



College of Information Studies

University of Maryland Hornbake Library Building College Park, MD 20742-4345

Content Management Systems

Week 5

LBSC 690

Information Technology

What's Wrong with the Web?

- HTML
 - Confounds structure and appearance (XML)
- HTTP
 - Can't recognize related transactions (Cookies)
- URL
 - Links breaks when you move a file (PURL)

What's a Document?

- Content
- Structure
- Appearance
- Behavior

History of Structured Documents

- Early standards were “typesetting languages”
 - NROFF, TeX, LaTeX, SGML
- HTML was developed for the Web
 - Too specialized for other uses
- Specialized standards met other needs
 - Change tracking in Word, annotating manuscripts, ...
- XML seeks to unify these threads
 - One standard format for printing, viewing, processing

eXtensible Markup Language (XML)

- SGML was too complex
- HTML was too simple
- Goals for XML
 - Easily adapted to specific tasks
 - Rendering Web pages
 - Encoding metadata
 - “Semantic Web”
 - Easily created
 - Easily processed
 - Easily read
 - Concise

Some XML Applications

- Text Encoding Initiative
 - For adding annotation to historical manuscripts
 - <http://www.tei-c.org/>
- Encoded Archival Description
 - To enhance automated processing of finding aids
 - <http://www.loc.gov/ead/>
- Metadata Encoding and Transmission Standard
 - Bundles descriptive and administrative metadata
 - <http://www.loc.gov/standards/mets/>

Really Simple Syndication (RSS)



```
<?xml version="1.0"?>
<rss version="2.0">
<channel>
  <title>Lift Off News</title>
  <link>http://liftoff.msfc.nasa.gov/</link>
  <description>Liftoff to Space Exploration.</description>
  <language>en-us</language>
  <pubDate>Tue, 10 Jun 2003 04:00:00 GMT</pubDate>
  <lastBuildDate>Tue, 10 Jun 2003 09:41:01 GMT</lastBuildDate>
  <docs>http://blogs.law.harvard.edu/tech/rss</docs>
  <generator>Weblog Editor 2.0</generator>
  <managingEditor>editor@example.com</managingEditor>
  <webMaster>webmaster@example.com</webMaster>
  <ttl>5</ttl>
  <item>
    <title>Star City</title>
    <link>http://liftoff.msfc.nasa.gov/news/2003/news-starcity.asp</link>
    <description>How do Americans get ready to work with Russians aboard the International Space Station? They take
      a crash course in culture, language and protocol at Russia's Star City.</description>
    <pubDate>Tue, 03 Jun 2003 09:39:21 GMT</pubDate>
    <guid>http://liftoff.msfc.nasa.gov/2003/06/03.html#item573</guid>
  </item>
</channel>
</rss>
```

See example at <http://www.nytimes.com/services/xml/rss/>

Atom Feeds

```
<?xml version="1.0" encoding="utf-8"?>
<feed xmlns="http://www.w3.org/2005/Atom">
  <title>Example Feed</title>
  <subtitle>A subtitle.</subtitle>
  <link href="http://example.org/feed/" rel="self"/>
  <link href="http://example.org/" />
  <updated>2003-12-13T18:30:02Z</updated>
  <author>
    <name>John Doe</name>
    <email>johndoe@example.com</email>
  </author>
  <id>urn:uuid:60a76c80-d399-11d9-b91C-0003939e0af6</id>
  <entry>
    <title>Atom-Powered Robots Run Amok</title>
    <link href="http://example.org/2003/12/13/atom03"/>
    <id>urn:uuid:1225c695-cfb8-4ebb-aaaa-80da344efa6a</id>
    <updated>2003-12-13T18:30:02Z</updated>
    <summary>Some text.</summary>
  </entry>
</feed>
```


XML: A Family of Standards

- Definition: DTD
 - Known types of entities with “labels”
 - Defines part-whole and is-a relationships
- Markup: XML
 - “Tags” regions of text with labels
- Markup: XLink
 - Defines “hypertext” (and other) link relationships
- Presentation: XSL
 - Specifies how each type of entity should be “rendered”

XHTML: Writing HTML as XML

```
<?xml version="1.0" encoding="iso-8859-1"?>
<html xmlns="http://www.w3.org/TR/xhtml1" >
<head>
  <title> Title of text XHTML Document </title>
</head>
<body>
<div class="myDiv">
  <h1> Heading of Page </h1>
  <p> here is a paragraph of text. I will include inside this paragraph
    a bunch of wonky text so that it looks fancy. </p>
  <p>Here is another paragraph with <em>inline emphasized</em>
    text, and <b> absolutely no</b> sense of humor. </p>
  <p>And another paragraph, this one with an  image, and a <br /> line break. </p>
</div>
</body></html>
```

Some Basic Rules for XML

- XML is case sensitive
- XML declaration is the first statement
 - `<?xml version="1.0"?>`
- An XML document is a “tree”
 - Must contain one root element
 - Other elements must be properly nested
- **All** start tags must have end tags
- Attribute values must have quotation marks
 - `<item id="33905">`
- Certain characters are “reserved”
 - For example: < is used to represent `<`

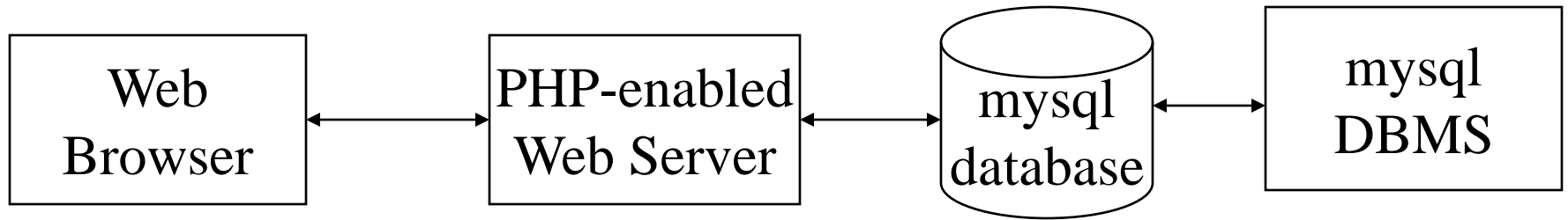
HTML Editors

- Goal is to create Web pages, not learn HTML!
- Several are available
 - Macromedia Dreamweaver available commercially
 - Microsoft Word (Page->Edit with Word in IE 7)
 - Many more options available on www.tucows.com
- Tend to use physical layout tags extensively
 - Detailed control can make hand-editing difficult
- You may still need to edit the HTML file
 - Some editors use browser-specific features
 - Some HTML features may be unavailable
 - File names may be butchered when you upload

HTML Validators

- Syntax checking: cross-browser compatibility
 - <http://validator.w3.org>
 - Try it on <http://www.umd.edu> ☹️
- Style checking: Section 508 compliance
 - <http://www.powermapper.com/products/sortsite/checks/accessibility-checks.htm>
 - Try it on <http://terpconnect.umd.edu/~oard> ☹️

Creating the “Deep Web”



Content Management Systems

- Separation of content and appearance
- Separation of roles
- Standard “design patterns”
 - Login and password recovery
 - Headlines and drill-down
 - Site map
 - Search
 - Shopping cart

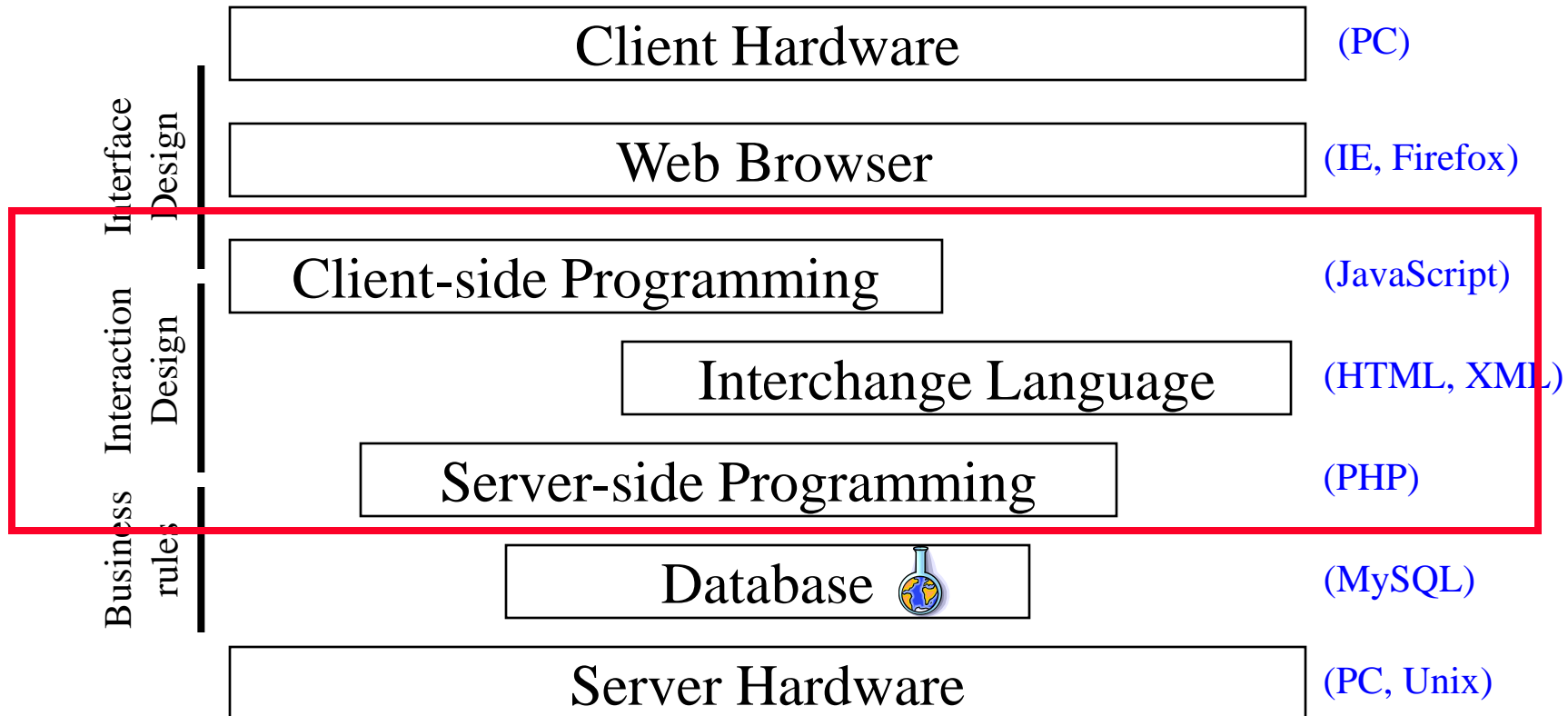
Roles

- Information architecture design
- Task assignment
- Content generation and modification
- Approval for “publication”
- Publication
- Error correction
- Tracking task progress

CMS Technology

- Database to store content
 - Also stores access control data and parameters
- PHP to control user experience
 - Reads database, generates HTML
 - “Canned” settings provide standard behaviors
- XHTML to convey user experience
- Allows limited interactivity
 - Most user actions require a server response
 - JavaScript may be used for form validation

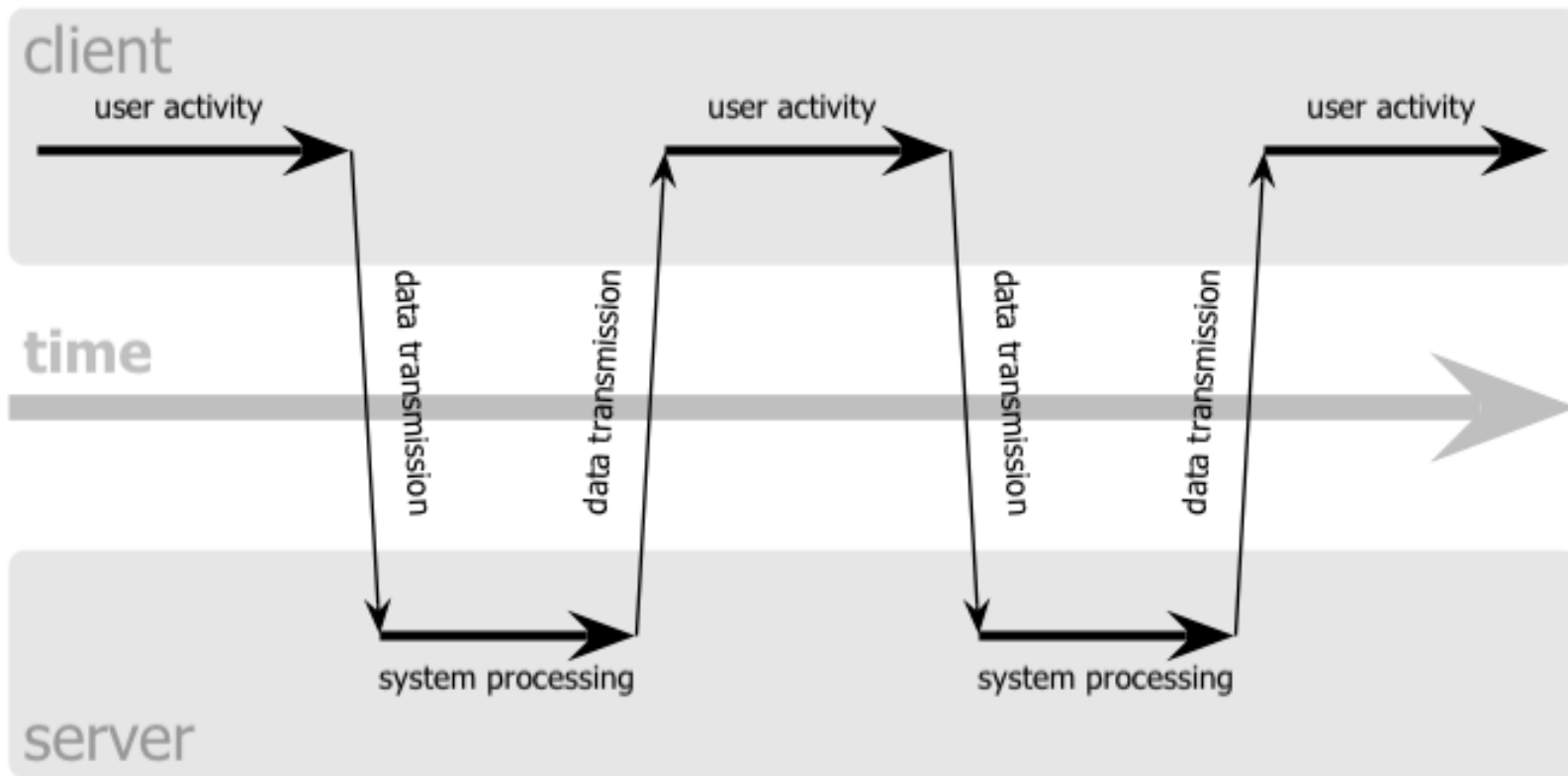
- Relational normalization
- Structured programming
 - Software patterns
- Object-oriented design
- Functional decomposition



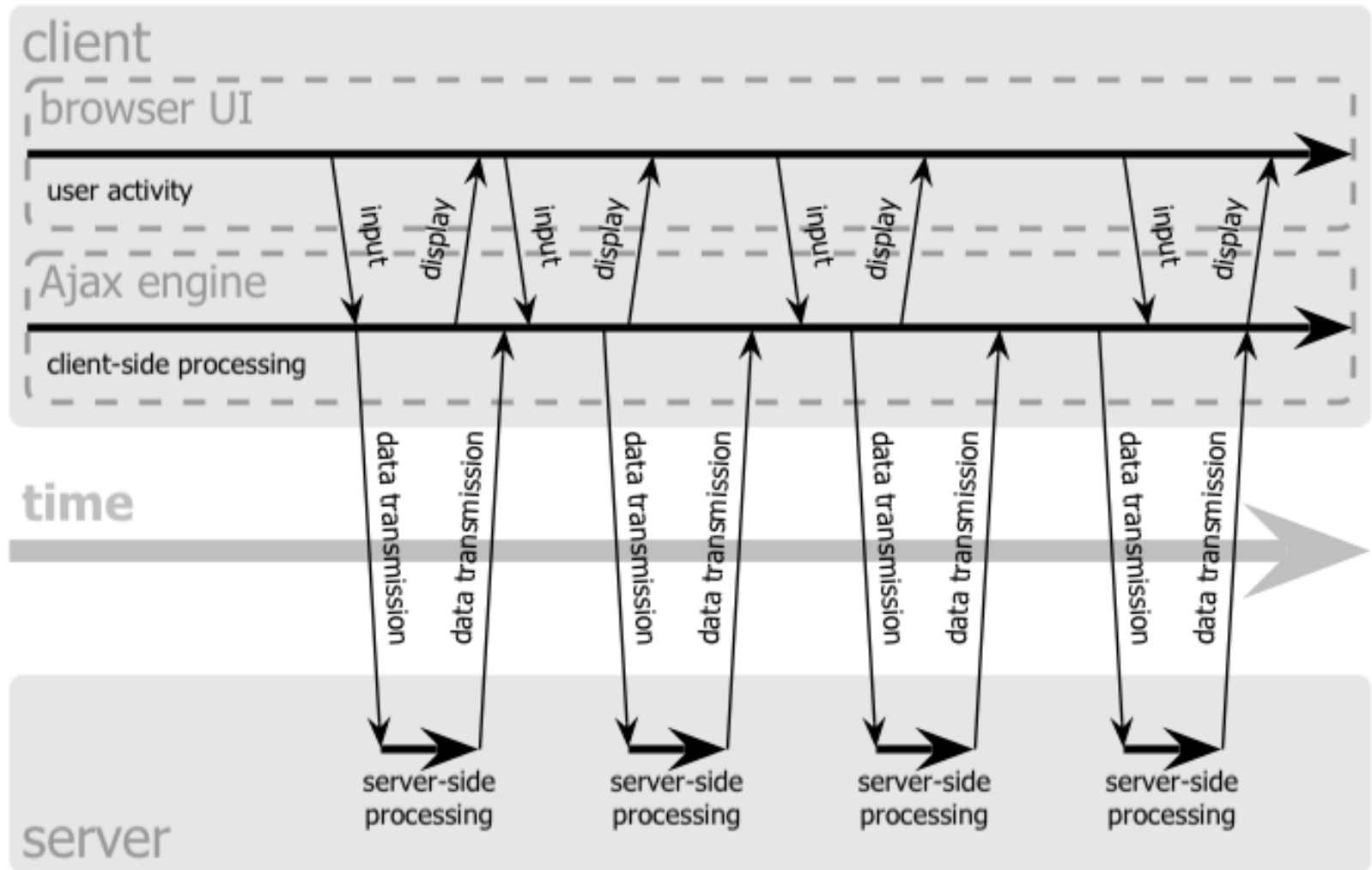
Ajax Applications

- Google Maps
 - <http://maps.google.com>
- Google Suggest
 - <http://www.google.com/webhp?complete=1&hl=en>
- Sajax Tables
 - <http://labs.revision10.com/?p=5>
- Sajax
 - <http://www.modernmethod.com/sajax/>

classic web application model (synchronous)



Ajax web application model (asynchronous)



Before You Go

On a sheet of paper, answer the following (ungraded) question (no names, please):

What was the muddiest point in today's class?