Methods for access to databases through Internet

Introduction

A growing number of methods are available to provide (end-) users access to the contents of databases

• through the international Internet
  or
• through an institutional intranet
Methods for access to databases through Internet

Using telnet

Using telnet: steps involved

The end-user applies a client program that supports the telnet protocol to connect to a telnet server that provides access to some program on a server computer to retrieve data from the database.
Using telnet: scheme

Using telnet: example

Using the Windows 95 telnet client program
Using telnet: comments (Part 1)

• + Advantage:
  » Relatively cheap, old client workstations can be used; in other words, more expensive, recent and powerful workstations are not required.

Using telnet: comments (Part 2)

• - Disadvantages:
  » The user interface in telnet is character-based only, provides not some graphical user interface that most users prefer.
  » Starting the telnet client from the Netscape web browser requires some installation.
  » A telnet client provides only a general user interface. Besides this the database server has to provide additional components to the user interface for retrieval.
Using telnet: comments (Part 3)

- Using telnet is considered as obsolete by many, due to the popularity of the web-approach; telnet access is replaced by access through web browsers / clients.
- Nevertheless, it is still applied.

Methods for access to databases through Internet

Using proprietary client software
Using proprietary client software:
steps involved

The end-user applies not an open standard client program, but a proprietary client program that is specifically developed to retrieve data through the Internet from the database on a server computer. (Such a retrieval client program is developed in most cases by or for the distributor of the database.)
Using proprietary client software: examples

• SPIRS
  = client software for ERL,
  by SilverPlatter

• Probase
  = client software for Data-Star

Using proprietary client software: comments (Part 1)

• + Advantages:
  » The proprietary client program for retrieval is designed specifically for the target databases, so that it can provide a more suitable user interface than a more general, standard client program.
  » Some metadata related to the specific databases on the server can reside on the client together with the client software, so that the end-user can access these offline, for instance to prepare a search action or to select a particular database from a set of databases that are accessible.
Using proprietary client software: comments (Part 2)

- Disadvantages:
  » This approach requires installation, maintenance and learning of specific software on each client computer; that costs time and disk space.
  » Using a graphical interface as well as metadata on the client system requires a relatively powerful client workstation.

Using proprietary client software: comments (Part 3)

- This approach is obsolete due to the popularity of the web approach; it is more and more replaced by access through “universal”, general web clients.
Methods for access to databases through Internet

Using the web and a database server

Using WWW and a database server: steps involved

The end-user applies a general, “universal” web client program that supports by definition the hypertext transfer protocol (http) to connect to a web (http) server that provides access to some program on a server computer to retrieve data from the database.
Using WWW and a database server: scheme

![Diagram showing the scheme]

Using WWW and a database server: example

![Example of a database search interface]
Using WWW and a database server: comments (Part 1)

- CGI = Common Gateway Interface = a standard
- **+ Advantages:**
  - Requires only common, well-known, “universal” WWW browser software on the client computers.
  - A graphical interface is possible with a graphical web client program.

Using WWW and a database server: comments (Part 2)

- **- Disadvantages:**
  - Using a graphical interface in the web client requires a relatively powerful client workstation.
  - A web client provides only a general user interface. Besides this the database server has to provide additional components to the user interface for retrieval.
  - Requires not only a web server, but also a database server.
Using WWW and a database server: comments (Part 3)

• An increasing number of classical, character-based commercial, public access database hosts offer access in this way:
  Examples:
  » Cambridge Scientific, Data-Star, Dialog, Ovid, SilverPlatter, ….
• This approach is very common since the end of the 1990s.

Using WWW and a database server: extensions

• In more complicated systems, a
  » Java applet, or
  » script (JavaScript, VBScript, …)
may be sent from the WWW server to the client computer to support further interactions between the client computer and the database server (even besides HTTP).
• Example: access to databases distributed by Ovid and SilverPlatter
Methods for access to databases through Internet

Using Z39.50 protocol

Using Z39.50 directly from the client computer: steps involved

The end-user applies directly on the client workstation one of the available client programs that function according to the open protocol Z39.50, that is developed to retrieve data through the Internet from one Z39.50 compliant database server computers, or even from several Z39.50 compliant database server computers at the same time.
Using Z39.50 directly from the client computer: scheme

Using Z39.50 directly from the client computer: comments (Part 1)

- This approach is mainly used in some libraries for access to catalogues and other databases.

- + Advantage:
  » More than one database can be accessed at the same time.  
  This allows time savings when searching for rare records.
Using Z39.50 directly from the client computer: comments

- Disadvantages:
  » Requires installation, maintenance, and learning of Z39.50 client software on the client computers; that costs time and disk space.
  » Not all databases can be accessed using the Z39.50 protocol.
- The approach is not common, in spite of the sophistication and the suitability for retrieval purposes.

Using Z39.50 through a WWW gateway: scheme
Using Z39.50 through a WWW gateway: comments

- Applied to search the Library of Congress catalogue and many other bibliographic databases.
- + Advantages:
  » Does not require installation, maintenance, and learning of Z39.50 client software on the client computers.
  » Requires only common, “universal” WWW browser software on the client computers.
- Some WWW servers provide a gateway that offers access to more than one Z39.50 server and their databases at the same time.

**Example**

Using Z39.50 through a WWW gateway: example
Using Z39.50 through a WWW gateway: example

• The EEVL subject directory on the WWW that is specialised in engineering includes since 2005 EEVL Xtra, which allows simultaneous searching of open access databases in the field of engineering.
Methods for access to databases through Internet

Using Windows server software as terminal server

Using ICA with database(s) on a Windows server: scheme
Using ICA with database(s) on a Windows server: comments

- ICA = Intelligent Console Architecture network protocol
- Back to the mainframe?
- Not common

Using ICA with database(s) on a Windows server: comments (2)

- + Advantage:
  » Can be used for instance for in-house access to non-client-server retrieval software and databases running on a Windows server
  (examples: Chemical Abstracts on CD-ROM,…).

- -Disadvantages:
  » Requires a powerful server “to do the work”
  » Requires additional payment for terminal server software
Using ICA with database(s) on a Windows server: advantages

+ Advantages:
  » Less maintenance and control required of client-pc park in an organisation.
  » Minimal “thin” client hardware and software is sufficient to implement this system.
  » Can be used for instance for in-house access to non-client-server retrieval software and databases running on a Windows server (examples: Chemical Abstracts on CD-ROM,…).

Using ICA with database(s) on a Windows server: disadvantages

- Disadvantages:
  » Requires a powerful server “to do the work”.
  » Requires additional payment for terminal server software.
  » Requires additional expertise among personnel on the implementation and maintenance of this kind of terminal server systems.
  » Multimedia applications may not work as well as by using a more “thick” client.
Methods for access to databases through Internet

Using point and click interfaces

Using “special” methods for Internet database access

- Interactive atlases.
  Geographical Information Systems (GIS).

- “Select, point and click” user interfaces to “databases”
Using “special” methods: example

Result of a “search”:
Using “special” methods: example

Methods for access to databases through Internet

Using a priori export of a database to HTML pages
Using a priori export of a database to HTML pages: scheme

- Advantages
  » Only a universal WWW browser is required at the client side.
  » Useful to access records from different databases merged together, through one interface at the same time.
  » A simple solution for small and simple databases.
Using a priori export of a database to HTML pages: disadvantages

- Disadvantages:
  » The database is not searchable anymore
    = individual records cannot be retrieved.
  » The data in the HTML format is not always as up to date as the original database.
  » Not suitable for large and complicated databases.
  » Not suitable for transaction processing.

Using a priori export of records to HTML pages: scheme
Using a priori export of records to HTML pages: advantages

- +Advantages
  » The database is still searchable = records can be retrieved.
  » Only a universal WWW browser is required at the client side.
  » Useful to access records from different databases merged together, through one interface at the same time.
  » HTML/WWW retrieval programs are familiar to many users.

Using a priori export of records to HTML pages: disadvantages

- - Disadvantages:
  » Each database record becomes an HTML page/file.
  » Consumes more disk space than the original database(s).
  » An HTML/WWW index and retrieval program is required (such as the package used by AltaVista).
Methods for access to databases through Internet

Using the protocol for metadata harvesting of the Open Archives Initiative

Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH)
!! Task - Assignment !!

Read:
Breeding, Marshall
Understanding the protocol for metadata harvesting of the Open Archives Initiative.

!! Task - Assignment !!

Read:
Open Archives Forum
OAI for Beginners
- the Open Archives Forum online tutorial. [online]
1. OAI for Beginners: Overview
2. History and development of OAI-PMH
?? Question ??

In the case that databases are produced where Internet bandwidth is low, then which approach is probably more appropriate to provide simultaneous access to more than one databases:

- Z39.50 or
- Open Archives Initiative
- protocol for metadata harvesting?
  Why?

?? Question ??

Explain advantages and disadvantages of the protocols/methods to provide access to more than one database in one action:

- Z39.50 versus
- Open Archives Initiative
- protocol for metadata harvesting?
• You are free to copy, distribute, display this work under the following conditions:
  » Attribution:
    You must mention the author.
  » Noncommercial:
    You may not use this work for commercial purposes.
  » No Derivative Works:
    You may not change, modify, alter, transform, or build upon this work.

• For any reuse or distribution, you must make clear to others the license terms of this work.