# QUERY STRING FORMAT

In <u>World Wide Web</u>, a **query string** is the part of a <u>Uniform Resource Locator</u> (URL) that contains data to be passed to web applications such as <u>CGI</u> programs.

The <u>Mozilla</u> URL location bar showing an URL with the query string title=Main_page&action=raw

When a <u>web page</u> is requested via the <u>Hypertext Transfer Protocol</u>, the server locates a file in its <u>file system</u> based on the requested <u>URL</u>. This file may be a regular file or a program. In the second case, the server may (depending on its configuration) run the program, sending its output as the required page. The query string is a part of the URL which is passed to the program. Its use permits data to be passed from the HTTP client (often a <u>web browser</u>) to the program which generates the web page.

**Structure**

A typical URL containing a query string is as follows:

      http://server/path/program?query_string

When a server receives a request for such a page, it runs a program (if configured to do so), passing the query_string unchanged to the program. The question mark is used as a separator and is not part of the query string.

A link in a web page may have a URL that contains a query string. However, the main use of query strings is to contain the content of an HTML form, also known as <u>web form</u> In particular, when a form containing the fields $field_1$, $field_2$, $field_3$ is submitted, the content of the fields is encoded as a query string as follows:

      $field_1=value_1$&$field_2=value_2$&$field_3=value_3$...

- The query string is composed of a series of field-value pairs.
- The field-value pairs are each separated by an <u>equal sign</u>.
- The series of pairs is separated by the <u>ampersand</u>, '&' or <u>semicolon</u>, ';'.

For each <u>field</u> of the form, the query string contains a pair field=value. Web forms may include fields that are not visible to the user; these fields are included in the query string when the form is submitted

This convention is a <u>W3C</u> recommendation. W3C recommends that all web servers support <u>semicolon</u> separators in the place of <u>ampersand</u> separators.

Technically, the form content is only encoded as a query string when the form submission method is GET. The same encoding is used by default when the submission method is POST, but the result is not sent as a query string, that is, is not added to the action URL of the form. Rather, the string is sent as the body of the request.

**URL encoding**

Main article: URL encoding

Some characters cannot be part of a URL (for example, the space) and some other characters have a special meaning in a URL: for example, the character # can be used to further specify a subsection (or fragment) of a document; the character = is used to separate a name from a value. A query string may need to be converted to satisfy these constraints. This can be done using a schema known as URL encoding.

In particular, encoding the query string uses the following rules:

- Letters (A-Z and a-z), numbers (0-9) and the characters '.', '-', '~' and '_' are left as-is
- SPACE is encoded as '+'
- All other characters are encoded as %FF hex representation with any non-ASCII characters first encoded as UTF-8 (or other specified encoding)

The encoding of SPACE as '+' and the selection of "as-is" characters distinguishes this encoding from RFC 1738.

Example

If a form is embedded in an HTML page as follows:

<form action="cgi-bin/test.cgi" method="get">
 <input type="text" name="first">
 <input type="text" name="second">
 <input type="submit">
</form>

and the user inserts the strings "this is a field" and "was it clear (already)?" in the two text fields and presses the submit button, the program test.cgi will receive the following query string:

first=this+is+a+field&second=was+it+clear+%28already%29%3F

If the form is processed on the server by a CGI script, the script may typically receive the query string as an environment variable named QUERY_STRING.