

Obtener Información De Páginas Blancas

Integrante: María Fernanda Segovia Chacón

23 de Octubre

2016

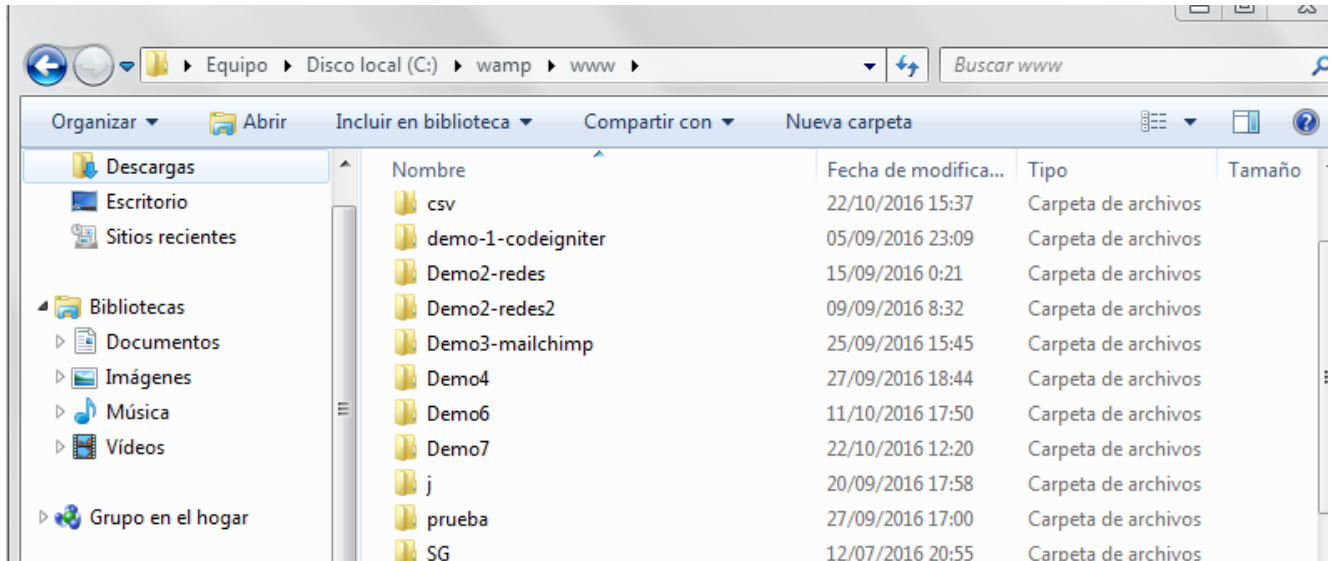


1. Requerimientos:

- Contar con un servidor WAMP o XAMP.
- Haber descargado el proyecto de “Demo7”

2. Desplegar el Proyecto:

1. El proyecto que se ha descargado se tiene que poner en la carpeta www de wamp server, en mi caso es C:\wamp\www, en xampp la dirección sería c:\xampp\htdocs\.



2. Luego de haber incluido la carpeta podemos ingresar al proyecto por medio de sublimetext para cambiar la conexión a la base de datos. Ingresamos a “Demo7/application/config/database.php” y aquí cambiaremos los parámetros de `$db['default']['username'] = 'root', $db['default']['password'] = ''`, `$db['default']['database'] = ''` por el usuario, contraseña y nombre de nuestra base de datos. En este proyecto no estamos utilizando la base de datos así que dejaremos vacíos algunos campos.



```
$active_group = 'default';
$query_builder = TRUE;

$db['default'] = array(
    'dsn' => '',
    'hostname' => 'localhost',
    'username' => 'root',
    'password' => '',
    'database' => '',
    'dbdriver' => 'mysqli',
    'dbprefix' => '',
    'pconnect' => FALSE,
    'db_debug' => (ENVIRONMENT !== 'production'),
    'cache_on' => FALSE,
    'cachedir' => '',
    'char_set' => 'utf8',
    'dbcollat' => 'utf8_general_ci',
    'swap_pre' => '',
    'encrypt' => FALSE,
    'compress' => FALSE,
    'stricton' => FALSE,
    'failover' => array(),
    'save_queries' => TRUE
);
```

3. Luego podemos acceder a la página desde nuestro buscador y podremos ver la página web mediante el link “<http://localhost/Demo7/>”.




4. Al hacer click en el botón de Seleccionar archivo y deberás cargar un archivo de Excel .xls otro formato no lo aceptara. Este archivo debe contener los números de teléfonos de la siguiente manera.



	A	B	C	D
1	15555299			
2	12457342			
3	12213430			
4	44237598			
5	15320098			
6	44323716			
7	14867608			
8	14221452			
9	15480382			
10	66528567			
11	74328120			
12	15634763			

5. Luego de cargar el archivo le das a subir luego de un momento te saldrá una tabla al costado de los números que si fueron encontrados.

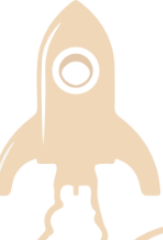
 Triangle

[f](#)
[t](#)
[g+](#)
[v](#)
[in](#)

HOME

Como funciona:

Solo tienes que crear un archivo .txt con los numeros que deseas buscar



Datos

Nombres y Apellidos	Direccion Completa	Provincia/ Distrito	Telefono Completo	Telefono Filtrado
De La Cruz Arana Ronald Jefferson	AV Colonial 5104 Pl. 1 Int. 5104-0253 022 Urb. Up Up Indus	Callao Cercado 1 -	5555299	15555299
Vinatea Recoba Luis Manuel	AV Santa Teresita 169 601 Pl. 6	San Isidro	1 - 2213430	12213430
Caspio Vasquez Jhonatan Alexis	Mz. D Lt. 10 An Villa Huanchaco 218 Urb. An Villa Huanchaco	Trujillo	44 - 237598	44237598
Fernandez Vasquez Irvin Jhonatan	Mz. 1 Lt. 24 Urb. Up Up Huanchaquito Alto	Trujillo	44 - 323716	44323716
Pizarro Bustamante Gladys Urbana	Mz. I Lt. 26 Urb. Ur Los Sureños	Puente Piedra	1 - 5480382	15480382

Sube tu archivo

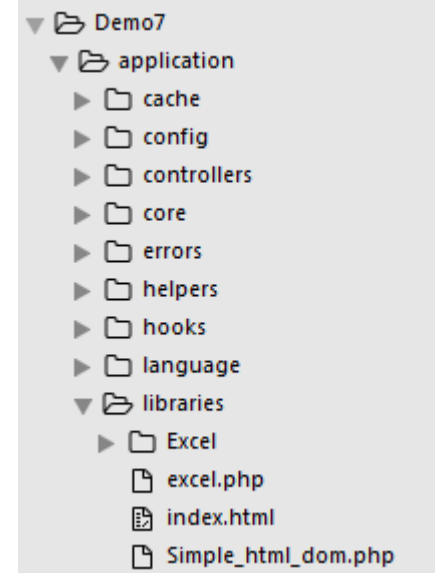
Ningún archivo seleccionado

© Your Company 2014. All Rights Reserved.



3. MVC CodeIgniter

Para este proyecto hay que hacer uso de la librería Simple_html_dom para poder parsear el código html y la librería Excel para php para subir los archivos xls. Pueden sacar las librerías del proyecto, se encuentran en la carpeta libraries de application en el proyecto.



En el controlador upload llamo a la librería upload del propio framework y también a la librería Excel en la función upload_file() que es en donde se obtienen los datos de los números telefónicos sacados de la vista de páginas blancas. Para poder sacar los datos de paginas blancas primero se tiene que mandar una cabecera a la página para poder hacer las consultas.

```
public function upload_file(){
    //Config the parameters to upload the file to the server.
    //Configuramos los parametros para subir el archivo al servidor.
    $config['upload_path'] = realpath(APPPATH.'../files');
    $config['allowed_types'] = 'xls';
    $config['max_size'] = '0';

    //Load the Upload CI library
    //Cargamos la libreria CI para Subir
    $this->load->library('upload', $config);

    if ( ! $this->upload->do_upload('file') ){
        //Displaying Errors.
        //Mostramos los errores.
        $data['error'] = $this->upload->display_errors();

        $this->load->view('upload', $data);
    }
    else{
        //Uploads the excel file and read it with the PHPEXcel Library.
        //Subimos el archivo de excel y lo leemos con la libreria PHPEXcel.
        $data = array('upload_data' => $this->upload->data());
        $this->load->library('excel');
        $excel = $this->excel->read_file($data['upload_data']['file_name']);

        $this->load->library("Simple_html_dom");

        $comienzo = microtime(true);
        $encontrados = 0;
        $no_encontrados = 0;

        $options = array('http' =>
            array(
                'method' => 'GET',
                'header' => "Content-type: text/plain; charset=UTF-8" . PHP_EOL .
                'Referer: http://www.forosdelweb.com' . PHP_EOL .
                'User-Agent: Mozilla/5.0 (Windows; U; Windows NT 5.1; es-ES; rv:1.9.0.6) Gecko/2009011913 Firefox/3.0.6' . PHP_EOL .
                'Cookie: user=FDW; actividad=programacion;' . PHP_EOL
            )
        );
    }
}
```



Luego se crea un array donde se llenaran los datos encontrados de cada numero. Cada que se mande la petición con un numero se retornara una vista de páginas blancas de donde extraeremos los datos.

```
$datos_llenar[] = array("Nombres y Apellidos","Direccion Completa","Provincia / Distrito","Telefono Completo","Telefono Filtrado");

foreach($excel as $row){
    foreach ($row as $col){

        $telefono_a_consultar = $col;

        $pagina_inicio = file_get_contents('http://www.paginasblancas.pe/telefono/'.$telefono_a_consultar,false, $context);
        $html = str_get_html($pagina_inicio);

        if($html === FALSE){
            //echo "La pagina que en blanco :".$telefono_a_consultar."\n";
        }else{
            $valor_duro = $html->find('div[class=m-result-business--recomendation-adv-data is-hidden]', 0);

            $valorTextArea = ($valor_duro != null) ? $valor_duro->innertext : '';

            $h3 = str_get_html($valorTextArea);

            if ($h3 === FALSE) {
                /*echo "No se encontro datos para este telefono: ".$telefono_a_consultar."\n";*/
                $no_encontrados++;
            }else{
                $valorh3 = $h3->find('h3', 0)->innertext;
                $valorspan1 = $h3->find('span', 0)->innertext;
                $valorspan2 = $h3->find('span', 1)->innertext;
                $valorspan3 = $h3->find('span', 2)->innertext;

                $datos_llenar[] = array($valorh3,$valorspan1,$valorspan2,$valorspan3,$telefono_a_consultar);
                $encontrados++;
            }
        }
    }
}
}
```

Y por último mandaremos a nuestra vista principal.

```
$data['excel'] = $excel;
$data['datos_llenar'] = $datos_llenar;
$this->load->view('upload', $data);
}
```