|  |  |  |
| --- | --- | --- |
| The four parts of a URL |  | This presentation is on URLs. URL sounds exotic – it stands for universal resource locator – but it is just a geek term for a Web address. A URL is the address of a resource on the Web.  You’ve typed URLs into Web browsers many times, but may not have thought about the meanings of each of the four parts of a URL. We will review them and also mention shortened URLs.  In order to understand the parts of a URL, we will also cover domain names, directories and file names. |
| Communication technology |  | URLs are a technical topic, but understanding them will also help you use the Web, so we include this presentation at two places in our class outline. |
| URL – a Web address |  | To retrieve a page from a Web site, you type its URL into the box at the top of the browser and hit enter. The browser then retrieves the page from the server and displays it. URL is just a fancy way to say the Web address of the page.  In this example, a short professional biography that is stored at the listed URL is retrieved and displayed. |
| The four parts of a URl |  | If you look at URLs, you see that they can be broken down into four parts. Note that some of the parts may be omitted at times – we’ll talk about that later. For now, let’s take a look at each of the four parts. |
| Part 1 – this is a Web page |  | The URL prefix “http://” indicates that this is the locator for a Web document that should be retrieved and displayed. It turns out that Web browsers can also function as clients for other services like file transfer, but that is seldom done. For that reason, one can often omit the “http://” since most browsers will plug it in for you. |
| Part 2 – the domain name of the server |  | The next part of our example URL reads “som.csudh.edu”. That is the domain name of the server. Domain name is another fancy geek word – it is the name of the computer the server program is running on. We will say more about domain names in another presentation. For now, it’s enough to say that the domain name identifies a unique computer on the Internet. |
| Part 3 – the directory the page is in on the server |  | The next part of the URL is the directory or folder containing the page we are looking for. Every computer has many files in storage. We would quickly become confused in trying to locate a single file in a list of say 100,000 names. For that reason, we organize our storage into hierarchical directories, sub-directories, sub-sub-directories and so forth.  Here we have three sub-directories – for staff, faculty and administrators. Since I am a faculty member, my directory, named “lpress,” is a subdirectory of the fac directory. Each faculty member – jsmith, jdoe, and the rest would have directories in the fac directory. |
| Part 4 – the name of the file on the server |  | The final part of the URL is the name of the file we wish to retrieve, in this case “shortbio.htm.” Since shortbio.htm is in the lpress directory, the Web server will send that file back to the client, which will display it. The user can then read my short biography.  Note that the file name has two parts separated by a dot. The suffix or “extension” indicates what type of file it is. The suffix “htm” means that this is an HTML file, in other words a Web page. (We will say more about HTML later). |
| Shortened URLs |  | Let’s look at a couple more things concerning URLs.  For a start, you can save space by shortening a URL. If you try them out, you will see that both of the links shown here are to the same page. There are many URL shortening services on the Internet. I used one called bit.ly to create this example. |
| Shortened URLs can be risky. |  | Be cautious when clicking on shortened URLs. Shortened URLs hide the actual domain name, so you cannot tell where they link to by looking at them.  Click the link shown here to see why you should only go to shortened URLs from trusted sources. |
| Most Web clients assume *http://* by default |  | As mentioned previously, if you delete the http://, the client will assume the message is for a Web server and request the page as if the *http://* had been included. |
| Default file names |  | If you delete the file name, the Web server will look for a file with a default name, which is specified by the server administrator.  Traditional default names are *index.htm, Index.html, default.htm,* and *default.html.* |
| Summary |  | Here we summarize the meaning of the four parts of a URL. In addition to learning about the parts of a URL, we saw that they may be shortened using a service like Bit.ly and that some parts of the URL were optional.  We also learned about domain names, directories and file names. |
| Self-study questions |  |  |