Web-based Clinical Information Systems

James Harrison, M.D., Ph.D.
Associate Director of Pathology Informatics
Department of Pathology
University of Pittsburgh
jhrsn@pitt.edu
jhh.cbmi.upmc.edu

NAACCR Annual Meeting
Informatics Workshop
Toronto, Canada
June 9, 2002
Agenda

- Building blocks of the Web: Web addresses, HTTP, Cookies, HTML, browser differences
- Page design comments
- Introducing XML
- XML and HL7
- Security and encryption
- Structure of Web-based clinical systems
  - Pitt: MARS and Special Vision
Application: Public Information and Marketing

- University of Pittsburgh Cancer Center
- Texas Cancer Registry
- NCI SEER Program
- OncoNavigator (Registries)
Application: Data Delivery

- NCI SEER Program
  - Overview
  - Incidence statistics
  - Graphical output
- Cancer Surveillance Online
  - Overview
  - Incidence by province
- NAACCR
  - CiNA
  - Data selection and specification
  - Graphical output
- Delivery of pre-set reports and data files for analysis
Application: Data Abstraction

- Belgian Thyroid Cancer Registry
  - Task Selection
  - Forms Index
Building Blocks of the World Wide Web

- A global network of servers that communicate via HTTP
- Transmit documents encoded in HTML
- HTML browsers on client machines

Independent request-response paradigm
"Connectionless," stateless client-server
Uniform Resource Locators (URL/URI)

A standard nomenclature for Internet addresses

http://path.upmc.edu/divisions/informatics.html
http://path.upmc.edu/divisions/index.html
ftp://path.upmc.edu/resources/cases/caseList.doc

Protocol server directory(ies) document

gopher://
telnet://
mailto:
file:///
Hypertext Transport Protocol (HTTP)

Request
GET url HTTP/1.0 [POST, PUT]
Accept: text/plain
Accept: text/html
User-Agent: Mozilla/4.0
Cookie: cookie data, if any

Response
HTTP/1.0 200 Document follows
Date: Mon, 09 Oct 1995 17:59:15 GMT
Server: NCSA/1.4.2
Content-type: text/html
Last-modified: Sun, 08 Oct 1995 17:49:26 GMT
Content-length: 1772

<HTML>
... the document ...
</HTML>
Maintaining State on the Web Server

- Simulates sessions on a connectionless server
- Pass state data back to the user's browser after each connection
- "Cookies:" a standard way to store and retrieve state data as text strings within the user's browser
- Allows the user to maintain a history of interaction with the server
- Cookies can be stored between user sessions, but may then be lost or accessed by non-authorized users
World Wide Web Documents

- Integrated display of many data types (assembly of multiple files)
- Constructed using Hypertext Markup Language (HTML)
- Document-to-document linkage
- Point-and-click interface
- Connectivity to other systems at the server
A Hypertext Markup Language Document

<html>
<head>
<title>My Document</title>
</head>
<body>
<h1>This is an example HTML document</h1>
<ol>
<li>First paragraph</li>
<li>Second paragraph</li>
</ol>
<p>This is the first paragraph. It doesn't really say very much. Now is the time for all good men to come to the aid of their country.</p>
<p>This is the second paragraph. The quick brown fox jumped over the lazy dog.</p>
</body>
</html>

Text editors for simple documents
Web management software:
FrontPage, Dreamweaver, GoLive
Interactivity: HTML Web Forms

- Create interface elements to accept input
- Return input text to the server as "fieldname=value" pairs:

  > http://jhh.cbmi.upmc.edu/cgi-bin/kbm/admin.pl?Username="harrison" &Password="2bornot2b"&kb_id="1"

```html
<form action="/cgi-bin/kbm/admin.pl" name="theform" method=post>
  <b>Username:</b><input type=text name="Username"><p>
  <b>Password:</b><input type=password name="Password"><p>
  <b>Choose FAQ to view:</b><select name="kb_id">
    <option value="1" selected>Pathology FAQ
    <option value="2">Pharmacy FAQ
    <option value="3">Radiology FAQ
  </select><input type=submit value="Login">
</form>
```
HTML Document Design

- Documents can be longer than the screen (scrolling is OK), but the content of the document should be clear in the first screen
- Allow the user as much control as possible (fonts, etc.)
- Documents are ideally simple to use and information-rich
  - 1000 text characters ~ 1K
  - 0.25" square full color ~ 1K (will compress, but…)
  - Know your audience: images dramatically slow page access
  - Each file access (including image files) counts as a server "hit"
- Simple documents are compatible across more browsers, less expensive to create, and easier to maintain
- Web management software packages encourage complex pages
- Brochures and reference books are styled differently for good reason
Differences in Web Browsers

• Most browsers handle simple documents similarly
• Complex documents may appear different and may required multiple versions for different browsers
• Browsers may differ between developers, between versions with a single developer, and between platforms with a single developer
• Browsers may handle html errors differently:
  > CiNA: looks OK in Internet Explorer but errors in HTML are apparent in some other browsers
HTML as a Markup Language

- HTML is a lightweight, standard, simple way to specify document structure (and appearance)
- HTML tags are tightly tied to visual display
  > Tag function is hard-coded into HTML browsers
  > Complex formatting leads to complex HTML
  > Limited ability to create interactive user interfaces
  > Available tags are specified by the HTML DTD (cannot create new, special-purpose tags)
    > Cannot represent data structures from databases
- HTML documents are difficult to maintain in large numbers
  > Inefficient to index and search
- HTML does not provide control over data entry
  > Any text is OK
From HTML to XML...

- HTML originally specified structural components of documents
  - HTML has evolved to become a presentation syntax
  - HTML was created based on the SGML framework

- XML is a simplified markup language framework
  - Allows creation of special-purpose markup languages
  - Can represent a variety of data structures and semi-structured data
  - Arbitrary tag nesting, recursion and markup detail
  - Human-readable and machine readable
  - Expected to be useful for creation of vertical-market data-interchange standards
  - HL7 v3 uses XML-based messages; SOAP; GTDS in Germany
  - High quality software to send, receive and parse XML is available at low cost to both users and vendors
Document Markup -- HTML

- Markup is restricted to display information only.
- Clinical information must be contained in free text.
- Available tags are specified in a rigid HTML "DTD".

```html
<h2>Patient Information</h2>
<ul>
  <li><b>Name:</b> Henry Levin, the 7th</li>
  <li><b>MRN:</b> 123456789</li>
  <li><b>DOB:</b> May 13, 1923</li>
</ul>

<h2>Clinical Data</h2>
<p>History of smoking for 40 years.</p>

<h2>Procedure</h2>
<p>Chest X-ray</p>

<h2>Findings</h2>
<p>Comparison is made with a chest-x-ray ...</p>

<h2>Impressions</h2>
<p>RLL nodule, suggestive of malignancy. Compared with a prior CXR from 6 months ago, nodule size has increased.</p>

<h2>Recommendations</h2>
<p>I notified the ordering physician of this finding by phone.</p>

Radiology Report - Chest X-Ray

**Patient Information**
- **Name:** Henry Levin, the 7th
- **MRN:** 123456789
- **DOB:** May 13, 1923

**Clinical Data**
History of smoking for 40 years.

**Procedure**
Chest X-Ray

**Findings**
Comparison is made with a chest x-ray ...

**Impressions**
RLL nodule, suggestive of malignancy. Compared with a prior CXR from 6 months ago, nodule size has increased.

**Recommendations**
I notified the ordering physician of this finding by phone.
Document Markup -- XML

<RadiologyReport>
  <PatientInfo>
    <Name>Henry Levin, the 7th</Name>
    <MRN>123456789</MRN>
    <DOB>May 13, 1923</DOB>
  </PatientInfo>
</RadiologyReport>

<ClinicalData>History of smoking for 40 years.</ClinicalData>

<Procedure>Chest X-ray</Procedure>

<Findings>
  Comparison is made with a chest-x-ray ...
</Findings>

<Impressions>
  RLL nodule, suggestive of malignancy. Compared with prior CXR from 6 months ago, nodule size has increased.
</Impressions>

<Recommendations>
  I notified the ordering physician of this finding by phone.
</Recommendations>

Radiology Report - Chest X-Ray

Patient Information
  • Name: Henry Levin, the 7th
  • MRN: 123456789
  • DOB: May 13, 1923

Clinical Data
  History of smoking for 40 years.

Procedure
  Chest X-Ray

Findings
  Comparison is made with a chest x-ray ...

Impressions
  RLL nodule, suggestive of malignancy. Compared with a prior CXR from 6 months ago, nodule size has increased.

Recommendations
  I notified the ordering physician of this finding by phone.

• Markup can identify specific data elements.
• Markup is flexible to accommodate new data elements and additional metadata.
• Metadata can be contained within tags.
• New DTDs can be created.
<procedure cpt="1234">
  <pat_phys pnum="abcd">
    <firstName>Elmer</firstName>
    <lastName>Fudd</lastName>
    <degree>M.D.</degree>
  </pat_phys>
  <proc_name>Upper endoscopy of gizzard</proc_name>
  <proc_date>09/09/1999</proc_date>
  <location name="ER"/>
</procedure>
HL7 2.3 Message Format

Figure 7-4. OBR attributes

<table>
<thead>
<tr>
<th>SEQ</th>
<th>LEN</th>
<th>DT</th>
<th>OPT</th>
<th>RP/ #</th>
<th>TB#</th>
<th>ITEM #</th>
<th>ELEMENT NAME</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4</td>
<td>SI</td>
<td>C</td>
<td></td>
<td></td>
<td>00237</td>
<td>Set ID - OBR</td>
</tr>
<tr>
<td>2</td>
<td>22</td>
<td>El</td>
<td>C</td>
<td></td>
<td></td>
<td>00216</td>
<td>Placer Order Number</td>
</tr>
<tr>
<td>3</td>
<td>22</td>
<td>El</td>
<td>C</td>
<td></td>
<td></td>
<td>00217</td>
<td>Filler Order Number +</td>
</tr>
<tr>
<td>4</td>
<td>200</td>
<td>CE</td>
<td>R</td>
<td></td>
<td></td>
<td>00238</td>
<td>Universal Service ID</td>
</tr>
<tr>
<td>5</td>
<td>2</td>
<td>ID</td>
<td>X</td>
<td></td>
<td></td>
<td>00239</td>
<td>Priority</td>
</tr>
</tbody>
</table>

From Rishel, "HL7 Modeling & Methodology and Control-Query Comm
HL7 v3 Message Format

Tag names from HL7 RIM

From Rishel, "HL7 Modeling & Methodology and Control-Query Com
Security

• Secure Sockets Layer (SSL)
  > Public/private key encryption/digital signature
  > Digital certificates in browser and server, certificate brokers
  > Secures all protocols (HTTP, FTP, Telnet, etc.)
  > Note different needs for e-commerce and clinical systems

• Other issues: browser caching/history, automated logoff
Web "Front Ends" to Clinical Systems

- Common Gateway Interface (CGI) and related protocols (Perl, etc.)
- Java servlets
- Application/middleware toolkits (e.g., PHP, Zope, Cold Fusion, WebSphere, Weblogic)
Logical Design of Clinical Web Systems
"Three-tier" design

Presentation Layer
- Web browser
- Client software
- Java applet

Middleware
- Web server
- Application server
- Interface engine
- Business rules
- Database connectivity

Data Sources
- DB managers
- Legacy systems
- MARS system
Technical Structure of Web-based Clinical Systems

- Standard client-server hardware
- Shared network
- Familiar graphical interface
- Web forms provide limited interface capability
- Java running in browsers can provide more functionality
- "Just-in-time" software distribution
- Currently relatively slow performance
- Specialized "browsers" may also provide improved interface function

Workstations with standard browsers (provide application shell)
Review

- The Web provides a relatively low cost and capable information system front end that is broadly accessible
- The primary expenses in Web development are expertise and time
- Registries are currently using the Web for marketing and data distribution
- Simple, information-rich page design has merit
- Registry software is likely to include Web-based components, particularly as useful XML syntaxes for cancer registries develop
- Managing secure information on the Web requires care and expertise
- Web servers and associated software serve as middleware to mediate between users and clinical data management systems