Running your own CGI Programs on the University Web Server
IT Documentation Team. Reviewed July 2015.

1. Introduction

CGI (Common Gateway Interface) programs provide a way to generate Web pages "on the fly" or to do behind-the-scenes processing of Web forms. CGI programs are commonly written using scripting languages such as Perl or compiled languages such as C. Nowadays, CGI programs have almost entirely been replaced for most Web applications by scripting languages such as PHP (www.php.net). Unless you know you need to develop CGI-style programs, using PHP instead is highly recommended.

Scripting languages such as Perl, Python or PHP will mostly run quite happily on the University’s Web server. Notice the servers may not be running the most update to date versions and not all available add-ons may be installed.

Compiled CGI programs must be in a format compatible with the University’s RedHat Linux Web servers. Programs intended for use on Windows or Mac cannot be used. Programs compiled on the central Solaris Unix server (sysa.abdn.ac.uk) cannot be used. IT Services does not provide a Linux server that can be used to compile C-language or similar programs.

A much easier to way to provide interactivity or other processing of web-based data is to use the PHP scripting language that’s a core part of the University's Web servers. Any file with an extension .php stored in your Web pages folder (h:/public_html) will be treated as PHP code.

In many Web sites, the running of CGI programs is restricted to those programs installed in the central Web server’s own cgi-bin disk folder. However, the University’s Web service provides a means for individual users to run CGI programs, stored in their own filespace, using a wrapper program (cgiwrap). This ensures that CGI programs may access only files that belong to their authors or ones that have specific access permissions set on them.

Your CGI programs must be stored in a sub-folder, cgi-bin, of the folder holding your HTML Web pages (this will be public_html).

This guide shows the changes you may need to make to get CGI programs, acquired elsewhere, to work here. It assumes you have some knowledge of Linux.

Some general information on CGI programming can be found at http://www.htmlgoodies.com/beyond/cgi/.

2. What you need to do

1. First visit the Web page at homepages.abdn.ac.uk/local/publish/ and complete the publishing process there.

   This process creates a folder (that will be on your H: drive) named public_html. This folder will hold all your Web pages. Inside that folder will be another folder named cgi-bin which will hold your CGI scripts. Place your executable files in the cgi-bin folder.

2. You can then use the following URL to run your CGI program:
   homepages.abdn.ac.uk/cgi-bin/cgiwrap/login/program
   replacing login with your central systems username and program with the filename of your CGI program.

3. Other URL formats

CGI programs can have data passed to them by adding extra information to the URL than invokes them. For example, the Linux environment variables PATH_INFO and QUERY_STRING are standard CGI variables and can assigned values as in these examples, the first sets QUERY_STRING and the second PATH_INFO.

   homepages.abdn.ac.uk/cgiwrap/login/script?this-is-somedata
   sets QUERY_STRING to “this-is-some-data”

   homepages.abdn.ac.uk/cgiwrap/login/script/this-is-somedata
   sets PATH_INFO to “this-is-some-data”

As CGI programs are executed on the Web server indirectly via the cgiwrap program, the URL format required will probably differ from any descriptions given in the CGI program’s own documentation.

Overleaf are some examples of the formats needed by the central Web server and the Linux environment variables they create. The CGI script being run is called test-cgi, the system running the browser is set as mypc.abdn.ac.uk.
Example One
Documented format (returns PATH_INFO):
http://www.mysite.edu/cgi-bin/test-cgi/dir1/dir2

Local format:
http://homepages.abdn.ac.uk/cgi-bin/cgiwrap/login/test-cgi/dir1/dir2/

Environment variables:
- QUERY_STRING: ''
- SCRIPT_NAME: '/cgi-bin/cgiwrap/login/test-cgi'
- PATH_INFO: '/dir1/dir2/'
- PATH_TRANSLATED: '/info/www-data/dir1/dir2'
- REMOTE_USER: '<NULL>'
- REMOTE_HOST: 'mypc.abdn.ac.uk'
- REMOTE_ADDR: '139.133.77.111'

Example Two
Documented format (returns QUERY_STRING):
http://www.mysite.edu/cgi-bin/test-cgi?query_string

Local format:
http://homepages.abdn.ac.uk/cgi-bin/cgiwrap/login/test-cgi?query_string

Environment variables:
- QUERY_STRING: 'query_string'
- SCRIPT_NAME: '/cgi-bin/cgiwrap/login/test-cgi'
- PATH_INFO: ''
- PATH_TRANSLATED: '/info/www-data'
- REMOTE_USER: '<NULL>'
- REMOTE_HOST: 'mypc.abdn.ac.uk'
- REMOTE_ADDR: '139.133.77.111'

Example Three
Documented format (returns PATH_INFO & QUERY_STRING):
http://www.mysite.edu/cgi-bin/testcgi/dir1/dir2?data1=value1&data2=value2

Local format:
http://homepages.abdn.ac.uk/cgi-bin/cgiwrap/login/testcgi/dir1/dir2?data1=value1&data2=value2

Environment variables:
- QUERY_STRING: 'data1=value1&data2=value2'
- SCRIPT_NAME: '/cgi-bin/cgiwrap/login/test-cgi'
- PATH_INFO: '/dir1/dir2'
- PATH_TRANSLATED: '/info/www-data/dir1/dir2'
- REMOTE_USER: '<NULL>'
- REMOTE_HOST: 'mypc.abdn.ac.uk'
- REMOTE_ADDR: '139.133.77.111'

4. Debugging your CGI Programs

When your CGI program is run by the Web server, it can be hard to discover what has gone wrong if you don’t get the output you expected. The only error you are likely to see is similar to this:

Server Error
The server encountered an internal error or misconfiguration and was unable to complete your request.

If the page you requested is generated by a CGI program (its URL must contain the character string cgi-bin) then the CGI program is faulty.

You should check with the author of the program or page as to the cause of the problem.

In the unlikely event that the server itself is wrongly configured please report the problem to the Service Desk and inform them of the time the error occurred, and anything you might have done that may have caused the error.

This is most likely a fault in your program rather than a configuration problem with the Web server it is running on. You should check your program using the cgiwrapd tool described next.
5. cgiwrapd

Cgiwrapd is another version of cgiwrap that helps you debug your programs.

To use it, replace the word cgiwrap in the URLs given above with the word cgiwrapd.

The result will be the normal output from your CGI program, preceded by a description of how cgiwrap processed your program and a list of the Linux environment variables and values that were made available to it.

6. Safe CGI

You should be aware that CGI programs can be insecure. They need to be written carefully to ensure that whatever data they process (eg data collected from a Web form) is validated for correctness so that no unexpected actions are then carried out by your program.

Any CGI program you create will run with the same system privileges you have - if you can delete accidentally one of your own files then so can your CGI program. So write your CGI program with care!

8. Common errors

There are a few errors that are often made when linking CGI programs to your pages:

Incorrect home directory:
Any reference you make in your programs to your own files (but not URLs) must use the full Linux pathname, not just the filename. For example, a file called xxxx.dat in your public_html (or public.htm) directory should be referred to as:

/home_directory/public_html/xxxx.dat

where home_directory should be replaced by the value given by the Linux command:

echo ~your_login

(replacing your_login by your own username).

It will look something like: /home/nof-a/tea123

The value returned by the Linux pwd command may depend on the machine you are logged into, so its value should not be used. You also cannot use locations such as ~/login/dir1/dir2/file in a CGI program, despite such locations being acceptable at the Linux command prompt.

Incorrect program locations:
CGI programs copied from other sites often make assumptions about the locations of system programs (such as Perl, Awk or Nawk). You will need to alter such programs before they will work (Perl is in /usr/bin/perl, gawk in /local/bin/gawk Python in /usr/bin/python).

Assumptions about relative URLs:
There is no “base URL” for a CGI program. If a CGI program tries to reference another page (URL) stored in the same directory, it will fail unless the reference is an absolute URL (i.e one that begins with a slash (/)).

Using Windows Notepad to edit a Linux script file:
Most Perl, Gawk or similar programs running on Linux have a first line that defines the location of the interpreter that will process the program. Using Notepad to edit such a file will append a Windows-style end of line marker to the location – resulting in an error message such as "Command not found" when running the program.