

Google Analytics vs. AWStats

Entendiendo el software de estadísticas de tráfico web

Willem L. Middelkoop

Aug. 20, 2018



Este mes he estado trabajando en estadísticas de sitios web, rastreando el tráfico usando diferentes tecnologías. Algunos de mis clientes usan Google Analytics, otros usan AWStats, y algunos usan ambos. Cuál es mejor es a menudo debatido, pero pocas personas realmente entienden las diferencias. Tiempo de arrojar algo de luz sobre la magia de las estadísticas web.

Comparando AWStats con Google Analytics

Es posible que espere que ambos programas de análisis web reporten la misma cantidad de tráfico del sitio web, pero este no es el caso. Intentaré explicar por qué los informes de Analytics y AWStats difieren (y por qué creo que realmente no importa).



Google Analytics y AWStats - software de estadísticas de tráfico web

Google Analytics

Como parte de Google Marketing Platform, Google Analytics es uno de los programas de análisis de tráfico web más utilizados disponibles. Google lanzó el servicio en 2005 después de adquirir varias compañías de software especializadas en análisis, en particular Urchin y Adaptive Path.

Cómo funciona Google Analytics

Google Analytics funciona colocando un pequeño fragmento de código de programación en su sitio web (o dentro de su aplicación). Esto inicia el seguimiento y la elaboración de perfiles que utiliza Google para generar estadísticas web. El código hace que su dispositivo de internet se conecte a los servidores de Google, que luego cuentan todas estas conexiones para generar estadísticas de uso.

Disponibilidad de Google Analytics

Google ofrece Analytics sin cargo: cuando utiliza Google Analytics, acepta compartir sus datos de tráfico con Google. Esto le permite a Google combinar datos de Analytics de diferentes sitios web para crear perfiles precisos de las personas que navegan por la web. Si bien estos perfiles son excelentes para entregar anuncios personalizados, también son motivo de preocupación por la privacidad. Google Analytics se utiliza en sitios web, dentro de aplicaciones y juegos.

AWStats

Creado originalmente por el francés [Laurent Destailleur](#) en 2000, AWStats es un programa de servidor ampliamente utilizado que analiza los archivos de registro del servidor web. El nombre AWStats es una abreviatura de "Advanced Web Statistics".

Cómo funciona AWStats

El registro del servidor web contiene el registro completo del tráfico del sitio web. Dentro del archivo de registro, se registran las conexiones individuales entre el servidor web y el dispositivo de Internet (teléfono, computadora, tableta). Después de analizar este registro de actividad, AWStats proporciona información detallada sobre el tráfico web.

Disponibilidad de AWStats

Todos pueden usar AWStats, ya que está disponible como software libre, incluido su código fuente. Anteriormente, escribí sobre [helping people with free software](#), explicando los beneficios del software libre (no se trata solo de dinero). AWStats debe tener acceso a los archivos de registro del servidor, lo que a menudo se realiza instalando AWStats en el propio servidor. Debido a esto, AWStats tiene un alcance limitado. No puede rastrear usuarios individuales en diferentes sitios web. AWStats se utiliza para servidores web, de correo, transmisión y archivos.

Diferencia: unidad de medida

Probablemente la diferencia más importante entre los sistemas es la unidad de medida. Google Analytics trabaja principalmente con "Sesiones" mientras que AWStats utiliza visitas (únicas):

- **Sesión de Google Analytics:** una serie de vistas de página desde el mismo navegador en un sitio web con no más de 30 minutos entre clics, excepto a la medianoche. Ver [explanation from Google](#).
- **Visita (única) de AWStats:** una dirección IP (única) que accede a una serie de páginas en un servidor web dentro de una hora. Si la serie se extiende más allá de la hora, la visita se cuenta como dos. Ver [explanation in the AWStats docs](#).

Diferencia: dónde se recopilan los datos

Además de la unidad de medida, el lugar donde se recopilan los datos también es diferente. AWStats utiliza el servidor web como lugar para medir. Literalmente, el software cuenta con qué frecuencia el servidor recibe una solicitud para entregar una página específica. El servidor responde a los visitantes humanos, pero también a las solicitudes automatizadas.

Bots

Estas solicitudes automatizadas las realizan bots, programas informáticos que navegan automáticamente por la web. A menudo, estos bots tienen una tarea específica, como indexar un sitio web para la lista de motores de búsqueda. La mayoría de los bots se identifican por su dirección IP de origen o el llamado UserAgent (una especie de texto de identificación que indica el origen del bot; como "GoogleBot"). La mayoría de los bots no se incluyen en las visitas de AWStats, sino que se especifican en una lista separada.

Robots/Spiders visitors	Hits	Last visit
139 different robots	17.671	20 Aug 2018 - 03:48
Googlebot-Image	15,863+1489	20 Aug 2018 - 04:27
bingbot	11,926+3019	20 Aug 2018 - 04:02
Seznambot	13,724+289	20 Aug 2018 - 04:36
Ahrefsbot	13,328+23	20 Aug 2018 - 04:30
empty user agent string	13,146+135	20 Aug 2018 - 04:30
Googlebot	12,891+249	20 Aug 2018 - 04:26
Yandexbot	6,330+4257	20 Aug 2018 - 03:25
Semrushbot	5,396+918	20 Aug 2018 - 04:37
DotBot	2,461+3359	20 Aug 2018 - 04:35
YandexImages	4,321	20 Aug 2018 - 19:22
Facebookexternalhit	3,498	20 Aug 2018 - 02:57
Mediapartners-Google	3,188	20 Aug 2018 - 22:21
MJ12bot	2,088+827	20 Aug 2018 - 00:39
Unknown robot identified by bot*	2,414+89	20 Aug 2018 - 14:15
Google Web Preview	2,473	20 Aug 2018 - 21:26
robot	2,358+16	20 Aug 2018 - 14:15
BaduSpider	1,700+48	20 Aug 2018 - 18:31
CCBot	1,238+64	20 Aug 2018 - 06:17
BLEXBot	1,125+141	20 Aug 2018 - 06:17
HubSpot Webcrawler	1,118+1	20 Aug 2018 - 20:29
MegaIndex.ru	1,095+7	20 Aug 2018 - 20:29
Preview	960	20 Aug 2018 - 06:49
Avicon	932	20 Aug 2018 - 18:38
idaprovider.com	814+33	20 Aug 2018 - 18:00
Freixox version 10 and lower - various robots	692+107	20 Aug 2018 - 18:00
robots.txt	784	20 Aug 2018 - 04:46
	447+254	20 Aug 2018 - 04:46
	573	20 Aug 2018 - 15:50
	538	20 Aug 2018 - 18:11
	518+6	20 Aug 2018 - 11:39
	0+483	20 Aug 2018 - 00:58
	272+185	20 Aug 2018 - 00:41
	304+82	20 Aug 2018 - 02:08
	308+28	20 Aug 2018 - 02:40
	167	20 Aug 2018 - 04:08
		20 Aug 2018 - 00:29
		20 Aug 2018 - 05:24
		15 Apr 2018 - 21:18
		19 Aug 2018 - 06:22
		19 Aug 2018 - 01:39
		13 May 2018 - 19:06
		20 Aug 2018 - 10:38
		11 Mar 2018 - 03:50
		2018 - 03:49

Robots navegando por la web: tráfico de bots listado por AWStats

Bots malos

Desafortunadamente, no todos los bots se comportan bien. Algunos de ellos trabajan para piratas informáticos para encontrar posibles debilidades de seguridad. Otros, como algunos robots de indexación de motores de búsqueda, se comportan como humanos para probar si el sitio web muestra la misma página tanto a humanos como a bots. Servir contenido diferente según el agente de usuario es una mala práctica utilizada por sitios web estúpidos para manipular la clasificación de los motores de búsqueda.

Reconociendo bots

AWStats intenta reconocer el comportamiento de los bots teniendo en cuenta la secuencia y la velocidad de las solicitudes de página (como un humano normal y saludable, no debería poder leer más de dos páginas por segundo). Pero esto no puede evitar que algunos bots se disfrazen con éxito como humanos, lo que hace que AWStats los cuente.

```

62.45.166.21 - [20/Aug/2018:20:49:38 +0200] "GET /en/upload/ HTTP/2.0" 200 46773 "https://willem.com/en/" "Mozilla/5.0 (X11; Linux x86_64; rv:52.0) Gecko/20100801 Firefox/52.0"
62.45.166.21 - [20/Aug/2018:20:49:47 +0200] "GET /blog/2018-08-02_helping-people-with-free-software HTTP/1.1" 200 43819 "Mozilla/5.0 (compatible; Googlebot/2.1; http://www.google.com/bot.html)"
62.45.166.21 - [20/Aug/2018:20:41:31 +0200] "GET / HTTP/2.0" 201 396 "https://willem.com/en/upload/" "Mozilla/5.0 (X11; Linux x86_64; rv:52.0) Gecko/20100801 Firefox/52.0"
62.45.166.21 - [20/Aug/2018:20:41:36 +0200] "GET /blog/ HTTP/2.0" 200 3232 "https://willem.com/en/" "Mozilla/5.0 (X11; Linux x86_64; rv:52.0) Gecko/20100801 Firefox/52.0"
62.45.166.21 - [20/Aug/2018:20:41:37 +0200] "GET /blog/2018-08-02_creating-animal-iphone-stand/ HTTP/2.0" 200 80682 "https://willem.com/blog/" "Mozilla/5.0 (X11; Linux x86_64; rv:52.0) Gecko/20100801 Firefox/52.0"
62.45.166.21 - [20/Aug/2018:20:41:42 +0200] "GET /blog/2018-08-08_working-from-a-busy-frame/ HTTP/2.0" 200 46928 "https://willem.com/blog/" "Mozilla/5.0 (X11; Linux x86_64; rv:52.0) Gecko/20100801 Firefox/52.0"
62.45.166.21 - [20/Aug/2018:20:41:46 +0200] "GET /blog/2018-07-18_enjoy-a-festival-without-alcohol/ HTTP/2.0" 200 40665 "https://willem.com/en/" "Mozilla/5.0 (X11; Linux x86_64; rv:52.0) Gecko/20100801 Firefox/52.0"
62.45.166.21 - [20/Aug/2018:20:41:52 +0200] "GET /blog/2018-05-09_lunch-by-plane/ HTTP/2.0" 200 63388 "https://willem.com/blog/" "Mozilla/5.0 (X11; Linux x86_64; rv:52.0) Gecko/20100801 Firefox/52.0"
62.45.166.21 - [20/Aug/2018:20:41:56 +0200] "GET /blog/2018-01-25_something-the-ameritech-will-never-have-partly/ HTTP/2.0" 200 40388 "https://willem.com/blog/" "Mozilla/5.0 (X11; Linux x86_64; rv:52.0) Gecko/20100801 Firefox/52.0"
62.45.166.21 - [20/Aug/2018:20:42:00 +0200] "GET /blog/2018-07-28_helping-people-with-free-software/ HTTP/2.0" 200 63919 "https://willem.com/blog/" "Mozilla/5.0 (X11; Linux x86_64; rv:52.0) Gecko/20100801 Firefox/52.0"
62.45.166.21 - [20/Aug/2018:20:42:05 +0200] "GET /blog/hacking/ HTTP/2.0" 200 41997 "https://willem.com/blog/2018-07-28_helping-people-with-free-software/" "Mozilla/5.0 (X11; Linux x86_64; rv:52.0) Gecko/20100801 Firefox/52.0"
62.45.166.21 - [20/Aug/2018:20:42:08 +0200] "GET /blog/2017-07-31_working-from-g-turned-off-my-lan-and-went-g-only/ HTTP/2.0" 200 44542 "https://willem.com/blog/hacking/" "Mozilla/5.0 (X11; Linux x86_64; rv:52.0) Gecko/20100801 Firefox/52.0"
62.45.166.21 - [20/Aug/2018:19:42:10 +0200] "GET /blog/2017-07-31_the-day-i-killed-my-lan/ HTTP/2.0" 200 61792 "https://willem.com/blog/2017-07-31_working-from-g-turned-off-my-lan-and-went-g-only/" "Mozilla/5.0 (X11; Linux x86_64; rv:52.0) Gecko/20100801 Firefox/52.0"

```

Parte de un archivo de registro de un servidor web, cada línea corresponde a una solicitud de página - ¿ves el bot?

JavaScript

En lugar de ejecutar software en el servidor como AWStats, Google Analytics funciona ejecutando código (JavaScript) en el dispositivo del visitante. Dependiendo de cómo esté

integrado el JavaScript de Analytics, la medición a menudo se retrasa porque lleva algún tiempo cargar JavaScript. A veces, solo comienza después de que la página completa (incluidas las imágenes y las fuentes) se haya cargado por completo. Las personas que abandonan la página antes de que se cargue por completo, o las personas que tienen JavaScript deshabilitado, por lo tanto, no se miden. Debido a que Google Analytics requiere JavaScript, la mayoría de los bots no se miden porque es difícil para los bots simular un navegador habilitado para JavaScript.

```

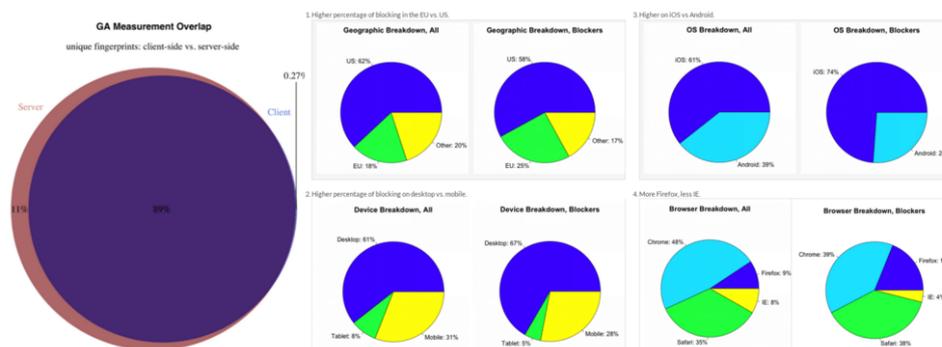
1 <script>
2 (function(i,s,o,g,r,a,m){i['GoogleAnalyticsObject']=r;i[r]=i[r]||function(){
3 (i[r].q=i[r].q||[]).push(arguments)},i[r].l=1*new Date();a=s.createElement(o),
4 m=s.getElementsByTagName(o)[0];a.async=1;a.src=g;m.parentNode.insertBefore(a,m)
5 })(window,document,'script','https://www.google-analytics.com/analytics.js','ga');
6
7 ga('create', 'UA-XXXXXXX-1', 'auto');
8 ga('send', 'pageview');
9
10 </script>

```

Fragmento típico de código JavaScript de Google Analytics - se puede ver claramente la referencia a los servidores de Google

Bloqueo de Google Analytics

Debido a las preocupaciones de privacidad con respecto a Google Analytics, algunas personas bloquean activamente el JavaScript utilizado por Google Analytics. Hay complementos de navegador y bloqueadores de anuncios que hacen esto automáticamente, sin requerir ningún conocimiento técnico del usuario. Otro desarrollador web inteligente utilizó una configuración de servidor especialmente diseñada para descubrir cuántas personas bloquean Google Analytics. En su experimento, [Jason Packer](#) descubrió que alrededor del 8-11% de los usuarios bloquean Google Analytics.



Porcentaje de personas que bloquean Google Analytics (2016), quantable.com - Jason Packer

Conclusión

Tanto la unidad de medida como el lugar donde se realiza la medición provocan las diferencias en las estadísticas de Google Analytics y AWStats. Incluidos los bots, AWStats cuenta demasiado, mientras que Google Analytics cuenta muy poco debido a problemas con la carga y el bloqueo de JavaScript.

Pero al final, realmente no importa, ya que debe utilizar estos programas estadísticos principalmente para analizar el rendimiento relativo. Las inexactitudes constantes en los números absolutos no influyen realmente en las respuestas a preguntas como estas:

- ¿Mi sitio web funciona mejor que ayer, el mes pasado o el año pasado?
- ¿Qué página, publicación de blog o producto de tienda web es más popular?
- ¿Está aumentando el tráfico medio? ¿Mi campaña publicitaria tiene algún efecto?

Si realmente quiere saber *quién* está visitando su sitio web, debería considerar pedirle al visitante que se identifique. Eso es mucho más preciso (¡y educado!) que confiar en Google Analytics o AWStats para identificar personas únicas. De esa manera, no tiene que preocuparse de que las personas tengan varios dispositivos y accedan a su sitio web desde diferentes ubicaciones, como el trabajo, el hogar o mientras viajan (tren, autobús).

Utilice software de estadísticas para realizar un seguimiento del rendimiento relativo, pero considere conocer a su visitante conectándose con él en lugar de recopilar datos de él. Construya relaciones en lugar de perfiles de datos. Ahí es donde está el valor.