

Google Analytics vs. AWStats

Inzicht in web traffic statistieken software

Willem L. Middelkoop

20 aug. 2018



Deze maand heb ik gewerkt aan webstatistieken, waarbij ik verkeer heb bijgehouden met verschillende technologieën. Sommige van mijn klanten gebruiken Google Analytics, anderen gebruiken AWStats, en sommige gebruiken beide. Welke beter is, wordt vaak bediscussieerd, maar weinig mensen begrijpen echt de verschillen. Tijd om wat licht te werpen op de magie van webstatistieken.

Vergelijking van AWStats met Google Analytics

Je zou verwachten dat beide webanalyseprogramma's dezelfde hoeveelheid websiteverkeer rapporteren, maar dit is niet het geval. Ik zal proberen uit te leggen waarom rapporten van Analytics en AWStats verschillen (en waarom ik denk dat het niet echt uitmaakt).



Google Analytics en AWStats - web traffic statistieken software

Google Analytics

Als onderdeel van het Google Marketing Platform is Google Analytics een van de meest gebruikte webverkeersanalyseprogramma's die beschikbaar zijn. Google lanceerde de dienst in 2005 na de overname van meerdere softwarebedrijven die gespecialiseerd zijn in analytics, met name Urchin en Adaptive Path.

Hoe Google Analytics werkt

Google Analytics werkt door een klein stukje programmeercode op je website (of in je app) te plaatsen. Dit initieert de tracking en profilering die Google gebruikt om webstatistieken te genereren. De code zorgt ervoor dat je internetapparaat verbinding maakt met de servers van Google - die vervolgens al deze verbindingen tellen om gebruiksstatistieken te genereren.

Beschikbaarheid van Google Analytics

Google biedt Analytics gratis aan: wanneer je Google Analytics gebruikt, ga je ermee akkoord om je verkeersgegevens met Google te delen. Hierdoor kan Google Analytics-gegevens van verschillende websites combineren om nauwkeurige profielen te creëren van mensen die op het web surfen. Hoewel deze profielen geweldig zijn om gepersonaliseerde advertenties te leveren, zijn ze ook reden tot bezorgdheid over privacy. Google Analytics wordt gebruikt op websites, in apps en games.

AWStats

Oorspronkelijk gemaakt door de Fransman [Laurent Destailleur](#) in 2000, is AWStats een veelgebruikt serverprogramma dat webserverlogbestanden analyseert. De naam AWStats is een afkorting voor "Advanced Web Statistics".

Hoe AWStats werkt

Het webserverlogboek bevat de volledige registratie van websiteverkeer. In het logbestand worden individuele verbindingen tussen de webserver en het internetapparaat (telefoon, computer, tablet) geregistreerd. Na analyse van deze activiteitenregistratie biedt AWStats gedetailleerde inzichten in het webverkeer.

Beschikbaarheid van AWStats

Iedereen kan AWStats gebruiken omdat het beschikbaar is als gratis software, inclusief de broncode. Eerder schreef ik over [helping people with free software](#), waarin ik de voordelen van gratis software uitleg (het gaat niet alleen om geld). AWStats moet toegang hebben tot de serverlogbestanden, wat vaak gebeurt door AWStats op de server zelf te installeren. Hierdoor heeft AWStats een beperkte scope. Het is niet in staat om individuele gebruikers te volgen op verschillende websites. AWStats wordt gebruikt voor web-, mail-, streaming- en bestandsservers.

Vershil: meeteenheid

Waarschijnlijk het belangrijkste verschil tussen de systemen is de meeteenheid. Google Analytics werkt voornamelijk met "Sessies" terwijl AWStats (unieke) bezoeken gebruikt:

- **Google Analytics-sessie:** een reeks paginaweergaven vanuit dezelfde browser op een website met niet meer dan 30 minuten tussen klikken, behalve om middernacht. Zie [explanation from Google](#).
- **AWStats (uniek) bezoek:** een (uniek) IP-adres dat binnen een uur een reeks pagina's op een webserver bezoekt. Als de reeks langer duurt dan een uur, wordt het bezoek als twee geteld. Zie [explanation in the AWStats docs](#).

Vershil: waar gegevens worden verzameld

Naast de meeteenheid is ook de plaats waar gegevens worden verzameld anders. AWStats gebruikt de webserver als meetplaats. De software telt letterlijk hoe vaak de server een verzoek ontvangt om een specifieke pagina te leveren. De server reageert op menselijke bezoekers, maar ook op geautomatiseerde verzoeken.

Bots

Deze geautomatiseerde verzoeken worden gedaan door bots, computerprogramma's die automatisch op het web surfen. Vaak hebben deze bots een specifieke taak, zoals het indexeren van een website voor zoekmachinevermelding. De meeste bots worden geïdentificeerd door hun oorspronkelijke IP-adres of zogenaamde UserAgent (een soort identificatietekst die de oorsprong van de bot aangeeft; zoals "GoogleBot"). De meeste bots zijn niet opgenomen in AWStats-bezoeken, maar worden vermeld in een aparte lijst.

Robots/Spiders visitors	Hits	Size	Last visit
Googlebot-Image	17.671	2.14 GB	20 Aug 2018 - 03:48
bingbot	15,863+1489	4.19 GB	20 Aug 2018 - 04:27
Seznambot	11,926+3019	2.98 GB	20 Aug 2018 - 04:36
Ahrefsbot	13,724+289	3.36 GB	20 Aug 2018 - 04:30
empty user agent string	13,328+23	1.07 GB	20 Aug 2018 - 04:30
Googlebot	12,891+249	576.62 MB	20 Aug 2018 - 04:35
Yandexbot	6,330+4257	300.78 MB	20 Aug 2018 - 02:57
Semrushbot	5,396+918	282.19 MB	20 Aug 2018 - 19:22
DotBot	2,461+3359	1.42 GB	20 Aug 2018 - 22:21
YandexImages	4,321	894.91 MB	20 Aug 2018 - 00:39
Facebookexternalhit	3,498	243.94 MB	24 May 2018 - 14:15
Mediapartners-Google	3,188	115.42 MB	20 Aug 2018 - 21:26
MJ12bot	2,088+827	262.81 MB	20 Aug 2018 - 03:31
Unknown robot identified by bot *	2,414+89	718.51 MB	17 Aug 2018 - 07:59
Google Web Preview	2,473	118.29 MB	26 Jul 2018 - 06:17
robot	2,358+16	101.83 MB	20 Aug 2018 - 06:17
BaduSpider	1,700+48	158.68 MB	20 Aug 2018 - 20:29
CCBot	1,238+64	79.05 MB	19 Aug 2018 - 06:49
BLEXBot	1,125+141	51.80 MB	19 Aug 2018 - 18:38
HubSpot Webcrawler	1,118+1	48.86 MB	04 Feb 2018 - 18:00
MegaIndex.ru	1,095+7	45.05 MB	10 Aug 2018 - 04:46
Previewer	960	48.86 MB	18 Aug 2018 - 15:50
avicon	932	124.54 MB	18 Aug 2018 - 18:11
idaprovider.com	814+33	26.81 MB	18 Aug 2018 - 11:39
Freixox version 10 and lower - various robots	692+107	47.08 MB	18 Aug 2018 - 00:58
robots.txt	784	48.20 MB	20 Aug 2018 - 00:41
	447+254	70.44 MB	20 Aug 2018 - 02:48
	573	21.02 MB	20 Aug 2018 - 02:08
	538	29.96 MB	20 Aug 2018 - 22:40
	518+6	22.03 MB	20 Aug 2018 - 04:08
	0+483	40.88 MB	15 Aug 2018 - 05:24
	272+185	20.64 MB	16 Apr 2018 - 21:18
	304+82	14.77 MB	19 Aug 2018 - 06:22
	308+98	4.43 MB	13 May 2018 - 01:39
	167	13.07 MB	20 Aug 2018 - 10:38
		4.42 MB	11 Jul 2018 - 03:50
		4.43 MB	20 Aug 2018 - 04:49

Robots die op het web surfen: bot traffic weergegeven door AWStats

Slechte bots

Helaas gedragen niet alle bots zich netjes. Sommigen van hen werken voor hackers om potentiële beveiligingsproblemen te vinden. Anderen, zoals sommige indexeringsrobots van zoekmachines, gedragen zich als mensen om te testen of de website dezelfde pagina toont aan zowel mensen als bots. Het aanbieden van verschillende inhoud op basis van de user agent is een slechte praktijk die wordt gebruikt door domme websites om de ranking in zoekmachines te manipuleren.

Bots herkennen

AWStats probeert botgedrag te herkennen door rekening te houden met de volgorde en snelheid van paginaverzoeken (zoals een normaal, gezond mens niet meer dan twee pagina's per seconde zou moeten kunnen lezen). Maar dit kan niet voorkomen dat sommige bots zich succesvol vermommen als mens, waardoor AWStats ze meetelt.

```

62.45.166.21 - [20/Aug/2018:20:42:38 +0200] "GET /en/upload/ HTTP/2.0" 200 46773 "https://willem.com/en/" "Mozilla/5.0 (X11; Linux x86_64; rv:52.0) Gecko/20100101 Firefox/52.0"
62.45.166.21 - [20/Aug/2018:20:42:38 +0200] "GET / HTTP/2.0" 200 390 "https://willem.com/en/upload/" "Mozilla/5.0 (X11; Linux x86_64; rv:52.0) Gecko/20100101 Firefox/52.0"
62.45.166.21 - [20/Aug/2018:20:42:38 +0200] "GET /blog/ HTTP/2.0" 200 3223 "https://willem.com/en/" "Mozilla/5.0 (X11; Linux x86_64; rv:52.0) Gecko/20100101 Firefox/52.0"
62.45.166.21 - [20/Aug/2018:20:42:37 +0200] "GET /blog/2018-06-02_creating-a-minimal-iphone-stand/ HTTP/2.0" 200 6062 "https://willem.com/blog/" "Mozilla/5.0 (X11; Linux x86_64; rv:52.0) Gecko/20100101 Firefox/52.0"
62.45.166.21 - [20/Aug/2018:20:42:37 +0200] "GET /blog/2018-06-02_working-from-a-bike-frame/ HTTP/2.0" 200 6092 "https://willem.com/blog/" "Mozilla/5.0 (X11; Linux x86_64; rv:52.0) Gecko/20100101 Firefox/52.0"
62.45.166.21 - [20/Aug/2018:20:42:36 +0200] "GET /blog/2018-07-18_enjoy-a-festival-without-alcohol/ HTTP/2.0" 200 6065 "https://willem.com/en/" "Mozilla/5.0 (X11; Linux x86_64; rv:52.0) Gecko/20100101 Firefox/52.0"
62.45.166.21 - [20/Aug/2018:20:42:35 +0200] "GET /blog/2018-05-09_lunch-by-plane/ HTTP/2.0" 200 6380 "https://willem.com/blog/" "Mozilla/5.0 (X11; Linux x86_64; rv:52.0) Gecko/20100101 Firefox/52.0"
62.45.166.21 - [20/Aug/2018:20:42:35 +0200] "GET /blog/2018-03-25_something-the-cause-walk-will-never-leave-patiny/ HTTP/2.0" 200 6038 "https://willem.com/blog/" "Mozilla/5.0 (X11; Linux x86_64; rv:52.0) Gecko/20100101 Firefox/52.0"
62.45.166.21 - [20/Aug/2018:20:42:30 +0200] "GET /blog/2018-07-20_helping-people-with-free-software/ HTTP/2.0" 200 63819 "https://willem.com/blog/" "Mozilla/5.0 (X11; Linux x86_64; rv:52.0) Gecko/20100101 Firefox/52.0"
62.45.166.21 - [20/Aug/2018:20:42:30 +0200] "GET /blog/hacking/ HTTP/2.0" 200 4197 "https://willem.com/blog/2018-07-20_helping-people-with-free-software/" "Mozilla/5.0 (X11; Linux x86_64; rv:52.0) Gecko/20100101 Firefox/52.0"
62.45.166.21 - [20/Aug/2018:20:42:18 +0200] "GET /blog/2017-07-31_working-from-g-turned-off-my-lan-and-went-g-only/ HTTP/2.0" 200 4462 "https://willem.com/blog/hacking/" "Mozilla/5.0 (X11; Linux x86_64; rv:52.0) Gecko/20100101 Firefox/52.0"
62.45.166.21 - [20/Aug/2018:20:42:13 +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/20100101 Firefox/52.0"

```

Onderdeel van een webserver logbestand, elke regel komt overeen met een paginaverzoek - zie je de bot?

JavaScript

In plaats van software op de server uit te voeren zoals AWStats, werkt Google Analytics door code (JavaScript) uit te voeren op het apparaat van de bezoeker. Afhankelijk van

hoe de Analytics JavaScript is geïntegreerd, wordt de meting vaak vertraagd omdat het enige tijd kost om JavaScript te laden. Soms begint het pas nadat de hele pagina (inclusief afbeeldingen en lettertypen) volledig is geladen. Mensen die de pagina verlaten voordat deze volledig is geladen, of mensen die JavaScript hebben uitgeschakeld, worden daarom niet gemeten. Omdat Google Analytics JavaScript vereist, worden de meeste bots niet gemeten omdat het moeilijk is voor bots om een JavaScript-browser te simuleren.

```

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>

```

Typisch stukje Google Analytics JavaScript code - je kunt duidelijk zien dat het verwijst naar de servers van Google

Google Analytics blokkeren

Vanwege de privacyproblemen met betrekking tot Google Analytics blokkeren sommige mensen actief de JavaScript die door Google Analytics wordt gebruikt. Er zijn browserplugins en adblockers die dit automatisch doen, zonder dat de gebruiker technische kennis nodig heeft. Een andere slimme webontwikkelaar gebruikte een speciaal gebouwde serverconfiguratie om te achterhalen hoeveel mensen Google Analytics blokkeren. In zijn experiment ontdekte [Jason Packer](#) dat ongeveer 8-11% van de gebruikers Google Analytics blokkeert.



Percentage van mensen die Google Analytics blokkeren (2016), quantable.com - Jason Packer

Conclusie

Zowel de meeteenheid als de plaats waar de meting wordt uitgevoerd, veroorzaken de verschillen in statistieken van Google Analytics en AWStats. Inclusief bots telt AWStats te veel, terwijl Google Analytics te weinig telt vanwege problemen met het laden en blokkeren van JavaScript.

Maar uiteindelijk maakt het niet echt uit, want je moet deze statistiekprogramma's vooral gebruiken om relatieve prestaties te analyseren. Consistente onnauwkeurigheden

in absolute aantallen hebben geen echte invloed op antwoorden op vragen zoals deze:

- Doet mijn website het beter dan gisteren, vorige maand of vorig jaar?
- Welke pagina, blogpost of webshop product is het populairst?
- Neemt het gemiddelde verkeer toe? Heeft mijn reclamecampagne effect?

Als je echt wilt weten *wie* je website bezoekt, moet je echt overwegen om de bezoeker te vragen zich te identificeren. Dat is veel nauwkeuriger (en beleefder!) dan te vertrouwen op Google Analytics of AWStats om unieke personen te identificeren. Op die manier hoeft je je geen zorgen te maken over mensen die meerdere apparaten hebben en je website vanaf verschillende locaties bezoeken, zoals werk, thuis of onderweg (trein, bus).

Gebruik statistieksoftware om relatieve prestaties te volgen, maar overweeg om je bezoeker te leren kennen door verbinding met hen te maken in plaats van gegevens van hen te verzamelen. Bouw relaties op in plaats van gegevensprofielen. Daar ligt de waarde.