Anatomy of an URL (Uniform Resource Locator)

An URL is a standardized way to describe how to locate a resource (like a web page or a file).
It is mainly use to find a document or a page on the Internet. An URL has multiple components. Here are a few examples:

Description	Protocol	Separator	FQDN (Fully Qualified Domain Name)					Separator	Path			Separator	Page		
			Host	Separator	domain	Separator	TLD (Top Level Domain)		Directory	Separator	Sub-directory		Page Name	Separator	File type
URL	https	://	www	.	apalala	.	com	/	en	/	news-views	/	events	.	html

Example 1

Description	Protocol	Separator	User & password to authenticate on the server				FQDN (Fully Qualified Domain Name)					Separator	TCP Port	Separator	Path				Page		
			Username	Separator	Password	Separator	Host	Separator	domain	Separator	TLD (Top Level Domain)				Directory	Separator	Sub-Directory	Separator	Page Name	Separator	File type
URL	ftp	://	Johndoe	:	Pa££W0rd!	@	ftp	.	apalala	.	be	:	5000	/	en	/	archives	/	file	.	zip

Example 2

Description	Protocol	Separator	Fully Qualified Domain Name					Separator	Page			Separator	Parameters						
			Host	Separator	domain	Separator	TLD		Page Name	Separator	File type		Parameter	Separator	value	separator	Parameter	Separator	Value
URL	http	://	my	.	apalala	.	com	/	index	.	php	?	page	=	index	&	session	=	56

Example 3

Description	Protocol	Separator	Fully Qualified Domain Name						Separator	Page		Separator	Bookmark
			Host	Separator	Subdomain	Separator	domain	Separator	TLD		Page Name	File type		
URL	http	://	www	.	login	.	apalala	.	be	/	index.htm	#	Chapter1		

Example 4

An URL for accessing the content of a physical book could be:

Morse_code://107.Albertstreet.Brussels.Belgium/Livingroom/library/URLfordummies.book#Chapter1

Anatomy of an URL (Uniform Resource Locator)

- **The protocol**:
  When surfing on the web, you will most likely encounter HTTP (Hyper Text Transport Protocol) and his encrypted more secure version, HTTPS. The protocol is a standardized way of communicating between the host (also called server) and the client (you).
- **The host:**
  It's the name (or alias) given to a server (or sometimes a set of servers) belonging to a specific domain (or entity). The most common name for the web server of a domain is *www*, although it is not mandatory. If we use the analogy of the location of buildings, that's the equivalent of the number of the building in the street.
- **The subdomain(s)**:
  Sub-domains are subdivision of a domain. They are often created to enable delegation of authority to a subsidiary of a department. Multiple level of subdomain can be created, like *team.department.enterprise.com.*
- **The domain:**
  The domain is the identifier of an entity. Most often, it's the name of the enterprise or organization. The domain is always linked to a Top Level Domain (or TLD, see below). Google is likely one of the most known Domain name and the company, Google, has associated it with a series of TLD like .com, .be, .fr or .uk. using the same analogy as above, it is equivalent to the street name.
- **The Top Level Domain:**
  The Top Level Domains are defined and maintained by the IANA (Internet Assigned Numbers Authority). They are meant to group domain based on a category: COM is for commercial entreprises; ORG for non-profit organization; GOV for US government; NET for Internet Access Providers and Internet Services Providers; BE, FR, UK, IT, and so on for country related domains, and so on. While there was originally very few TLD, IANA has recently allowed the creation of a very large list of TLD (visible on IANA's site)
- **A Fully Qualified Domain Name:**
  A FQDN is made of the TLD, the domain, the possible sub-domain(s) and the host if any is mentioned. It is the information used by DNS (Domain Name Service) servers to point you to the physical address of the host you are willing to reach (using the IP address that is associated with this FQDN in the DNS server)
- **Directory & subdirectory:**
  are essentially the same thing you have on your computer's disk. A way to organize information. Although, directory and subdirectory on web servers do not always reflect the logical organization of the file on the host.
- **Page name & file type:**
  Here also, we are close to what you use on a PC. Documents have a name and a file extension. On a web server, file extension are associated with the kind of content they provide and how they are processed. HTML for web page content, CSS for stylesheet, PNG & JPEG for pictures and so on.
- **The interrogation mark (?)**:
  The interrogation mark is a specific separator that allow to pass parameters to the host in order to receive a dynamic content (meaning that change with the context, like when you type a work to search for in Google). After the mark, you will often see a few pair of *name and value (Name=value) separated by & symbol.*
- **Bookmark:**
  Bookmarks are a reference within a web page, allowing you to go directly to the specific content in the page.