

The Interweb

I have called this the Interweb but, in reality, there is no such thing; it is an amalgamation of two words: the Internet and the World Wide Web. The Internet can be thought of as like a system of roads, the WWW as a fleet of busses. The Internet can exist without the WWW. Many other things can run along roads as well as busses, such as cars and lorries and similarly other things such as email and file transfers can use the Internet. The WWW cannot exist without the Internet (or I suppose something very similar) just as a bus is pretty useless without a road. Please feel free to ignore the next section on how the Internet works, or any of the rest, some may find it interesting!

How it Works

Connecting a wire between two computers is all very well, but to be useful you have to be able to transfer information along it. Systems like Morse Code are simple, but limited in scope and slow compared to the rate at which computers can work. The simplest way to send a signal along a wire is to switch a voltage on and off. When the voltage is on that can be thought of as a '1', when off as a '0' so switching on and off sends a stream of 1s and 0s along the wire. You have to be able to distinguish between a steady stream of 1s or 0s, so the switching is done according to a clock which is a steady stream of 10101010,,, on a separate wire which carries no information except to denote how long each 'bit' lasts.

The continuous stream of 1s and 0s has to have some form of structure to carry an intelligible message, it also has to carry such messages from a number of different sources to a number of different destinations. The definition of such structures is known as a **protocol**. These are generally defined in layers, one on top of or using a lower level one, so that a lower level one can be exchanged for some other mechanism, provided that it implements the same protocol. So, for example the lowest level protocol of voltages on wires (usually the one known as **ethernet**) can be exchanged for wireless signals (usually the one known as wifi) and everything else that is built on top of that stays exactly the same. The first step is to split the stream up into chunks of 8 bits, known as bytes or to network nerds as octets. Each byte can represent a (small) number expressed in binary, or as a pattern representing an alphanumeric character, punctuation or other control symbol; they can also be combined into larger numbers or more complex characters (e.g. Chinese).

A number of bytes is then treated as a packet, with a recognisable start and stop. Since a full message will involve many packets there will be a sequence number, and, since many different applications may be sending messages, a number known as a port which says which application sent the message and which should receive (e.g. normal email uses 25 for both.) Networks that work by sending packets of data between a number of machines are often referred to as **packet-switching** networks. On the Internet, the protocol that defines how

messages are split up into packets, sent off in sequence then recombined at the other end, including acknowledging receipt and resending lost packets is known as the Transmission Control Protocol (TCP.)

The final step is to add onto each packet a pair of addresses, which is something to identify where the message came from and where it should go. The basic form of an Internet address is four numbers, each between 0 and 255 inclusive, which is the range of numbers that can be held in one byte. This could be written as (e.g.) 78.129.207.235 . When the Internet was first developed it was thought that there would be no chance of running out of numbers. Now, with every home computer, every tablet, every mobile phone and in some houses the fridge and every lightbulb having its own address, they have run out. Since a lot of addresses were handed in out in blocks to companies there are still some unused ones around and some that can be reused, but a move to a new six number address (IP6 as opposed to IP4) is under way. Some of the IP addresses can be reused for a number of machines at the same time by using a **subnet**. For example every IP address of the form 192.168.*.* can be used this way. In a typical house, the router is tightly coupled to a computer at the exchange which has its own fixed IP address and all traffic has to go along that path. Most routers by default are given the IP address 192.168.1.1 (occasionally 192.168.0.1) and will give out numbers 192.168.1.* numbers to all the computers, phones and tablets that connect to it. When packets are sent the router changes the from address to that of the machine at the exchange and adds something to the packet so when a response appears it can forward the packet to the correct device.

Numeric addresses are not very memorable, so to make them more understandable and to help impose a hierarchy that makes the system easier to manage the concept of **domains** is introduced. This means an address can take the form of (e.g.) **shrewsburyu3a.org.uk** . Translation to and from the numeric address is done by a group of specialized computers known as Domain Name Servers (DNS). When you type a domain address into your browser, the first step is to contact the most convenient DNS server which will give you the proper numeric address. The last part of the domain address is a **Top Level Domain (TLD)**, there is one per country and others such as **com, org, edu, mil, net and gov**. These have recently been joined by many others such as **online, coffee, doctor** etc. These top level domains are managed, along with other aspects of the Internet by the Internet Corporation for Assigned Names and Numbers (ICANN) which is a nonprofit corporation based in Los Angeles. Each country will then manage the subdomains of its TLD, and second or third level domains will be handed out to companies or organisations who will hand them out in turn to subdivisions until eventually you get to one specific computer.

Various other protocols run on top of the Internet protocol such as SMTP (Simple Mail Transfer Protocol), FTP (File Transfer Protocol) and HTTP (HyperText Transfer Protocol).

Some History

In 1963 a certain J. C. R. Licklider introduced in a paper the concept of the "Intergalactic Computer Network" and shortly afterwards he was appointed to ARPA (sometimes DARPA) the Advanced Research Projects Agency of the US Department of Defense. Over the next few years the concept of packet-switching networks was developed and by 1969 the ARPANet was being demonstrated, connecting together the DoD computers with those of some of its major contractors in the Los Angeles area and then bringing in several universities and spreading to sites on the East Coast. The first transatlantic link was actually by satellite to Norway, followed by other European countries. The two protocols that still form the backbone of the system, TCP/IP (Transmission Control Protocol/Internet Protocol) were developed in 1974. A few separate systems were developed, such as JANET (Joint Academic NETwork) in the UK, but by the late 70s/early 80s they had all merged into what we currently just call the Internet. I first started using systems connected to JANET in 1982 when access was mainly restricted to academics and researchers in a few companies. One of the major early requirements was the ability to connect documents. Technical documents tend to have a lot of cross references and interdependencies so the ability to switch quickly from one document to another and back was very important. A number of universities were researching into these hypertext systems, I was peripherally involved with one called Microcosm being developed at Southampton University by Wendy Hall (now Dame Professor Wendy Hall) but the one that made the final cut in 1990 was the one developed by Tim Berners-Lee at CERN, (now also at Southampton) who developed HTTP (HyperText Transfer Protocol), a markup language to insert hyperlinks into document called HTML (HyperText Markup Language) and developed the first web browser to display the documents and follow the links. This system became known as the World Wide Web. It is managed by an international organisation known as the **World Wide Web Consortium (W3C)** with Tim Berners-Lee still there as director.

The Web has carried on developing as its uses have gone beyond research into commerce and everyday life. It still uses HTML, now at version 5. Web pages have become dynamic, that is the browsers are able to run programs internally to make things happen on a page. This was first accomplished by running small programs known as **applets** written in a language called **Java** (so called because its developers drank a lot of coffee!) Nowadays programs are written as scripts within the HTML of a web page using a language called (incorrectly) **javascript**. Much prettier pages are made using a system called **CSS** (Cascading Style Sheets.) Many web pages are not written directly in HTML but instead run a program on the web server that generates HTML on the fly. Finally pages become more responsive using a system called **Ajax** which lets the browser and the server communicate and update a page without having to reload it all.

Web Browsers

The Web Browser is the piece of software that enables the user to access pages on the WWW and navigate between them; browsing or “surfing” the net. It will also handle some other net related activities such as file transfers. For many people it is just about the only software they use. A Chromebook has an Operating System that is built around Google’s Chrome browser.

The first browser to become popular, and help popularise use of the WWW was called Mosaic and came out in 1993 but was soon replaced by an improved browser called first Mosaic Netscape, then just Netscape, and finally Netscape Navigator. Microsoft had been slow in realising the public demand for the Internet and WWW browsers but by 1995 it had brought out Internet Explorer. IE was provided free with Microsoft Windows from Windows 95 on, so it quickly became the most popular browser; Netscape did not become free until 1998 and since then virtually all browsers are provided free. The so called “browser wars” hotted up as Microsoft was accused of misusing its dominant market position to disallow competition among browsers. Eventually MS had to allow users of Windows a free choice of browsers, but still managed to push its own browsers to maintain a dominant position. The Mozilla corporation that produced Netscape later developed Firefox and in 2008 Google brought out Chrome, which rapidly became the most popular browser.

A big problem has always been that the different browsers behave slightly differently when rendering a web page, so great efforts have been made to standardize the behaviour. While MicroSoft had a dominant position with not just its free browser but also a free server IIS (Internet Information Services). It used its own mechanism for dynamic pages, DirectX and had its own scripting language, VBScript and even modified the HTML and CSS so that the many companies that opted to use the Microsoft server delivered web pages that could only be properly rendered by a MS browser. Even companies that used other servers had to produce what was in effect two versions of every page, one for Microsoft browsers and one for everybody else. Fortunately the pernicious dominance of Microsoft has waned, but even its modern browsers while claiming to obey the standards actually continue to behave idiosyncratically. Many developers, particularly in small companies, have reached the stage of refusing to make all the small changes necessary to make the pages work properly in IE or Edge and if customers call support they are told to install one of the other browsers. Microsoft finally realised that it was losing custom and the latest versions of Edge are basically Google Chrome with some modifications. Unfortunately the modifications to try and maintain backward compatibility with previous versions and to supply their own “look and feel” have resulted in an inferior product, so you are better sticking to the actual Chrome rather than its poor relation.

The browser works by typing the full address of a site into the address bar at the top. Modern browsers will fill in part of it for you, and make guesses at what you are trying to type. The full

address is known as a URL (Uniform Resource Locator), not to be confused with a URI (Uniform Resource Indicator) which may look exactly the same but does something subtly different!

<https://www.facebook.com/groups/shrewsburyu3a/>

This has three parts, the protocol being used, which on the web will usually be http or https (the more secure version), followed by the domain name (www.facebook.com) and then some path information leading to a specific file on the designated server.

Web Sites

If you want to create a website you need a domain name, and a server computer to host your site. Domains are rarely bought directly from the agencies that manage them, but from intermediary companies such as GoDaddy, 1&1 and namecheap; they typically cost between £10 to £25 per year depending on the popularity of the top level domain, some of which have restrictions on who they accept. These companies will often also offer to host the domain for you, or you may choose to rent a server or space on a server for amounts from £10 per month or so upwards. Connecting the domain to the server is a bit fiddly unless you stick to the one place for both, but there is plenty of help available online.

The site can be built by writing code directly in HTML, CSS and javascript, plus a server language such as PHP, or you can use one of the common Content Management Systems (CMS) such as Wordpress which make things a lot easier but still need a bit of effort to administer, or you can use the built your own site sites such as 1&1, wix or weebly. These are much simpler to use but the payback is that you are much more restricted on the sort of site you can build.

This is a very simple site just to illustrate what the plain html looks like.

<http://mullsoft.co.uk/u3atalk.html>

The current Shrewsbury U3A website is an example of a build it yourself site.

<https://shrewsburyu3a.weebly.com/>

The new Shrrwsbury U3A website uses Wordpress, This is still a prototype, it will not be fully available until September when the URL will change.

<http://mullsoft.co.uk/shrewsburyu3a>

Social Media

It is worth making a special mention of social media. Nothing seems to divide people so much, often on generational lines. Some people spend a lot of time on social media maintaining contact with a wide variety of people who may be scattered all over the world. Some thing it is

the worst thing that ever happened and people should stick to email if they really must use electronic communication. If anyone is interested I wrote a paper about measuring how “close” people are when they may be in constant touch but physically a long way apart.

<https://delphi.lexxion.eu/article/DELPHI/2019/1/4>

It is true that the main company, Facebook, is not a particularly pleasant company, and your personal data, what little you allow them to have, will probably be sold on to dubious companies using push advertising, but the same could be said of most of the companies that we deal with from banks to grocers. It is worth remembering that Facebook and other social media sites are very quickly putting newspapers and other forms of mainstream media out of business, so reports are not exactly unbiased.

Conclusion

Nowadays it seems everybody needs to have a presence on the web from a local trader and small clubs up to news providers, banks and big ecommerce sites like Amazon and Ebay. The amount of business and commerce done on the Internet/Web is ever increasing and we see results in, for example, the loss of shops in town centres and the disappearance of newspapers. Whether this is a good thing depends on your viewpoint. For those of us over 60ish who grew up in an era without the Internet this can all seem almost disastrous; no shops, no newspapers and young people who only talk to each other using a mobile phone and never socialize “properly”. To those who have grown up with the web this is all perfectly natural; they cannot see what the fuss is and cannot wait to see more change and the final disappearance of the dinosaur society. Change has always happened and often we forget that in the whole span of human societies things like newspapers and even town centres full of shops are quite recent arrivals. The web continues to develop, the latest development, rather incorrectly termed the Semantic Web, seeks to connect individual items of data rather than complete documents. You can see the effect when, for example, you look at a map on your phone which knows your location and shows you all the items of potential interest in the immediate vicinity.