

The free web system

“Simplicity”

as implemented on

SpaceTelescope.org

— Technical Documentation —

By

Lars Holm Nielsen

Lars Lindberg Christensen, ESA/ESO/ST-ECF

Erik Nordström Andersen

<http://www.spacetelescope.org/projects/web>

Last modified: 23 May 2004

Table of Contents

1	Introduction	3
2	Basic Definitions.....	4
2.1	Web page design and structure.....	4
3	Prerequisites/Reading guide	5
4	Dreamweaver.....	8
4.1	Templates	8
4.1.1	Templates Basics.....	8
4.1.2	Editing template parameters	9
4.1.3	Forcing an update of an html page or the whole site	10
4.1.4	Making a new section.....	11
4.1.5	Creating/Updating the submenu.....	13
4.1.6	Creating a normal page.....	15
4.2	Cascading Style Sheets	16
4.3	Synchronizing local and remote site	17
4.4	Special cases	17
5	The data	18
5.1	How the different image display sizes are made	18
5.2	How the different video formats are made	20
6	The metadata	21
6.1	Editing a metadata file with Excel	22
6.2	Priorities and sorting	23
6.3	Creating an Excel macro	23
7	Perl scripts.....	25
7.1	Archive scripts	25
7.2	Maintenance scripts	26
7.3	Common scripts	26
7.4	Generator scripts	27
7.5	Templates	29
7.6	Automatic front-page generation (frontpage.pl)	29
7.7	Editing Perl scripts on Windows.....	30
7.8	Source code for Perl scripts.....	30
8	.esoacc files	31
8.1	Redirects	31
8.1.1	Making a redirect.....	31
8.2	"File Not Found"-page and "Authorization Required"-page.....	32
8.3	Password protection.....	32
8.4	Example of an advanced .esoacc file with redirects and password	

"Simplicity" on SpaceTelescope.org

protection.....	33
9 Searching.....	34
10 Link checker	35
11 Crontab	36
11.1 Editing the crontabfile	36
12 Index	37

"Simplicity" on SpaceTelescope.org

1 Introduction

Simplicity is a web system originally designed for the production and maintenance of popular scientific web sites that contain large quantities of data (images, videos, posters, wallpaper and the like) and metadata (image captions, descriptions, sizes, priorities etc.). However the system may have a much broader use.

The necessary information for its implementation can be found at:

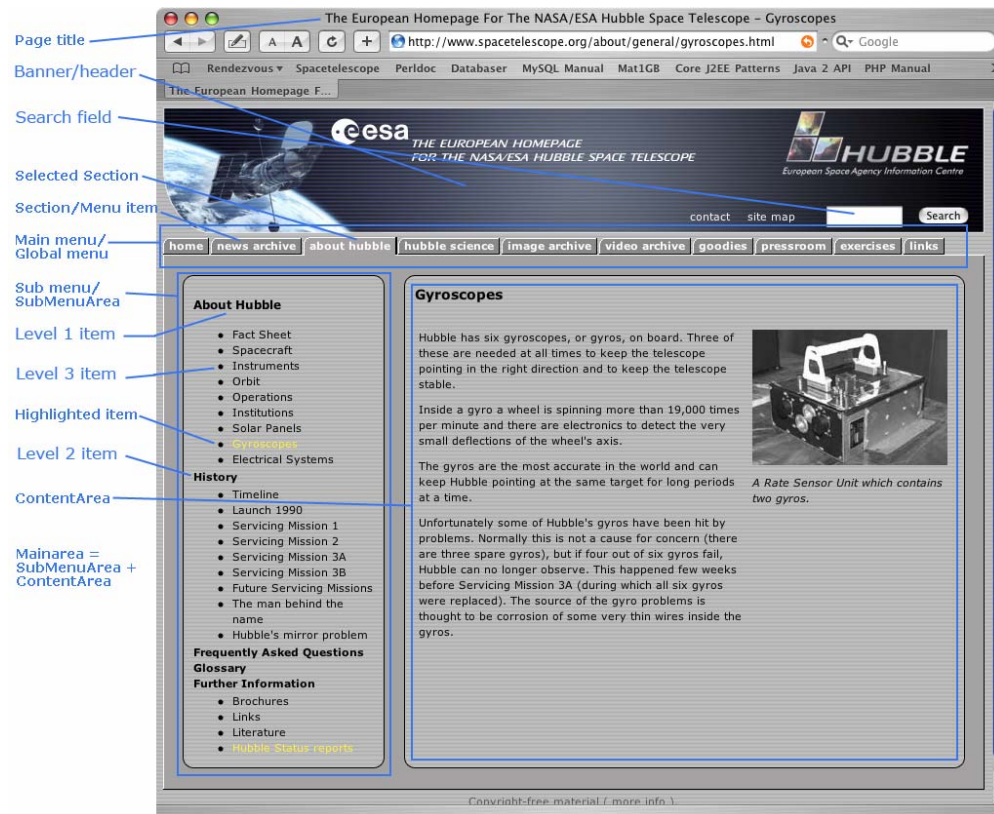
<http://www.spacetelescope.org/projects/web>

"Simplicity" on SpaceTelescope.org

2 Basic Definitions

2.1 Web page design and structure

The screen-shot below depicts a *sub page* of the web site. For the reading of this document it is important to note the names of the specific parts of a page.



"Simplicity" on SpaceTelescope.org

3 Prerequisites/Reading guide

Some preliminary advice and information:

- A basic working knowledge of **Macromedia Dreamweaver**, **Microsoft Excel** and **UNIX** is desirable to make the fullest use of this documentation.
- If possible, look at the files themselves while reading.
- Names of files, parameters etc. are typed in **Bold**, pieces of code with *Courier New* and commands, file paths with *Italics*.

The web site Spacetelescope.org has been used as an example for the implementation of the *Simplicity* web system throughout. Although most of the information here can be applied to any web site using *Simplicity*, there are some issues that are specific to Spacetelescope.org , [such as...](#)that serve only as specific examples of the consequences of decisions taken during the implementation process.

The spacetelescope.org web site has the following sitemap structure:

"Simplicity" on SpaceTelescope.org

Home

- Copyright Information

News

- List by year
 - 2004
 - 2003
 - 2002
 - 2001
 - 2000
 - 1999

About Hubble

Fact Sheet
Spacecraft
Instruments

ACS
WFPC2
STIS
NICMOS
FGS
(COSTAR)
FOC
WFPC1
FOS
HSP

Orbit
Operations
Institutions
Solar Panels
Old Solar Panels
Gyroscopes
Electrical Systems

- History

Timeline

Launch 1990
Servicing Mission 1
Servicing Mission 2
Servicing Mission

3A

Servicing Mission

3B

New solar panels
Replacement of the PCU
Nicmos returns!
A new coat for the telescope
A little boost
Future Servicing Missions
The man behind the name
Hubble's mirror problem

- Frequently Asked Questions
- Glossary
- Further Information
 - Brochures
 - Links
 - Literature

Image Archive

- Hall of Fame
- Zoomable Images
- Wallpapers
- View All
- Top 100

Video Archive

- Hall of Fame
- Video News Releases
- View All

Hubble Science

- The Deep Fields
- The age and size of the Universe.
- Stellar Evolution
- Our Solar System

- Black Holes, Quasars and Active Galaxies
- Formation of Stars
- The Composition of the Universe
- Gravitational Lensing
- Europe & Hubble

Goodies

- Calendar 2004
- Hubble Posters
- Wallpapers
- Print Layouts
- Best of Hubble Images
- Interactive Hubble
- How Hubble Images Are Made
- Hubble Image Experience

Pressroom

- Press Kits
- Embargo Site (requires

login)

- Interview Possibilities
- Video formats

Exercises

Projects

- Our Web System
 - Background article about

Simplicity (html)

Background article about

Simplicity (PDF)

Simplicity User Manual

(PDF)

Download components!

- FITS Liberator
 - Specifications (html)
 - Specifications (PDF)
 - Download components!
 - Educational data sets
 - Data archives
 - Submit your images
- IAU Working Group
- Science Communication Guide
- PR and Virtual Observatories
- An astronomy A/V Lab
- A Danish Communication Strategy

About Us

- HEIC
 - MISSION STATEMENT
 - World-class abilities
 - European-class abilities
 - HEIC Products
 - Our group
- Logo archive
- Evaluation
 - Press clippings
 - Web stats
 - Web news sites
- Contact Us

"Simplicity" on SpaceTelescope.org

Before constructing a new web site it is important to:

- Undertake appropriate market research to understand customer needs
- Plan the functionality of the web site. For example, an Image Archive should:
 - o Enable fast intuitive searches
 - o Support mini-thumbnails (~100 per page).
 - o Allow category-search with a clickable image map
 - o Place simple text search directly on the first page with radio buttons/drop down menus the most important options
 - o Provide an advanced search option for nearly all fields
- Plan the sitemap (i.e. which **Sections** to include)
- Identify the **archives** to be produced (images, videos etc.)
- Plan the **file formats** for data in the archives (tiff, jpg, gif, .mov, .mpeg) and **display sizes** (small, medium, large etc.) to be included
- Identify the metadata to be included with the archives - the fields (columns) to be included with the records (rows).

"Simplicity" on SpaceTelescope.org

4 Dreamweaver

4.1 Templates

The web system relies heavily on nested templates. These are all located in */Templates*, and have the following structure.

- **Global Design Template** (Contains the main design: header, global menu and main area)
 - **Global Subpage Template** (Contains the design for a normal sub page, i.e. a sub menu area and a content area)
 - **About Hubble Page** (Contains the menu for the About Hubble subsection)
 - **Exercises Page**
 - **Goodies Page**
 - **Hubble Science Page**
 - **Image Archive Page**
 - **Links Page**
 - **News Archive Page**
 - **Pressroom Page**
 - **Video Archive Page**

4.1.1 Templates Basics

Areas that can be edited by the user are defined and template parameters are set when a template is created.

An editable area:

```
<!-- TemplateBeginEditable name="MainArea" -->
<!-- TemplateEndEditable -->
```

Template parameters (three examples):

```
<!-- TemplateParam name="SelectedSection" type="text" value="home" -->
->
<!-- TemplateParam name="selectedleft" type="URL"
value="/design/menu_selected_left.gif" -->
<!-- TemplateParam name="unselectedleft" type="URL"
value="/design/menu_unselected_left.gif" -->
```

Template parameters are used in the template to define content based on the value of the parameter. In the following example:

"Simplicity" on SpaceTelescope.org

```

```

The image source is set to the value of the "selectedleft" template parameter if the value of the template parameter "SelectedSection" equals "goodies", otherwise it is set to the value of the template parameter "unselectedleft". In this way it is possible to create a template with a menu, where each menu item can be highlighted by setting the template parameter on a page that uses the template.

A nested template is a template that is itself based on another template. For instance, the **Global Subpage Template** is a nested template of the **Global Design Template**.

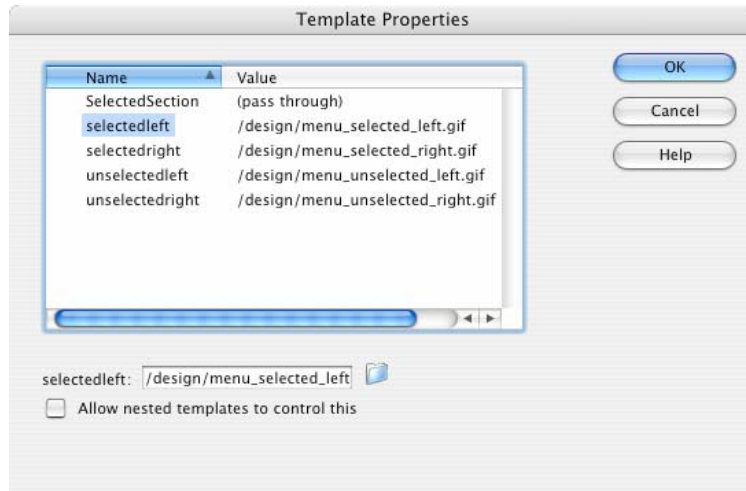
Parameters are can only be edited one level below the template. If, for instance, a template parameter **Param1** is defined in **Template 1**, and **Template 3** is based on **Template 2** that is itself based on **Template 1** then **Param1** can be changed in **Template 2**, but not in **Template 3**. If **Param1** is to be edited in **Template 3**, say, then the template parameter should be defined using the option `passthrough="true"`:

```
<!-- TemplateParam name="SelectedSection" type="text" value="home"  
passthrough="true" -->
```

4.1.2 To edit template parameters

- Open an html page based on the template.
- Choose the menu item Modify > Template Properties. A window lists all defined parameters.
- Edit the parameters or "Allow nested templates to control" the parameter by checking the option.
- Press OK. The template will now be updated.
- See image below.

“Simplicity” on SpaceTelescope.org



4.1.3 Force an update of an html page or the whole site

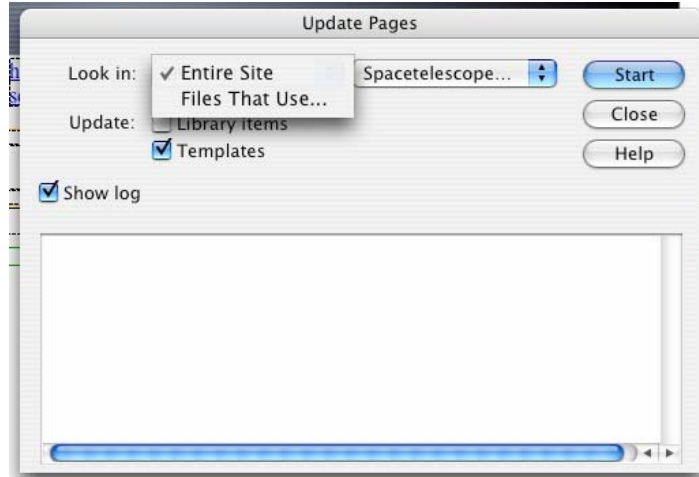
When a template is saved, either all pages using the template can be updated or single html pages maybe updated manually with the following:

- Open the page to be updated
- Choose the menu item *“Modify > Template > Update Current Page...”*

All pages in the site can be updated by choosing the menu item *“Modify > Template > Update Pages...”*. This allows work in progress to be saved, but will only update the pages that use the template when work is complete..

When using *“Modify > Template > Update Pages...”* the window below will appear. The entire site or only files that use a specific template may be updated.

“Simplicity” on SpaceTelescope.org



4.1.4 Making a New Section

To add another level 1 menu point or **Section**, first add the new section to the **main navigation** (together with other main sections such as **News Archive**, **About Hubble**, **Hubble Science**, **Image Archive**, and **Goodies**,.).

In **Dreamweaver** edit the **Global Design Template** and switch to “*Code View*”:

- Find the code block with several entries such as the one below (see also the image below):

```
<td></td>
<td class="mainMenuItem@@@( (SelectedSection == "about") ?
"selected" : "unselected" )@@@"><a href="/about/index.html"
class="mainmenu">about hubble</a></td>
<td></td>
```

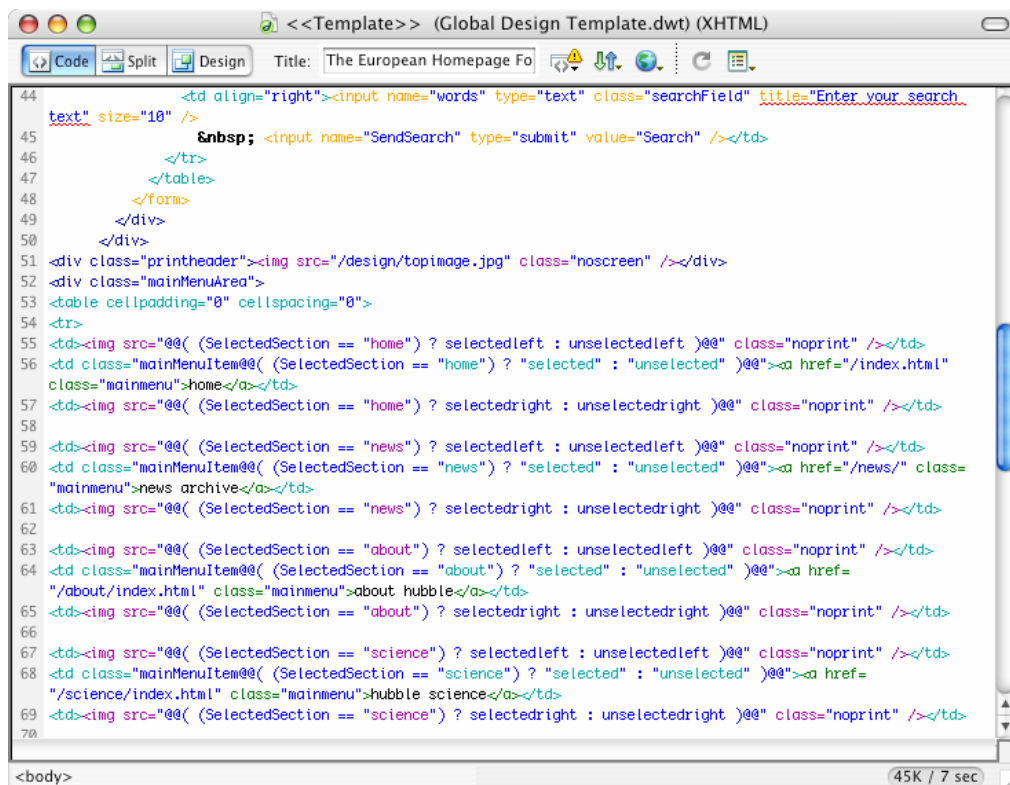
Each block of code like this is a Level 1 menu item, or **section**. All items are separated by one blank line.

- At the place where the section is to be inserted, make a copy of a code block from another section, ensuring that menu items are separated by one blank line in the HTML code.

"Simplicity" on SpaceTelescope.org

To edit the piece of copied code:

- Change "about" in the three places that say "SelectedSection == "about"" to a unique name - most logically the name of the section. Remember this name, as it is needed below.
- Change the link in "<a href="/about/index.html" class...".
- Change the text just after the link that says "about hubble" to the text that should appear as the level 1 menu item in the browser.
- Save the template and so update all pages that use this template. Since this is the **Global Design Template** all pages in the site use it, and will all be updated.



```
44 <td align="right"><input name="words" type="text" class="searchField" title="Enter your search
text" size="10" />
45 <td align="right"><input name="SendSearch" type="submit" value="Search" /></td>
46 </tr>
47 </table>
48 </form>
49 </div>
50 </div>
51 <div class="printhead"></div>
52 <div class="mainMenuArea">
53 <table cellpadding="0" cellspacing="0">
54 <tr>
55 <td></td>
56 <td class="mainMenuItem@"( SelectedSection == "home" ) ? "selected" : "unselected" )@"><a href="/index.html"
class="mainmenu">home</a></td>
57 <td></td>
58 </tr>
59 <tr>
60 <td></td>
61 <td class="mainMenuItem@"( SelectedSection == "news" ) ? "selected" : "unselected" )@"><a href="/news/" class=
"mainmenu">news archive</a></td>
62 <td></td>
63 </tr>
64 <tr>
65 <td></td>
66 <td class="mainMenuItem@"( SelectedSection == "about" ) ? "selected" : "unselected" )@"><a href=
"/about/index.html" class="mainmenu">about hubble</a></td>
67 <td></td>
68 </tr>
69 <tr>
70 <td></td>
71 <td class="mainMenuItem@"( SelectedSection == "science" ) ? "selected" : "unselected" )@"><a href=
"/science/index.html" class="mainmenu">hubble science</a></td>
72 <td></td>
73 </tr>
74 </table>
75 </div>
76 </div>
77 </body>
```

The new section now needs its own template.

To create a new html page based on the **Global Subpage Template**:

"Simplicity" on SpaceTelescope.org

- Edit the template parameter "SelectedSection" to highlight the menu-pane for the new section (see section 4.1.3). The value should be the name used in point 3 above.
- Edit the title (for instance by adding a "- Section Name" to the title).

There are two areas that can be edited: **ContentArea** and **SubMenuArea**. **ContentArea** contain the content of the pages.

- Add at least one editable area by choosing "Insert > Template Objects > Editable Region". **SubMenuArea** contains the sub-menu (see section 4.1.5 for how to create a sub menu)
- Save the file as a template ("File > Save as Template...").

4.1.5 Creating/Updating the submenu

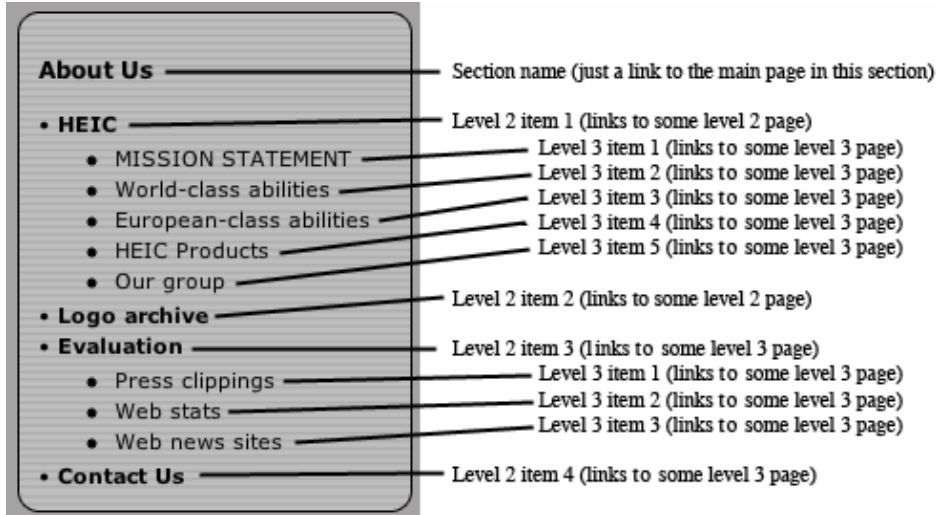
Creating or updating the submenu can be done essentially in "Design View" after the initial statement is inserted in "Code View". . The easiest way to create a new submenu is by copying and pasting the HTML-code from another submenu.

- Switch to "Code View" and insert the following line in top of the editable **SubMenuArea**:

```
<!-- TemplateParam name="SelectedMenuItem" type="text"
value="Default Name" -->
```

- Switch back to "Design View" and create the menu using the normal Dreamweaver tools. A menu should look like this:

"Simplicity" on SpaceTelescope.org



When using Return to go to the next line HOLD DOWN *Shift*, unless creating a list (level 3 items). This creates a `
` tag instead of a `<p>` tag.

Do not be concerned with the font size/style etc. These are changed by **Cascading Style Sheets (CSS)**.

- Now switch to "Code View" and go to the editable **SubMenuArea**.
- Edit each link using the following examples as a guide (menu items can be highlighted from a page based on this template):

The **Section** link should look like this (add the *class* attribute):

```
<a href="../link.html" class="submenuHeadline">About  
Hubble</a><br />
```

A **level 2** item link should look like this (add the class attribute, change the two "Menu Item 1" and add `• ` in front of the link):

```
&bull;&nbsp;<a href="../link.html" class="submenuItem@@"  
(SelectedMenuItem == "Menu Item 1") ? "Selected" : "" )@"> Menu  
Item 1</a><br />
```

"Simplicity" on SpaceTelescope.org

A `` tag should look like this:

```
<ul class="submenuList">
```

A **level 3** item should look like this (add the class attribute, and change the two Submenu Item 1s):

```
<li><a href="../link.html" class="submenuSubItem@@(
(SelectedItem == "Submenu Item 1") ? "Selected" : " ")@@">
Submenu Item 1</a></li>
```

Look at an existing **submenu** as a guide.

4.1.6 Creating a normal page

Choose "File > New" from the menu, and select the right template and press *OK*.

- Edit the title if needed (add a "- Page Name" to the title).
- Edit the template parameter "SelectedItem", to highlight the **submenu** item (see Section To edit template parameters 4.1.2). The value should be the unique name for the submenu item. Refer back to the template that the page is based on for the exact name if necessary, switch to "Code View" and go down to the link and locate the value. For instance in this submenu item:

```
<li><a href="../fact_sheet.html" class="submenuSubItem@@(
(SelectedItem == "Fact Sheet") ? "Selected" : " ")@@">Fact
Sheet</a></li>
```

the value is "Fact Sheet" and it is located in the "SelectedItem == "Fact Sheet" part.

- Edit the **content area**.
- Save the page.

Note: The page will not appear correctly in Dreamweaver as Dreamweaver cannot load the style sheet.

"Simplicity" on SpaceTelescope.org

4.2 Cascading Style Sheets

The web system relies heavily on Cascading Style Sheets. There are two style sheets, one for normal viewing and one for printing, so special care should be taken when updating the style sheets. The files are located in */style/style.css* (for normal viewing) and */style/printstyle.css* (print).

The HTML-pages should contain the following in the header, to tell the browser which to use for the screen and which for printing:

```
<link href="/style/style.css" rel="stylesheet" type="text/css"
media="screen" />
<link href="/style/printstyle.css" rel="stylesheet" type="text/css"
media="print" />
```

Note the `media` attribute. The `visibility` property whether items are seen on screen but not in print. For instance a `<div>`-tag containing a top image with lots of graphics can be hidden on the print sheet in the following way:

The `div`-tag in the html-file:

```
<div class="header">...</div>
```

CSS class in `style.css`:

```
.header {
    position: relative;
    width: 842px;
    height: 129px;
    background-image: url(../design/topimage.jpg);
}
```

CSS class in `printstyle.css`:

```
.header {
    visibility: hidden;
    position: absolute;
    top: 0px;
    left: 0px;
    width: 0px;
    height: 0px;
}
```

Note that in the print stylesheet, the `div`-tag is moved to the top-left corner and has a width/height of zero so that it doesn't interfere with visible elements.

"Simplicity" on SpaceTelescope.org

4.3 Synchronizing local and remote site

One of the advantages of **Dreamweaver** is its excellent local/remote file handling, or *synchronization*. This keeps the local set of web files and the remote set on the 'live' server the same. An alternative is to use *ftp*, but **Dreamweaver** appears more user friendly. There are different options for synchronizing and users should experiment to find the method that suits them best.

Note two important issues:

1. When synchronizing, one of Dreamweaver's features is to change the timestamp on the local files to the time of upload. This makes it difficult to keep track of file versions etc., but cannot, to our knowledge, be avoided.
2. If some of the folders in the web site contain a large number of files (more than 10,000 or 50,000), special care must be taken to avoid synchronizing these since it would take too long (hours). A feature called *cloaking* can make certain directories invisible, but must be done locally, and thus the remote directories (with presumably just as many files) are searched during the compare (which may take hours). The only workaround we know is to *mark* and *synchronize* only parts of the web site (e.g. all directories with less than 50,000 files).

4.4 Special cases

The cost/benefit issue of supporting Netscape 4.x users on Spacetelescope.org was investigated. It was found that an unreasonable amount of work was required to make the web site work on this platform, largely due to known bugs in Netscape 4.x' and because the browser does not conform to globally accepted HTML-standards. It was decided not to support Netscape 4.x and instead to redirect these users to */netscape.html* (as Netscape 4.x actually crashes if the browser window is resized while viewing the web site).

"Simplicity" on SpaceTelescope.org

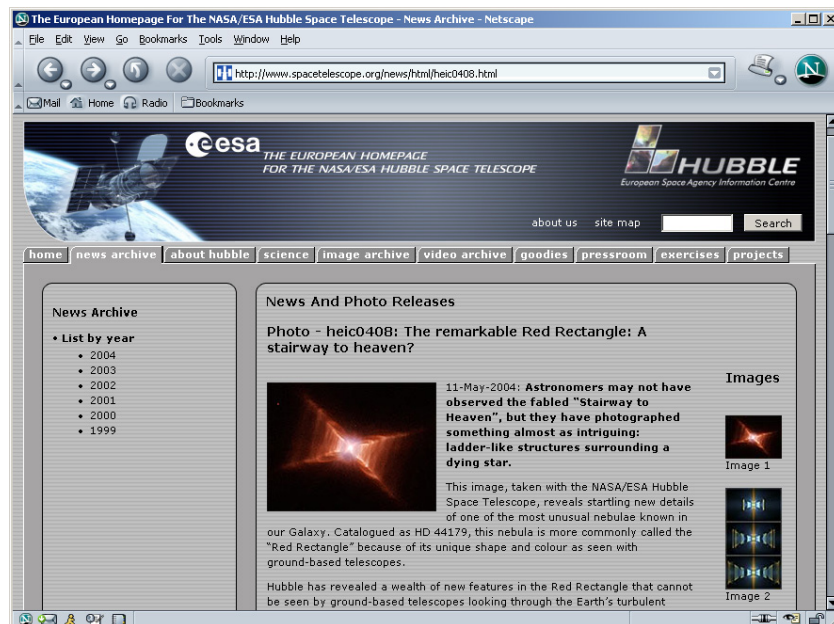
5 The data

The following types of data exist on Spacetelescope.org:

- Images placed in */images/*
- Videos placed in */videos/*
- News placed in */news/*
- Brochures placed in */about/further_information/brochures*
- Logos placed in */about_us/logos*
- Children's drawings in */kids_corner*
- Calendars in */goodies/calendars/*
- CD-ROMs in */goodies/cdroms/*
- Posters placed in */goodies/posters/*
- Print layouts in */goodies/printlayout/*

5.1 How the different image display sizes are made

The web site Spacetelescope.org uses 16 different image display sizes.



Here two of the different image display sizes are seen (News to the left, and Newsmini to the right).

"Simplicity" on SpaceTelescope.org

The different display sizes that are generated from the original images are (W x H in pixels):

1. **original:** "Fullsize original" (tiff/jpg/gif)
2. **large:** "Large jpeg" (orig x orig), jpeg quality 80. ImageReady.
3. **screen:** "Screensize jpeg" (1280 x *), jpeg quality 60. ImageReady.
4. **medium:** "Medium jpeg" (320 x *), jpeg quality 60, unsharp mask radius 1, Threshold 0, Amount 50. ImageReady.
5. **thumbs:** "Thumbnail jpeg" (max(x,y) = 122), jpeg quality 45, sharper, unsharp mask radius 1, Threshold 0, Amount 31. ImageReady.
6. **mini:** "Mini jpeg" (max(x,y) = 50), jpeg quality 45 sharper, unsharp mask radius 1, Threshold 0, Amount 50. ImageReady.
7. **wallpaper1:** "Wallpaper 1024 x 768", jpeg quality 80. Photoshop.
8. **wallpaper2:** "Wallpaper 1280 x 1024", jpeg quality 80. Photoshop.
9. **wallpaper3:** "Wallpaper 1600 x 1200", jpeg quality 80. Photoshop.
10. **postcard:** (500 x 333), jpeg quality 60. Photoshop.
11. **round:** image for the front page (82 x 82), transparent gif, perceptual, 256 colour, no dither, pattern dither transparency. ImageReady.
12. **news:** (180 x *), jpeg quality 53, sharper, unsharp mask radius 1, Threshold 0, Amount 50. ImageReady.
13. **zoomview:** (400 x 400), 2DhyperNav, qual.: 8 high, tilesize 256, Optimize tables. Photoshop.
14. **wallpaperthumbs:** "Thumbs 122 pix" created from wallpaper1 jpegs, jpeg quality 45, sharper, unsharp mask radius 1, Threshold 0, Amount 31. ImageReady.
15. **newsmini:** (60 x *), jpeg quality 45, sharper, unsharp mask radius 1, Threshold 0, Amount 31. ImageReady.
16. **hofthumbs:** For Hall of Fame (95 x *), jpeg quality 45, sharper, unsharp mask radius 1, Threshold 0, Amount 31. ImageReady.

These different display sizes are generated with either **Adobe ImageReady CS** droplets (for the simple resizing formats) or with **Photoshop CS** droplets that use Photoshop's built-in **Javascript** facility (for the formats that need 'intelligent' choices or logic).

"Simplicity" on SpaceTelescope.org

Some of these formats are also used in other archives (see above).

5.2 How the different video formats are made

The web site uses five different types of main videos:

1. Small QT: 180 x 144 pixels, Quicktime Sorensen-3 encoded (needs Quicktime 4 or 5)
2. Small MPEG: 180 x 144 pixels, MPEG-1 encoded
3. Medium QT: 360 x 288 pixels, Quicktime Sorensen-3 encoded (needs Quicktime 4 or 5)
4. Medium MPEG: 360 x 288 pixels, MPEG-1 encoded
5. Broadcast: 720 x 576 pixels, Uncompressed Quicktime or AVI, Broadcast quality (CCIR 601 PAL). Zipped with Winzip. This format can be imported by any broadcast video editing system and edited.

The video formats are made in 'pipeline-fashion' with **Cleaner XL** from **Discreet**. It is possible to use different input formats, and a huge number of output formats, sizes, effects etc. can be chosen. It is possible to make a batch list of files to process overnight and the programme is therefore very effective. In addition the programme is very fast, and has excellent preview functions.

"Simplicity" on SpaceTelescope.org

6 The metadata

Metadata for the data content of the archives (the images, videos etc.) such as id, title, caption etc. are stored in Microsoft Excel compatible comma-separated .csv files (hereafter called CSV files). The CSV files are located in */home/web/spacetelescope/docs/csvfiles/*. The different CSV files are:

- *newsdata.csv*
- *imagedata.csv*
- *videodata.csv*
- *posterdata.csv*
- *kidsdrawingdata.csv*
- *logodata.csv*
- *cdromdata.csv*
- *calendardata.csv*
- *brochuredata.csv*

Imagedata.csv on Spacetelescope.org contains the following fields for each record:

- id
- alternative id/OPO id
- Release number
- Release type
- Priority
- Title
- release date/local time Munich (CET or CEST)
- Teaser, Short description
- Description, Caption
- Credit
- Object name
- Object type
- Mission
- Instrument
- Year
- Chrono, subnumber
- Wallpaper
- Zoomable
- original pixel height
- original pixel width
- Comment
- Caption link
- Press link

"Simplicity" on SpaceTelescope.org

CSV files should be edited with Excel (or similar), and saved in CSV format. There is one important issue. As commas often appear in normal text it is not files possible to use commas as separators. Instead the pipe character "|" is used.

In order to use "|" instead of commas, set the List Separator to "|" in "Control Panel > Regional Settings" (Choosing Advanced button). This refers to Windows only.

Release ID	Title	release date/local	Teaser	Description	Caption	Credit	Object name	Object ID	Mission	Instrument	Year	Chrono	Viewing	Zoom	original	original	Comm					
hec0406	News	0	Hubble: 09/03/2004 15:30			Galaxies, galaxies in NASA, ESA, and Hubble Ultra	Cosmos	Hubble	ACS	2004	6	yes	yes									
hec0405	Photo	0	V393 M 04/03/2004 12:00			"Starry Night", Vincenzo	NASA, the Hubble	Erasing star	Star, nebula	Hubble	ACS	2004	5	yes	no							
opo0308	News	0	Hubble Mosaic of the Majestic	Somb		NASA's Hubble Space	NASA and The Hubble	Herbig Galaxy	Hubble	ACS	2003	26	yes	yes								
hec0305	News	0	A perfect storm of turbulent	gases		Like the fury of a raging	European Space	Messier 17, the	Nebula	Hubble	WFPC2	2003	5	yes	no							
opo0301	News	0	Hubble Looks Through Cosmic	Zoom		The Advanced Camera	NASA, N. Berthe	Abell 1689	Galaxy	Hubble	ACS	2003	1	yes	yes							
hec0301	News	0	The Boomerang Nebula - the	coolest		The Boomerang Nebula is	European Space	Boomerang	Nebula	Hubble	WFPC2	2003	1	yes	no							
hec0206	News	0	A "wallpaper" of distant	galaxies		is e	Against a stunning	background	NASA, Holland F	Tadpole	Galaxy	Hubble	ACS	2002	6	yes	yes					
hec0206	News	0	Hubble's newest camera takes	a de		epicturing a night	NASA, Holland F	Mice Interact	Galaxy	Hubble	ACS	2002	6	yes	yes							
hec0206	News	0	Hubble's newest camera images	gh		Presenting a night	NASA, Holland F	Cone	Nebula	Nebula	Hubble	ACS	2002	6	yes	yes						
opo0201	News	0	Thackeray's Globules in IC	2944		Strangely glowing dark	NASA and The	Thackeray's	Nebula	Hubble	WFPC2	2002	1	yes	no							
hec0114	News	0	Nebula NGC 2080, nicknamed	the		"Ghost Head Nebula" (ESA, NASA, & N	Nebula	NGC	Nebula	Hubble	WFPC2	2001	14	yes	no							
opo0110	News	0	The Heart of the Whirlpool	Galaxy		The Whirlpool galaxy, M51, NASA and The	F.M.51, ICAS	1, Galaxy	Hubble	WFPC2	2001	10	yes	yes								
hec0109	News	0	The Red Spider Nebula -	Surfing		in S-Credit: ESA & Garrett	Mell	ESA & Garrett	M	The Red Spider	Nebula	Hubble	WFPC2	2001	9	yes	no					
hec0108	News	0	Hubble images remarkable	double		image Credit: ESA, NASA, ESA, and	NGC 1950, O-Star	Cl	Hubble	WFPC2	2001	6	yes	no								
opo0066	News	0	Light and Shadow in the	Carina		Nebula/Previously unseen details	NASA, The Hubble	Carina	Nebula	Nebula	Hubble	WFPC2	2000	6	yes	yes						
opo0942	News	0	The Triple Nebulae -	Stellar		Sibling Rivalry This NASA/ESA	Hubble	SNAS and	Jeff	NGC 6514, Tr	Nebula	Hubble	WFPC2	1999	42	yes	no					
opo0941	News	0	A Grazing Encounter	Between		Two in the direction of the	core	NASA and The	NGC 2207, IR	Galaxy	Hubble	WFPC2	1999	41	yes	no						
opo0926	News	0	Hubble images a Swarm	of		Accretion. This stellar system is	M80	The Hubble	Herit	M80, NGC 60	Star	Cl	Hubble	WFPC2	1999	26	yes	yes				
opo0925	News	0	Magnificent Details in a	Dusty		Spiral In 1995, the majestic	spiral	Hubble	Herit	NGC 4414, IR	Galaxy	Hubble	WFPC2	1999	25	yes	no					
opo0912	News	0	Multiple Generations of	Stars		in the "in the most active	star	Hubble	Herit	Hodge 301, T	Nebula	Hubble	WFPC2	1999	12	yes	no					
opo0901	News	0	Looking Down a Barrel	of		Gas at a (The NASA/ESA	Hubble	S	Hubble	Herit	M57, The	Rin	Nebula	Hubble	WFPC2	1999	1	yes	no			
opo0926	News	0	Saturn in Natural	Colours		The NASA/ESA	Hubble	S	Hubble	Herit	Saturn	Solar	S	Hubble	WFPC2	1999	26	yes	no			
opo0901	News	0	The Hubble Deep Field			Several hundred never	by	Robert	Williams	(Part	of	the	H	Galaxy, Hubble	WFPC2	1996	1	yes	yes			
opo0945	News	0	Maelstrom of Star	Birth		This spectacular color	pa	C.J.R. O'Dell	(Rice	M42, NGC 19	Nebula	Hubble	WFPC2	1995	45	yes	yes					
opo0944	News	0	Pillars of Creation			Undersee comet? Enchan	Jeff	Hester	and	IM16, The	Eng	Nebula	Hubble	WFPC2	1995	44	yes	yes				
opo0404	News	1	An Abrasive Collision	Gives		One Ge		A collision of two	gal	NASA and The	F.M.4, Messier	Galaxy	Hubble	WFPC2	2004	4	yes	no				
hec0402	Photo	1	Supernova 03/02/2004 15:00			"The nearby dwarf	gal	European	Space	NGC 1569, A	Galaxy, Hubble	WFPC2	2004	2	yes	no						
hec0401	News	1	Supernova 1993J	exploding		(artist's new observations with	ESA and	Justin	Supernova	11	Star	Hubble	Illustrat	2004	1	yes	no					
opo0331	News	1	Hubble Photographs	Turbulent		Neig	"A small portion of the															
opo0330	News	1	Firestorm of Star	Birth		Seen in a Loc	"This festively colorful															
opo0324	News	1	Celestial Composition			"And a backdrop of																
opo0322	News	1	Hubble's Sharpest	Ever		Color View	"This view of	Mars, the	sharpest	photo	Mars	Solar	S	Hubble	2003	22	yes	no				
opo0321	News	1	Too Close for Comfort			"This Hubble Space	Te															
opo0320	News	1	Celestial Fireworks			"Presenting the	putts															
opo0316	News	1	Supernova Shock	Wave		Partly	Co															
opo0315	News	1	Young and Old Stars	Form		And	The	deepest	visible	light	NASA, ESA and	M31, NGC 22	Galaxy	Hubble	ACS	2003	15	yes	yes			
hec0312	News	1	Mege starbirth cluster is	largest,		but	This	illustration	shows	an	European	Space	Artist's	impre	Star	Cl	Hubble	NIA	2003	12	yes	no

A screenshot of how the metadata look in Excel.

The metadata files can be read-protected, so that the entire set of metadata cannot be downloaded.

6.1 Editing a metadata file with Excel

As mentioned above the metadata files are best edited with Excel as this program works well with columns and rows of information. It has built-in

"Simplicity" on SpaceTelescope.org

functions such as *spell-checking*, "*fill-Down*" and others that make maintenance work for the web site easier.

However there is still room for 'silly mistakes' and some care should be taken that the data are consistent. One common mistake is to include a space character in the object id. This is a mistake that is difficult to find, and the only trace is that the Perl script will return nothing when queried.

For text fields such as captions it is possible to use html-formatting codes, such as <P> or
 for vertical spaces, for bold and <i> for italics, as the Perl scripts reproduce content verbatim.

6.2 Priorities and sorting

Perl scripts return output from the metadata files in the order that it retrieves them so it makes sense to include a priority field in the metadata, and to sort on this before saving the CSV file. A macro can also be created that does this (see Section 6.3).

To run an existing sort macro, go to the menu item "*Tools > Macro > Macros...*" and choose the appropriate macro.

6.3 Creating an Excel macro

As an example we will create a macro that sorts the content of a CSV file. First make sure the CSV file is open.

1. Go to a randomly chosen cell.
2. Choose "*Tools > Macro > Record New Macro...*" from the menu (a new window opens).
3. Name the macro (without white space), and choose the option "*Personal Macro Workbook*" under "*Store macro in:*". You can choose to make a short-cut for the macro.
4. Press *OK* (notice that a little window with a stop-button appears).
5. Perform a normal sort of the data: Press *Ctrl-A* to select all cells, choose "*Data > Sort...*" from the menu, choose the columns to sort by (and if it should be ascending or descending), and press *OK* in the sort window.

"Simplicity" on SpaceTelescope.org

6. Press the *stop* button in the little window that opened in stage 4.

The macro has now been created. NOTE: Save the changes in "*Personal Macro Workbook*" when Excel closes.

See Section 6.2 on how to run the macro.

"Simplicity" on SpaceTelescope.org

7 Perl scripts

All Perl scripts are located in */bin*. Development is carried out in */bin/sandbox*. To run a Perl script from a command line, type "perl <script>.pl param1=value1", where param1=value1 is a parameter.

There are four types of Perl scripts; **maintenance scripts** for generating static html pages, **common scripts** with subroutines, **archive scripts** for generating search results and **generator scripts** for intelligent decisions on when to generate new static html pages. Maintenance scripts are only used in maintenance runs, and the archive scripts are executed live on the site.

All Perl scripts start with **configuration variables** that define where the **metadata** (Excel/CSV) files are placed, which **html-template** to use etc. All configuration variables are written in capitals and have meaningful names.

7.1 Archive scripts

These scripts are optimized for fast performance, to minimize CPU usage and to provide a fast response for the user. They are the direct interface to the archives and are called every time a user searches an archive, or clicks on the 'previous-next' bar below the images. The layout of the page can be changed using various variables. The scripts offer different search functionalities and possibilities. Only **overview pages** (i.e. lists of output pages) are handled by these scripts. The individual output pages are **static html page** generated by the **maintenance scripts** (see Section 7.2).

On Spacetelescope.org the following different archive scripts exist:

- brochures.pl
- calendar.pl
- cdroms.pl
- drawings.pl
- images.pl
- logos.pl
- news.pl
- printlayouts.pl
- posters.pl
- videos.pl

"Simplicity" on SpaceTelescope.org

7.2 Maintenance scripts

The **maintenance scripts** are executed if **static html pages** within the archives need to be updated. They do not need to execute particularly fast, as they are executed infrequently and never accessed directly by the users. They are executed automatically by **generator scripts** that decide intelligently which html pages to update. The **common scripts** are used widely by the maintenance scripts. This introduces a slight overuse of CPU resources when the maintenance scripts are executed, but greatly reduces the code volume and increases readability.

On Spacetelescope.org the following maintenance scripts exist (located in */scripts/staticscripts*):

- staticbrochures.pl
- staticcalendar.pl
- staticcdroms.pl
- staticdrawings.pl
- staticimages.pl
- staticlogos.pl
- staticnews.pl
- staticprintlayouts.pl
- staticposters.pl
- staticvideos.pl

7.3 Common scripts

These scripts contain code common to all the above scripts. These scripts are called from within the above scripts.

On Spacetelescope.org the following maintenance scripts exist (located in */bin*):

- filesize.pl
- findimages.pl
- findnews.pl
- findvideos.pl
- formatdate.pl
- freesearch.pl
- massivewarning.pl
- printgoogle.pl
- printstatus.pl

"Simplicity" on SpaceTelescope.org

[ERIK: maaske vi boer have en enkelt linie for hver om hvad de goer? + Et eksempel med kommentering]

7.4 Generator scripts

In order to reduce the CPU load if the website has many hits simultaneously, static pages are generated for all entries in the archives (for instance an html page for a news item, image or video).

On Spacetelescope.org the following generator scripts exist (located in */scripts*):

- *gennews.pl*
- *genimages.pl*
- *genvideos.pl*
- *gencalendars.pl*
- *gensimplearchives.pl* – This script is common for a series of simple archives that do not need any special treatment.

Some scripts run every 5 minutes to check if they need to generate a new set of static pages for a given archive. We only want to generate a static page if it is absolutely necessary, since we for instance don't want to generate 1700 HTML pages each time the CSV file for the image archive is updated.

To do this:

- Check if any files that affect the pages have been changed or an embargo timestamp has been passed (this allows embargoed news to be published at a predefined time). The files on which a page typically depends are:

- */scripts/staticscripts/static<archive>.pl*. Used to generate the static html pages. All pages must be regenerated if this changes.

- */csvfiles/<archive>data.csv*. The data file.

- */bin/template_<section>.html*. This is the HTML template for the static page. All pages must be regenerated if this has changed.

- */bin/template_<embargo section>.html*. This is the HTML template for the static embargo protected pages. All embargo pages must be regenerated if this has changed.

"Simplicity" on SpaceTelescope.org

- */scripts/templates/esoacc_template*. This is the embargo protection template file.

- If only some pages are to be generated, each entry in the CSV file is checked to see if it has changed, or whether an embargo date has passed. This is done by computing a checksum - a number that is computed based on the date a page is generated and stored, and compared with a recomputed checksum on subsequent runs. If old and new numbers do not match, an entry has been changed.

The computed checksums, embargo timestamps and timestamps for the files mentioned above are all located in */scripts/data/<archive>*.

To trigger a check of some pages to be regenerated, type: `"cd /home/web/spacetelescope/docs; touch docs/csvfiles/newsdata.csv; scripts/gennews.pl"`. (This example uses the News Archive). For simple archives use: `"cd /home/web/spacetelescope/docs; touch csvfiles/<archive>data.csv; scripts/gensimplearchive.pl <archive>"` instead, where *<archive>* is replaced by the archive name.

The generator scripts also generate an *.esoacc* file that denies access to embargoed entries like news items, images or videos. It is generated on the basis of a template located in */scripts/templates/esoacc.tpl*. The template looks like:

```
<Files ~ "(IDLIST)\.(EXTLIST)">
AuthType Basic
AuthName "Spacetelescope Embargo Page"

AuthUserFile /path/to/htpasswordfile
require user <webuser>
</Files>
```

IDLIST is replaced by a list of embargoed entry ids, and *EXTLIST* is replaced by a list of extensions (typically html, txt, doc, pdf, jpg, gif, tif and zip) specific for each archive.

If an entry (news/image/video etc.) is removed from the CSV file, the corresponding static html-page is also deleted. All non-generated files in output directories will also be deleted, unless they start with a full stop like *.esoacc*.

"Simplicity" on SpaceTelescope.org

7.5 Templates

All scripts use a type of html template to generate the html output. All template files are located in `/bin/template_<section>.html`. The templates are edited with Dreamweaver and basically contain everything (design, menus, header, footer etc.) except the output produced on-the-fly by the script.

Note, that these templates are not the same as **Dreamweaver templates**. In Dreamweaver terminology these templates are just normal html pages.

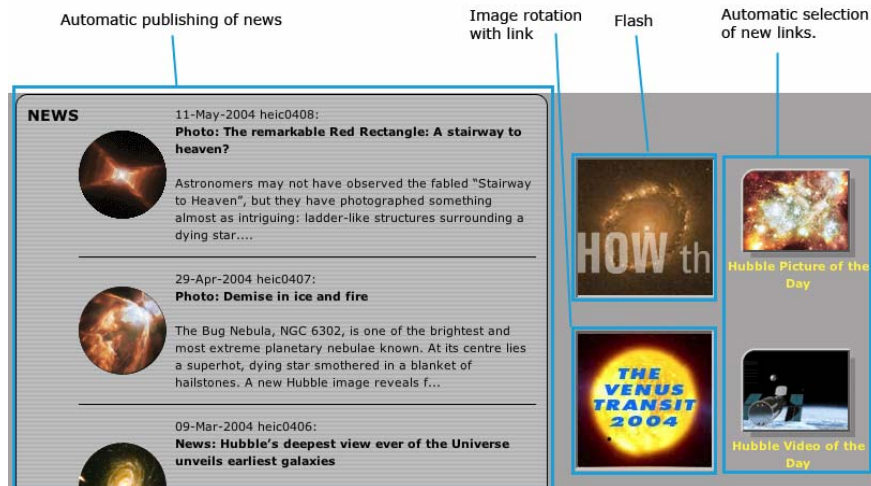
Do not edit these templates, unless you know exactly what you are doing.

7.6 Automatic front-page generation (frontpage.pl)

A static front page (`/index.html`) is generated every 5 minutes. This is done instead of serving a dynamic on-the-fly calculated front page for all visitors, and saves CPU time for the web server (the content of the page rarely changes). The script rotates Flash and image advertisements on a daily basis, and places the latest three news items on the front-page. A new front-page can be generated manually by typing `"cd /home/web/spacetelescope/docs; bin/frontpage.pl > index.html"`.

The front-page scripts rotate Flash from the directory `/flash/dynamic`. Images are rotated from `/design/dynamic_images`. In Flash links are made directly in the Flash-file, but images require a file placed in `/design/dynamic_images` called `<image name>.link`. So, to have an image named `venus_transit.jpg`, the link file would be named `venus_transit.link`. The content of the link file should just be a normal text file with one line, containing an absolute link to the website (e.g. `/bin/videos.pl`).

"Simplicity" on SpaceTelescope.org



7.7 Editing Perl scripts on Windows

To edit the Perl scripts under Windows, note the type of line breaks used. For instance in Notepad, the file is saved using a Windows line break, but before the Perl script can be executed on the server, these line breaks have to be converted to UNIX line breaks.

Use the following command to convert all Windows line breaks to UNIX in a Perl script:

```
dos2unix script.pl script.pl
```

And make the file executable with the command:

```
chmod 755 script.pl
```

7.8 Source code for Perl scripts

For security reasons the Perl scripts are not listed here. Contact Lars Lindberg Christensen by sending an email to lars@eso.org to obtain the source code for the Perl scripts.

“Simplicity” on SpaceTelescope.org

8 .esoacc files

.esoacc files are used to redirect requests, deny access to CSV files, deny directory listings, request authorization for embargo pages and change the default “File Not Found”-page.

This is specific to ESO systems. Normally they are called *.htaccess* files.

8.1 Redirects

There are many redirects in the .esoacc-files on Spacetelescope.org. Most of them are single pages from the old web site that are mapped to the right pages in the new site.

To make the site more friendly to search engines, there are some redirects to trick the search engines to believe that they are indexing normal pages, when they are actually indexing perl scripts (most robots will not index perl scripts). These redirects are typically of the form:

```
/news/archive/<searchtype>/<string>/<viewtype>/<from> maps to  
/bin/news.pl?searchtype=<searchtype>&string=<string>&viewtype=<viewtype>  
&from=<from>.
```

For instance a request to `/news/archive/year/2004//` would show the list of news from 2004.

8.1.1 Making a redirect

To redirect requests to <http://www.spacetelescope.org/goodies/old.html> to <http://www.spacetelescope.org/goodies/new.html>:

- Create an .esoacc-file in `/goodies` (if it's not already there). The following two lines will do this

```
RewriteEngine On #(this line should only be one time in a file)  
RewriteRule ^old\.html new.html [R]
```

The `RewriteEngine On` turns on the module in Apache that makes the redirects. Make the actual redirect in the next line. The first part, `^old\.html`, indicates that requests made to <http://spacetelescope.org/goodies/old.html> should be redirected. If the `^` in the front is omitted, the line will match the

"Simplicity" on SpaceTelescope.org

old.html pages in the current directory and in any subdirectories. The last part new.html [R] tells Apache to redirect to new.html. If [R] is omitted, then the URL-address in the browser will **not** be changed to <http://www.spacetelescope.org/goodies/new.html>. This is similar to renaming new.html to old.html and moving it to /goodies if it's not already located there.

8.2 "File Not Found"-page and "Authorization Required"-page

If the web server can't find a file request by a client, the page `/error_404.html` is shown. Likewise if a client fails to authenticate itself when requesting an embargo page, the page `/error_401.html` is shown.

8.3 Password protection

From ESO's pages:

- Create directory `$HOME/.webauth` and set access rights:

```
mkdir $HOME/.webauth; chmod o+x $HOME/.webauth; chmod o-r $HOME/.webauth.
```

- Create a file `.esoacc/.htaccess` in the directory that is to be protected with basic authentication and add something to `.esoacc/.htaccess`. like

```
AuthName "string describing the protected Web area"
AuthType Basic

AuthUserFile /path/to/my/homedirectory/.webauth/passwordfile

require valid-user
```

To give password access to selected users only in AuthUserFile use:

```
require user username1 [username2 username3...]
```

Make sure that the file is readable by others (`chmod o+r .esoacc`) and always give the full path to the passwordfile in the parameter AuthUserFile.

- Create the **passwordfile** (parameter AuthUserFile) in `.webauth` that will contain name:password records for the protected Web area, using the

"Simplicity" on SpaceTelescope.org

htpasswd command (often found in /usr/server/bin/):

- o htpasswd -c "AuthUserFile" pippo1
Warning: Never use your Unix login password.
- o add pippo1 to the .esoacc file
- o chmod o+r "AuthUserFile"

To add more users just type:

- htpasswd "AuthUserFile" pippo2
- Add pippo2 to the .esoacc file

8.4 Example of an advanced .esoacc file with redirects and password protection

```
RewriteEngine on

# The condition (first two lines) must be true for making
# the rewrite. %1 in the RewriteRule references the text in
# parentheses in the first or second line.
RewriteCond %{QUERY_STRING} ^string=(heic[0-9]+)$ [OR]
RewriteCond %{QUERY_STRING} ^&string=(heic[0-9]+)$
RewriteRule ^news\.pl /news/html/%1.html? [R]

# The condition (first two lines) must be true for making
# the rewrite. %1 in the RewriteRule references the text in
# parentheses in the first or second line.
RewriteCond %{QUERY_STRING} ^string=(heic[0-9]+)$ [OR]
RewriteCond %{QUERY_STRING} ^&string=(heic[0-9]+)$
RewriteRule ^embargo\.pl /news/html/%1.html? [R]

<Files embargo.pl> # Password protect the file embargo.pl
AuthType Basic
AuthName "Spacetelescope Embargo Page"

AuthUserFile /path/to/passwordfile
require user test
</Files>

ErrorDocument 401 /error_401.html # "Authorization required"-error
page
ErrorDocument 404 /error_404.html # "File not found"-error page
```

"Simplicity" on SpaceTelescope.org

9 Searching

To make the html-pages searchable, a utility called **Htdig** indexes the web pages every Monday, Wednesday and Friday at 1:01 AM. To run Htdig manually type `"/home/web/bin/rundig.hubble"`

The configuration file for Htdig is located at

`/home/web/spacetelescope/docs/search/htdig/htdig-hubble.conf.`

HTML templates for the output from Htdig are located in following files (edited with Dreamweaver):

- `search/htdig/templates/wrapper.hubble.html` (default layout)
- `search/htdig/templates/nomatch.hubble.html` (is shown when no match is found)
- `search/htdig/templates/short.html` (short description of found item)
- `search/htdig/templates/long.html` (long description of found item)

The `/Videos` and `/Images` archives are not indexed.

"Simplicity" on SpaceTelescope.org

10 Link checker

To check for broken links, link checking of all web pages (internal as well as external) is done on the 3rd of each month at 1:01 AM. The crontab entry (see next section) to run it looks like this:

```
01 01 03 * * * cd /home/web/spacetelescope/docs/search;linbot -b
http://www.spacetelescope.org
```

The result is located at <http://www.spacetelescope.org/search/linbot/>.

[Webcheck 1.0](#)
[Site Map](#)
[Bad Links](#)
[Images](#)
[What's Old](#)
[What's New](#)
[What's Slow](#)
[No Titles](#)
[External Links](#)
[Not Checked](#)
[Problems \(By Author\)](#)



The European Homepage For The NASA/ESA Hubble Space Telescope

Bad Links

Link	http://www.stsci.edu/hst/	
Status	503: Service Unavailable	
	The European Homepage For The NASA/ESA Hubble Space Telescope - Links	None
Parents	The European Homepage For The NASA/ESA Hubble Space Telescope - Litterature	None
	The European Homepage For The NASA/ESA Hubble Space Telescope - Links	None
	The European Homepage For The NASA/ESA Hubble Space Telescope - Links	None
Link	http://www.stsci.edu/hst/proposing/documents/cp_cy12/primer_cover.html	
Status	503: Service Unavailable	
	The European Homepage For The NASA/ESA Hubble Space Telescope - Links	None
Parents	The European Homepage For The NASA/ESA Hubble Space Telescope - Litterature	None
	The European Homepage For The NASA/ESA Hubble Space Telescope - Links	None
	The European Homepage For The NASA/ESA Hubble Space Telescope - Links	None
Link	http://www.spacetelescope.org/about/about/history/launch_1990.html	
Status	404: Not Found	
Parents	The European Homepage For The NASA/ESA Hubble Space Telescope - Links	None
	The European Homepage For The NASA/ESA Hubble Space Telescope - Links	None

"Simplicity" on SpaceTelescope.org

11 Crontab

To add a new *cron job*, edit */scripts/templates/crontabfile*, and type `crontab -l /path/to/crontabfile` to activate it. Below is an example of a crontab file. Note that there should be no line break between lines that are not separated with at least one blank line.

```
01 01 * * 1,3,5 /home/web/bin/rundig.hubble 1>/dev/null 2>/dev/null
00,05,10,15,20,25,30,35,40,45,50,55 * * * * cd
/home/web/spacetelescope/docs/; bin/frontpage.pl > index.html
00,05,10,15,20,25,30,35,40,45,50,55 * * * * /home/web/bin/perl
/home/web/spacetelescope/docs/scripts/gennews.pl 1>/dev/null
```

11.1 Editing the crontabfile

Use a text editor to open */scripts/templates/crontabfile*. Insert the following line:

```
* * * * * command 1>/dev/null 2>/dev/null
```

The asterisks (*) specify when to run the command, and denotes, from left: minutes, hours, day of month, month and day of week (0 is Sunday, 6 is Saturday). Each of them can be replaced by a list of times and ranges of time. For instance to run a command every 10th minute in the time range 8:00-10:00 and 18:00-20:00, Monday to Friday, every month, use the following line:

```
00,10,20,30,40,50 08-10,18-20 * * 1-5 command 1>/dev/null
2>/dev/null
```

The last part `"1>/dev/null 2>/dev/null"` says that normal output "1" should be thrown away, likewise with any errors "2". If the errors are to be kept then omit the `"2>/dev/null"` part.

"Simplicity" on SpaceTelescope.org

12 Index

Simplicity

5