

## **B.11 Intranets and the internet**

- [B.11 Intranets and the internet](#)
- [B.11.1 Intranets](#)
- [B.11.2 The internet](#)
- [Getting connected](#)
- [Firewalls](#)
- [Web space](#)
- [Useful websites](#)
- [B.11.3 Metadata for electronic resources curated by HERs](#)
- [Useful websites](#)
- [Panel 7: Elements of the Dublin Core](#)
- [B.11.4 What is Z39.50?](#)
- [Applications of Z39.50 involving the UK](#)
- [What does Z39.50 do?](#)
- [How does it work?](#)
- [Implementing Z39.50](#)
- [XML](#)
- [The Future](#)
- [Useful websites](#)

## **B.11 Intranets and the internet#**

### **B.11.1 Intranets#**

Intranets are closed networks that are established by organisations to serve the computing needs of their staff. An intranet may be small, for example a number of computers connected to a Local Area Network (LAN) within a building, but can be very large, for example the computing networks for a series of buildings connected to a Wide Area Network (WAN). In a WAN the buildings may be widely separated, as is the case in organisations with regional offices, such as English Heritage or multinational companies. The technology that links the computers is similar to that used in the internet. The difference is that use of an intranet is restricted to those with security clearance and a valid password. Intranets are enclosed inside a firewall to secure the information held on corporate systems from unauthorised access.

### **B.11.2 The internet#**

The internet is a shortening of 'Inter Networking' and offers a very different user environment consisting of millions of computers linked into a global communications network. Users include both private individuals and organisations from across the world all connected to the internet through an Internet Service Provider (ISP) or by establishing a web server. ISPs maintain web servers and other communications equipment enabling their clients to access a number of services from the internet. Access to the World Wide Web and electronic mail or 'e-mail' are normally considered as the most important internet services.

### **Getting connected#**

Many HERs are already linked into an intranet, for example the network run by their local authority. Some HERs run their own small networks with a database server allowing access to a small number of simultaneous authorised users. A number of HERs have access to the internet either through a network connection provided by their organisation or by establishing a direct connection from a PC using a modem and an ISP.

## Firewalls#

When intranets are connected into the global internet a protective security screen or firewall is normally established to control access to locally held resources from unauthorised outside users. The level of access to internet services allowed to individual users may also be controlled. This is generally achieved by setting up a proxy server that deals with internet requests from users connected to the internal network. Such a set up generally protects individual users but mobile computing is increasingly adding to security problems in that machines such as laptops and notepads may be used both within and outside of an organisational firewall with the possibility of introducing ? infections? into a system. A personal firewall on such machines is increasingly important particularly in light of the growing number of web based attacks (see [B.8.2](#)).

## Web space#

Access to the internet either through a corporate web server, co-ordinated by a web master, or from an ISP gives access to web space. This is space on the web server that may be used for file storage and for a website to advertise the HER. Websites have an address or URL which locates a hypertext file on the World Wide Web. Hypertext files can incorporate text, images and other digital data that are linked to the intranet. Websites may simply display information about an HER and its services with links to other websites of interest and facilities to allow users to record comments or send emails to the HER. Advanced sites can enable on-line searching of databases, incorporate video, sound and virtual-reality elements and many other features. HERs are increasingly presenting subsets of their data online as part of the drive towards greater accessibility. (See also [F.8](#))

## Useful websites#

Brief history of the internet <http://www.isoc.org/internet/history/brief.shtml>

Firewalls: [http://searchsecurity.techtarget.com/sDefinition/0,,sid14\\_gci212125,00.html](http://searchsecurity.techtarget.com/sDefinition/0,,sid14_gci212125,00.html)

Intranets: [http://searchwebservices.techtarget.com/sDefinition/0,,sid26\\_gci212377,00.html](http://searchwebservices.techtarget.com/sDefinition/0,,sid26_gci212377,00.html)

Proxy servers: [http://searchsmallbizit.techtarget.com/sDefinition/0,,sid44\\_gci212840,00.html](http://searchsmallbizit.techtarget.com/sDefinition/0,,sid44_gci212840,00.html)

World Wide Web (W3C) Consortium: <http://www.w3.org/>

Site design: Ability Net: <http://www.ability.org.uk/>

Site design: RNIB: [http://www.rnib.org.uk/professionals/webaccessibility/designbuild/Pages/design\\_build.aspx](http://www.rnib.org.uk/professionals/webaccessibility/designbuild/Pages/design_build.aspx)

Web Accessibility initiative: <http://www.w3.org/WAI>

## B.11.3 Metadata for electronic resources curated by HERs#

HERs hold a variety of electronic resources including databases, text files, spreadsheets, CAD files and GIS. Metadata should be recorded for each of those resources to provide a catalogue to the HERs' holdings as an aid to resource discovery. HERs are recommended to follow the Dublin Core for electronic resources other than GIS (see [panel 6](#)), which provides a standard content-description model widely used on the internet (see [Miller and Greenstein 1997](#)). HERs might want to consult the e-Gov metadata standard (eGMS) which adds several elements to the core Dublin Core schema. The UK GEMINI (Geo-spatial Metadata Interoperability Initiative) standard has replaced the NGDF standard previously recommended. UK GEMINI was developed through a collaboration of the Association for Geographic Information (AGI) and the Cabinet Office e-Government Unit, with additional representation from national and local government, and the academic community.

## Useful websites#

Dublin Core <http://dublincore.org/>

e-Gov <http://www.esd.org.uk/standards/egms/>

Metadata [http://guides.archaeologydataservice.ac.uk/g2gp/CreateData\\_1-2](http://guides.archaeologydataservice.ac.uk/g2gp/CreateData_1-2), <http://www.ukoln.ac.uk/metadata/>

UK GEMINI <http://www.agi.org.uk/uk-gemini/>

## Panel 7: Elements of the Dublin Core#

Element	Description
Title	The name given to a resource by its creator.
Creator	Person(s) or organisation(s) involved in creating the original resource, for example the Environment Department might be responsible for creating a corporate GIS layer of SSSIs. As an aid to resource discovery it may be helpful to record both named individuals and their organisations, for example Fred Bloggs and Sam Spade of English Heritage might be involved in creating a CAD file.
Subject	The topic of the resource, in keywords or phrases that describe the subject or content of the resource.
Description	A text description of the contents of the resource, including abstracts in the case of document-like objects or content descriptions in the case of visual resources.
Publisher	The publisher, distributor, department or other entity(ies) who are responsible for making the resource available, for example the Archaeology Data Service might provide access to a database created by Westshire County Council.
Date	Dates when the resource was created, published, released, maintained or updated. A number of different types of dates may need to be recorded. Not to be confused with the date of the contents of the resource.
Type	The general type of the resource, such as text, image, database etc.
Format	The format in which the data is represented, for example text, HTML, ASCII, executable application, JPEG image etc. The size of the resource and the medium on which it is held may also be recorded.
Identifier	A unique identifier for the resource, for example URLs, ISBN number or internal identification number.
Source	The work from which the resource is derived, for example the source of an HTML page might be a paper volume.

Language	The language(s) of the intellectual content of the resource.
Relation	The relationship to other resources, for example images in a document.
Coverage	The spatial and temporal coverage of the information contained in the resource.
Rights	Statement of copyright or other rights relating to the resource.

### **B.11.4 What is Z39.50?#**

Z39.50 is the name of a computer-to-computer communications protocol that is promoted by the National Information Standards Organization (NISO), an American National Standards Institute (ANSI) standards developer that serves the library, information, and publishing communities. Z39.50 is being implemented worldwide because it makes it easier for users to search and retrieve information simultaneously from groups of related databases on the internet.

#### **Applications of Z39.50 involving the UK#**

Much of the development of Z39.50 took place in the library sector, as a way of simultaneously interrogating the catalogues maintained in different libraries. The recent development of associated standards such as the Bath and CIMI profiles have made it easier to employ Z39.50 outside of library cataloguing. These provide support for more generalised record syntaxes such as Dublin Core and enable users to ask questions such as what, when, where or to conduct queries based on coordinate systems; undoubtedly more suited for those with an interest in Heritage. Suitability led the HEIRNET (Historic Environment Information Records Network) consortium to develop HEIRPORT; a Z39.50 system for searching geographically distributed databases of a number of heritage organisations including:

- English Heritage
- RCAHMS NMRS
- Durham SMR
- Portable Antiquities Scheme
- HEIRNET Register
- SCRAN
- ADS Catalogue

Using a standard web browser, a remote user can access the HEIRPORT gateway to interrogate its various targets. The results are then passed back to the user's browser through the origin, where they are displayed in HTML.

Recently a European Z39.50 gateway was launched as part of the [ARENA](#) (Archaeological Records of Europe - Networked Access) project. Partners currently include:

- Denmark: National Agency for Cultural Heritage
- Iceland: Fornleifastofnun Íslands, The Institute of Archaeology
- Norway: The Museum Project, Oslo
- Poland: Poznan Archaeological Museum
- Romania: cIMeC, Institute for Cultural Memory
- United Kingdom: Archaeology Data Service

English Heritage, NMRs and HERs are also contributing to these projects through the provision of subsets of their data to the ADS.

## **What does Z39.50 do?#**

Z39.50 supports information retrieval in a client and server environment where a computer operating as a client submits a search request (that is, a query) to another computer acting as an information server. Software on the server performs a search on one or more databases and passes a result set of records to the client for processing.

The power of Z39.50 is that it separates the user interface on the client side from the information servers, search engines, and databases. It provides a consistent view of information from a wide variety of sources, offering the ability to integrate information from different databases without imposing rules on the structure of those databases. Z39.50 can be implemented on any platform and enables different computer systems (with different operating systems, hardware, search engines, database management systems) to inter-operate and work together seamlessly.

This means that if an HER database is mounted on the internet using Z39.50, it can be searched from both the HER website and any number of Z39.50 gateways without the necessity for any changes to its structure providing they support shared standards such as CIMI and Bath profile.

## **How does it work?#**

Z39.50 standardises the way in which the client and the server communicate with each other, specifies procedures and provides a common language for the activities of both. In information retrieval, a series of messages passes between the client (referred to as the 'origin') and server (referred to as the 'target') to establish a connection, initiate a Z39.50 session, and negotiate expectations and limitations on the activities that will occur (such as the maximum size of the records). After this, the client may submit a generic search request (and to any number of usually remote targets). The request specifies query type and query term, for example, undertake a 'what' search for 'round barrow'. Configuration files map 'what' to the local equivalent which might be a 'monument\_type' field in a 'monument' table and defines a local query. The server executes the search against one or more databases, and creates a result set. The client can then ask for records from the result set and, upon receipt, display records to the user. An advantage of Z39.50 is that each database can have unique characteristics, for example differences in the way they store data, in the access points available for searching and in the software platform on which they reside. The objective of Z39.50 is to support computer-to-computer communication in standard and mutually understandable terms and support the transfer of data between the systems independent of the structure, content, or format of the data in a particular system. In short Z39.50 is a tool for virtual searching.

## **Implementing Z39.50#**

Implementing Z39.50 has hitherto been complicated, but it is getting progressively simpler as the standards have developed. Two things are essential:

- First, the appropriate software has to be installed to handle communications between one computer and another.
- Second, the data within the target's databases have to be mapped on to the metadata standard(s) defined in shared profiles. Mapping to metadata allows for a description of the data to be passed from client to server in a format that both can manage. Information can be passed from target to origin in a number of standard formats.

Advice on implementing Z39.50 is available from both the ADS and the mda.

## **XML#**

XML (Extensible Markup Language) is a standard for creating markup languages which describe the structure of data. It is not a fixed set of elements like HTML, but rather, it is like SGML (Standard Generalised Markup Language) in that it is a metalanguage, or a language for describing languages. XML enables authors to define their own tags. XML is a formal specification of the World Wide Web Consortium. (More detailed definition is available from: [http://searchwebservicestechtarget.com/sDefinition/0,,sid26\\_gci213404,00.html](http://searchwebservicestechtarget.com/sDefinition/0,,sid26_gci213404,00.html))

The e-Government Interoperability Framework (e-GIF) mandates the adoption of XML and the development of XML schemas as the cornerstone of the government interoperability and integration strategy. A key element in the development of XML schemas is an agreed set of data standards. (<http://webarchive.nationalarchives.gov.uk/+http://www.cabinetoffice.gov.uk/govtalk/schemasstandards/e-gif.aspx>).

*Useful web sites:* <http://www.w3.org/XML/>

XML UK: the United Kingdom Forum for Structured Information Standards <http://www.xmluk.org/>

## **The Future#**

In the past systems such as HEIRPORT have been described as portals. Today, however, a number of standards bodies have begun work on developing portal standards. Significant amongst these is Java Specification Request (JSR) 168 which defines a set of Java Application Program Interfaces for developing portlets which can be used in compliant portals. Portlets provide ?Web Services?. Portlets can be designed to interface with remote Z39.50 targets or indeed just about anything that is network accessible. Users can ?pick and mix? which portlets and thus services they wish to use within a personalised portal environment. Thus someone living in the Northeast might activate a Durham HER portlet, a local weather portlet and local newsfeed portlet within their portal environment. The CREE (Contextual Resource Evaluation Environment) consortium is looking at providing portlet interaction with various services including HEIRPORT targets within a formal (JSR-168 compliant) portal environment.

## **Useful websites#**

ARENA: <http://ads.ahds.ac.uk/arena/>

CREE: <http://www.hull.ac.uk/cree/>

HEIRNET: <http://www.britarch.ac.uk/HEIRNET/>

HEIRNET register: <http://ads.ahds.ac.uk/heirnet/index.cfm>

Web Portals: <http://www.w3.org/2002/07/DIAT/posn/sun-portal.html>

Z39.50 for all: <http://www.ariadne.ac.uk/issue21/z3950>