



SUMMER 2005 NEWSLETTER

URL SYNTAX DEMYSTIFIED – When you type a web address like “www.kwom.com” into your web browser, you are specifying a “URL”. URL stands for “Uniform Resource Locator” (pronounced “you are ell”, not “Earl”). A URL specifies a resource as well as how to reach that resource. URLs also appear as clickable hyperlinks in webpages and documents. Usually a URL takes you to a website, but they can also be used for other resources like email addresses, Usenet newsgroups, or ftp sites.

The general syntax for a URL is “<scheme>:<scheme-specific parameters>” where the scheme specifies an access method, and the parameters after the colon depend on the scheme. Examples of schemes are file, mailto, news, telnet, ftp, http, and https. Note that a valid URL will always contain a colon.

Some schemes (e.g. mailto and news) do not directly connect to a resource on a network. Typically you have already set up your computer with a server (e.g. a mailserver) that contacts the network resource on your behalf. All you need to specify after the colon is the information the server needs to locate the resource. A mailto URL needs an email address, e.g. “mailto:help@kwom.com”. A news URL needs the newsgroup name, e.g. “news:chi.internet”.

URLs for schemes that connect directly to a resource will look like “<scheme>://<hostname>/<path>”. This is where the dreaded “colon-slash-slash” comes from. The colon is the separator after the scheme, and the double slash indicates that a path to a resource on a TCP/IP network will follow. A special case is when the resource exists on your own computer, so the hostname is absent, e.g. “file:///C:/filename”.

A more typical example is “http://www.kwom.com/index.html” which specifies the use of HTTP (HyperText Transfer Protocol) to access the file “index.html” at hostname “www.kwom.com”. But you probably just type “www.kwom.com”. Why don’t you need to type the full URL?

Most web browsers (e.g. Internet Explorer, Netscape, Safari, Firefox) assume you want to access websites using HTTP unless you tell them otherwise. So if you leave out the scheme and the colon, the web browser will add “http://” at the beginning of the URL. Most of the time this is what you want, but if you are going to a secure website you may have to manually type “https://” at the beginning of the URL.

So your web browser will automatically change “www.kwom.com” to “http://www.kwom.com”. Technically this is still not valid URL syntax because there is a missing slash after the hostname. But this is such a common mistake that webservers will automatically redirect you to “http://www.kwom.com/”.

Note that we have still not specified a path (filename and optional directory path). Whenever you point to just a hostname, or to just a directory on that host, the webserver will search for a default homepage file in that directory. Typically the default file searchlist will contain entries like index.html, index.htm, default.html, home.html. If no default file is found, some webservers will display the directory contents.

So when you type “www.kwom.com”, your web browser and the webserver fill in the missing syntax and actually take you to “http://www.kwom.com/index.html”.

What about “www”, does it have a special meaning – that you are going to a website? Actually, no it doesn’t. “www” is just a hostname at the domain “kwom.com” (see more about domains below). If we wanted to have our website at “www2.kwom.com” or “fred.kwom.com” we could.

So can you leave out the “www” and just type the URL as “kwom.com”? Most of the time this will work, but only because the domain owner has created a DNS record (see more about DNS below) pointing a missing hostname to the same place as hostname “www”. It is best to always type the hostname (e.g. www) and not assume that using the domain name as a URL will take you where you want to go. For example, until recently “kwom.com” would take you to our mailserver not our website.

URLs can also specify usernames, passwords, TCP port numbers, etc. The general syntax is “<scheme>://<user>:<password>@<hostname>:<port>/<path>;<param=value>?<query>#<anchor>”. An example is a website that uses a different TCP port number than the default HTTP port 80. To remotely administer a Linksys router, the URL might be “http://172.16.1.1:8080”. In this case, if you just type “172.16.1.1:8080” your web browser sees a colon and doesn’t try to fill in the missing scheme for you. So when specifying a non-standard TCP port number, you will have to type the http:// at the beginning of the URL.

DOMAIN NAMES – Devices with IP addresses are referred to as hosts, with a hostname that can be looked up to determine the numeric IP address of the host. Hostnames are read hierarchically from right to left, and usually have the form “host.domain.tld”, where “tld” stands for “Top Level Domain”.

The original TLDs were .com (commercial), .net (internet), .org (nonprofit organization) and .edu (education). New TLDs have been added including .biz, .info, .name and .ws. There are also country-code TLDs like .us, .ca, .uk and .ru. Many people think .tv stands for television, but it is actually the country-code for Tuvalu. For a list of country-code TLDs see <http://www.iana.org/cctld/cctld-whois.htm>.

Each TLD has a registry listing the owner and authoritative DNS servers (see below) for each domain under that TLD. There will also be one or more registrars who (for a fee) can register a domain for you and maintain the registry entries. Domain registration typically costs between \$8 and \$35 per year, depending on the TLD, the registrar, and how many years you pay for at once. Network Solutions, GoDaddy and Register.com are some popular registrars.

If you are thinking about registering a domain, in most cases you will want the TLD to be “.com”. If your website is www.widgetsrus.net, some people will type www.widgetsrus.com out of habit. Domain names can only contain letters, numbers and hyphens, and as a practical matter should not be longer than 15-20 characters or people will get frustrated trying to type them correctly. Don’t assume that a domain name like canariesandparakeets.com will get you indexed under the keywords canaries and parakeets in the search engines. Search engines look at the text on the page, the title metatag, and in some cases the keywords metatag, but they don’t particularly care about the domain name.

Some people think that paying to register a domain also gets you website design and hosting. Not true! You also need to budget for a web designer and for webhosting. The hardest part though is usually deciding on the information you want on your website and how to organize it. Some people assume they can design the website themselves. This is certainly possible if you have the time and talent, but it is more involved than most people realize. If you are thinking about purchasing web design software, the most popular is Dreamweaver from Macromedia, another choice is GoLive! from Adobe. You may also need graphic editing software such as Photoshop.

DNS – DNS stands for “Domain Name System”. DNS servers (also called nameservers) are like online phonebooks for Internet hostnames and IP addresses. All routing on the Internet is done by numeric IP address, so when you type www.kwom.com into the Address bar of your web browser, your computer must first contact a DNS server to look up the corresponding IP address (69.49.197.130). DNS servers also do reverse lookups to find the hostname corresponding to an IP address, and MX record lookups to determine what mailserver accepts mail for a domain.

Each domain must have at least 2 authoritative DNS servers that all other DNS servers can contact to look up hostnames in that domain. Your registrar or hosting provider can provide DNS for your domain.

Your computer needs to know the IP address of a DNS server to look up hostnames. For dialup or DSL connections, it is usually best to have your computer obtain the DNS server addresses automatically when it connects using the PPP protocol. Occasionally we get a call from a customer who is having slow or intermittent performance looking up addresses, and it turns out they have their computer set up manually to use another ISP’s DNS servers. This will usually work, but may cause puzzling problems whenever the other ISP has DNS server problems. KWOM’s DNS servers are listed on our website.

If you ever need to determine your IP address, go to: http://www.kwom.com/my_ip.pl

EMAIL ADDRESS SYNTAX – All email addresses have the form user@domain.tld. If it doesn’t contain an @ sign and at least one dot, it isn’t a valid email address. Some people try to send email to a website, e.g. www.kwom.com. This doesn’t work – you can’t send email to a website address. Another mistake is leaving out the tld, e.g. sending email to help@kwom. The proper email address is help@kwom.com.

EMAIL NOT ADDRESSED TO ME – If we deliver an email to your mailbox, it was addressed to you, even if you don’t see your email address on the “To:” line. Our mailserver would not deliver mail to you unless you were specified as one of the recipients, although you may not have been the primary recipient. Probably some spammer included you in the blind copy (bcc:) list.

GOOGLE MAPS – Now in addition to Yahoo Maps and Mapquest, there is a beta of Google Maps at <http://maps.google.com>. (Remember we said the hostname doesn’t have to be www?) The user interface is very cool, especially if you have a broadband connection, and you can switch between map and satellite image modes.