06 10 2016   **SEWN 2016 labsheet 1 (assessed coursework)**
                  Search engine optimisation and syndicating dynamic content

This labsheet requires you to design, create and publish a website that is to be included in search engine results pages (SERPs) for a specific search query, and also to syndicate some dynamic content from an on-line data source. Later in labsheet 2 you will your rank the data sourced from your web service using a popular text weighting metric in information retrieval.

Part 1: Ranking web pages in SERPS

Your first task is to understand how to make web sites and web pages *visible* to web search engines.

1. Research how search engines crawl and rank web pages, and find out how to optimise a web page, in terms of its design and structure, so that it will appear in SERPS, particularly Google and Bing. The following page is a good start, but do try to find some additional information too:

    http://www.searchenginewatch.com/showPage.html?page=2167961

2. Using the information found during your research for (1), plan the design, structure and content of your home page. Your aim is to ensure that Google or Bing return your home page as high as possible in their listings when you search for the keywords:

    <yourfirstname> <yourlastname> MSc (or BSc if appropriate)

    Before you begin you may want to refresh your memory on the basics of HTML and Web page design: pages http://www.w3schools.com/html/ and http://www.htmlcodetutorial.com/ may be useful for this.

3. Next, create your own personal home page to be published on the college web site. You can use any text/html/web editor for this purpose, e.g. Notepad++, TextPad or other text editor. Your page can be as simple or as complex as you choose in terms of content, but using JavaScript or CSS to decorate your page will not gain additional marks. Save your page as **index.html**.

4. Create a separate page, linked from your home page, optimised for the search phrase:

    web page optimisation for SERPS Birkbeck

    We want you to be able to enter the term above into Google and Bing, so that you can see your optimised page highly ranked in SERPS. Therefore, please use any and all *above-board* techniques that you have found out about to achieve a high ranking. Also, with regard to the use of letter *s* in *optimisation*, you are free to use the alternative spelling of *optimization*.

5. Publish your website to folder **public_www** in your **H: drive** and save all your website files to it. Your web site should now be viewable at:

    http://titan.dcs.bbk.ac.uk/~<your username>

    Where <your_username> is replaced with your DCSIS username.

    If you get a *Forbidden* message when trying to view your pages then you will need to contact Systems Group to allow your pages to be read by all users. Systems Group can be found in room MAL261 or contacted at http://www.dcs.bbk.ac.uk/about-us/staff/systems-group/ or by email at sg@dcs.bbk.ac.uk.

6. To make your pages visible on the WWW, as soon as your site ready, i.e. you can use the URL above with your username to view your pages, send a message to martin@dcs.bbk.ac.uk with the URL. A link containing the URL will then be added to page:

    http://www.dcs.bbk.ac.uk/~martin/sewn/sewn_2016_students.html

    Making your pages accessible on the WWW.

7. Check Google and Bing with the query text above, and when your pages appear send another message to martin@dcs.bbk.ac.uk with the following details:

a. The name of the search engine(s) on which it appears.
b. The ranking (1st, 2nd, 3rd, etc.) of your pages in the results list for each of the two search engines.

Part 2: Syndicating dynamic content

Frequently, content from many web services is *aggregated* on-line into a single *feed* which is *syndicated* to users via devices known as *readers*, e.g. Flipboard or Quartz News. We require you to select and display, i.e. syndicate, dynamic content from a single web service for this part of the labsheet. You are free to use an RSS feed or other data source to satisfy the requirements for either (1) or (2) below:

1. RSS ('Really simple syndication' or 'Rich site summary') feeds provide an open method of making website content machine-readable and easily sharable, i.e. RSS is used to syndicate website content. RSS feeds most commonly consists of frequently updated works such as blog entries, news headlines, audio and video media and HTML. More information is available at:

   - http://blogs.law.harvard.edu/tech/rss (on the RSS v2.0 Specification).
   - http://www.faganfinder.com/search/rss.shtml (on RSS in general).

   Add a new page to your website that includes the content syndicated from an RSS feed of your choice. There are several ways to do this: you could try one of the following or another tool:

   - Use the JavaScript RSS Box Viewer at http://p3k.org/rss/ to create a script to include on your web page.
   - Use FeedWind at http://feed.mikle.com/en/

2. Many people and organisations maintain accounts, or post content, to social networks like Twitter, Google+, Facebook or other web services. So we would like to find out about their APIs (Application programmable interfaces) in order to display content from one of them on a new page on your web site. More information is available at:

   - https://dev.twitter.com/docs/embedded-tweets
   - https://developers.google.com/+/api/latest/activities/list#examples

For this part of the assignment, it may be that not all services will work in all browsers and some RSS links on pages may only work when served by a web server. It is also possible that some of the above will require use of JavaScript or accounts with the social network providers concerned. It is best to try your pages from the PCs in the laboratories at the college.

The web service you choose **must provide a significant quantity of data.** This is because you are required to rank data sourced from it in labsheet 2 using a popular text weighting metric in information retrieval.

What to hand in:
1. Notify martin@dcs.bbk.ac.uk with the URL link to your home page, as soon as the page is live (Part 1.6).
2. A further message(s) as soon as your pages appear in the Google and Bing results (Part 1.7).
3. A short (of no more than an A4 page) report in .doc(x) or .pdf format. The report should briefly outline the SEO techniques you used to make your pages rank highly for Part 1 above, and for Part 2, describe why you chose your web service, and potential problems and possible solutions of combining content from several web services into a single display. Illustrate your points with an example. Any references to books or on-line material for this labsheet should be included in your report.

Submission deadline:
1. The report by 20 10 2016, to be submitted via Moodle using the **Labsheet 1 Search engine optimisation and syndicating dynamic content** drop box.
2. E-mail notifications of (1) and (2) for Part 1 when appropriate.

Late assignments: No extensions are available as for this labsheet and any late submissions will be graded as per the guidelines of the relevant course being studied.