

# Web Team Code Documentation

## Formmail.ini

```
;For forms that use /WRAP_SDDU/BL-INFO-WEBS/SoiC Website/_php/formmail/formmail.php
;This file wants to live at
;webservet.uits.indiana.edu:/ip/soic2/formmail.ini
;webserve.uits.indiana.edu:/ip/soicweb/www/formmail.ini
;/WRAP_SDDU/BL-INFO-WEBS/SoiC Website/_php/formmail/formmail.ini.txt
;Please update it in all spots.
;
;It is here for backup/recordkeeping purposes
```

```
[email_addresses]
soic = "soic@indiana.edu"
soicugrd = "soicugrd+recruiting@indiana.edu"
soiccsgr = "soiccsgr@indiana.edu"
web = "soicweb@indiana.edu"
help = "soichelp@indiana.edu"
infograd = "infograd@indiana.edu"
ilsmain = "ilsmain@indiana.edu"
iugrenger = "iugrenger@indiana.edu"
iseengr = "iseengr@indiana.edu"
soicengr = "soicengr@indiana.edu"
iucs = "iucs@indiana.edu"
```

## php.ini (development server)

```
#name the file /ip/soic2/wwws/php.ini
include_path=".:ip/soic2/wwws/:ip/soic2/wwws/_php/"
```

## php.ini (production server)

```
#name the file /ip/soicweb/wwws/www.soic.indiana.edu/php.ini
include_path=".:ip/soicweb/www/www.soic.indiana.edu"
```

## sitemap generator

These scripts live in /ip/soicweb/bin  
and are run by cron, edit with crontab -e

```
Current cron:
15 00 * * * /usr/local/bin/php /ip/soicweb/bin/sitemap/create_soic_sitemap.php >> /ip/soicweb/bin/sitemap-soicweb.log
15 00 * * * /usr/local/bin/php /ip/soicweb/bin/sitemap/create_cs_sitemap.php >> /ip/soicweb/bin/sitemap-csweb.log
15 00 * * * /usr/local/bin/php /ip/soicweb/bin/sitemap/create_informatics_sitemap.php >> /ip/soicweb/bin/sitemap-informaticsweb.log
15 00 * * * /usr/local/bin/php /ip/soicweb/bin/sitemap/create_ils_sitemap.php >> /ip/soicweb/bin/sitemap-ilsweb.log
15 00 * * * /usr/local/bin/php /ip/soicweb/bin/sitemap/create_ise_sitemap.php >> /ip/soicweb/bin/sitemap-iseweb.log
*/4 * * * /usr/bin/wget -O - -q -t 1 http://www.soic.indiana.edu/\_php/onestart-xml-updater.php
```

## create \_XXXX\_sitemap.php

```
<?php
```

```
ini_set('display_errors', 'On');
error_reporting(E_ALL | E_STRICT);
```

```
/******\
* Author : Binny V Abraham *
* Website: http://www.bin-co.com/ *
* E-Mail : binnyva at gmail *
* Get more PHP scripts from http://www.bin-co.com/php/ *
*****\
* Name : PHP Google Search Sitemap Generator *
* Version : 1.00.A *
* Date : Friday 17 November 2006 *
* Page : http://www.bin-co.com/php/programs/sitemap\_generator/ *
* *
* You can use this script to create the sitemap for your site automatically. *
* The script will recursively visit all files on your site and create a *
* sitemap XML file in the format needed by Google. *
* *
* Get more PHP scripts from http://www.bin-co.com/php/ *
\*****/
```

```
// Please edit these values before running your script.
```

```
//////////////////// Options //////////////////////////////////////
$url = "http://www.cs.indiana.edu/"; //The Url of the site - the last '/' is needed
```

```
$root_dir = '/ip/soicweb/www/www.soic.indiana.edu/about/departments/computer-science-informatics/computer-science/';
//Where the root of the site is with relation to this file.
```

```
$file_mask = '*.html'; //Or *.html or whatever - Any pattern that can be used in the glob() php function can be used here.
```

```
//The file to which the result is written to - must be writable. The file name is relative from root.
$sitemap_file = 'sitemap.xml';
```

```
// Stuff to be ignored...
//Ignore the file/folder if these words appear in the name
$always_ignore = array(
    '.htaccess', 'php.ini', 'local_common.php', 'images', '.htaccess.txt', 'php.ini.txt'
);
```

```
//These files will not be linked in the sitemap.
$ignore_files = array(
    '400.shtml', '401.shtml', '403.shtml', '404.shtml', '500.shtml'
);
```

```
//The script will not enter these folders
$ignore_folders = array(
    '_containers', '_examples', '_feeds', '_global', '_includes', '_internal', '_php', '_iu-brand', 'css', 'error', 'fonts', 'img', 'js', 'search', 'toolkit', 'topics', 'audience',
    'search'
);
```

```
//The default priority for all pages - the priority of all pages will increase/decrease with respect to this.
/* use this for web page--> $starting_priority = ($_REQUEST['starting_priority']) ? $_REQUEST['starting_priority'] : 70;*/
$starting_priority = 70;
```

```
//////////////////////////////// Stop editing now - Configurations are over //////////////////////////////////
```

```
// Run the site map generation
generateSiteMap();
```

```
////////////////////////////////////
function generateSiteMap() {
    global $url, $file_mask, $root_dir, $sitemap_file, $starting_priority;
    global $always_ignore, $ignore_files, $ignore_folders;
    global $total_file_count, $average, $lowest_priority_page, $lowest_priority;
```

```
//////////////////////////////////// Code //////////////////////////////////////
chdir($root_dir);
$all_pages = getFiles("");
```

```
$xml_string = '
<?xml version="1.0" encoding="UTF-8"?>
<urlset
xmlns="http://www.google.com/schemas/sitemap/0.84"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:schemaLocation="http://www.google.com/schemas/sitemap/0.84
http://www.google.com/schemas/sitemap/0.84/sitemap.xsd">
```

```

';

$modified_priority = array();
for ($i=30;$i>0;$i--) array_push($modified_priority,$i);

$lowest_priority = 100;
$lowest_priority_page = "";
//Process the files
foreach ($all_pages as $link) {
    #if(strpos($link, '&')) continue;
    //Find the modified time.
    $handle = fopen($link,'r');
    $info = fstat($handle);
    fclose($handle);
    $modified_at = date('Y-m-d\Th:i:s\Z',$info['mtime']);
    $modified_before = ceil((time() - $info['mtime']) / (60 * 60 * 24));

    $priority = $starting_priority; //Starting priority

    //If the file was modified recently, increase the importance
    if($modified_before < 30) {
        $priority += $modified_priority[$modified_before];
    }

    if(preg_match('/index\.w{3,4}$/', $link)) {
        $link = preg_replace('/index\.w{3,4}$/', "", $link);
        $priority += 20;
    }

    //These priority detectors should be different for different sites :TODO:
    if(strpos($link, 'example')) $priority -= 30; //If the page is an example page
    elseif(strpos($link, 'demo')) $priority -= 30;
    if(strpos($link, 'tuorial')) $priority += 10;
    if(strpos($link, 'script')) $priority += 5;
    if(strpos($link, 'other') !== false) $priority -= 20;

    //Priority based on depth
    $depth = substr_count($link, '/');
    if($depth < 2) $priority += 10; // Yes, I know this is flawed.
    if($depth > 2) $priority += $depth * 5; // But the results are better.

    if($priority > 100) $priority = 100;
    $loc = $url . $link;
    if(substr($loc, -1, 1) == '/') $loc = substr($loc, 0, -1); //Remove the last '/' char.

    if(!isset($total_priority)) {$total_priority=0;}
    $total_priority += $priority;
    if($lowest_priority > $priority) {
        $lowest_priority = $priority; //Find the file with the lowest priority.
        $lowest_priority_page = $loc;
    }

    $priority = $priority / 100; //The priority is given in decimals

    $xml_string .= "
<url>
<loc>$loc</loc>
<lastmod>$modified_at</lastmod>
<priority>$priority</priority>
</url>\n";
}

$xml_string .= "
</urlset>";
if(!$hndl = fopen($sitemap_file, 'w')) {
    //header("Content-type:text/plain");
    print date("Y-m-d H:i") . " : Can't open sitemap file - '$sitemap_file'.\nDumping result to screen...\n
<br />
<br />
<br />\n\n\n";
    print "
<textarea rows="25" cols="70" style="width:100%">'. $xml_string.'</textarea>";
} else {
    echo date("Y-m-d H:i") . " : Sitemap was written to ' . $url.$sitemap_file . '
Total files in sitemap: ' . count($all_pages) . '
' . PHP_EOL;

```

```

fputs($hndl,$xml_string);
fclose($hndl);
}

$total_file_count = count($all_pages);
$average = round(($total_priority/$total_file_count),2);
}

//////////////////////////////////// Functions //////////////////////////////////////
// File finding function.
function getFiles($cd) {
$links = array();
$directory = ($cd) ? $cd . '/' : ''; //Add the slash only if we are in a valid folder

$files = glob($directory . $GLOBALS['file_mask']);
foreach($files as $link) {
//Use this only if it is NOT on our ignore lists
if(in_array($link,$GLOBALS['ignore_files'])) continue;
if(in_array(basename($link),$GLOBALS['always_ignore'])) continue;
array_push($links, $link);
}
//asort($links); //Sort 'em - to get the index at top.

//Get All folders.
$folders = glob($directory . '**',GLOB_ONLYDIR); //GLOB_ONLYDIR not available on windows.
foreach($folders as $dir) {
//Use this only if it is NOT on our ignore lists
$name = basename($dir);
if(in_array($name,$GLOBALS['always_ignore'])) continue;
if(in_array($dir,$GLOBALS['ignore_folders'])) continue;

$more_pages = getFiles($dir); // :RECURSION:
if(count($more_pages)) $links = array_merge($links,$more_pages); //We need all thing in 1 single dimensional array.
}

return $links;
}

```