

# The web

## HTTP

HTTP = HyperText Transfer Protocol

HTML = HyperText Markup Language

HTTP specifies how documents/resources are transferred (downloaded/uploaded) over the web. They can be any sort of file.

URL = Uniform Resource Locator – the address of some file (page) on the web.

Example: `https://liucs.net/cs101f17/`  
          <sup>^1</sup>      <sup>^2</sup>      <sup>^3</sup>

1. is the protocol portion of the URL (usually just http or https but there are others)
2. is the host name (might just be domain name). Domain name is what is independently purchasable from a domain registrar. Host name can include additional info to the left: `www.liu.edu` indicates the domain is `liu.edu`, and `www` is the host portion. `my.liu.edu` indicates the same domain, but `my` is the host portion. Host name can also just be IP address, although it doesn't always work.
3. is the path. It indicates what resource, **on that host**, we would like to access.

An example with a non-HTTP protocol: `tel:+17184881274` The protocol is “tel” which is for telephone.

HTTP is how your web browser requests resources from a web server. There are different “verbs” used:

- GET – access/download the resource (view page)
- POST – send some data that may have an effect (submit form)
- others

There are also different response codes:

- 200 means “OK”
- 3xx is some sort of “redirect”
- 400 means request was malformed
- 403 means “forbidden”
- 404 means “not found”
- 5xx means server error

HTML is the language in which we specify the content and layout of web pages. It also can include CSS (for style) and JavaScript (for scripts).

## HTML

Below is the code of a sample page that I demonstrated in class. Copy and paste it to a file saved with the .html extension (you can use Notepad, Notepad++, or TextEdit for this). Then you can edit and open in browser simultaneously. Save changes in the editor then refresh the browser.

You can see a [live version of this page](#) in your browser.

```
<html>
  <head>
    <style type="text/css">
      body { margin: 4ex 20%; }
      h1 { background: #fcc; border: 1px solid #f33 }
    </style>
  </head>
  <body>
    <h1>A silly page by Chris League</h1>
    <p>Welcome to my home page!</p>

    <h2>My Hobbies</h2>
    <p>I like to:</p>
    <ol>
      <li>Travel around the world</li>
      <li>Take pictures</li>
      <li>Eat all kinds of foods</li>
      <li>Play piano</li>
      <li>Write computer programs</li>
    </ol>

    Hyperlinks can be deceiving because the URL and the linked text
    are distinct. They don't need to match. Here's a dubious link
    to: <a href="http://en.wikipedia.org/wiki/Phishing">Citibank</a>

    <h2>Resources</h2>
    <ul>
      <li><a href="http://www.w3schools.com/">w3schools</a> is a
        great reference for HTML and CSS.</li>
      <li>I ask and answer programming questions
        on <a href="http://stackoverflow.com/">StackOverflow</a></li>
    </ul>

  </body>
```

</html>

## Miscellany

- One possible advantage of **View Source**
- An employment ad that will autonomously vet candidates for high levels of accuracy, tenacity and attention to detail

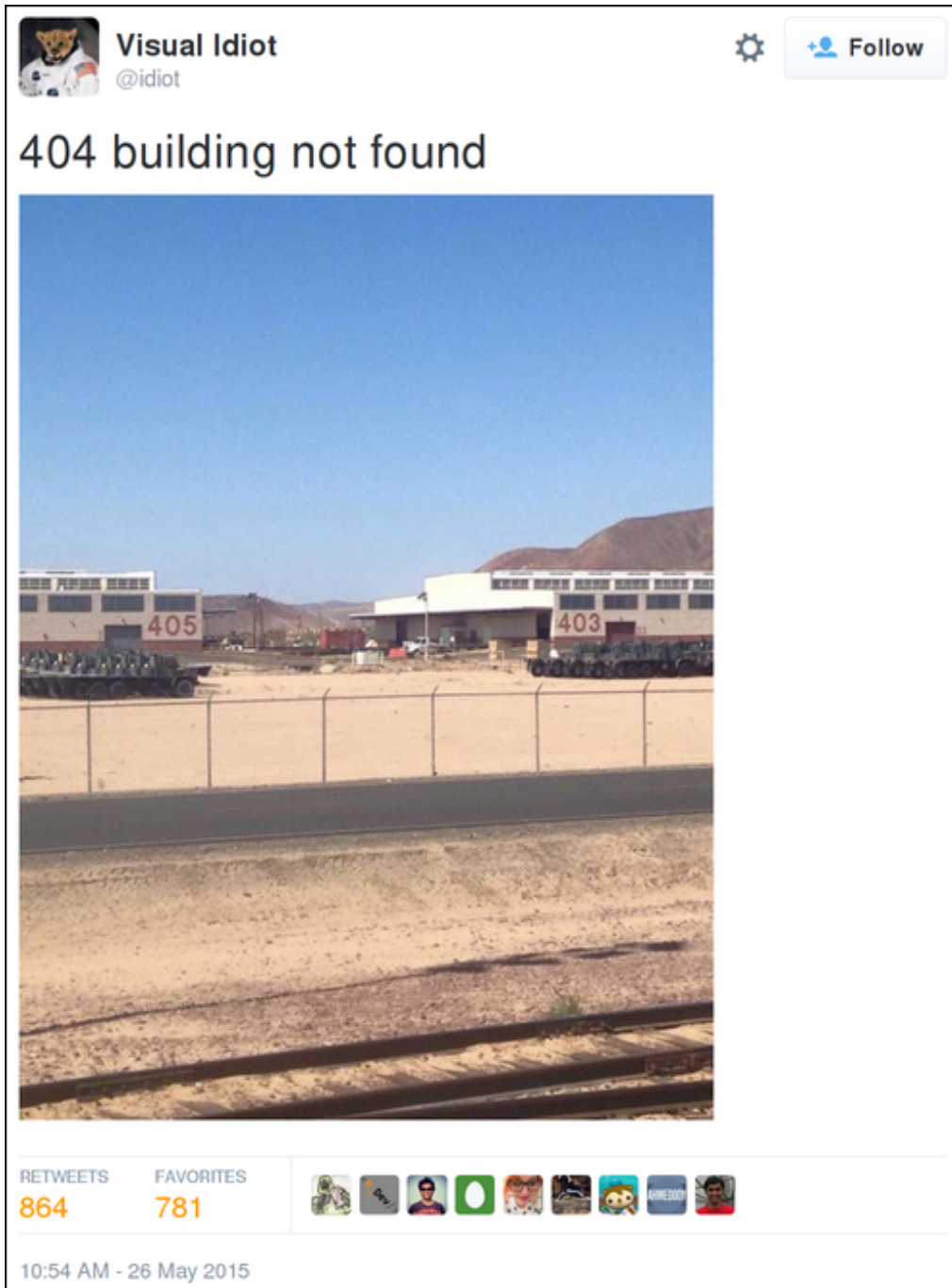


Figure 1: @idiot on Twitter