

Background

Google Analytics is a free web analytics package provided by Google. It reports on Page Impressions and Unique Browsers, and is optimised to plug into Google's AdWords and AdSense paid search tools.

Because it is zero-cost (for smaller volumes of traffic - see caveat 3 below), Google Analytics is used by a large variety of media owners.

Fig. 1 Sample Google Analytics Dashboard

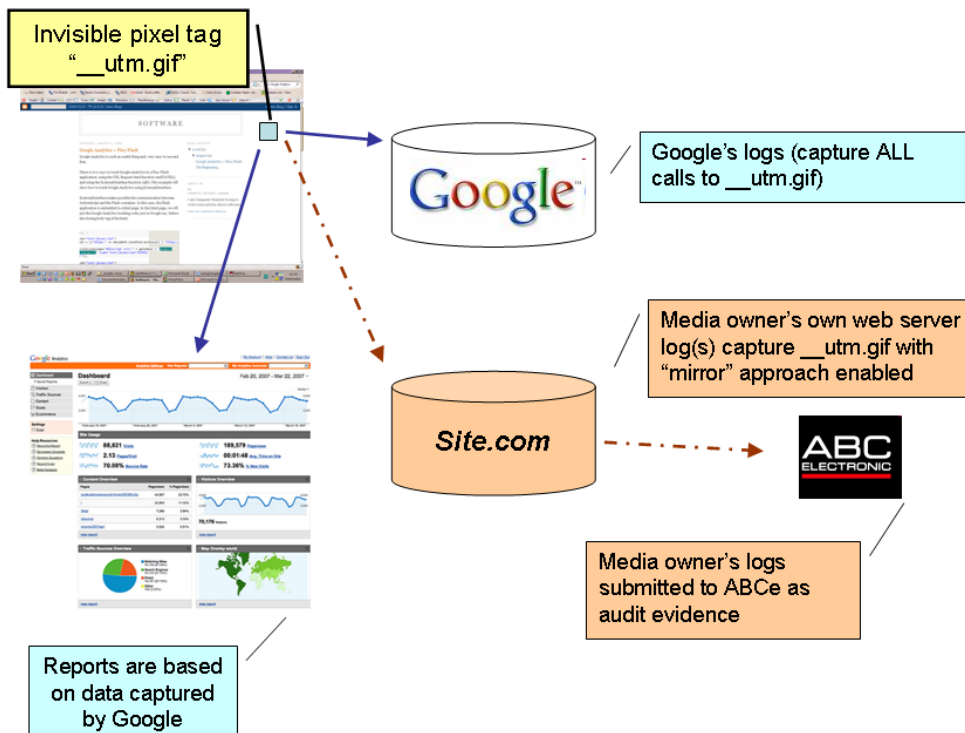


Google Analytics is a Javascript page tagging service based on Urchin, an analytics tool that was acquired by Google. Media owners wishing to run the code must incorporate some HTML into each page they wish to tag. The code makes a call to Google's servers, which host the code. The media owner does **not** host the code on their own web-servers.

Note that Google Urchin is a separate tool self-hosted by the media owner and is **not** covered by this paper; for details on ABCe auditing using Google Urchin please contact your ABC Account Manager.

The ABCe website audit requires the submission of supporting logfile data. For legal and privacy reasons, Google have hitherto not released their data for any purpose, and so ABCe has suggested that media owners "mirror" the logfile data. In other words, the media owner should configure their Google Analytics tag (or write their own supplementary code) to ensure that every request sent to Google has a matching request logged by the media owner's own server.

Fig. 2 Google Analytics Logging



Technical Assistance

The basic Google Analytics code can be found at the URL
<http://code.google.com/apis/analytics/docs/tracking/asyncTracking.html> .

The older version of the code based on "ga.js" can be found at
<http://code.google.com/apis/analytics/docs/tracking/gaTrackingOverview.html> .

We suggest you check Google's own support pages for the latest updates.

Guidance on how to set up the mirroring approach appropriately to the version of the Google Analytics code used on the media owner's site is available from
<http://code.google.com/apis/analytics/docs/tracking/asyncMigrationExamples.html#local> (see the section on "Using a Local Server").

The mirroring approach will direct requests for __utm.gif to a different set of server logs, which should be accessible to the media owner. The media owner can then filter the logs to show only the __utm.gif calls and submit this data to ABCe in support of the Page Impression and Unique Browser claims reported by Google Analytics.

Media owners who are expecting a high level of traffic should consider directing all __utm.gif requests to a dedicated server handling only these requests (for example, reportingserver.site.com). This may require additional technical configuration such as using a load-balancer to route all __utm.gif requests to such a server. The benefit of this is that the media owner is able to collect the log files from this server (or group of servers) easily, and hence to supply ABCe with a copy of the data on which the Google Analytics reporting is based for auditing purposes (see Caveat 4 below).

ABCe recommends that media owners who have issues configuring Google, and who require a full expert opinion, follow Google's advice and request the services of a Google Analytics Consultant, usually a third party consultancy accredited by Google, via the following URL: <http://www.google.com/intl/en/analytics/partners.html> .

Caveats

1. ABCe cannot provide consultancy services beyond suggesting the above approach. Furthermore, the cookie identifier used by Google Analytics ("utma") for the counting of Unique Browsers contains the domain information. Therefore, any network comprising more than one top-level domain will not be counted in compliance with the industry-agreed standard, since this requires all cookies to be persistent across all certified domains. Google Analytics can be configured to count across a "tree" of subdomains, but not across more than one top-level domain.
Google provide instructions on configuring across subdomains to report Unique Browsers compliantly at the URL
<http://code.google.com/apis/analytics/docs/tracking/gaTrackingSite.html>
2. In the light of this, some media owners have used the Google Analytics tag but chosen to count Unique Browsers via a different cookie implemented on their own servers collecting the mirror data, which can be logged because they are using their own servers. Alternatively, IP and User-Agent combination can be used to count Unique Browsers and will **not** be subject to the domains issue outlined above.
However, in **both** of these cases, the Unique Browser numbers reported by Google will probably **not** be substantiated by the data provided for audit – even though the Page Impressions may be.
3. The media owner's log files will need to comply with industry-agreed reporting standards for the minimum fields required –
 - Date
 - Time
 - IP Address
 - URL and Query (Domain will be found within the query string of the __utm.gif, as will Referrer of the page)
 - User-Agent
 - Method
 - Status
 - Cookie (so ensuring that _utma is logged)
 - Referrer (which will show the URL of the page, since it is the Referrer of utm.gif)

and will need to be a single uniform format for ALL web servers. Furthermore, ABCe recommends that if the media owner expects to count more than 20 million Page Impressions in the month, the Google Analytics .gif calls should be captured in a separate dedicated log file. This will decrease the time and cost required to isolate these .gif calls.

Note also that Google state the following in
<http://www.google.com/support/analytics/bin/answer.py?answer=55476&ctx=sibling>:

"Google Analytics is a free service that offers users up to 5 million pageviews a month. If your site generates more than 5 million pageviews per month, you



will need to link it to an active AdWords account in order to avoid interruption of your Google Analytics service.”

4. ABCe audits of website traffic figures counted by Google Analytics have currently only certified the reported Page Impression and Unique Browser totals. In particular, Google Analytics does not calculate Visits or Visit Duration in accordance with industry-agreed standards.

Google Analytics uses a cookie to calculate Visits which expires either after 30 minutes or, crucially, whenever the browser closes.

The industry-agreed Visit metric is an open standard, which does not require sites to use cookies and is based on the interval between Page Impressions. This does not take account of situations where the user closes their browser (which causes the Google session cookie to end) and then commences another session on the same site within 30 minutes.

Therefore Google Analytics will typically report more Visits than can be certified. Other analytics tools - including most of the ABCe 2-star Associate Subscribers listed on <http://www.abc.org.uk/Corporate/AboutABCe/associates.aspx> - report Visits in accordance with the industry-agreed standards.

Visit Duration is affected by this issue. Furthermore, Google Analytics assigns single-page Visits (so-called “bounces”) a duration of 0 seconds. The JICWEBS Reporting Standard for Visit Duration does **not** count single-page Visits. For both these reasons, the Visit Duration reported by Google Analytics will typically be shorter than the industry-agreed standard as certified by ABCe.

ABCe



www.abc.org.uk

