Internet Basics

- URL: Uniform Resource Locator
    - Scheme (HTTP:)
    - Double slash (optional but required for FTP and HTTP)
    - Domain name
    - Port number (optional)
    - Path (/pubs)
- HTTP: Hypertext Transfer Protocol (1.1, 2.0, 3.0)
- Web Server: Retrieves and Replies to HTTP requests, doesn't know specifics of a page
- HTML: Hypertext Markup Language
    - First founded in CERN by Tim Burners-Lee to distribute papers faster
- WWW Consortium

HTML Combined

- Tags are case insensitive, attributes may be sensitive,
- Comments <!-- and end with →
- and <ol> Ordered lists show Roman numerals, ul is bullet points
- <dl> definition list and <menu>
- <table>
    - <caption> to title the table
    - <th></th> and <td></td> for table header cell and table data cell
    - <tr> identifies container for a row of cells
- HTML uses a Universal Character Set
- Href: Hypertext reference
- <Link> only used in header, can be used to link other versions of the website
- Image maps are used so some parts of the picture are clickable <usemap> and <map>
- <meta name = "robots" content = "noindex, follow">

HTML Style Sheets

- CSS3 (Cascading Style Sheets)
    - Inline, <style> </style>, <link>
    - Classes are ., ID is #
- <span> is inline (Like only for one word for example) as opposed to <div>
- Precedence

- More specificity takes precedence
- DOCTYPE tells the browser what HTML version you are using
- Pseudo Elements
    - :link, :visited:, :hover,:active

```
pseudo classes
 - :first-child, Selects every <p> elements that is the first
   child of its parent
 - :hover, Selects links on mouse over
 - :active, selects the active link
 - :focus, selects the input element which has the focus
 - :lang, selects every <p> element with a lang attribute
pseudo elements
 - :first-line, add a special style to the first line of a text
 - :first-letter, add a special style to the first letter of a
   text
 - :before, to insert some content before the content of an
   element
 - :after, to insert some content after the content of an element
```

- Vendor prefixes are used when CSS features are not yet released globally yet

JS Basics

- Javascript can be minified (Don't care about spaces, indents, etc.)
- All numbers are treated as floating point
- Immutable strings
- Var is local to the function
- Let or const would make it global outside of functions
- Array in Javascript can be any type of anything

```
There are many ways to iterate over an array
 - for loop; for (i=0; i < len; i++) {. . . }
 - for in loop; for (x in person) { . . . }
 - while loop; while (condition) {. . . }
```

- Objects: Key value pairs, can be accessed using dot
- Alert, Confirm, and Prompt are built in Javascript for Popups

```
JavaScript now controlled by the ECMA standard body
ECMA stands for European Computer Manufacturers
Association
```

JSON

- Javascript Object Notation

XMLHttpRequest will be covered in more detail in the AJAX lecture

The XMLHttpRequest code:

```
var req = new XMLHttpRequest();
req.open("GET", "file.json", true); // "asynchronous" operation
req.onreadystatechange = myCode; // the callback
req.send(null);
```

The JavaScript callback: eval() parses JSON, creates an object and assigns it to variable doc

```
function myCode() {
  if (req.readyState == 4) {
    if (req.Status == 200) {
      var doc = eval('(' + req.responseText + ')');
    }
  }
}
```

Using the data:

```
var menuName = doc.getElementById('menu'); // finding a field menu
doc.menu.value = "my name is"; // assigning a value to the field
```

How to access data:

```
doc.commands[0].title // read value of the "title" field in the array
doc.commands[0].action // read value of the "action" field in the array
```

- Has same origin policy
    - Can use <script> to get around it
- Use JSON.parse instead of eval() due to security
- JSONP: Json with Padding

Python

- Strings are immutable
- .strftime() for displaying time

```
from flask import abort

#retrieve a task
@app.route('todo/api/v1.0/tasks/<int:task_id>', methods=['GET'])
def get_task(task_id):
    task = [task for task in tasks if task['id'] == task_id]
    if len(task) == 0:
        abort(404)
    return jsonify({'task': task[0]})
```

- Django

Dom

- Dom: Document Object Model
- XML Dom is made to be used with any language
- document.getElementByTagName("img")
- innerHTML
- .appendChild instead of += for innerHTML

Forms and CGI Mechanism

- \<textarea\>
- \<input\>

  ```
  SELECTED
    • Menu item is pre-selected in the list
  VALUE="text"
    • Text specifies the value to be sent to the
      script if the option is selected
    • By default, the text following the OPTION
      element is sent
  DISABLED
    • Specifies a "grayed", non-selectable item
  ```
- **HTML5 adds the REQUIRED attribute**
- Fieldset to group forms \<legend\>
- CGI basically enables scripts to create pages
  - CGI allows web servers to call external programs or scripts to handle dynamic content generation. This enables websites to respond to user

input, generate personalized content, and perform various other tasks beyond just serving static files. (Mostly PHP) invoked from anchor tags
- Environment Variables: Created by the web server and immediately set before gateway execution
    - Non-Request Specific
    - Request Specific
- Server Directive
    - Inform the server on the type of input
    - Content Type, Location, Status
- Show_vars.php

Web Server

Enables browser requests

Mainly provides

- Support for retrieving hypertext documents
- Manages access to the Web site
- Provides several mechanisms for executing server-side scripts
    - Common Gateway Interface (CGI)
    - Application Program Interface (API)
    - Direct Module Interfaces (SAPI)
- provides log files and usage statistics

- Multithreads or copies the server for other requests
- Does business logic and connects to databases
- Apache, NGINX
- Document tree root is given to the web server at the start
- Configuring a server
    - Location of the server (Server root)
    - Location of document to deliver (document root)
    - Location of CGI Scripts or server-side components to execute
- Host Filtering: Limit to limit directories

- Virtual hosting is a way of setting up your server so it can appear to be multiple Web sites at once
    - Address-based or name-based

JS Advanced

- Window (window.close(), window.open(), etc)
- Navigator: Built-in object that describes the browser
- addEventListener()
- Functions can be stored in variables
- PassPhrases
    - Encryption key that you memorize and use Diceware
- Strict mode

HTTP Protocol

- MIME: Multipurpose Internet Mail Extensions
    - Internet standard for email
- Intermediaries
    - Proxy
        - Caching Proxy
    - Gateway
        - Convert HTTP Request to another protocol
    - Tunnel
        - VPN
- TCP uses a persistent connection (takes a while)
- Entity Tag (Etag)
    - Makes caching more effective (fingerprinting)
- 400s Client Error, 500s Server Error
- HSTS (HTTP Strict Transport Security): Forces you to use HTTPS
- CORS: Allows resource sharing
- Ipconfig
- tracecert (finding a route)
- Ports
- NAT (Network Address Translation): Allows a computer to have a private IP address
- Subnets
    - Subnet mask
    - Takes few bits away to make this subnet

Assume a host has an IP address of **74.125.127.104** with a
subnet mask of **255.255.255.192** (also written as
74.125.127.104/26)

- What is the network address? (i.e., what class IP address is
  this?)
- What is the subnet mask in binary?
- How many hosts can be in the subnet?
- 74.125.127.104 = 01001010 01111101 01111111 01101000
- 255.255.255.192 = 11111111 11111111 11111111 11000000

Answers:

- **Network Address** – this is a Class A address, so the first 8
  bits are allocated for the network: **74.0.0.0**
- **Subnet Mask** in binary will have the first 26 bits as 1's with
  the last 6 bits as 0's
- **# Hosts** – the last 6 bits are reserved for hosts, giving 2^6-
  2=**62** host addresses for this subnet
  - Remember that the address with all 0s and all 1s are
    reserved

- To get the network/subnet address from an IP address and
  subnet mask, perform a **logical AND** operation between them

```
137.229.154.221     = 1000 1001  1110 0101  1001 1010  1101 1101
255.255.224.0       = 1111 1111  1111 1111  1110 0000  0000 0000
137.229.128.0       = 1000 1001  1110 0101  1000 0000  0000 0000
```

- The **network/subnet address** is 137.229.128.0
- This **subnet mask** provides **19** bits for the network address
  and 13 bits for the host address
- Another way we could have written the IP address with the
  subnet mask is 137.229.154.221/**19**

Private IPv4 addresses are in the following ranges
- **Class A** Private: **10.0.0.0 – 10.255.255.255**
  - 2^0=1 network with 2^{24} addresses
- **Class B** Private: **172.16.0.0 – 172.31.255.255**
  - 2^4=16 networks with 2^{16} addresses
- **Class C** Private: **192.168.0.0 – 192.168.255.255**
  - 2^8=256 networks with 2^8 hosts on each network

DHCP: Dynamic Host Control Protocol: Automatically assigns IPs to other computers in the network

Verge Paper

- Robots Exclusion Protocal
- Google Scrapes websites
- AI companies scrape but websites get no benefits