 23 16:38  text/plain       sources-test-faq  26Kb
Oct 29  1994  text/plain       ssn-privacy  31Kb
Oct  7  1994  text/plain       standards-faq  85Kb
-- press space for next page --
  Arrow keys: Up and Down to move. Right to follow a link; Left to go back.
 H)elp O)ptions P)rint G)o M)ain screen Q)uit /=search [delete]=history list
```

```
* Introduction
* 1 General Questions
* 1.1 What is PostScript?
* 1.2 How can I tell how many pages my document will have?
* 1.3 How can I print just one page or a few pages from a big
* 1.4 How can I print more than one page per sheet of paper?
* 1.5 How can I edit a PostScript picture?
* 1.6 How can I print PostScript on a non-PostScript printer?
* 1.7 What other graphics languages are there?
* 2 Printers
* 2.1 How can I get my printer to talk back to me?
* 2.2 Should I leave my printer on?
* 2.3 How do I suppress the power-on start page?
* 2.4 How do I make a downloaded font ``persistent''?
* 2.5 How do I remove a persistent (``permanent'') downloaded font?
* 2.6 How do I reset the printer without power-cycling?
* 2.7 About saving files
* 2.8 What's the control-D business?
* 2.9 Why does the printer say ``still busy'' when my document is
* 2.10 How should I set up my spoolers?
* 2.11 What are PPD files?
* 3 Formats and Conversions
* 3.1 How can I convert PostScript to some other graphics format?
* 3.2 How can I convert DVI to PostScript?
* 3.3 How can I convert HP Laserjet language (PCL) to PostScript?
* 3.4 How can I convert TeX PK format font to PostScript?
* 3.5 How do I embed PostScript into troff?
* 3.6 How do I embed PostScript into LaTeX or TeX?
* 3.7 How can I convert an image to PostScript?
* 3.8 How can I convert ASCII text to PostScript?
* 3.9 How can I convert PostScript to ASCII?
* 3.10 How do I convert between PFB, PFA, Mac, PFM, and AFM?
* 4 Fonts
* 4.1 What are .PFB and .PFA files?
* 4.2 How can I convert a PostScript font to TeX's PK format?
* 4.3 Why are Adobe fonts hidden?
* 4.4 How do I get bitmap representations of Adobe fonts?
* 4.5 What are some good ftp sites for fonts?
* 4.6 How can I re-encode a font?
* 4.7 What's the difference between a Type 1 and a Type 3 font?
* 4.8 What vendors sell fonts for PostScript printers?
* 4.9 What are ATM fonts?
* 4.10 What are Multiple Master Fonts?
* 4.11 Do I need a Level Two printer to use Multiple Master Fonts?
* 4.12 What are Type 4 fonts?
* 4.13 What are Type 5 fonts?
```

CAPITOLO 4

FTP - FILE TRANSFER PROTOCOL

Lo strumento TCP/IP per trasferire file nella rete Internet prende il nome di **FTP**.

E‘ una applicazione interattiva.

Si stabilisce una sessione interattiva con un nodo remoto.

```
FTP
FTP> open host
FTP> user   [username]
FTP> get fileremoto filelocale
FTP> put filelocale filenodoremoto
FTP> pwd
FTP> ascii
FTP> binary
```

4.1 modalita‘ trasferimento

Testo	ascii
uuencoded	ascii
binhex (mac .hqx)	ascii
postscript	ascii
tar	bin
compressi (.Z .zip .gz .z)	bin
eseguibili	bin
immagini (.gif .jpg)	bin

4.2 Ftp anonymous

Alcuni siti (FTP server) forniscono un servizio detto **ftp anonymous**

Questo consente un accesso *molto limitato* ad un'area pubblica di memoria.

```
$ ftp indirizzoserver

user:          anonymous

passwd:        nome@nodo.dominio
```

Solitamente l'accesso e‘ garantito solo per recuperare i files e non per depositarli.

L'accesso puo‘ essere ulteriormente regolamentato.

- README
- INDEX
- contents.index
- 00Contens
- /pub
- get file -

CAPITOLO 5

GOPHER

Il servizio gopher fornisce la possibilita‘ di accedere ai dati sulla rete internet attraverso un’ intererfaccia relativamente semplice e che prevede l’uso di risorse ’povere’.
Protocollo semplice che non carica la rete.

Le informazioni sono strutturate sotto la forma di menu successivi.

Grazie al gopher e‘ possibile accedere al servizio di ricerca per Titoli e Files, detto **Veronica**

Il client Gopher e‘ installato sul nodo ALE2TS della Lan locale e richiede le informazioni al server GOPHER della sissa.

```
                   Internet Gopher Information Client 2.0 pl6

                      Root gopher server: gopher1.sissa.it

 -->  1.  SISSA ISAS - INTNL SCHOOL FOR ADVANCED STUDIES - TS ITALY/
      2.  ----------------- Local Information --------------------.
      3.  SISSA infos/
      4.  Other Scientific Institutions in Trieste Area/
      5.  -------------------- E-prints ---------------------------.
      6.  Scientific Preprints/
      7.  ----------- Gopher Servers Infos and Services -----------.
      8.  USENET and Network News /
      9.  Other Gopher and Information Servers/
      10. Gopher's Services/




Press ? for Help, q to Quit, u to go up a menu                      Page: 1/1
```

```
                   Internet Gopher Information Client 2.0 pl6

                              Gopher's Services

      1.  Library On-Line Catalogs/
      2.  American English Dictionary <?>
 -->  3.  Searching gopherspace using Veronica <?>
      4.  Searching gopherdirs using Veronica <?>






Press ? for Help, q to Quit, u to go up a menu                     Page: 1/1
```

CAPITOLO 6

RICERCA NEL WWWSPACE

```
                      The Lycos Home Page: Hunting WWW Information (p1 of 3)

                             CMU LYCOS (TM)
                       THE CATALOG OF THE INTERNET

   A 1995 GNN Best of the Net Nominee. Over 20 million queries answered.

     _________________________________________________________________

     * Carnegie Mellon and CMG @Ventures to form Lycos, Inc..

     * Search the Lycos catalog with NlightN, for more details see the
       Carnegie Mellon press release.

     * Search the big Lycos catalog (4.24 million web pages), or if you
       have forms, use the Lycos search form to set the number of hits,
       and other options. You can also try the experimental Windows
       NT-based big Lycos catalog.

     * Search the small Lycos catalog (505k web pages), or if you have
       forms, use the Lycos search form to set the number of hits, and
-- press space for next page --
New mail on node VSTST0 from VSTST0::MRGATE::"VSTS03::C=GB::A= ::P=UK.AC::O=mail
base::web-support-request" M)ain screen Q)uit /=search [delete]=history list
```

```
                                                           Lycos search:  (p1 of 2)
                                         LYCOS SEARCH:

   Load average: 0.43: Lycos June 7, 1995 small catalog, 505748 unique
   URLs (see Lycos News)



   Enter one or more words or word prefixes in any order. The searcher
   will prefer documents that match more of your search terms, that match
   your term more closely, that have more occurrences of any one term,
   and occurences earlier in the document.

   Lycos Search Language description
   Form-based search with options (set number of hits, terse mode)
   Register your own URLs with Lycos or Delete your own URLs
   Lycos: Frequently Asked Questions
--More--  This is a searchable index.  Use 's' to search
New mail on node VSTST0 from VSTST0::MRGATE::"VSTS03::C=GB::A= ::P=UK.AC::O=mail
base::web-support-request" M)ain screen Q)uit /=search [delete]=history list
```

```
                                                     Lycos search:  (p1 of 2)
                                LYCOS SEARCH:

   Load average: 0.43: Lycos June 7, 1995 small catalog, 505748 unique
   URLs (see Lycos News)



   Enter one or more words or word prefixes in any order. The searcher
   will prefer documents that match more of your search terms, that match
   your term more closely, that have more occurrences of any one term,
   and occurences earlier in the document.

   Lycos Search Language description
   Form-based search with options (set number of hits, terse mode)
   Register your own URLs with Lycos or Delete your own URLs
   Lycos: Frequently Asked Questions
Enter a database search string: postscript
New mail on node VSTST0 from VSTST0::MRGATE::"VSTS03::C=GB::A= ::P=UK.AC::O=mail
base::web-support-request" M)ain screen Q)uit /=search [delete]=history list
```

```
                                      Lycos search: postscript (p1 of 16)
                          LYCOS SEARCH: POSTSCRIPT

   Load average: 0.29: Lycos June 7, 1995 small catalog, 505748 unique
   URLs (see Lycos News)



   Found 6901 documents matching at least one search term.
   Printing only the first 10 of 6901 documents with at least scores of
   0.010.

   Matching words (number of documents): postscript (6851), postscript2
   (1), postscriptbutton (1), postscriptdatei (5), postscriptdevice (1),
   postscriptdoc (1), postscriptdrucker (2), postscripted (2),
   postscriptem (1), postscriptf (1), postscriptfil (4), postscriptfile
--More--  This is a searchable index.  Use 's' to search
New mail on node VSTST0 from VSTST0::MRGATE::"VSTS03::C=GB::A= ::P=UK.AC::O=mail
base::web-support-request" M)ain screen Q)uit /=search [delete]=history list
```

CAPITOLO 7

URL

In Web l'URL ovvero Uniform Resource Locator e' l'indirizzo che consente di identificare e recuperare un documento su Internet. Un URL si presenta costituito da tre parti:

- protocollo (http, ftp, gopher, telnet...)
- riferimento alla macchina, all' area (dominio) internet a cui appartiene e al porto su cui opera il protocollo in questione.
- puntatore ai documenti detto *path*

protocollo://macchina.dominio:porto/path

```
http://www-dft.ts.infn.it:6163/~alex/me.html
----   ****************** ----**************
 |              |          |        |
 |              |          |        |
prot.     ind.macchina   porto     path
```

Ecco alcuni altri esempi di URL:

```
file://localhost/eprint/

gopher://gopher.sunet.se/

ftp://nic.umass.edu/pub/ednet

tn3270://auducacd.duc.auburn.edu/

telnet://hytelnet@access.usask.ca:23/
```

Ogni Url richiama una ed una sola pagina. Questo non impedisce che nel tempo la pagina venga modificata, sostituita, o persino eliminata.