



# Webalizer

Guía para personas usuarias



# Contenido

## 0. ¿Qué es Webalizer?

### 1. Uso de Webalizer

#### 1.1 Activación

#### 1.2 La interfaz

### 2. Entender la información

#### 2.1 Terminología principal

- Hits

- Files

- Sites

- Visits

- Pages

- Kbyte

#### 2.2 Definiciones comunes

- Site

- URL

- Referrers

- Search strings

- User agents

- Entry/exit pages

- Countries

- Response codes



# O

## ¿QUÉ ES WEBALIZER?

Webalizer es una herramienta para obtener estadísticas de vuestras webs. Es un software libre, rápido y gratuito, que analiza archivos de registro (.log) de un servidor de Internet, en este caso el servidor Pangea.

Ofrece un alto detalle de los movimientos de un sitio web, es fácilmente configurable y presenta los informes en formato HTML para poder visualizarlos fácilmente con un navegador web. Webalizer produce estadísticas anuales, mensuales, diarias y horarias. En los informes mensuales, puede mostrar estadísticas varias de uso general, de uso por día y hora, de uso para sitios visitados, URL, agentes de usuario (navegadores), referenciantes y país.

Webalizer es altamente configurable desde el servidor, por lo que cualquier configuración específica que necesita la realizaremos desde Pangea y nos lo tenéis que pedir.

El código fuente de The Webalizer está disponible bajo la Licencia Pública General GNU.

Summary by Month										
Month	Daily Avg				Monthly Totals					
	Hits	Files	Pages	Visits	Sites	kB F	Visits	Pages	Files	Hits
<a href="#">Jan 2020</a>	9778	9239	5983	475	1775	2605500	2376	29918	46199	48894
<a href="#">Dec 2019</a>	7495	6826	3895	649	7555	13834686	20126	120746	211610	232351
<a href="#">Nov 2019</a>	8604	7985	4959	459	6737	11981637	13777	148772	239569	258127
<a href="#">Oct 2019</a>	7602	7028	3530	463	7347	13004683	14373	109447	217879	235680
<a href="#">Sep 2019</a>	9676	9033	2896	409	6121	15553554	10643	75307	234872	251601
<b>Totals</b>						<b>56980060</b>	<b>61295</b>	<b>484190</b>	<b>950129</b>	<b>1026653</b>



# 1 USO DE WEBALIZER

## 1.1 Activación y acceso

Para activar vuestro Webalizer debéis avisar a [suport@pangea.org](mailto:suport@pangea.org) para que lo active para el dominio que necesitáis. Si disponéis de varios dominios, tendréis estadísticas para cada uno de ellos.

Una vez activo, debéis ir a consultarlo desde el navegador web, indicando una URL similar a esta (sustituyendo '**vuestrodomio.ni**' por el nombre correspondiente, por ejemplo 'pangea.org'):

- **<https://www.vuestrodomio.ni/webalizer>**

(En nuestro ejemplo, sería <https://pangea.org/webalizer>)

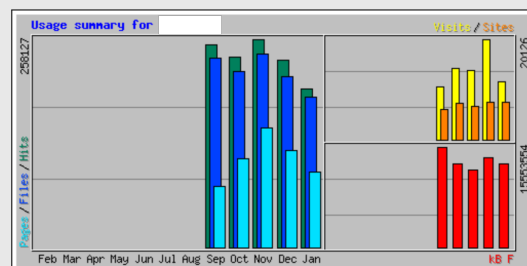
## 1.2 La interfaz

Este será el entorno que veremos:



Stats of

Summary by Month  
Generated 23-Jan-2020 00:04 UTC



Month	Daily Avg					Monthly Totals				
	Hits	Files	Pages	Visits	Sites	KB F	Visits	Pages	Files	Hits
Jan 2020	8924	8446	4262	530	7504	12988291	11669	93777	185826	196328
Dec 2019	7495	6826	3895	649	7555	13834686	20126	120746	211610	232351
Nov 2019	8604	7985	4959	459	6737	11981637	13777	148772	239569	258127
Oct 2019	7602	7028	3530	463	7347	13004683	14373	109447	217879	235680
Sep 2019	9676	9033	2896	409	6121	15553554	10643	75307	234872	251601
<b>Totals</b>						<b>67362851</b>	<b>70588</b>	<b>548049</b>	<b>1089756</b>	<b>1174087</b>

Generated by [Webalizer Version 2.23](#)

Si hace poco tiempo que pedisteis la activación de Webalizer, estos cuadros aparecerán vacíos o con muy poca información.

- El cuadro de arriba corresponde al resumen gráfico por meses.
- El cuadro de abajo, el resumen de accesos a su web en forma de tabla de datos. Desde este cuadro podemos acceder al detalle estadístico por meses, haciendo clic en cada uno de ellos, por lo que aparecemos en una página como esta:



Stats of

Summary Period: January 2020  
Generated 23-Jan-2020 00:04 UTC

[Daily Statistics] [Hourly Statistics] [URLs] [Entry] [Exit] [Sites] [Search] [Countries]

Monthly Statistics for January 2020		
Total Hits		196328
Total Files		185826
Total Pages		93777
Total Visits		11669
Total kB Files		12988291
Total Unique Sites		7504
Total Unique URLs		1161
	Avg	Max
Hits per Hour	371	3667
Hits per Day	8924	15254
Files per Day	8446	14756
Pages per Day	4262	11469
Sites per Day	341	760
Visits per Day	530	884
kB Files per Day	590377	808232
Hits by Response Code		
Undefined response code	0.01%	23
Code 200 - OK	94.65%	185826
Code 204 - No Content	0.21%	411
Code 206 - Partial Content	0.27%	529
Code 301 - Moved Permanently	1.11%	2174
Code 302 - Found	0.33%	643

Desde el menú de la parte superior podemos seleccionar el informe que deseamos:

- |                      |              |
|----------------------|--------------|
| 1. Daily Statistics  | 5. Exit      |
| 2. Hourly Statistics | 6. Sites     |
| 3. URLs:             | 7. Search    |
| 4. Entry             | 8. Countries |

Dentro de cada cuadro (gráfico y/o tabla) encontraréis datos como los que os aclaramos en el apartado de Terminología.

## NOTA

Tenga en cuenta que habrá datos estadísticos que no podrá obtener si previamente no los ha previsto, ya que requieren configuraciones específicas. Por ejemplo, si desea ampliar el listado de Top URL más consultadas, no podrá hacerlo con efecto retroactivo, ya que webalizer no habrá registrando estos datos.

## IMPORTANTE

Cualquier configuración específica debe consultarse y pedir a [suport@pangea.org](mailto:suport@pangea.org)



# 2 ENTENDER LA INFORMACIÓN

## 2.1 Terminología principal

**HITS** Los *hits* representan el número total de solicitudes realizadas en nuestro servidor durante un período de tiempo definido (mes, día, hora, etc.). Cada vez que alguien accede a alguna página o elemento de vuestra web está realizando una solicitud al servidor, por tanto registrando un *hit*.

**FILES** Los *files* o archivos representan el número total de *hits* (solicitudes) que efectivamente dieron como resultado que el usuario recibiese un poco de información (que se le abriera la página, que se le mostrara una imagen, que se le descargara el archivo solicitado ...). No todas las visitas envían datos, como las solicitudes 404-Not Found y las solicitudes de páginas que ya están en la memoria caché del navegador (ya las has abierto en alguna ocasión desde el mismo navegador).

**CONSEJO:** si te fijas en la diferencia entre los hits y los files, puede obtener una indicación aproximada de los visitantes que repiten, ya que cuanto mayor es la diferencia entre los dos, más personas solicitan páginas que ya han almacenado en caché (ya han visto). Esto es un dato a tener en cuenta ya que valora la 'fidelización' que vuestra web alcanza.



**SITES** Los *sites* o lugares son el número de direcciones IP únicas (como el DNI de cada ordenador) o nombres de *host* (máquinas anfitrionas) que hicieron solicitudes en nuestro servidor. Se debe tener cuidado al usar esta métrica, ya que arroja datos que pueden dar pie a engaño: puede parecer que muchos visitantes provienen de un solo lugar, y también puede parecer que provienen de muchas direcciones IP, por lo que se debe utilizar simplemente como una medida aproximada de la cantidad de visitantes de su web.

**VISITS** Las visitas se dan cuando un sitio remoto (o persona visitante) hace una solicitud de una página en nuestro servidor cada primera vez. Mientras el mismo lugar siga haciendo solicitudes dentro de un período de tiempo de espera determinado, todos se consideran parte de la misma visita. Si el tiempo transcurrido desde la última solicitud es mayor que el período de tiempo de espera especificado (el valor predeterminado es 30 minutos), se inicia y cuenta una nueva visita, y la secuencia se repite. Dado que sólo las páginas activarán una visita, los sitios remotos que enlazan con URL gráficas y otras URL que no sean de página no se contarán en el total de visitas, lo que reduce el número de visitas falsas.

**PAGES** Las *pages* o páginas son aquellas URL consideradas páginas reales de su web, que alguien solicita, y no cada uno de los elementos individuales que la componen (como





gráficos o clips de audio ...). Hay quien llama a esta métrica Page views (Vistas de página) o Page impresiones (Impresiones de página). Su valor predeterminado es cualquier URL que tenga una extensión .htm, .html o .cgi.

Un **KByte (KB)** equivale a 1024 bytes (o 1 Kilobyte). Se usa para mostrar la cantidad de datos transferidos entre nuestro servidor y la máquina remota (o persona visitante de su web), en función de los datos encontrados en el registro de servidor.

## 2.2 Definiciones comunes

**SITE** Un *site* o lugar es una máquina remota (el ordenador de la persona que está mirando tu web, o un servidor) que realiza solicitudes en nuestro servidor. Se basa en la dirección IP/nombre del host o anfitrión de cada máquina remota.

**URL** Llamado «Localizador uniforme de recursos» (por sus siglas en inglés *Uniform Resource Locator*). Todas las solicitudes realizadas a un servidor web deben solicitar algo.

Un URL es esa cosa y representa un objeto en algún lugar de nuestro servidor, al que puede acceder el usuario remoto, o que genera un error (es decir: 404 - No se encontró).

Las URL pueden ser de cualquier tipo (HTML, audio, gráficos, etc.).



**REFERRERS** (O 'referentes') Los referentes son aquellas URL que conducen a un usuario en vuestra web o hacen que el navegador solicite algo de nuestro servidor. La gran mayoría de las solicitudes se realizan desde sus propias URL, ya que la mayoría de las páginas HTML contienen enlaces a otros objetos de la propia web, como archivos gráficos. Si una de las páginas HTML de su web contiene enlaces a 10 imágenes, cada solicitud que una persona visitante haga a esta página HTML producirá 10 consultas más (*hits*), especificando como referente la URL de su propia página HTML.

**SEARCH STRINGS** Las cadenas de búsqueda o *search strings* se obtienen al examinar la cadena del *referrer* (referente) en busca de patrones conocidos de varios motores de búsqueda. Los motores de búsqueda y los patrones de búsqueda deben ser especificados por el usuario dentro de un archivo de configuración (por eso debe pedirnos esto si desea rastrear las cadenas de búsqueda). Si no, el valor predeterminado mostrará la mayoría de los patrones más habituales.

**USER AGENTS** Los *user agents* o agentes de usuario son un nombre elegante para los navegadores (Firefox, Opera, Konqueror, etc.). Cada navegador (agentes de usuario) se muestra de una manera única en nuestro servidor. Sin embargo, tenga en cuenta que

<<< **NOTA:** Sólo disponible si esta información está contenida en los registros del servidor.

<<< **NOTA:** Sólo disponible si esta información está contenida en los registros del servidor.



muchos navegadores permiten al usuario cambiar su nombre 'público', por lo que se encontrará algunos nombres falsos muy evidentes en la lista.

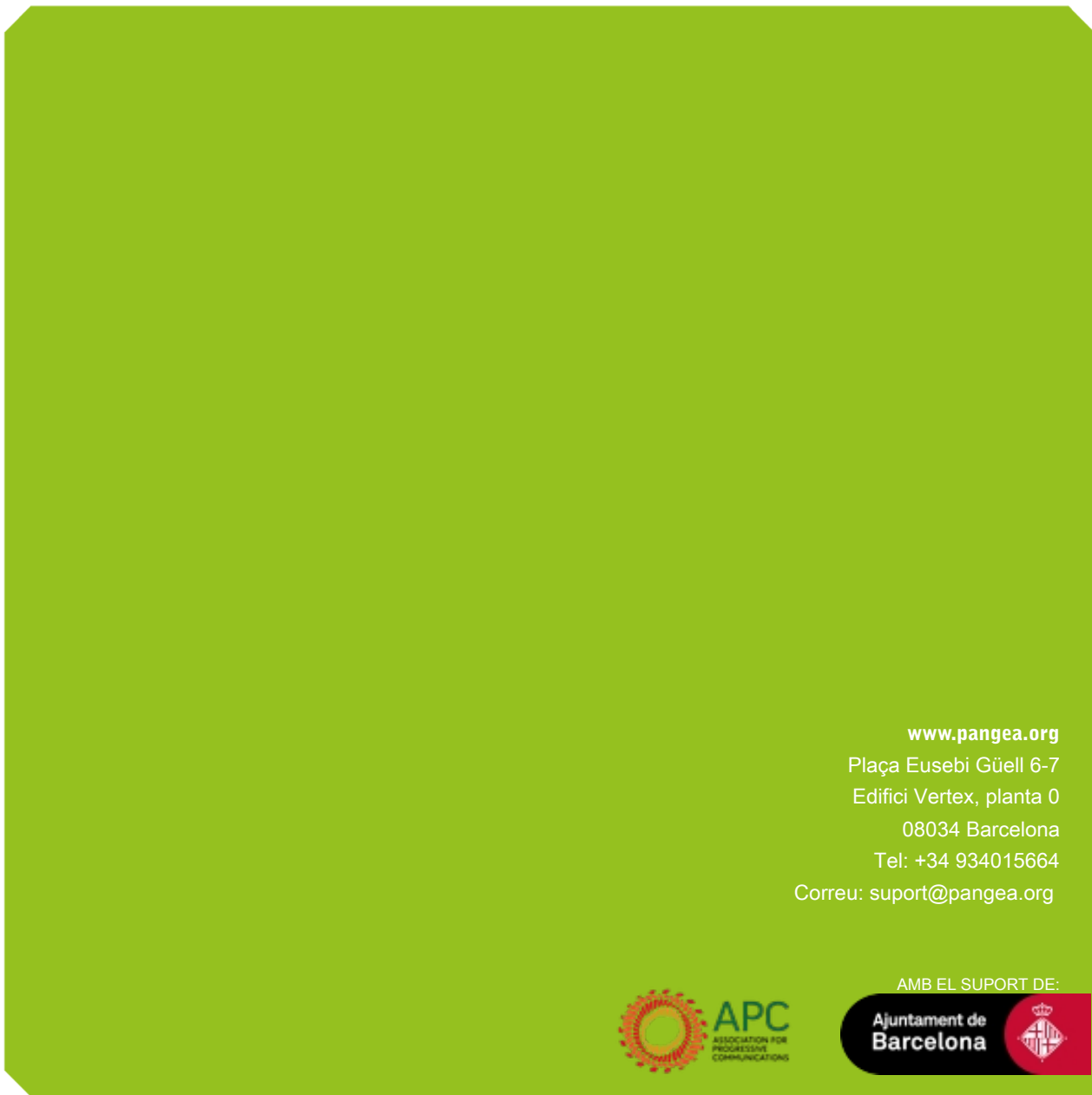
**ENTRY / EXIT PAGES** Las páginas de Entrada/Salida son aquellas páginas que fueron solicitadas por primera vez en una visita (Entrada), y las últimas solicitadas (Salida). Estas páginas se calculan utilizando la lógica de visitas que le hemos explicado anteriormente (ver **Visits**). Cuando una persona usuaria visita por primera vez algo en su web, la página solicitada se cuenta como una página de entrada, y cualquiera que sea la última URL solicitada (desde la que la persona visitante 'se marchó' de su web), se cuenta como una página de salida.

**COUNTRIES** Los países (countries) se determinan en función del dominio de nivel superior del sitio solicitante. Sin embargo, es cuestionable, ya que ya no hay una fuerte aplicación de los dominios como había en el pasado: un dominio .COM puede residir en EEUU o en otro país; un dominio .IL puede estar realmente en Israel, sin embargo, también puede estar ubicado en los EE.UU. o en otro lugar. Los dominios más comúnmente consultados son .COM (Comercial de EE.UU.), .NET (Red), .ORG (Organización sin ánimo de lucro) y .EDU (Educativo). Un gran porcentaje también puede mostrarse como No



resuelto/Desconocido, ya que bastantes accesos telefónicos y otros puntos de acceso de clientes no se resuelven en un nombre y se dejan como una dirección IP.

**RESPONSE CODES** Los códigos de respuesta se definen como parte del protocolo HTTP / 1.1 (RFC 2068). Estos códigos son generados por nuestro servidor web e indican el estado de finalización de cada solicitud que se le haya hecho.



[www.pangea.org](http://www.pangea.org)

Plaça Eusebi Güell 6-7

Edifici Vertex, planta 0

08034 Barcelona

Tel: +34 934015664

Correu: suport@pangea.org



AMB EL SUPORT DE:

Ajuntament de  
**Barcelona**



**Pangea**  
*.org*

< INTERNET  
ÈTIC I SOLIDARI >



Aquesta guia està subjecta a la llicència de Reconeixement-CompartirIgual 4.0 Internacional de Creative Commons. Si voleu veure una còpia d'aquesta llicència accediu a <http://creativecommons.org/licenses/by-sa/4.0/> o envieu una carta sol·licitant-la a Creative Commons, PO Box 1866, Mountain View, CA 94042, USA.