|  |  |  |  |
| --- | --- | --- | --- |
| **Title:** | Initial Outline System Specification for Secure Facilities Sharing Tool | | |
| **From:** | Adam Field, Christopher Gutteridge | **Date:** | 13/03/2013 |

### Document History

|  |  |  |  |
| --- | --- | --- | --- |
| **Version** | **Date** | **Comments** | **Contributors** |
| 0.1 | 08 Oct 2012 | Document created. | Adam Field |
| 0.2 | 09 Oct 2012 | Modifications from CJG email correspondance. | Adam Field, Christopher Gutteridge |
| 0.3 | 31 Oct 2012 | Data specification tables split off into new document | Adam Field |
| 0.4 | 02 Nov 2012 | Clarification of Open Data. | Adam Field |
| 0.5 | 13 Mar 2013 | Added curl example and Scheduling Uploads section. | Adam Field |

### Background

A tool is to be developed to provide a service to the S5 group of universities (Cambridge, Imperial, Oxford, UCL and Southampton) as a means to share information about RCUK Costed facilities with a view to share usage of them.

### Data Formats

See “Data Specification for Secure Facilities Sharing Tool.”

### Data Submission

A simple web front-end to the tool will be constructed to allow the submission of the data via an HTTP POST. Users will be able to submit using command-line tools (e.g. curl) or with a simple web form on the site. The site will run over HTTPS, and will use HTTP basic authentication for security. Some simple checks will be made on the data at time of upload, and an appropriate HTTP response will be returned.

### Data Retrieval

A table containing facilities data from all submitters will be available for download from the web front-end, either from a simple web-page or via command-line tools from a predictable URL. It is also important that the original data from any given institution would also be downloadable in a similar way. Downloading these csv files would require a user account on the system. Note that only CSV is provided by the service.

### Open Data

We have included a column in both the submission and the retrieval table to indicate whether the data about the facility should be considered public. This would be very strictly ‘opt-in’, and would allow the data to be reused under the OGL or CC0 lecenses. Any rows submitted as such would be aggregated into a table of open data which would be publically available.

### Security

As there will be few users of the system (we envisage each institution getting a single username and password), HTTP basic authentication is likely to be sufficient. We could also restrict access to S5 Institutions’ IP ranges if it was deemed desirable.

### URLs

A neutral domain name will need to be chosen for this service. Once that is done, we propose to use the following paths below that URL:

|  |  |
| --- | --- |
| URL | Description / Notes |
| /index.php | The front page of the service, containing a form to allow the uploading of data, links to allow the downloading of data, and some brief documentation (including contact details). Also processes CGI requests. |
| /template.csv | The URL from which an empty table (with column headings) can be downloaded. |
| /southampton.csv | The URL to which a data table can be submitted with an HTTP POST (or PUT). Also the URL from which an institutions’ data can be downloaded |
| /status.csv | Provides a CSV file containing details about when each institution last successfully uploaded data. |
| /all\_data.csv | The URL from which the aggregated table can be downloaded. |
| /public\_data.csv | The URL from which the aggregated table of open data can be downloaded. |

Alternatively, CGI parameters can be used with the index.php script:

|  |  |  |
| --- | --- | --- |
| Parameter | Request Type | Description / Notes |
| file | GET | The file to retrieve, one of:   * template.csv – Download a blank template. * status.csv – Download a list of institutions and the date on which they last submitted data. * [institutionid].csv – Download an institution’s latest submitted CSV file. * combined.csv – Download the aggreagate file of all institutions’ data. |
| file | POST | The file being uploaded. |
| institution | POST | The institution for which the file is being uploaded. |

### Example: Curl

An example of the use of the curl command to upload a file is given below:

curl -u <username>:<password> -F file=@</path/to/file.csv> -F institution=<institution> 'https://url\_of\_service.com/index.php'

The service uses a standard CGI POST request whose arguments are 'institution' and 'file'. There are a vast number of technologies that can interface with this, so the above should be considered only one way of doing it.

### Scheduling Uploads

There is a very small chance of a malformed spread sheet being downloaded if a request is made for the Combined Data table at the same time as a new file is uploaded. If daily updates are made by scripts at predictable times, it may be a good idea to set aside a time window purely for the uploading of files in order to avoid scheduling conflicts. For general 'by hand' uploading and downloading, it is vanishingly unlikely that this will ever happen.