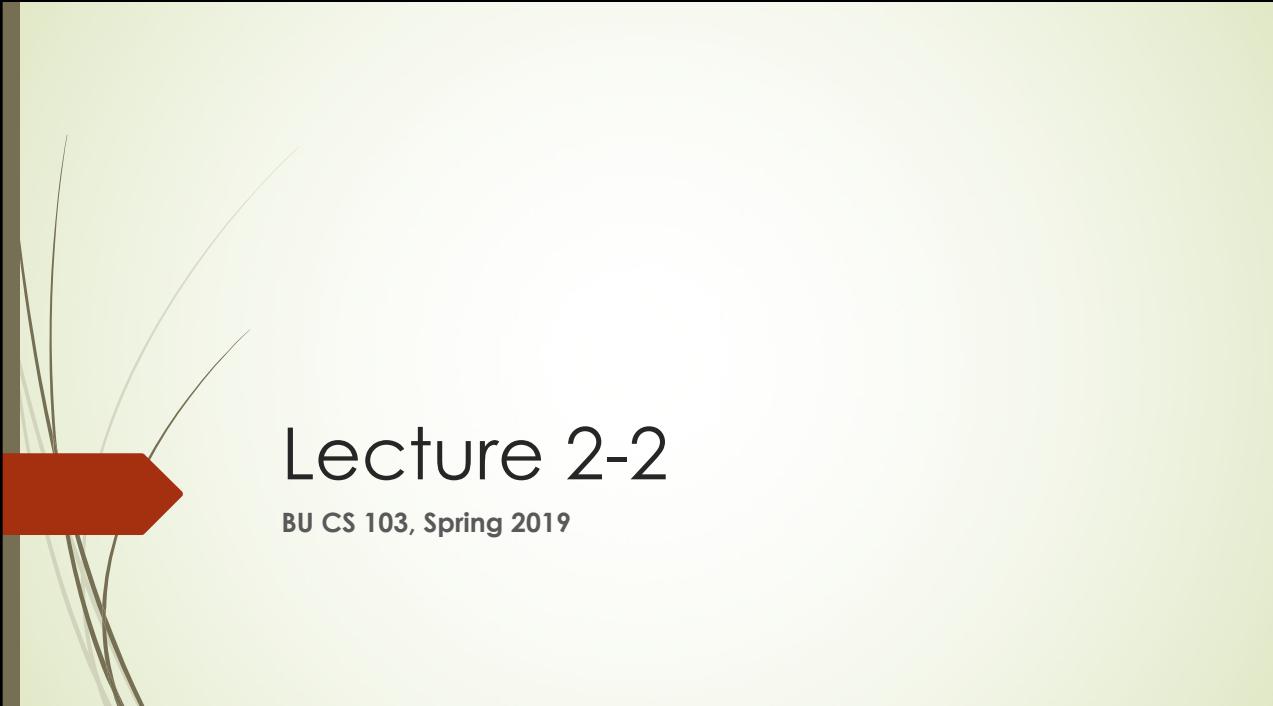


Introduction to Internet Technology and Web Programming

Computer Science 103
Boston University
Vahid Azadeh-Ranjbar

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



Lecture 2-2

BU CS 103, Spring 2019

Today's Materials

- ❖ Review from last lecture
- ❖ Introduction to HTML
- ❖ Creating an HTML page
- ❖ Internet security (Part 2)
 - ✓ Password theft and phishing

Deadlines

- ✓ Lab 0 Assignment: due was today at 6:00PM but it is extend till Friday at 6:00 PM
 - There is a no late submission period for lab 0 assignment
- ✓ Lab 1 Challenge: due Thursday 01/31 at Midnight
- ✓ Lab 1 Assignment: due Thursday 02/06 at 6:00 PM



Review from last lecture



Which part of *http://www.bu.edu/students* is the domain name?

- A. http://
- B. WWW
- C. bu
- D. edu
-  E. bu.edu
- F. www.bu.edu
- G. www.bu.edu/students
- H. http://www.bu.edu/students



In the acronym "gTLD", what does the "g" stand for?

- A. general
-  B. generic
- C. geographic
- D. global



Which of these is not a type of gTLD?

-  A. General
- B. Restricted
- C. Sponsored
- D. Unrestricted

Questions on Assigned Reading

- ❖ According to the reading from **Hubspot**, when is it a good idea to use a domain name ending in **.org**?
- ❖ When is it NOT a good idea to use a domain name ending in **.org**?

Questions on Assigned Reading

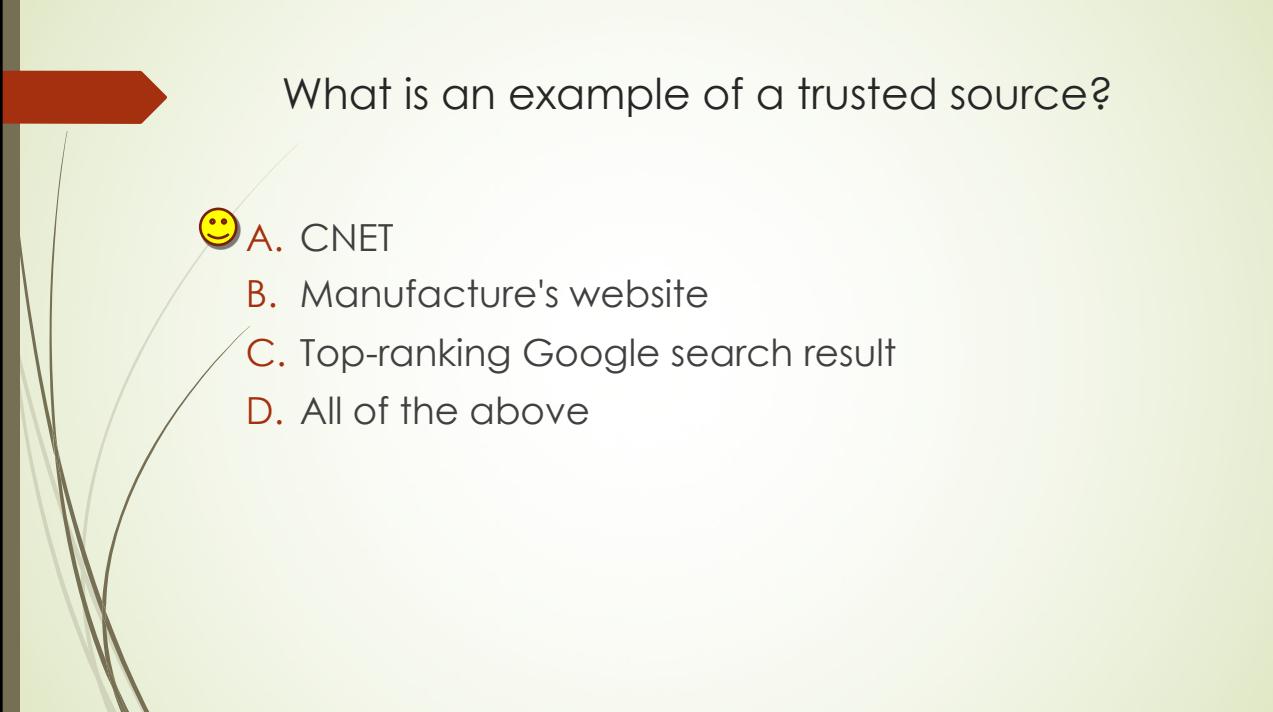
- ❖ According to the reading from **Hubspot**, when is it a good idea to use a domain name ending in **.org**?
- ✓ If you are a **nonprofit**, especially if you are reasonably well known, by all means use a **.org** domain name.

Questions on Assigned Reading

- ❖ When is it NOT a good idea to use a domain name ending in **.org**?
 - ✓ The .org extension is technically open to all, but is not recommended for **business ventures**, as there is an expectation that **.org = nonprofit**.

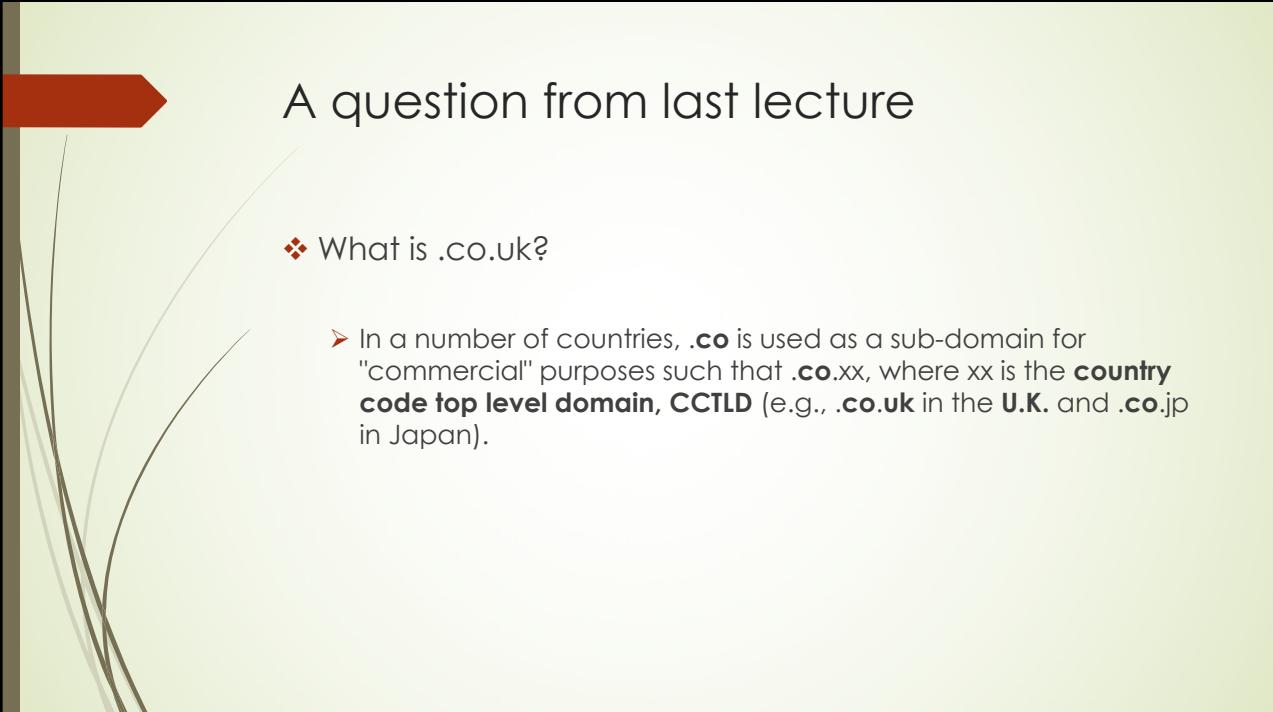
Questions on Assigned Reading

- ❖ According to Symantec.com, the **.accountant** TLD has many more "bad" sites (sites that distribute malware, spam, or phishing) than the **.realtor** TLD.
 - Why?



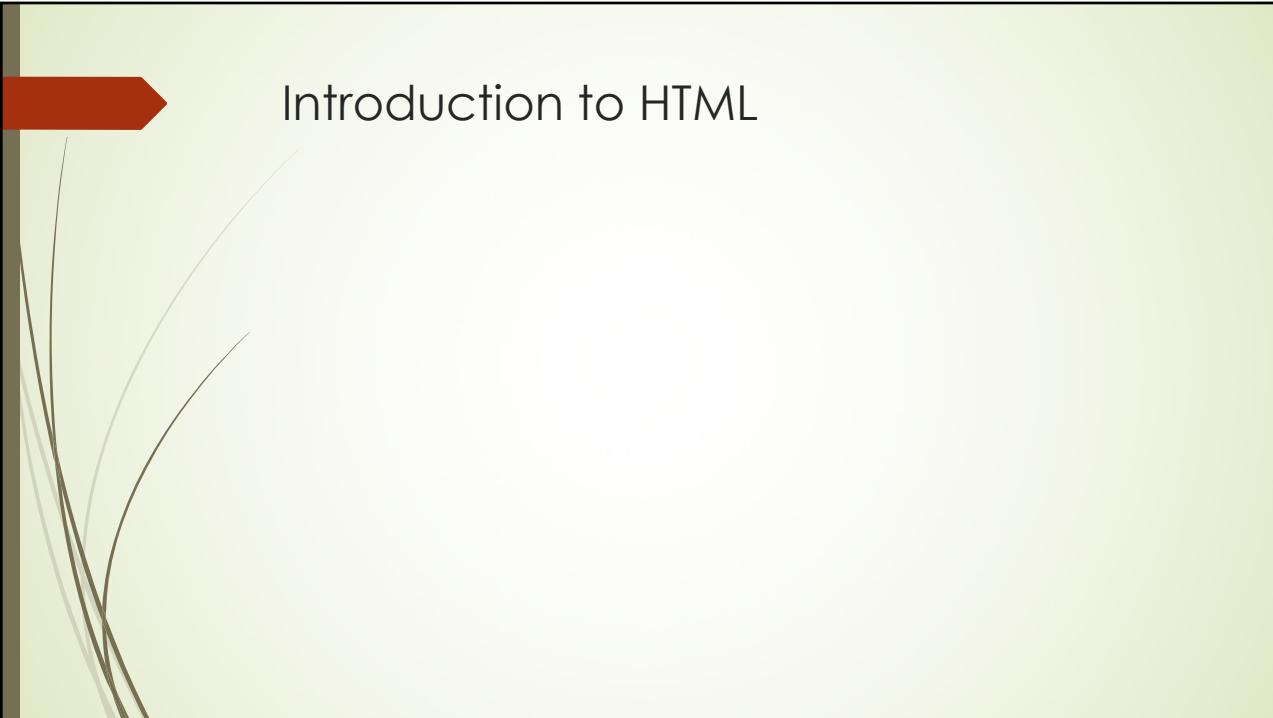
What is an example of a trusted source?

-  A. CNET
- B. Manufacture's website
- C. Top-ranking Google search result
- D. All of the above

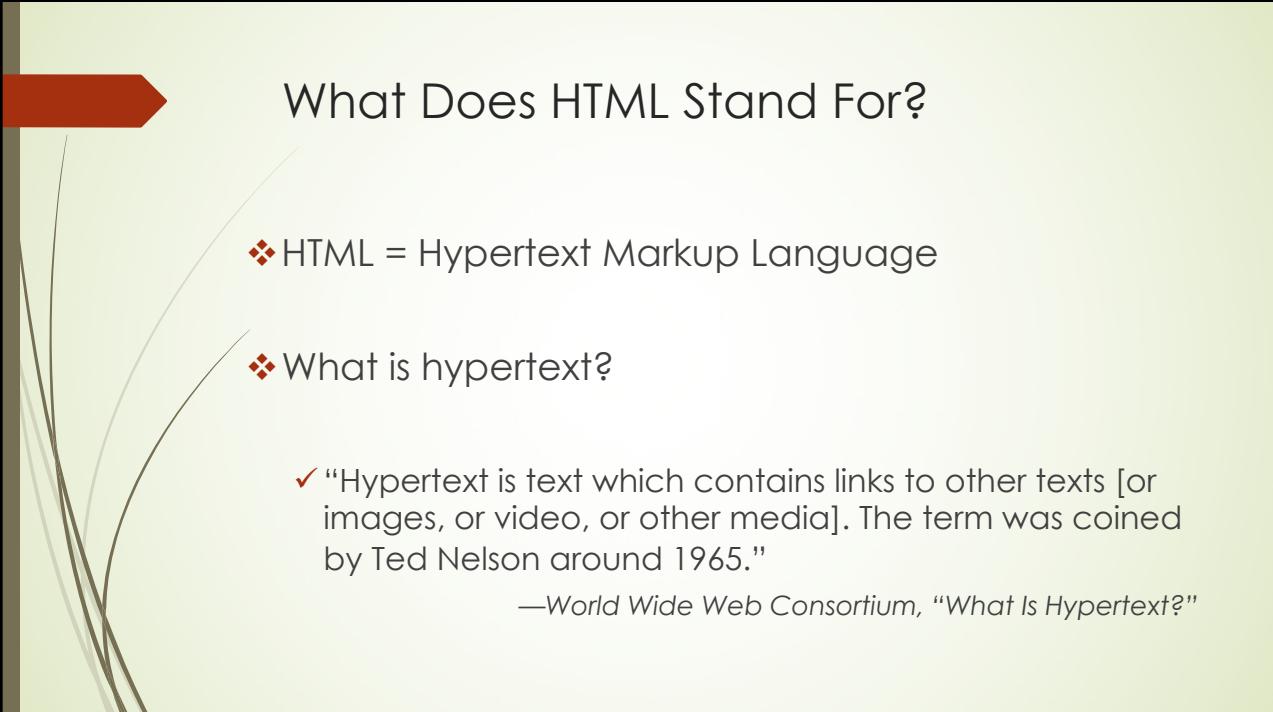


A question from last lecture

- ❖ What is .co.uk?
 - In a number of countries, .co is used as a sub-domain for "commercial" purposes such that .co.xx, where xx is the **country code top level domain, CCTLD** (e.g., .co.uk in the **U.K.** and .co.jp in Japan).



Introduction to HTML



What Does HTML Stand For?

- ❖ HTML = Hypertext Markup Language
- ❖ What is hypertext?
 - ✓ “Hypertext is text which contains links to other texts [or images, or video, or other media]. The term was coined by Ted Nelson around 1965.”
—World Wide Web Consortium, “What Is Hypertext?”

What Is a Markup Language?

- ❖ “A markup language is a computer language that uses **tags** to define elements within a document.”

—Tech Terms

Sample Markup Languages

- ❖ HTML
 - ✓ HyperText Markup Language
 - ✓ Used for web pages
 - ✓ Identifies paragraphs, headings, lists...
- ❖ XML
 - ✓ eXtensible Markup Language
 - ✓ Used for data exchange
 - ✓ Precise set of rules (the “schema”) determines which elements are required, which elements are children of one another, etc.

More Sample Markup Languages

- ❖ XHTML

- ✓ HTML that follows the rules of XML

- ❖ MathML

- ✓ Mathematical Markup Language
 - ✓ Used for formulae within web pages

Creating an HTML Page

An HTML Document...

- ❖ Consists of **content (words and images)** and **HTML tags** indicating its structure.

1. Open Your Text Editor

- ❖ Recommended text editors:
 - ✓ **Notepad++** (PC) – download from link on Course Materials page of Blackboard
 - ✓ **BBEdit** (Mac) – download from link on Course Materials page of Blackboard

2. Save Your Text File

- ❖ Any filename is OK as long as it has no spaces in it
 - ✓ Most web servers run on the Unix or Linux operating system
 - ✓ Unix and Linux systems do not handle filenames with spaces well
- ❖ Filename must end in .html (or .htm), not .txt

2. Save Your Text File – Where?

- ❖ If you save your file on your computer, then **you** can view it in a web browser – but no one else can see it, and you can see it only from your computer
- ❖ To be accessible to the WWW, your file must be saved on your web host
- ❖ And not just anywhere on your web host, but in a particular folder.
- ❖ For most web hosts, this is the **public_html** folder (for Unix/Linux machines) or the **Inetpub/wwwroot** folder (for Windows web servers)

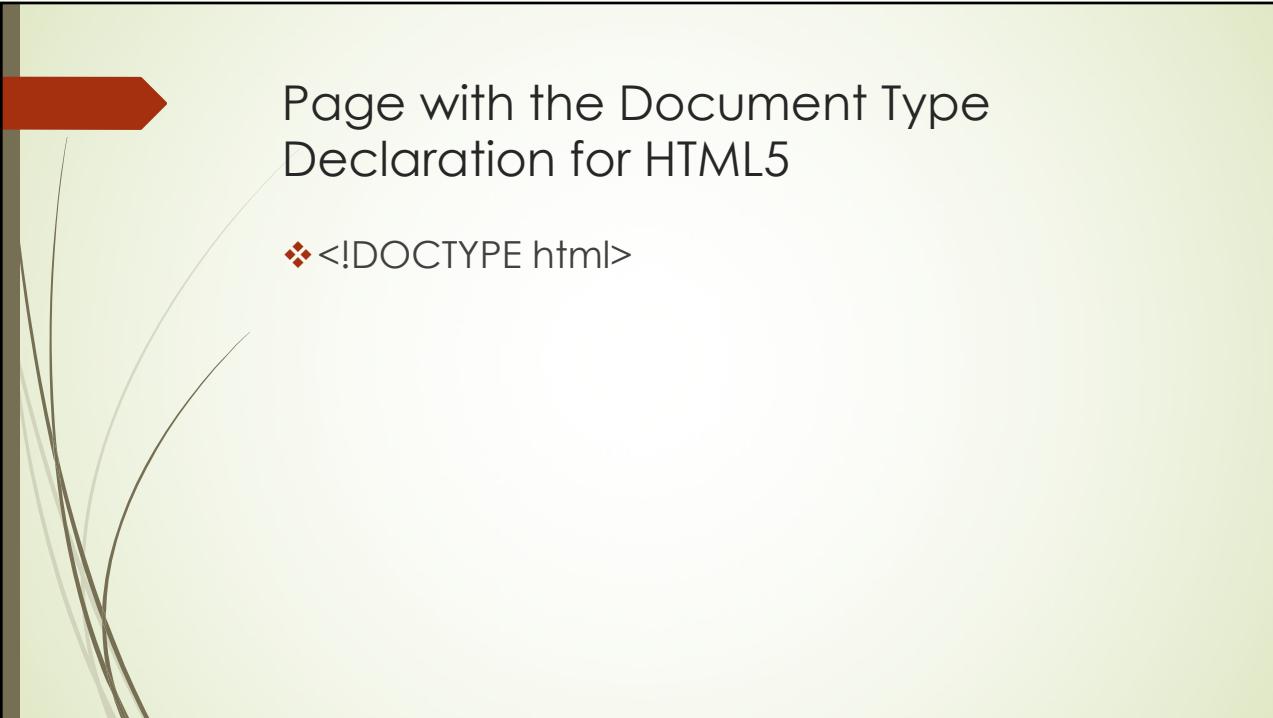
3. Choose an HTML Version

- ❖ HTML has different versions that were approved at different times.
- ❖ The version of HTML you are using is stored in the “document type declaration”
 - ✓ **the very first line in your html file.**
- ❖ The World Wide Web Consortium maintains the standard definitions of HTML, each of which has its own document type declaration

HTML Version	Document Type Declaration
HTML5	<!DOCTYPE html>
XHTML 1.0 Transitional	<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Transitional//EN" "http://www.w3.org/TR/xhtml1/DTD/xhtml1-transitional.dtd">

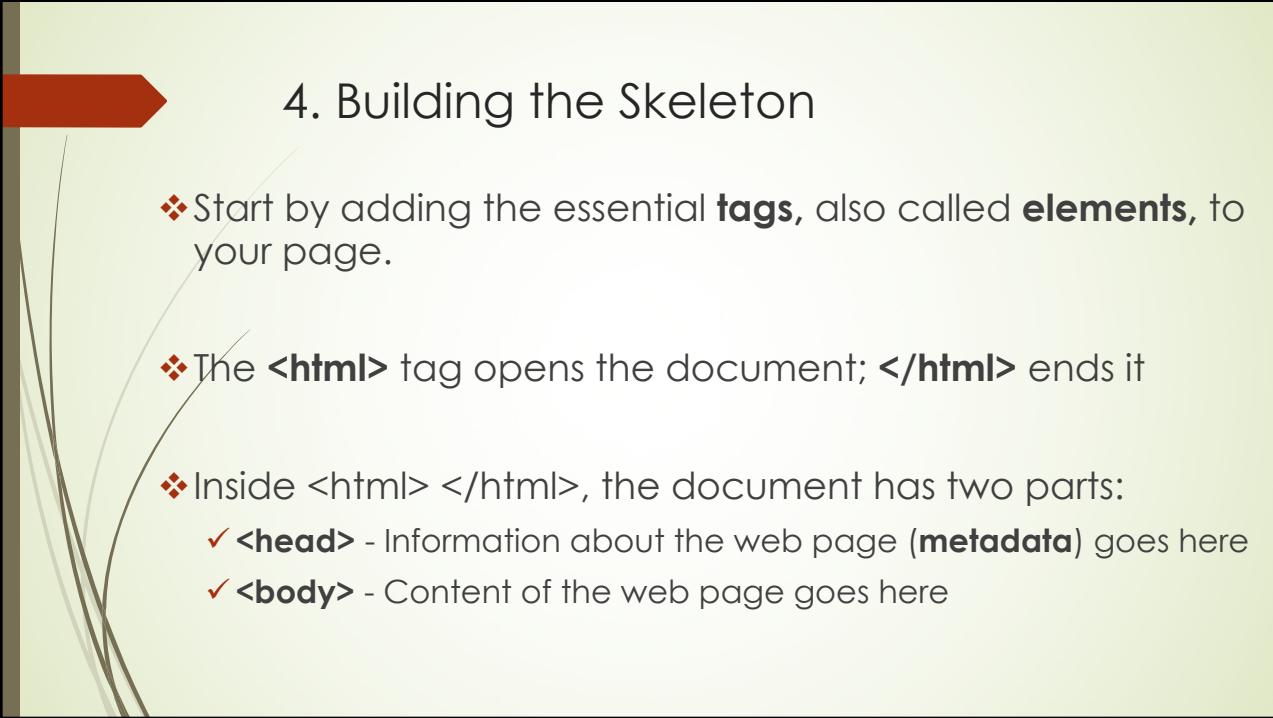
Versions of HTML

- ❖ According to **w3techs.com**, the most common versions of HTML today are:
 - ✓ HTML 5 (60%)
 - ✓ XHTML (approx. 28% of all websites)
 - ✓ Older versions of HTML (12%)
- ❖ Major differences between versions:
 - ✓ Syntax (punctuation)
 - ✓ Which tags are supported, which are deprecated, and which are no longer supported
 - ✓ Which tags are required and which are optional



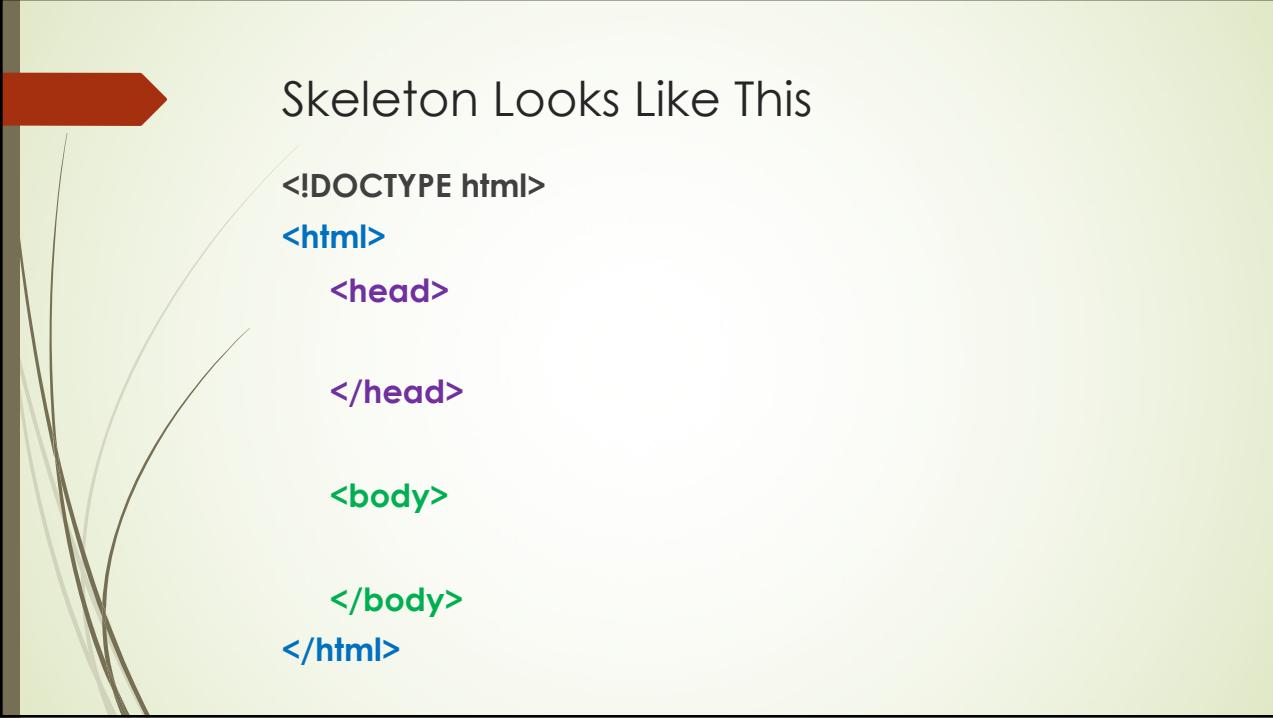
Page with the Document Type Declaration for HTML5

- ❖ <!DOCTYPE html>



4. Building the Skeleton

- ❖ Start by adding the essential **tags**, also called **elements**, to your page.
- ❖ The **<html>** tag opens the document; **</html>** ends it
- ❖ Inside **<html> </html>**, the document has two parts:
 - ✓ **<head>** - Information about the web page (**metadata**) goes here
 - ✓ **<body>** - Content of the web page goes here



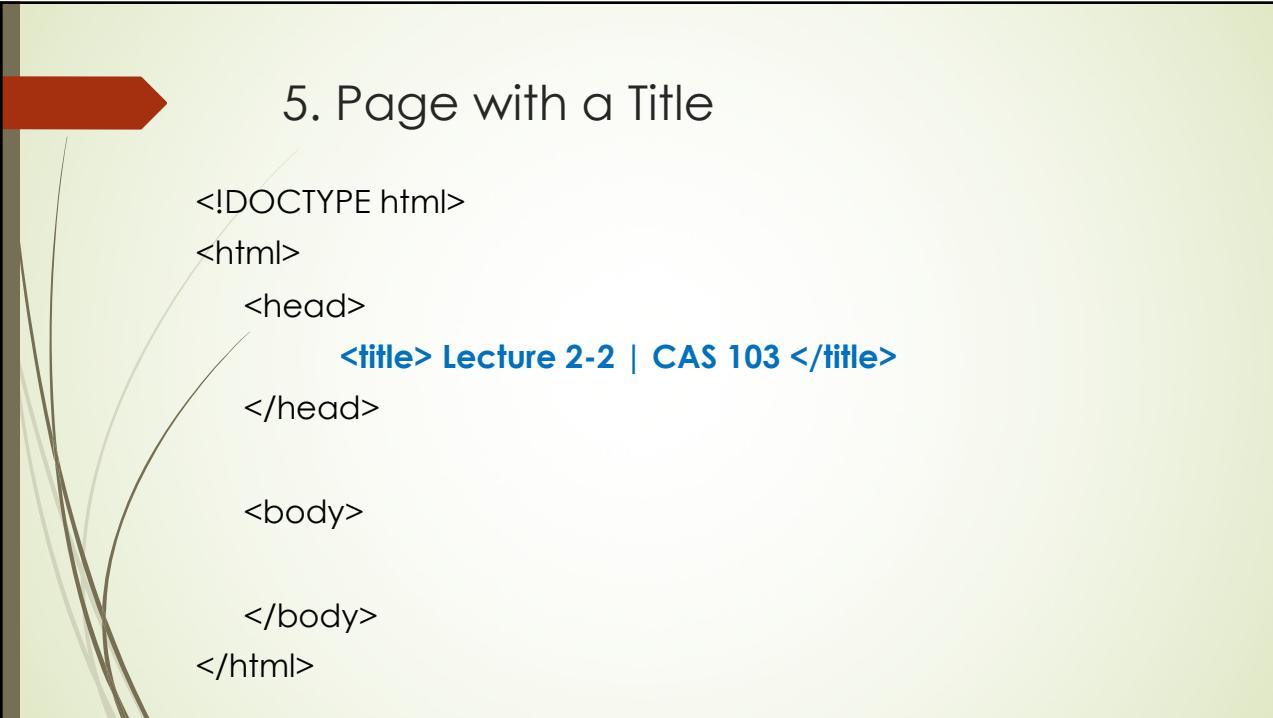
Skeleton Looks Like This

```
<!DOCTYPE html>
<html>
  <head>

  </head>

  <body>

  </body>
</html>
```



5. Page with a Title

```
<!DOCTYPE html>
<html>
  <head>
    <title> Lecture 2-2 | CAS 103 </title>
  </head>

  <body>

  </body>
</html>
```

Title Element

- ❖ The <title> tag does not appear on the web page itself, but it does show in the browser frame or tab (depending on browser settings)
- ❖ Should include both the name of the web page and the name of the website
- ❖ The title is critical to both **usability** and **findability**
- ❖ If a page has no title, the URL or filename will be displayed in the browser frame instead

It's time to experience it!

6. Add Your Content to the Body

- ❖ Add the content of the page between the two body tags.
- ❖ Use HTML tags to identify the elements of the page.
- ❖ All content should be inside a pair of elements like these.
 - ✓ **Paragraph** < p > </ p >
 - ✓ **Top-level heading** < h1 > </ h1 >
 - ✓ **Second-level heading** < h2 > </ h2 >
 - ✓ Third-level, fourth-level, fifth-level, and sixth-level headings
 - ✓ **Unordered list** < ul > </ ul >
 - ✓ **Ordered list** < ol > </ ol >
 - ✓ **List item within a list** < li > </ li >

Let's build a page with some contents!

7. Add Hyperlinks

- ❖ Hyperlinks are what make the World Wide Web unique
- ❖ Why would we want to link to another web page?

Reasons to Hyperlink

- ❖ Navigation
 - ✓ Help the user get between pages on your own site
- ❖ Help the User
 - ✓ Help the user get to information that you don't "own"...
 - ✓ ... or don't want to provide yourself
- ❖ Do the right thing
 - ✓ Acknowledge your sources
 - ✓ Link to Creative Commons license
 - ✓ Promote other sites that share your goals

Creating Hyperlinks

- ❖ ` Boston University `
- ✓ **href** is an **attribute** of the `<a>` tag – it provides information necessary for the `a` tag to work. Here, that information is the designation (or target) of the link.
- ✓ Note that **href** must be enclosed in straight quotation marks (curly quotes won't always work)
- ✓ The **href** URL must begin with **http://** (or **https://**).
- ✓ The text between the `<a>` and `` tags is called the **link text**.

Let's add a link to our page!

Further Resources

- ❖ When learning HTML, looking things up on your own is also essential
- ❖ Two good references:
 - ✓ HTML Dog (used in lab)
 - ✓ W3Schools.com – has a page explaining every HTML tag

Good HTML...

- ❖ Has all required tags
- ❖ Uses both opening and closing tags
- ❖ Nests tags correctly
- ❖ Spells tags correctly
- ❖ Provides attributes when required

Why Write Good HTML?

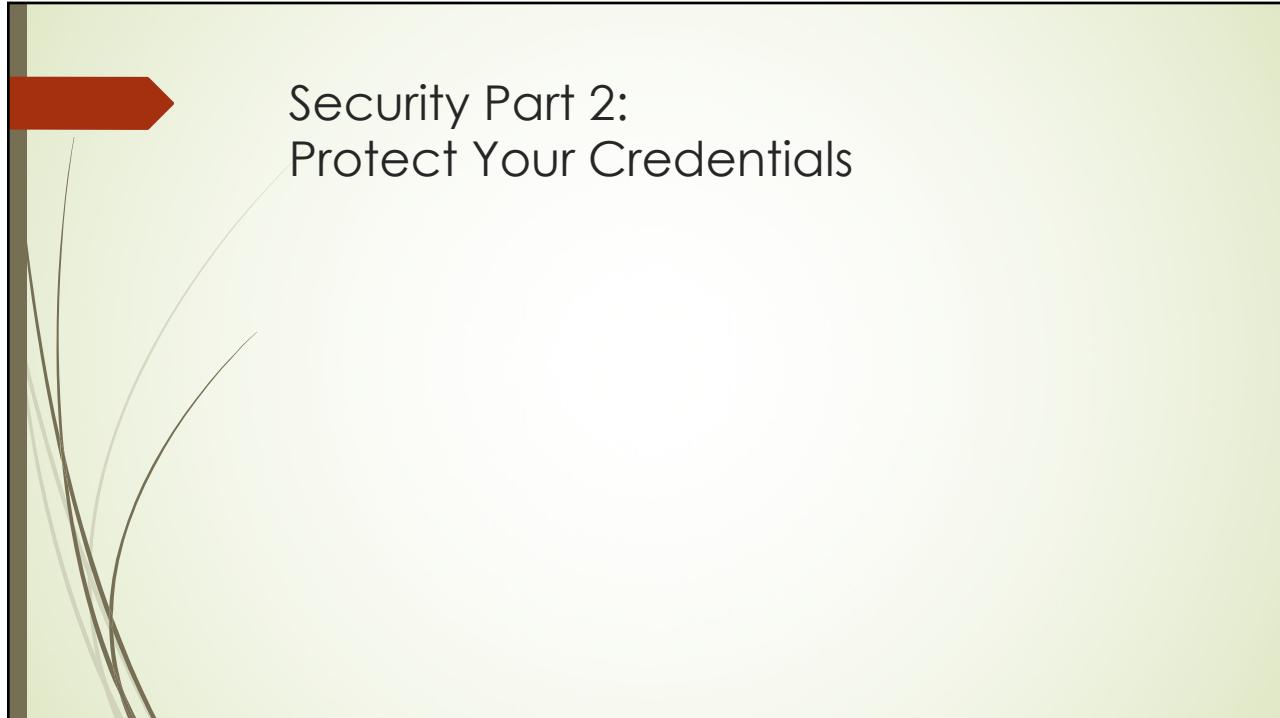
- ❖ Good HTML will work in all web browsers. (Browsers will fix some mistakes, but not all, and they don't always fix them the same way.)
- ❖ Good HTML is more accessible because assistive devices use headings and lists to help people navigate through pages.

In-class activity

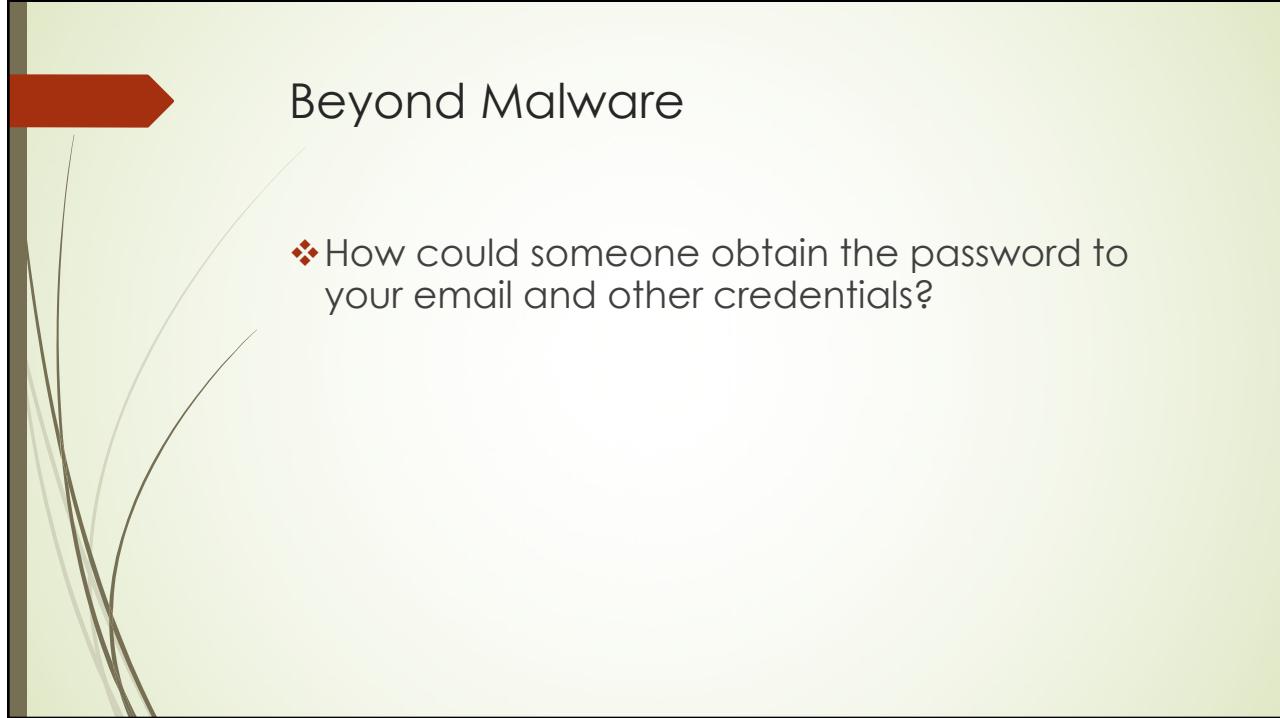
- Go to the following URL (CodePen) website and debug the given code. How many HTML errors are in that HTML code?

<https://codepen.io/azadeh83/pen/gQxQrz>

Depend on the way you count the bugs, it can be 4 or 5.



Security Part 2: Protect Your Credentials



Beyond Malware

- ❖ How could someone obtain the password to your email and other credentials?

Causes of Password Theft

- ❖ Someone **guessed** it using information they know about you.
- ❖ A cybercriminal used a **computer program** to try thousands of passwords until it found the one that worked.
- ❖ Someone guessed the answers to your **security questions** and used them to reset your password
- ❖ Another **site was hacked** and your password was stolen
- ❖ You entered your password on a computer infected with keylogger malware.
- ❖ Your password was intercepted when you entered it while on public WiFi
- ❖ You accidentally gave it away through **phishing**

What Is Phishing?

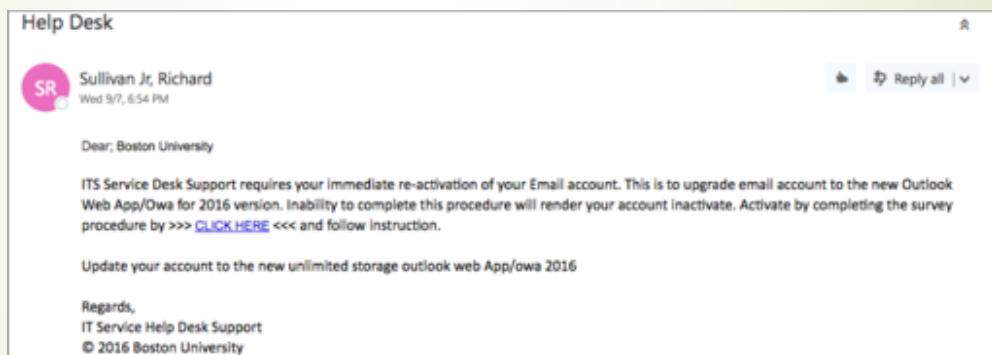
- ❖ Phishing is “the act of sending an **e-mail** to a user falsely claiming to be an established legitimate enterprise **in an attempt to scam the user into surrendering private information** that will be used for identity theft.”

—Webopedia

How to Spot a Phishing Email

- ❖ Sense of urgency or threat – something bad will happen if you don't respond right away
- ❖ The email asks you to send a password, or click a link to view, verify, or reset something – but the URL to which the link takes you is not a trusted website
- ❖ The “From” email address does not match the “From” name (but not always; can be spoofed)
- ❖ The “From” name or email address does not match a trusted source
- ❖ Errors in grammar, spelling, or punctuation

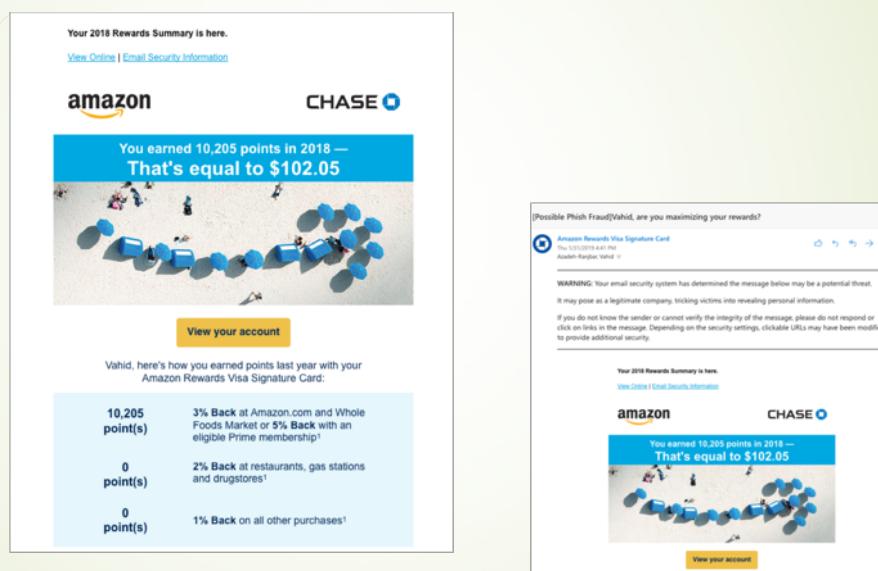
“Primitive” Phishing Email



Spear Phishing

- ❖ Phishing emails have become much more sophisticated
- ❖ “Spear phishing” targets a particular person using publically available information, such as the name of your company, school, or supervisor...
- ❖ ... or the domain name of your website

Spear Phishing Example



What to Do If You Suspect Phishing

- ❖ If anything looks suspicious, get the sender's phone number from the official source (its own website or a directory, not the email) to find out if it is legitimate
 - ✓ Do not click on the link
 - ✓ Do not respond to the email
- ❖ If the spam or phishing attempt seems to be coming from BU, or if you have been the victim of a scam, then please alert BU Information Services and Technology by email to:
abuse@bu.edu

To Avoid Phishing: Check the URL

- ❖ Hover over a link in an email or on a web page to see the URL in your web browser.
- ❖ This feature works automatically in most web browsers.
- ❖ The URL should begin with http:// or https://
- ❖ The domain name, including the TLD, should exactly match the website you expect to go to.

Phishing Using TLDs

- ❖ One of the Homework 0 assigned readings explained how cybercriminals were using a certain top-level domain to make phishing more effective. What was it?

Phishing Using TLDs

- ❖ One of the Homework 0 assigned readings explained how cybercriminals were using a certain top-level domain to make phishing more effective. What was it?
 - ✓ .om
- ❖ What is the name of this phishing attack?

Phishing Using TLDs

- ❖ One of the Homework 0 assigned readings explained how cybercriminals were using a certain top-level domain to make phishing more effective. What was it?
 - ✓ .om
- ❖ What is the name of this phishing attack?
 - ✓ **typosquatting**
- ❖ Cybersecurity firm Endgame noticed recently that someone was attempting to distribute malware through the URL **Netflix.om**.

For Next Time ...

- ❖ Enjoy the lab tomorrow!
- ❖ Do not forget to attend your own lab section.