

Labs will start the week after next.

Anatomy of a URL (cont.)

Recap: only the scheme (`http://`) and hostname (`stuff.com`) required, everything else is optional.

Scheme: Usually a protocol type like *http*. Sometimes something else, like *mailto*.

User Info: covered already

Hostname: Used to be able to tell the location of the server from the `_____` suffix (forget the name). This isn't the case anymore, since you can use any of them anywhere.

Port Number: Usually web browsing uses port number 80, but 8000 and 8080 are common for proxies.

Side note: The computers in the lab are not connected directly to the internet but to a proxy server instead which forwards the webpages to us. Port number 8000 (he thinks).

Path: You can specify a particular directory/file.

Side note: We all have a URL with our account a particular address. Mine is:

`cs1.ucc.ie/~nb17/`

The tilde is "an alias for a particular directory"(?). It points to your own directory. Everyone has a folder called `public_html`, this is what the `~` stands in for(?).

The file we see is called `index.html`, and this is the default the webserver looks for if no file is specified. If it doesn't find that

it looks for index.php. If it finds neither then it shows what files are there.

He recommends not calling it index.html because you can then find stuff more easily (he reckons he can't explain this properly until a lab, where we all have accounts and so on).

He recommends deleting index.html so that you can find stuff more easily. Other people won't recommend that.

Query: There can only be 1 question mark in the URL, and it must be here, after the path. These are parameters. You can put form data here. Always parameter_name=value, separated by &. (?page=3&lan=en)

Fragment: (#position) Scrolls down to a particular section in a webpage (end result).

In the book of modules, each module code link is a link back to the page you are on but with a new fragment.

Nowadays the browser is clever enough that it doesn't make a request again on clicking a link like this, and instead just scrolls to the appropriate position.

This is not specified in http or anything, though, and if the browsers weren't recognising that they shouldn't load the page again, they would load it again every time you click a hyperlink. You can also use this to specify a time point in a youtube video to start from.

The video he used to show it was a drone flight over Cork uploaded to youtube by Raymond Fogarty.

If the returned html depends only on the URL then we speak of a RESTful internet service.

REST stands for Representational State Transfer.

AJAX is a development 10-15 years ago which breaks this principle. It allows you to e.g. play games in the web browser by modifying the page but not changing the URL.

If this principle holds, you can show a URL to someone else and they will see the same stuff you will see.

Standards & Validation

There are (too) many flavours of html. The most common ones are HTML 4.0, XHTML, and HTML 5.0.