Electronics and
Computer Science

UNIVERSITY OF
Southampton

COMP3220 Web Infrastructure

Laboratory 1: HTTP

This laboratory is designed to give you practical experience of constructing HTTP requests and interpreting HTTP responses. In it, you will investigate the verbs offered by HTTP, and the behaviour of key HTTP headers. This laboratory is not assessed.

Tools

In order to complete this worksheet, you will need to use curl.

curl is a command line tool that allows the transfer of data using various protocols (including both HTTP and HTTPS). curl is installed on iSolutions Windows machines, and on the ECS uglogin server. curl needs to be run from a UNIX shell (on macOS, use terminal), or from the Windows command shell (available from the Start menu) or PowerShell. A tutorial for curl is available online at https://curl.haxx.se/docs/manual.html

In this worksheet, you will be inspecting HTTP requests and responses. The `-v` flag will display the whole request that was sent and the headers that were received. You should use this flag for the exercises below.

```
curl -v http://en.wikipedia.org/
```

If you are having difficulty scrolling through the output from curl, you can use the `-I` option to suppress the returned resource representation.

The `-X` flag will allow you to use HTTP methods other than GET

```
curl -v -X DELETE https://en.wikipedia.org/
```

The `-L` flag will follow 3xx redirects:

```
curl -v -L http://en.wikipedia.org/
```

The `–H` flag allows the sending of arbitrary headers. In the example below, sending a User-Agent header claiming to be a mobile device triggers the sending of mobile-friendly sites for some websites – try it on the exercises below.

```
curl -v -H "User-Agent: Mozilla/5.0 (iPhone; CPU iPhone OS 6_1_3 like Mac OS X)
AppleWebKit/<WebKit Rev> (KHTML, like Gecko) Chrome/<Chrome Rev> Mobile Safari/<WebKit
Rev>" -L http://en.wikipedia.org/
```

Exercises

In the following exercises, you will be constructing HTTP requests using curl for both HTTP and HTTPS. For each exercise, you should make a note of the method and headers that are sent, and the status code and headers that you receive.

1. Retrieve a representation of http://www.ox.ac.uk/
2. Retrieve a representation of https://en.wikipedia.org/
3. Retrieve a representation of http://www.debian.org/
4. Retrieve a representation of http://www.ox.ac.uk/frotz
5. Retrieve a representation of http://google.com/
6. Retrieve a representation of http://www.google.co.uk/
7. List the methods that are supported by http://www.google.co.uk/ (use the OPTIONS method)
8. Delete https://www.ecs.soton.ac.uk/
9. Create a new resource https://www.ecs.soton.ac.uk/frotz
10. Retrieve a representation of https://www.debian.org/ in British English (use the Accept-Language: header with -H)
11. Retrieve a representation of https://www.debian.org/ in German
    (use the Accept-Language: header with -H)
12. Retrieve a representation of https://www.google.com/teapot
13. Retrieve a representation of http://id.southampton.ac.uk/building/59
14. Retrieve a representation of http://data.southampton.ac.uk/building/59
15. Retrieve a representation of http://data.southampton.ac.uk/building/59 in application/rdf+xml format

Advanced Exercises

Construct requests for the exercises above using nc and openssl, as shown in lectures.

References

Hypertext Transfer Protocol – HTTP/1.1 – Semantics and Content
http://tools.ietf.org/html/rfc7231

curl manual
https://curl.haxx.se/docs/manual.html