

## CONSIDERATIONS FOR CREATING AND MAINTAINING A HOME PAGE ORIENTED WITH THE SAS® SYSTEM

Robert Schechter  
Web Master, Philadelphia SAS® User Group  
Zeneca Pharmaceuticals, Wilmington, DE

### ABSTRACT

According to the SAS Institute's web page (updated April 16, 1998) there are 27 SAS RUGs and LUGs with web pages. This paper will explain the design and coding considerations used in the development of the PhilaSUG site.

There are several fundamental considerations in designing a home page that transcend HTML (HyperText Markup Language, the code that describes the structure of a web document's content). These concerns include layout, design, and ease of use. The impacts of various browsers and HTML levels will be discussed. Finally, the basic HTML (level 2 - Frames, Style Sheets, and Dynamic HTML will not be covered) used to author the PhilaSUG site will be demonstrated using a popular browser.

### DESIGN

The key to a satisfactory Web page is good design, and design is not something that is done only once; rather, it is an ongoing process of maintaining a harmony between the appearance of the message and the content of the message itself.

### GOALS/PURPOSE

The first step in establishing any Web site is analysis of the intended content and of the audience that will access that material. Before anything is written, a clear understanding of the purpose is necessary, and from that understanding, the tasks to be undertaken through the Web site can be defined. Only after these things are done can a truly useful and organized site be developed. You probably want to:

- showcase your organization (state your organization's mission statement)
- provide a schedule of events (local, regional, national)
- present past and current newsletters
- make available past meeting agendas
- show the agenda, meeting date, map, and driving directions for your next meeting
- display a list of officers or steering committee members and contact information
- conduct user polls and surveys

### CONTENT

If the site is to be functional, it must contain information that is complete enough to answer the user's questions, current enough to be useful, and pertinent to the topic being sought out by those who seek access. Too much irrelevant information can make the essential data difficult to find and drive users away from the uninteresting clutter. Also, you should consider the types of information that you have access to that your site visitors may want to know. Consider different things that can be done to keep the Web site fresh and interesting, and thus to encourage return visits; perhaps features can be varied from time to time so that people will continue to visit in hopes of seeing a new page or department. Above all, keep the information up-to-date. Nothing is do discouraging to return visitors as a Web site with outdated information, and nothing is no indicative of neglect of the site. At the PhilaSUG site, not only do I keep the calendar of events current, I post the current local SAS Institute seminars and identify the current articles in Observations®<sup>2</sup> and SAS.COM®<sup>3</sup>.

Though it is permissible and recommended that phrases such as "for more information call/contact ..." be used, these phrases should not be used in lieu of providing detailed information. The very nature of the web demands current information, do not over use "under construction" signs and text. An alternative wording on the home page might be that "this web site is a work-in-progress and will continually be updated. Please check back regularly."

### STRUCTURE

By this time, the goals of the Web site and the topics to be addressed there should be well-defined; the next step is to consider the form the Web site is to take, and the structure that is to be imposed on it. And there are several possible choices of ways to do this.

A Web site can be laid out sequentially, just as one might lay out a newsletter; in that case, all topics would be addressed in sequence, with one subject leading logically to another. Some subjects may demand such an approach, but the potential of hypertext to join pages in the order the user desires makes great flexibility possible.

By constructing a careful outline of the topics and their interrelationships, one can create a hierarchy of documents and the links between them that have a logical flow for the user. There are no limitations in HTML on the number of levels or the complexity of

the links you can create. Begin with a menu including the major topics; each major topic will have its own menu of subtopics, which will be discussed or given menus of their own. It is important to keep the interface the same throughout all the levels (up, down, forward, back, return to previous menu, return to main menu).

Of course, it is possible to combine the two approaches and have, for example, a hierarchical structure with linear components. The PhilaSUG site started out as a hierarchical structure built from a rather rigid outline, but as I continued to cross-link various topics I soon ended up with a true web-like structure. If you want to combine different forms, and it is likely that you will, be sure to flowchart (or storyboard) and record the process you use. A clearly-designed structure will ensure easy access to information and logical presentation of the data.

## LAYOUT

The actual layout of the site has to be designed with ease of use in mind. It should never be forgotten that the typical user may not have the degree of experience with the subject that the person designing the Web site has. Thus, once the desired content has been decided upon, methods of convenient access to that content have to be developed.

No user should have to scroll and scroll through a long page to find a specific bit of information. However, there may be times, when the subject is a unified one, and the page is linked for easy navigation, when a long page may be quite acceptable. See PhilaSUG's Agenda page for an example of this technique.

It is usually a good idea to put a table of contents (TOC) on the home page. Each item in the table can be linked to the part of the document to which it refers. In this way, through hypertext, the TOC is used as both a tool for easy navigation of and as a map of the layout and the logic of the entire site. And so it is highly desirable that users be led from the general topics of the TOC to the progressively more detailed information that they need. Many sources recommend that the number of links from general to specific be kept under four; but, while that is a good goal, it may not always be possible when dealing with complex topics.

Design your layout to be understandable and readable. Each page should contain the same controls for navigation back and forth within the linked structure and for return to the higher levels. Often these controls are a set of clickable icons, but they may take some other form; the consistency is the important thing. Also, avoid technobabble for its own sake. If, because of the nature of the subject, your site is replete with acronyms or terms likely to be unfamiliar to some of your users, create links to their definitions. The first time SAS is mentioned link to their home page <sup>4</sup>, the same for NESUG <sup>5</sup>.

## PRESENTATION

A Web site should be visually attractive. One can find a nourishing meal set plainly on a plate, whereas the

same food would be more desirable if carefully presented with color and garnish. Even if a user is visiting only to find a specific bit of information, a site that is both easy to use and visually inviting is likely to tempt him to linger and look at the other things offered there.

However, care must be taken not to overdo or clutter the site, and it is almost always a serious mistake to obscure the genuine content of the document with unnecessarily complicated presentation material. One should also be mindful of the download time of the various segments and think carefully before including many unnecessary Java applets, graphics, audio files, or other items that slow the process out of proportion to their value. Many site designers are, I think, excessively insistent that content must prevail over the form of the site. These purists are right only if they are saying that a pretty site with little or no content will not be returned to; but many (and I count myself among them) feel that a bland site filled with plain but content-filled text will also fail to encourage users to return.

Positioning of the items on a page is a priority. The headers, the charts, the tables, and the text should all create a harmony of both meaning and visual logic. The site is there to provide data to the users, and the overriding concern of the designer is to see that the data is presented in as accessible a fashion as possible. A graphic that is poorly sized or poorly placed will detract from the likelihood of meeting this goal; great care should be taken with the placement and size of all the elements that compose the page.

And it must be remembered that not all browsers will display the site in the same fashion. True, most users do have one of the market standard browsers, but there are ten to twenty percent of users who do not, and these people should not be forgotten. Many people either have or choose to use a text-only browser mode, while others may lack many graphics forms or centering or the ability to read and create forms. Still others are restricted to Braille or synthesized speech. And so, no matter how excellent the presentation, one should remember that some people will not see anything but the data and design the site to make that data accessible to all, regardless of browser.

## GRAPHICS

Graphic images are expected on a well-designed Web page, and the designer must decide how many images are necessary, how many will add to the quality of the presentation, and how many may not be needed at all. A colorful graphic is almost always desirable on the home page, inviting and piquing the interest of the user. A graphic added to a message of welcome is usually desirable. But the pages devoted to the subtopics don't always need pictures, even though pictures can add to the page's appeal. The graphic should never be used merely because it is available; it should be used when it enhances the site.

The worst problems connected with graphics result from the large size of many of the files. Even in a highly compressed format, graphics files can make use or download of pages unwieldy. For this reason, it is

usually wise to use small versions (thumbnails) of any large graphics that are needed by the site. You can give the link to the larger file, preferably indicating the size of the file next to the link; the user can then decide if he wants to spend the time necessary for the download.

When a user decides to access a page with his browser, the procedure is standard. The text file is received along with the code indicating its layout. After the download of the text, the browser again makes calls to the server (the host computer of the Web site) to access and download the graphics files.

In older versions, the browser did not display the page on the computer until this entire process was complete. The delay was occasioned by the need of the browser to know the size of the images before the page could be laid out properly. Today, most browsers can interpret a special code the site designer places into the text file to transmit the size of the images. This code enables browsers to lay out and fill in the page before the actual image files are received and the graphics put in place. The designer must take the responsibility for seeing that this code is available to the browsers.

Image formats most commonly recognized by graphical web browser are GIF and JPEG. Newer formats, such as PNG are not widely supported yet. Formats commonly used on computers, such as TIF, BMP, PICT, and PCX tend to have fairly large file sizes (compared to GIFs and JPEGs), and would require significant download times if used. The GIF and JPEG formats are compressed files. The GIF format works best for line drawing and images having solid colors without blending of the colors. The JPEG format works best for photographs.

Many designers forget the availability of an ALT attribute for images. The ALT attribute allows the placement of alternative material, usually text, for any user whose browser does not or cannot view the graphics. In fact, even if there is no text necessary to explain the missing image, ALT can be used to suppress the presence of the image marker (desirable with text-only browsers) merely by using a blank for the value of the attribute (ALT="").

## PUBLICIZING

A Web site's existence does not guarantee it an audience. No one will visit if no one knows the site is there. There are several ways to promote a site, both on-line and off-line.

The first thing to do once the site is on-line is to have it listed by the SAS Institute<sup>6</sup>, and NESUG<sup>7</sup>. List the site's address on each issue your newsletter and announce it at your meetings. Register your site with the major Web data bases.

## CODING

HTML documents, also called Web pages, are the fundamental units of information organization and delivery on the Web. HTML (HyperText Markup Language) mixes ordinary text with special strings of

characters, called markup, that instruct browsers how to display HTML documents<sup>8</sup>.

HTML is a *markup language* that describes the structure of a Web document's content plus some behavioral characteristics. Web browsers can understand and interpret this language<sup>9</sup>.

The trick to understanding HTML lies in knowing how to separate the two components of the HTML file: the *content* (ordinary text) and the *controls* (controls are also known as *tags* or *markup*). If you look at an HTML source file, you see *markup* in the file that doesn't show up when your browser displays the page. Markup consists of the characters within the HTML bracket markers (< >) and controls how text appears on-screen<sup>10</sup>.

The first step in having a home page on the Internet is creating an HTML file. Use your text editor (e.g., NOTEPAD) to create a file called index.htm in a directory called my\_lug\_name. The first part of the name can be anything (though "index", "Homepage" or "home" are good identifiers), but the ".htm" or ".html" suffix is mandatory -- it identifies the file as a HyperText file that can be read by a Web browser. The first file in the directory should always be "index.htm" or "index.html" because most servers will look for this file first when the file is not included. Thus <http://members.aol.com/PhilaSUG1> becomes a shortcut to <http://members.aol.com/PhilaSUG1/index.htm>.

Many HTML controls are container tags, meaning there's text contained between a tag and its corresponding ending tag. The following is a list of some of the most common HTML container tags:

TAG	ENDING TAG	DESCRIPTION OF USE
<html>	</html>	<b>begins HTML document</b>
<head>	</head>	<b>begins a heading</b>
<body>	</body>	<b>body of text</b>
<h1>	</h1>	<b>level 1 heading</b>
<h2>	</h2>	<b>level 2 to 6 heading</b>
<ul>	</ul>	<b>a bulleted list</b>
<p>	</p>	<b>underlined item</b>
<b>	</b>	<b>paragraph</b>
<u>	</u>	<b>bolded text</b>
<i>	</i>	<b>underlined text</b>
<em>	</em>	<b>italicized text</b>
 	</br>	<b>emphasized text</b>
<nobr>	</nobr>	<b>inserts a line break</b>
<blink>	</blink>	<b>keeps text from breaking</b>
<center>	</center>	<b>text will blink</b>
<font>	</font>	<b>text will be centered</b>
<hr>	</hr>	<b>this tag allows the font to be changed according to several attributes</b>
<ol>	</ol>	<b>horizontal rule across page</b>
<li>	</li>	<b>creates indented lists with some sort of numbering scheme</b>
		<b>HTML method of</b>

```

identifying each item in
the list
<li>          </li>    first item
<li>          </li>    second item

```

For a relatively complete listing of the HyperText Markup Language elements, including their attributes and arguments, with their browser compatibility, I recommend *The Compendium of HTML*<sup>11</sup>.

Any good HTML reference book will explain the language elements from A - Z. What I will attempt to show you is the method I used to develop the PhilaSUG page from start to finish.

```

Type:
<html>
</html>

```

The HTML element identifies the file as containing HTML elements. An HTML document should have a HEAD and BODY elements between the HTML start and stop tags. A document type declaration can be placed at the start of an HTML document (before the opening HTML tag) which identifies it as a document conforming to a particular version and level of HTML. I.e. <!DOCTYPE HTML PUBLIC "-//IETF/DTD HTML LEVEL 0/EN">

```

<!DOCTYPE HTML PUBLIC "-//IETF/DTD HTML
LEVEL 3.0/EN">
<html>
<head>
</head>
<body>
</body>
</html>

```

The <head> section takes no free-flowing text but does allow other tags, one of which, <title>, is mandatory.

The <head>... </head> section contains information to describe the page to the reader (<title>), to search engines (<meta>), to establish its relationship to other pages (<link>) or to identify external scripting (<script>). It may also contain other information that is not considered part of the document (<base>, <isindex>, <style>, and <nextid>).

A printed book must always have some sort of title and chapters may have one as well. <TITLE> is the equivalent for HTML documents. It is a short phrase (6 words or so) that describes the content of the page. The information is displayed in the top of the window of the browser, or at the top of the screen. This tag is mandatory.

```

<!DOCTYPE HTML PUBLIC "-//IETF/DTD HTML
LEVEL 3.0/EN">
<html>
<head>
<title>PhilaSUG Local SAS Users Group Home
Page</title>
</head>
<body>
</body>
</html>

```

The <body> tag defines the "canvas" of the page. This is the place, on the reader's screen, where all text and graphics appear. It is possible to have multiple canvases by using frames or layers.

```

<!DOCTYPE HTML PUBLIC "-//IETF/DTD HTML
LEVEL 3.0/EN">
<html>
<head>
<title>PhilaSUG Local SAS Users Group Home
Page</title>
</head>
<body text="#000000" bgcolor="#FFFFFF"
background="philbk01.gif">
</body>
</html>

```

For the rest of this paper we will be adding additional elements between the <body> and </body> tags and will show only the new information.

Let's add a logo and page heading.

The <img> tag controls how that picture is displayed. There is no </img> tag. The align= attribute forces the alignment of the surrounding text with the image. Note the operative word here is text. The image sits where you place it, only the text moves, relative to the image. The alt= attribute provides a method of introducing text in the event that a browser cannot download in-line images and should always be used. The "height=" and "width=" attributes are considered quite important. They allow the browser to allocate sufficient space for the image without having to wait for the image to completely download before beginning to display the text on the page. The src= attribute specifies the source of the picture. It takes a "URL" as the argument.

The <hn> tag sets up a heading condition on the page. The size of the type is dictated by "n" and takes a value from 1 to 6. All sizes are relative to the specified (or default) base font.

```

<img align=left src=psa_logo.gif" alt="[PhilaSUG
Logo]" width="123" height="161">
<img align=right src=psa_logo.gif" alt="[PhilaSUG
Logo]" width="123" height="161">
<h1 align=center><font
color=#ff0000">PhilaSUG</font></h1>
<h2 align=center><font color=#ff0000">The
Philadelphia Area
<A
HREF="http://www.sas.com">SAS&reg;</A> Users
Group</FONT></h2>

```

I decided to hot link the word SAS to the SAS Institute home site. The <a> tag is used as a point at which a reader can click to "jump" to another place in the same page, another page or a specific place in another page. HREF attribute takes three arguments, "URL", "filename" or "#name" (all in quotes). Both cause the loading of an externally acquired file or when the reference points to a "#" (hash) entry, it points to target "#name" in the existing file. ISO-Latin-1<sup>12</sup> is available to express certain characters like the registered symbol (®).

Let's separate the heading from the rest of the body. The <HR> element specifies that a horizontal rule of some sort (the default being a shaded engraved line) be drawn across the page. The align= attribute forces the rule to display relative to the window margins. The size= attribute lets the author dictate how thick the rule should be and the width= attribute allows the author to dictate how long the rule should be.

<Center> </center> tags cause the text to be centered between the existing margins.

The <P> tag defines the start of a paragraph. The end tag is assumed if another block element is placed before another <p> or </p> tags. It should be noted that for CSS (Cascading Style Sheets) purposes it is wise to use the closing </p> tags. Some strange things might occur if this is not done.

The PhilaSUG Table of Contents could be presented as a list. HTML has special tags that permit you to create lists that include numbers, bullets, indentations, and other similar features.

HTML allows you to create a numbered, or ordered list. You can do this by using the <ol> and <li> tags. <ol> creates indented lists with some sort of numbering scheme. It supports the <lh> (list heading) and <li> (list item) tags. The <lh> provides a heading capability to lists, and the </lh> tag is optional. When one is creating a listing, it is imperative that each item in the list be identified. The <li> tag is the HTML method of identifying each item in the list. The </li> is optional. The <li> tag cannot work alone; it must be used in conjunction with <ol> or <ul> tags.

A numbered list, with a heading would be coded:

```
<ol>
<lh>PhilaSUG Table of Contents
<li>Call for Papers
<li>Calendar of Events
<li>The Current Newsletter
</ol>
```

HTML allows you to create a bulleted, or unordered, list with bullets automatically created. The technique is very similar to the numbered list technique shown above except that you would replace the <ol></ol> tags with <ul> </ul> tags. In HTML 3, you can replace the default bullet with a special graphic with the src= attribute, as in:

```
<ul src="ballcolr.gif">
<lh>PhilaSUG Table of Contents
<li>Call for Papers
<li>Calendar of Events
<li>The Current Newsletter
</ul>
```

For the PhilaSUG Table of Contents, I wanted the bulleted list with my own graphic, but I choose the unmarked list because it was compatible with HTML 2 and above. The <dl> tag is a form of annotated listing. The <dl> listing takes two additional tags: <dt> (the definition term) and <dd> (the definition description). I ended up coding:

```
<h2>PhilaSUG Table of Contents
<dl>
<li>
```

```
<b><a href="call4pap.htm">Call for
Papers</a></b></li>
<li>
<b><a href="calendar.htm">Calendar of
Events</a></b></li>
</dl>
```

For the local SAS Institute's seminars I had a collection of topics, locations, dates and times. HTML 3 supports tables, which allow you to control the rows and column placement of information on the Web page, by using the <table></table> tags. The <table> element is for the presentation of tabular material. A table is created by opening the table, creating rows within the table <tr> and cells <td> within the rows. If there were a "trick" to using tables, it would simply be "keep it neat and organized". The way I learned to use tables was to find a site with a table similar to what I wanted and view the source of that page.

## REFERENCES

1. [http://www.sas.com/events/seminars/seminar\\_dates.html](http://www.sas.com/events/seminars/seminar_dates.html)
2. <http://www.sas.com/service/doc/periodicals/obs/observations.html>
3. <http://www.sas.com/software/sas.com/current/>
4. <http://www.sas.com/>
5. <http://www.pace.edu/nesug/>
6. <http://www.sas.com/usergroups/link-form.html>
7. [Mailto:nesug-i@list.pace.edu](mailto:nesug-i@list.pace.edu)
8. Quoted from *HTML for Dummies*, 4
9. *Ibid.*, 24
10. *Ibid.*, 39
11. <http://www.htmlcompendium.org/sitemap.htm>
12. <http://www.letters.com/~gray/docs/latin.html>

## ACKNOWLEDGMENTS

I would like to extend my sincere thanks to JoAnn Matthews of the University of Pittsburgh Medical Center, for graciously allowing me to build on her prior paper, "Creating a Home Page on the Internet", and for her kind encouragement.

SAS is a registered trademark or trademark of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

## AUTHOR

The author may be contacted via e-mail at: [robert-schechter@worldnet.att.net](mailto:robert-schechter@worldnet.att.net) or at the address below:

Robert Schechter  
Zeneca Pharmaceuticals  
1800 Concord Pike, FOC CW3-30  
PO Box 15437  
Wilmington, DE 19850