



Introduction to Internet Technology and Web Programming

Computer Science 103
Boston University
Vahid Azadeh-Ranjbar

These notes are based on the lecture notes provided by [Professor Susan Worst](#).



Lecture 13-2

BU CS 103, Spring 2019



Presentation Week

- you will go to two different classrooms.
- The classrooms are at Building (**CAS B18 & SCI 117**).
- There will be about 35 students in each classroom.
- There will be 17 or 18 presentations on Mon. and Wed in each class.
- The presenters have 3 minutes to present their websites.



How to present?

- Start with introducing yourself, your background, your knowledge for computer science before taking CS 103, and any other introduction.
- Then, talk about your domain name and the reason you picked it up.
- Introduce the name of your website and what the website is about.
- Show us the home page and let us know if you've made your website from scratch, by using CSS templates or it is a Wordpress Website.
- Talk about the layout of the page. Is it a one column, two column or three column layout? Maybe the home page is one column and the internal pages are multicolumn.
- Talk about the responsivity of your pages. How did you make them responsive?
- After showing us the homepage, show us two of your internal pages and explain them. Pick up the most interesting internal pages to show to your audience if you want to get a good grade from them.
- Then, talk about the interactive features that are used in your website. Did you use the presentations in Show and Tell Day or find the features online?
- At the end, show us the credits page and appreciate your references.



My Extended Office Hours

- ❖ Helping you for your final project website
 - Thursday Apr. 25: 10:30 AM - 12:30 PM & 3:00 – 5:00 PM
 - Friday Apr. 26: 2:00 – 4:00 PM
 - Tuesday Apr. 30: 10:00 – 12:00 AM & 3:00 - 5:00 PM
 - Wednesday May 1: 5:00 – 6:00 PM
- ❖ Helping you for your final exam
 - Thursday May 2: 10:00 AM - 12:00 PM & 2:00 - 5:00 PM
 - Friday May 3: 10:00 AM - 12:00 PM & 2:00 - 4:00 PM



TF/CA's Office Hours

- They will hold their regular office hours until Thursday May 2 (The last day of classes).
- They also hold office hours during the lab sessions (this Thursday and next Thursday) at KCB building.



What will be Covered Today?

- ❖ Web analytics (The last topic for CS103 in this semester ☺)
- ❖ Review for Final Exam



Web Analytics



What Is Web Analytics?

- ❖ **Web analytics** is the measurement, analysis and reporting of web data for purposes of optimizing web usage.

Source: WAA Standards Committee. "Web analytics definitions." Washington DC: Web Analytics Association (2008). Cited in *Wikipedia, The Free Encyclopedia*, 14 Oct. 2015. Retrieved 4 Nov. 2015.



Why Web Analytics?

- ❖ Web analytics answers questions like:
 - Who is coming to my website?
 - What pages are they looking at?
 - How long are they spending?
 - Can they find the information they need?



Analytics

❖ Two approaches to analytics:

1. Analyze **web server logs**
2. Use **Google Analytics** or **Open Web Analytics**



Web Server Logs

- ❖ Every time a visitor accesses any file on your website, the access is recorded in a **web server log**
- ❖ Most web servers record logs in a common format called the **Common Log Format** (to make it easier to analyze them).

Where Are Web Server Logs Stored?

- ❖ Web server logs are not stored in the same folder as web pages.
- ❖ They are **not accessible to the public**.
- ❖ Shared web hosting services like DreamHost often do not allow you to navigate to the logs directly
- ❖ However, DreamHost allows you to download them

Data Captured in Common Log Format

- ❖ **Remotehost** - hostname or IP address
- ❖ **Rfc931** – the remote logname of the user
- ❖ **Authuser** – the username as which the user has authenticated
- ❖ **[date]** – date and time of the request
- ❖ **Request** – the file being requested
- ❖ **Status** – the HTTP Status Code
- ❖ **Bytes** – the size of the file transferred

Source: Logging Control In W3C httpd

<http://www.w3.org/Daemon/User/Config/Logging.html#common-logfile-format>



Other Fields Usually Provided in Logs

- ❖ **User Agent** – what browser and version visiting is using
- ❖ **Referer** – The web page containing the link that visitor clicked on to get to this page



More about Remotehost

- ❖ **Remotehost** - hostname or IP address
- ❖ **Rfc931** – the remote logname of the user
- ❖ **Authuser** – the username as which the user has authenticated
- ❖ **[date]** – date and time of the request
- ❖ **Request** – the file being requested
- ❖ **Status** – the HTTP Status Code
- ❖ **Bytes** – the size of the file transferred




More about Remotehost

- ❖ **Remotehost** stores the IP address of the requesting computer
- ❖ Sometimes **Remotehost** shows a domain name instead of the IP address



IP Address

- ❖ The IP address provides information about the **geographic location** of the computer, and sometimes more
- ❖ Most networks (including BU) and Internet Service Providers (Comcast, RCN) will assign you a different IP address every time you connect. So your IP address generally can't be traced to you.



Your Current IP Address

- ❖ Find your current IP address at


<http://ipaddress.com/>

- ❖ Check it from home, or tomorrow, and it will probably be different



Decoding IP Addresses

- ❖ IP address ranges are assigned to different countries and organizations
- ❖ Look up an IP address:
<http://ipaddress.com/ip-lookup/>
- ❖ Example: 172.217.21.206
- ❖ By looking up IP addresses, you can tell where some of your visits are coming from.



More about the HTTP Status Code

- ❖ **Remotehost** - hostname or IP address
- ❖ **Rfc931** – the remote logname of the user
- ❖ **Authuser** – the username as which the user has authenticated
- ❖ **[date]** – date and time of the request
- ❖ **Request** – the file being requested
- ❖ **Status** – the **HTTP Status Code**
- ❖ **Bytes** – the size of the file transferred



More about the HTTP Status Code

- ❖ What is HTTP?
 - Hypertext Transfer Protocol
 - The protocol used to request and deliver information on the web
- ❖ The status code tells you whether the visitor's request was successful, or, if not, why it failed



Common HTTP Status Codes

❖ Common HTTP status codes include:

- 200 = OK (request has succeeded)
- 301 = moved permanently
- 401 = not authorized
- 403 = forbidden
- 404 = file not found
- 500 = internal server error

❖ Complete list from the W3C

<https://www.w3.org/Protocols/rfc2616/rfc2616-sec10.html>



Alternative for Web Analytics: Google Analytics

❖ Two approaches to analytics:

1. Analyze web server logs
2. Use **Google Analytics** or **Open Web Analytics**



Alternative for Web Analytics: Google Analytics

- ❖ Google Analytics is a free service run by Google
- ❖ User data is collected through a JavaScript snippet on the web page
- ❖ Data is sent to Google who then shares it with you



Google Analytics Pros and Cons

- ❖ Pros
 - Very robust data
 - Free
 - Very commonly used in the industry
- ❖ Cons
 - Compromises privacy of visitors to your site




Opting Out of Google Analytics

- ❖ Google provides a browser add-on that you can use to keep your own visits to websites that use Google Analytics from being tracked.

<https://tools.google.com/dlpage/gaoptout?hl=en>

- ❖ It's a good idea to provide this link to your visitors so that they can opt out



Getting Started with Google Analytics

1. Create an account at <http://www.google.com/analytics>
2. For HTML sites: add the Google Analytics JavaScript to every page of your website.
3. For WordPress sites: use a plugin, such as **Yoast**, to add the Google Analytics code to every page.

➤ **Counts as one interactive feature**



Introduction to Google Analytics

- ❖ See "How to Read Google Analytics Data: The Basics (2015)" for a good introduction

https://www.youtube.com/watch?v=xZbd_yEN5nE




Preparing for the Final Exam



What's on the Final Exam?

❖ Final exam is **not cumulative**.

1. Lectures 9-1 to 13-2 will be covered in the final exam.
2. Plus everything about HTML/CSS coding
3. And essentials that every web developer should know.



Material from Lectures 9-1 to 14-1 (The most important topics for the final exam)

1. PHP & WordPress
2. JavaScript
3. Search Engine Optimization
4. Web Analytics and Google Analytics
5. JavaScript/CSS/HTML coding

1. PHP & WordPress


Questions may include, but are not limited to, the following:

- ❖ Static web pages are written entirely in _____?
 - HTML
- ❖ To make a page dynamic, we need _____.
 - a Programming Language such as PHP or JavaScript
- ❖ JavaScript called a _____.
 - client-side programming language
- ❖ PHP is called a _____ programming language.
 - server-side
- ❖ WordPress which is called a "_____" is using PHP.
 - content management system (CMS)
- ❖ Server-side languages run on the _____.
 - web server (e.g., DreamHost's computers)


1. PHP & WordPress (Continue ...)

❖ Questions may include, but are not limited to, the following:

- ✓ What the acronym "PHP" stands for?
- ✓ Where PHP runs?
- ✓ How does the server tell that a file is PHP?
- ✓ How does the server identify PHP within the file?
- ✓ Advantages of PHP over JavaScript?
- ✓ What are the advantages and disadvantages of server-side programming languages?




According to the lecture, which is NOT one of the 3 most common Content Management Systems right now?

- A. Drupal
- B. Joomla
-  C. Squarespace
- D. WordPress



The second "P" in "PHP" stands for...

- A. Postprocessor
- B. Preformatter
-  C. Preprocessor
- D. Processor

2. JavaScript

Questions may include, but are not limited to, the following:

- Client-side programming languages Live on _____.
 - ✓ the user's computer (the "client")
- Client-side programming languages are Incorporated into _____.
 - ✓ web browsers
- Is **JavaScript** the only common Client-side programming language right now? True/False
 - ✓ True

2. JavaScript (Continue ...)

❖ Questions may include, but are not limited to, the following:

- ✓ Where JavaScript runs?
- ✓ How does the browser identify a JavaScript file?
- ✓ How does the browser identify JavaScript within an HTML file?
- ✓ Minification – what is it and why is it done?
- ✓ What are the advantages and disadvantages?

PHP vs. JavaScript coding

PHP coding

```
<?php  
.....  
?>
```

JavaScript Coding

```
<script>  
...  
</script>
```

What Can JavaScript Do?

- ❖ Check data submitted via forms
- ❖ Create web page widgets
- ❖ Create pop-up windows
- ❖ Make photos in slideshows move
- ❖ Update part of a page
- ❖ Send data from a web page to Google Analytics




In the Lightbox file "**lightbox.min.js**", "min" stands for

- A. Mini
- 😊 B. Minified
- C. Minimal
- D. Minimized



The **<script>** tag can be placed in the head or the body of an HTML page


- 😊 A. True
- B. False



The <script> tag is used to insert JavaScript code, while the <link> tag is used to reference an external .js file

A. True

😊 B. False



3. Search engine optimization

Questions may include, but are not limited to, the following:

- 1) How Do Search Engines Find a Web Page?
- 2) How Web Pages are indexed?
- 3) How Web Pages are ranked in a specific topic?

3. Search engine optimization (Continue ...)

Questions may include, but are not limited to, the following:

- ✓ Stages of Google search, in detail (discovery, indexed & ranked)
- ✓ What are Sitemaps?
- ✓ What do Sitemaps do?
- ✓ What is a robots.txt file?
- ✓ What does a robots.txt do?
- ✓ What is a PageRank algorithm?
- ✓ How does a PageRank algorithm work?

3. Search engine optimization

Questions may include, but are not limited to, the following:

- ❖ In the **Discovery** process, _____ (small computer programs belonging to search engines) crawl the web, jumping from link to link and exploring each page.
 - Bots or Spiders
- ❖ a Sitemap is an _____ file that lists URLs for a site along with additional metadata about each URL
 - XML




Search engines "discover" pages primarily through links from other web pages.

- 😊 A. True
- B. False




SEO stands for Search Engine...

- A. Observation
- B. Operation
- 😊 C. Optimization
- D. Obfuscation



Unethical SEO techniques are called...

- A. Black glove
- 😊 B. Black hat
- C. Black guard
- D. Black mask




The most important information used in indexing your page is...

- A. The domain name
- B. The filename
- C. The <h1> heading
- 😊 D. The <title>





Which step in the search process does the robots.txt file affect?

-  A. Discovery
- B. Indexing
- C. Lookup and retrieval
- D. Returning results in order




The robots.txt file must be named "robots.txt" in order to work.

-  A. True
- B. False





Which step in the search process does the sitemap (primarily) affect?

-  A. Discovery
- B. Indexing
- C. Lookup and retrieval
- D. Returning results in order



The sitemap must be named "sitemap.xml" in order to work.

-  A. True
- B. False




Which step in the search process does the PageRank algorithm primarily affect?

- A. Discovery
- B. Indexing
- C. Lookup and retrieval
- 😊 D. Returning results in order




The PageRank of your web page is primarily based on


- 😊 A. How many other web pages link to your page
- B. How many webpages your page links to
- C. How much traffic your web page gets
- D. How much traffic your page sends to other sites



A, B, C are web pages accessible to anyone (not password-protected). Web Page A has a hyperlink to Web page B. The PageRank™ score for both pages is calculated.



Another page (call it Web page C) adds a link to Web page B. When PageRank™ scores are calculated again, will the PageRank™ of Web page B go up, go down, or stay the same?

-  A. Go up
- B. Go down
- C. Stay the same

4. Web Analytics

❖ Questions may include, but are not limited to, the following:

- ✓ Types of questions you can answer using web analytics?
- ✓ The two methods of collecting web analytics data?
- ✓ What data can be found in web server logs?
- ✓ Are web server logs public?
- ✓ What data is captured in Common Log Format?
- ✓ What can be discovered from an IP address?
- ✓ What are Common HTTP Status Codes?
- ✓ How Google Analytics can be Opting Out?

4. Web Analytics (Continue ...)

❖ Questions may include, but are not limited to, the following:

- ✓ Types of questions you can answer using web analytics?
 - Who is coming to my website?
 - What pages are they looking at?
 - How long are they spending?
 - Can they find the information they need?

4. Web Analytics (Continue ...)

❖ Questions may include, but are not limited to, the following:

✓ The two methods of collecting web analytics data?


- Analyze **web server logs**
- Use **Google Analytics** or **Open Web Analytics**

4. Web Analytics (Continue ...)

❖ Questions may include, but are not limited to, the following:

✓ What data is captured in Common Log Format?

- ❖ **Remotehost** - hostname or IP address
- ❖ **Rfc931** – the remote logname of the user
- ❖ **Authuser** – the username as which the user has authenticated
- ❖ **[date]** – date and time of the request
- ❖ **Request** – the file being requested
- ❖ **Status** – the HTTP Status Code
- ❖ **Bytes** – the size of the file transferred



It is possible to "opt out" of having your visits to websites recorded in the web server logs.

A. True

😊 B. False



Which of these is collected in the Common Log Format?

A. User location (city and state)

B. User's tracking preferences (on/off)

😊 C. User's IP address

D. User's name



A status code of "200" means

- A. Forbidden
- B. Internal server error
- 😊 C. OK
- D. Page moved permanently



The status code for "file not found" is

- A. 304
- 😊 B. 404
- C. 504
- D. 604
- E. 704

5. JS/CSS/HTML coding

❖ Making Variables, do operations and assigning values to variables

```
let a = 12;  
let b = 16;  
let c = 12 * 16;  
let d = 11 / b;  
let e = a - b;  
let f = a + b;
```

5. JS/CSS/HTML coding

❖ Other Printing methods

alert()

```
alert("Hello World");
```

prompt()

```
prompt("Print your name:");
```

document.write()

```
document.write("Hello World");
```

innerHTML

```
document.getElementById("main").innerHTML= "Hello World";
```

5. JS/CSS/HTML coding

❖ Conditional Statement

```
if (a < b) {  
    ...;  
}
```

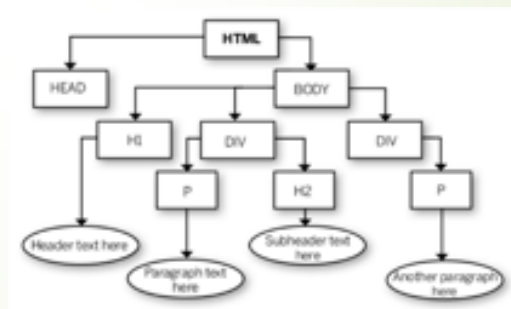
Example: Make two variables and assign them 2 and 5. print their subtraction in console if the subtraction is negative.

```
let a = 2;  
let b = 5;  
let e = a - b;  
if (e < 0) {  
    document.write(e);  
}
```

5. JS/CSS/HTML coding

❖ What is DOM?

Document Object Model



5. JS/CSS/HTML coding

❖ No question about loops


5. JS/CSS/HTML coding

❖ **Functions:** You need to know what is a function, but no need to know how to write a function.

❖ In the exam, you will be given a complete or a partial function.

The diagram illustrates the parts of the function definition `function addTwoNumbers (value1, value2) {
 return value1 + value2;
}`. Red text labels with arrows point to specific parts: "Create a function" points to the `function` keyword; "with this name" points to the function name `addTwoNumbers`; "that takes two values" points to the parameters `(value1, value2)`; and "and returns the sum of them" points to the `return` statement.

```
function addTwoNumbers (value1, value2) {  
  return value1 + value2;  
}
```



5. JS/CSS/HTML coding

Events:

❖ **onclick**

- User clicks on an HTML element


❖ **onmouseover**

- User moves the mouse over an HTML element




Questions about accessibility

- Designing for people who are color blind
- Designing for auditory disabilities





Most people who are color blind use what assistive technology to access the web?

-  A. None
- B. Special glasses that correct color vision
- C. Screen magnifier
- D. Screen reader
- E. Screen Color Editor



Most people who are color blind...

- A. See empty space where there is color
- B. See only in black and white
-  C. Have trouble distinguishing between certain colors
- D. None of the above



To make web pages accessible to people who are color blind...

- A. Don't use color on web pages
- B. Don't use color for anything important
- C. Avoid use of red and green
- 😊 D. Make sure color is not the only means of communicating important information



To make websites accessible to people who are deaf, provide:

- A. Nothing (not needed)
- B. Captions for video
- C. Sign language interpretation for video
- D. Transcripts for audio and video
- 😊 E. B and D
- F. C and D
- G. B, C, and D



Other topics from lecture 9-1

- Working with XHTML
- Web Fonts & Icon Fonts



What Will Not Be on the Exam?

- ❖ Internet security
- ❖ Copyright / Creative Commons licenses / fair use
- ❖ Styling tables
- ❖ Font stacks
- ❖ CSS colors
- ❖ Approaches to Mobile
- ❖ Styling Hyperlinks
- ❖ Good link text or good alternative text



For next time

- ❖ Do not forget to go to lab sessions if you have not met one of CS103 staffs for one-on-one meetings.
- ❖ Remember to check the Final Project folder and go to "your" room on Monday and Wednesday.
- ❖ I will hold many office hours. Stop by if you need help.
- ❖ See you on Monday for final project presentations!