

NAME

s.htm – template and configuration for s.cgi

SYNOPSIS

/usr/local/aspseek/etc/s.htm

DESCRIPTION

s.htm acts both as a configuration file and as a HTML template for **s.cgi(1)**. It looks like HTML file, and is divided into sections. **s.cgi(1)** takes some sections, makes meta variables substitution in them, and outputs the resulting HTML.

Section is a piece of **s.htm** file that is started with line `<!--sectionname-->` and ends with line `<!--/sectionname-->`. Section consists of HTML formatted text containing meta variables. Meta variable is one or two letters preceded by \$ sign (example: \$DX). Meta variable names are case-sensitive. **s.cgi(1)** substitutes meta variables with corresponding string.

Special “variables” section

This is a special section in which you define various configuration parameters in the form "**Parameter value**". Unlike all the other sections, it starts with the string `<!--variables` and ends with `-->`. The contents of this section is never included in **s.cgi(1)** output.

The following parameters are defined here.

DaemonAddress *xxx.xxx.xxx.xxx[:nnn]*

Defines the IP address (*xxx.xxx.xxx.xxx*) and port (*nnn*) of search daemon **searchd(1)**. If port is not set, ASPseek's default port is used.

You *should* put at least one **DaemonAddress**. You *can* put several **DaemonAddress** lines; in this case **s.cgi(1)** will connect to all **searchd(1)** daemons, retrieve results from all of it, and show merged results.

Note that you can't give DNS names (like **www.myhost.com**) as an argument to **DaemonAddress**, only numeric IP addresses are allowed here. This is done because name lookup needed in case of DNS name will slow down **s.cgi(1)** considerably.

MaxExcerpts *num*

Defines the maximum number of excerpts that are shown in results.

MaxExcerptLen *num*

Defines the maximum length (in characters) of each excerpt string.

PagesPerScreen *num*

Defines the maximum number of links to other search result pages to be shown if there are many results found.

ResultsPerPage *num*

Defines how many results will be shown on one page. Note that this value is overwritten by **PS** cookie if set, and **ps** parameter to **s.cgi(1)**. Less value makes search results page to appear faster, so don't set it too high.

Clones *no*

If this line is present, clones detecting and showing is disabled.

You can also redefine error messages that are used in cases **s.cgi(1)** encounters an error. The following error messages are used.

ER_STOPWORDS *string*

Displayed if only stopwords are used in query.

ER_EXTRASymbol *string*

Displayed if some extra symbols were found at the end of query.

ER_EMPTYQUERY *string*

Displayed if the query is empty.

ER_TOOSHORT *string*

Displayed if the pattern is used, and not enough characters are provided in it. See also the description of **MinFixedPatternLength** parameter in **searchd.conf(5)**.

ER_NOQUOTED *string*

Displayed if unmatched quote was found.

ER_NOPARENTHESIS *string*

Displayed if unmatched parenthesis was found.

ER_NOSEARCHD *string*

Displayed if **s.cgi(1)** was unable to connect to **searchd(1)**.

META VARIABLES

Meta variables described here are **not** position-dependent, that is to say they can be used in any template sections. Variables that can be used only in some sections described in the **TEMPLATE SECTIONS** below.

\$\$ This is actually not a meta variable, but a way to include \$ sign to **s.cgi(1)** output. So, if you want to put a \$ in template, write it as \$\$.

\$AV ASPseek version (like 1.2.10).

\$P String "&tmpl=" concatenated with value of **tmpl** parameter of **s.cgi(1)**, URL-escaped, or empty string if there was no **tmpl** parameter given.

\$c Current value of **cs** (charset) parameter of **s.cgi(1)**. Works only in Unicode version.

\$f Number of first document displayed on this page.

\$fm Current value of **fm** (word forms) parameter of **s.cgi(1)**.

\$ps Current value of **ps** (results per page) parameter of **s.cgi(1)**.

\$s Current value of **s** (sorting type) parameter of **s.cgi(1)**. Can be either *rate* or *date*.

\$np Current value of **np** (page number) parameter of **s.cgi(1)**.

\$gr Current value of **gr** (grouping by site) parameter of **s.cgi(1)**. Can be either *on* or *off*.

\$ad Current value of **ad** ("account distance") parameter of **s.cgi(1)**. Can be either *on* or *off*.

\$bd Current value of **bd** (search in body) parameter of **s.cgi(1)**. Can be either *on* or *off*.

\$ds Current value of **ds** (search in description) parameter of **s.cgi(1)**. Can be either *on* or *off*.

\$kw Current value of **kw** (search in keywords) parameter of **s.cgi(1)**. Can be either *on* or *off*.

\$tl Current value of **tl** (search in title) parameter of **s.cgi(1)**. Can be either *on* or *off*.

\$CC Template section **cached** if document is text/html or text/plain, or template section **textversion** if document was converted from other format.

\$DB Document size in kilobytes.

\$DC Document Content-type (like *text/html*).

\$DD Document description (taken from META DESCRIPTION tag).

\$DE URL-escaped document URL. Useful for building "cached link".

\$DK Document keywords (taken from META KEYWORDS tag).

\$DM Document Last-Modified date.

\$DN Document number (in order of appearance).

\$DS Document size in bytes.

\$DT Document title.

\$DR Document rating (as calculated by **ASPseek**).

\$DU Document URL, unescaped.

\$DX Document text. If excerpts are enabled and found, several excerpts from URL content, otherwise the first couple of lines, to give an idea of what the document is about.

\$DZ Contents of template section **sizeb** if document size is less than 1024, otherwise contents of template section **sizek**.

\$l Number of last document displayed on this page.

\$t Total number of documents found.

Search statistics

\$W Information about the number of word forms found. For example, if query was where first number is the URL count and second number is the total count). You can use **\$W** only in **restop** section.

\$Y Total time spent to perform query.

The following meta variables gets substituted by corresponding template sections, if some conditions are met.

\$w Contents of **inres** template section if more than one result is found. **inres** template section should define either checkbox "Search within results" or link to "search within results" page.

\$V Contents of **navigator** template section, if more than one page of results are found, otherwise empty.

\$M

\$Mn Contents of **moreurls** template section, if grouping by sites is enabled and more than *n* results are found from the site, where *n* is number following **\$M**, usually 2. If *n* is not specified, it is set to 1.

\$R Contents of **ressites** section, if more that one site is found; otherwise contents of **resurls** section.

Navigator is a piece of HTML code used for presenting links to other search result pages. The following variables are used.

\$NL Contents of **navleft** section, if current page > 0, otherwise contents of **navleft0** section.

\$NB Contents of **navbar1** template section, repeated *pages per screen* times. For page number equal to current, contents of **navbar0** section is printed instead.

\$NR Contents of **navright** section, if this page is not last one, otherwise contents of **navright0** section.

\$NH URL of the another results page, used as a value for "<A HREF" in **navleft** and **navright** sections.

\$NP Number of page, useful in **navleft***, **navbar*** and **navright*** sections.

\$E Text of error message. See description of **variables** section for info about error messages.

TEMPLATE SECTIONS

top This section is included first on every page. So, you should begin this section with standard HTML preamble: <HTML><HEAD>... and so on. Here in this section you also provide search form.

The following meta variables are used:

\$Q HTML-escaped search query.

\$QF URL-escaped search query, concatenated with previous query if search was made within results).

\$q The same as **\$QF** if it is encountered in one of **res*** section, otherwise HTML-escaped search query, concatenated with previous query if search was made within results).

\$A Argument for FORM ACTION tag with host name and protocol (example: `http://www.aspseek.com/cgi-bin/s.cgi`).

Example:

```
<!--top-->
<HTML>
<HEAD>
  <TITLE>ASPseek: $Q</TITLE>
</HEAD>
<BODY>
```



```
<!--/exclose-->
<!--hicolors-->
ffff66
ff66ff
66ffff
ff6666
6666ff
66ff66
<!--/hicolors-->
```

Clones

Clones are HTML documents with the same contents. **ASPseek** found such documents by comparing calculated MD5 checksums.

The following meta variables are used:

\$CL List of clones.

Cached pages

The following section names are used to customize the appearance of "cached" pages (e.g. the pages with page contents stored in ASPseek's database).

cachetop

This section is included first on every "cached" page. The following meta variables can be used: **\$DU**.

Displaying error messages

The following template sections are used to display various error messages.

notfound

As its name implies, this section is displayed in case when no documents are found. You usually give a little message saying that and maybe some hints how to make search less restrictive.

Below is an example of notfound section:

```
<!--notfound-->
Sorry, but search hasn't returned results.<P>
<I>Try to produce less restrictive search query,
or check words spelling</I>.
<!--/notfound-->
```

error This section is displayed in case some internal error occurred while searching. See the list of errors in description of **variables** section. Use the **\$E** meta variable to print error message text.

Example of error section:

```
<!--error-->
<FONT COLOR="#FF0000">An error occurred.</FONT><P>
<B>$E</B>
<!--/error-->
```

queryerror

This section is displayed in case if query contains error like unmatched quote or parenthesis. You should use **\$E** meta variable here.

Example of error section:

```
<!--queryerror-->
<FONT COLOR="#FF0000">Error in query: $E</FONT><P>
<B>$E</B>
<!--/queryerror-->
```

complexPhrase

This section is displayed in case if boolean expression is used inside phrase. Intended just to warn user, and probably give some hints.

complexExpression

This section is displayed in case if boolean expression is used in query. Intended just to warn user, and probably give some hints.

Including a file into template

include *filename*

If line `<!--include filename -->` is encountered, then **ASPseek starts loading template definitions from file, specified by *filename***, then resumes processing of current file. Example of use: if two or more template definitions differ only in **top** and **bottom** sections, then create one file with sections other than **top** and **bottom**, then create two different files with **top** and **bottom** sections.

Special OPTION SELECTED substitutions

In order to save the value for option user selected across the different search result pages, the following "selected magic" hack is used. **s.cgi(1)** scans the template for all of `<OPTION>` tags in `<SELECT>` which do have attribute **SELECTED** followed by = sign and some value, and removes all of the **SELECTED** fields with the attribute different of that in **VALUE**.

In the following example the string **SELECTED** will only be presented in line that have its **VALUE** equal to value of **\$ps** meta variable.

```
<SELECT NAME="ps">
<OPTION VALUE="10" SELECTED="$ps">10
<OPTION VALUE="20" SELECTED="$ps">20
<OPTION VALUE="50" SELECTED="$ps">50
</SELECT>
```

Defining different output formats

You can have several section with the same name in template. Normally, the first encountered section is used. This behavior can be overridden by supplying **o=n** parameter to **s.cgi(1)**. If value of *n* is more than zero, then "*n+1*"th sections are used. If number of occurrences of particular section is less then value of "*n+1*", then last section with the needed name is used.

The following meta variables are used:

\$o Substituted with the current value of **o** parameter of **s.cgi(1)**.

User-defined template sections

Besides the standard template sections, user-defined sections can be used. **s.cgi(1)** treats line `<!--sectionname-->` as the start of user-defined template section and line `<!--/sectionname-->` as the end of user-defined template. Note that name of section must begin with alphanumeric character, otherwise this line is treated as regular HTML comment belonging to the current template section, so it is important to put space after "`<!--`" in comments.

User-defined templates can be used from other templates with the following meta variable.

\$T*sectionname*

Gets substituted with user-defined template section named *sectionname*.

Random number generation

This is a feature of **s.cgi(1)** that allows you to include random numbers to resulting HTML, for example in order to get a random banner. Numbers are generated in the range `[0; max]`, where *max* is equal to value of **RandomN=max** in **variables** section (see above). So, you put meta variable **\$rN**, which will be substituted to generated random number. *N* should be in the range 0...127, that means you can use up to 128 different random number variables.

In the example below **\$r1** will be substituted with random number in the range from 0 to 100, **\$r2** - from 0 to 500.

```
<!--variables
....
Random1=100
```

```
Random2=500
-->
....
<A HREF="http://www.my.com/click?id=$r1"><IMG
SRC=http://www.my.com/getbanner.cgi?id=$r1></A>
<A HREF="http://www.other.com/url?n=$r2"><IMG
SRC=http://www.other.com/pic?n=$r2></A>
....
```

BUGS

This documentation is incomplete. Not all variables can be used in all template sections. If you are unsure or see some strange misbehavior, consult the source code (file `templates.cpp`).

SEE ALSO

`s.cgi(1)`, `aspseek(7)`.

AUTHORS

Copyright (C) 2000, 2001, 2002 by SWsoft.
Man page by Kir Kolyshkin <kir@asplinux.ru>