Audit Bureau of Circulations of South Africa

Reporting Standards
Web Traffic
Version 3 2015
Issued 1 October 2015

SECTION A: INTRODUCTION & MANDATORY METRICS

This document contains the ABC Reporting Standards for products registered with ABC and reporting Web Traffic. As Reporting Standards are updated periodically, please check the website www.abc.org.za to ensure you are using the latest applicable standards.

If you have any queries regarding how the Reporting Standards affect you or any specific queries please contact the ABC General Manager on +27 11 4474290 or email abc@abc.org.za

Category and Type	Metric Name	Metric Definition (See Appendix 1 for detailed definitions)
Web, Reach	**Unique Browser**	A unique and valid identifier (e.g. IP address +User-Agent and/or Cookie, UDID for Apps). These are reported as the de-duplicated net number of Unique Browsers for a given period.
Web, Volume	**Page Impression**	A file, or combination of files, sent to a valid browser as a result of that browser's request being received by the server.

SECTION B:GENERAL GUIDANCE

INTRODUCTION

This section sets out the data and information that is reported on the Web Traffic certificate.

PRINCIPLES

B1. You must report traffic for a defined Reporting Period

B2. You must report the mandatory metrics

B3. You must report the claimed inventory

B4. You must report a product name

REQUIREMENTS

B1. You must report traffic for a defined Reporting Period

a) You will report traffic generated in a defined Reporting Period.

b) The Reporting Period must be quarterly, six-monthly or annually.

B2. You must report the mandatory metrics

a) The mandatory metrics that must be reported are:

i) The average of the Daily Unique Browsers for each calendar month in the Reporting Period (i.e. de-duplicated by Unique Browser within each day but not between days). *For example: average of Daily Unique Browsers for January, February March, etc.*

ii) Page impressions.

B3. You must report the claimed inventory

a) You must report on the certificate a domain/inventory listing of those domains, URLs or content identifiers (e.g. Apps) that cover at least the top 95% of the total Page Impressions relating to the Unique Browsers being certified.

i) The listing will be in descending size order of Page Impressions.

B4. You must report a product name

a) You must specify the product name you wish to appear on the certificate.

i) If this name is a descriptor rather than a domain or URL then it must reasonably reflect the claimed inventory.

ii) If this name is a domain or URL that appears in the list of claimed inventory then it must be the largest by Page Impressions.

SECTION C: REPORTING REQUIREMENTS

This section sets out the requirements relating to the reporting of web traffic.

PRINCIPLES

C1. Traffic must comply with the Reporting Standards

C2. Traffic must be human-initiated and the associated content is intended to be seen by the user

C3. Evidence to support the claim must be retained and available for a minimum period.

C4. Each web traffic claim is audited to verify it is in accordance with the applicable Reporting Standards

REQUIREMENTS

C1. Traffic must comply with the Reporting Standards

No additional requirements.

C2. Traffic must be human-initiated and the associated content is intended to be seen by the user

a) You must exclude robotic traffic. *Examples of robotic traffic include that generated from search engines, personal spiders, automated site monitoring tools, offline browsers, automated content requests from PDA devices, web feed aggregators and other automated syndication agents.*

b) You must exclude traffic that has been 'pushed' into the user's browser (i.e. not the result of an intentional user request). *Examples of pushed traffic are subsited traffic, contextual linking and ISP page replacement.*

d) You must exclude any filetypes that are served in conjunction with a valid Page Impression. *For example: graphics or stylesheets.*

e) You must exclude traffic to URLs that are concurrently served with a valid Page Impression or do not provide content that is intended to be seen by the user. *For example framesets, pop-ups or departure pages (served with no visible content for site monitoring purposes).*

   i) However, if a user requests that a panel within a frameset (or a pop-up) is refreshed, then serving the refreshed panel or pop-up may be counted as a Page Impression since it has been requested.

   ii) You must exclude initial requests for PDF files (with a status code of 200) from Page Impressions if these are already certified as downloads (as they can't be counted twice).

f) You must exclude invalid HTTP transactions.

   i) You must exclude log records with invalid HTTP status codes - defined as any records that do **not** have the following codes: 200, 201, 202, 203, 204, 205 and

304. (Note that status code 206 must be excluded because it indicates a partial fulfillment of a request).

ii) You must exclude HTTP method requests that are **neither** "GET" nor "POST". Note: These requirements apply to data logs for browser-side page tagging tools by default because the page must be rendered successfully in order for the tag code to run.

C3. Evidence to support the claim must be retained and available for a minimum period.

a) You must retain and be able to provide all records supporting the claim. *For example: For User Account metrics evidence of registration and/or payment.*

b) The logged records provide the specific data regarding each file request processed by the server.

i) You must retain and be able to provide for audit all the logged records supporting the claim.

ii) You must retain the logged records for a period of 6 months following certification of the claim or until the audit of the certificate for the subsequent Reporting Period has been completed if sooner.

c) The logged records must contain sufficient information to identify the traffic to be counted and audited. This will be agreed with your auditor; but:

i) This will include the identification of the time and date of the transaction, identification of the Unique Browser and details of the URL/query parameters requested. Please refer to guidance for details of the information typically required. Note:

- This must include sufficient information to identify and exclude robotic traffic (typically IP address and User Agent)
- You must be able to differentiate between web and App traffic.
- Any anonymisation techniques applied to the logs must be agreed in advance with your auditor.

ii) You must ensure data collection servers are date and time synchronised, preferably to GMT, so that all their log files' date and time stamping are aligned.

iii) You should not change the format of your log files during the Reporting Period. Please contact your auditor regarding any planned changes.

C4. Each web traffic claim is audited to verify it is in accordance with the applicable Reporting Standards

a) The audit must be carried out by ABC Staff Auditors.

b) Requirements in relation to the auditor and audits will be covered by the ABC Rules, ABC Audit Programmes and contractual arrangements.

c) If following an audit we identify material problems with the Return Form or Certificate then we will propose to revise the claim. If a Certificate has already been issued we will issue a revised Certificate with an Audit Report that identifies the changes. This replaces your original Certificate and must be used in its place. The process is as follows:

    i) We will send you a letter detailing the reason/problem giving rise to the amendment.

    ii) You will have 10 working days from the receipt of this letter to provide any further information to us, or object to the revision of the claim.

    iii) If you wish to object to the revision of the claim you must do this in writing to the General Manager who will investigate and provide a decision within 10 working days.

    iv) Subsequent Certificates will not be issued until we have resolved all queries on a previous audit and issued the revised Certificate, if applicable.

REPORTED DATA

GENERAL PRINCIPLES

GC1. Traffic must be human-initiated and the associated content is intended to be seen by the user

a) *You must exclude robotic traffic.*

i) It is recommended best practice to use the industry-standard ABC/IAB Global Robots and Spiders List ("ABC/IAB Robots List") in the exclusion process.

ii) The following types of robotic user-agents are included in the standard exclusion process:

- **Personal spiders and offline browsers** can have significant and material effects on site traffic. Their activity levels are highly unpredictable over time and across sites. Hence, their User-Agents are NOT included in the standard ABC/IAB Robots List.

- **PDA devices, web feed aggregators and other automated syndication agents** are included in the ABC/IAB Robots List. The Page Impressions certified for your site should NOT include any PDA or web feed (e.g. RSS) aggregator traffic.

- **Records with unidentifiable User-Agents** (usually nulls, "-") are also deemed invalid, since there is a risk that the activity was made by a robot. Therefore, any record with a null User-Agent must be excluded along with the robots, unless the site can provide adequate justification for their inclusion.

iii) Traffic from anonymous proxies, via translation services or other third party tools, or locally cached pages, are not by default invalid. However the behaviour of such traffic may identify it as not human-initiated.

b) *You must exclude traffic that has been 'pushed' into the user's browser.*

i) **Subsited traffic**: Occurs when, upon a user requesting a page, a new browser window opens automatically on the user's device (most often as a pop-under) which carries a different page, usually from another site. This second browser window therefore generates a Page Impression for a page (and usually a site) different to that which was intended by the user's action.

ii) **Contextual linking**: Occurs when the activity of a user in a non-browser application (such as an Instant Messenger client) is analysed and a new browser window is opened (usually as a pop-under) containing a site considered relevant to the subject of the user's activity.

iii) **ISP page replacement**: Some ISPs serve a page from their own site containing their own content when the ISP's customer enters an invalid URL in their browser which would otherwise produce a standard DNS error page (e.g. "Server not found"). [This is different to the return of a standard 404 error page; in the 404's case, the server (domain) the user wanted has been found, but the page has not.] Since the user has not requested these pages such ISP page replacement is considered to be pushed traffic and therefore invalid.

Note: Automated Page Impressions following on from an intentional user request are not deemed to be pushed traffic and therefore they may be included.

c) *You must exclude traffic to URLs that are concurrently served with a valid Page Impression or provide content that is not intended to be visible to the user.*

i) Frameset exclusion can be complicated by the difficulty of distinguishing between wanted and unwanted records. The URLs associated with frames will all have valid file extensions. So, unlike the straightforward elimination of unwanted ".gif" or ".jpeg" records, it will not be possible to filter records for inclusion or exclusion simply on the basis of their file extension.

ii) Departure pages. *For example bounce-through pages, goto pages.*

- These can be used to count Referrals In or Clickouts and so you may wish to process the log records of departure pages in order to count these metrics.

GC2. Evidence to support the claim must be retained and available for a minimum period

a) *The logged records provide the specific data regarding each file request processed by the server.*

i) Logged records may be generated by web servers, page tag servers (typically collecting graphic requests generated by browser-side measurement tools) or packet sniffers.

c) *The logged records must contain sufficient information to identify the traffic to be counted and audited.*

i) For web browser traffic, each log record will typically need to contain the following data fields:

- Date and time stamp of the request, including any adjustment to the time
- IP Address of the originating **user** (NB – Dotted-quad format aa.bbb.cc.ddd, NOT the hostname)
- Full request-URI, including:
  - Domain (Host)
  - Requested URL

  - Any applicable query parameters
- Full unmodified User-Agent string
- Referrer URL
- Unique Browser Identifier (e.g. Cookie) if not logged in other fields
- Additionally, for web server logs:
  - HTTP Status code (200, 302, 404 etc)
  - HTTP method of the request (GET, POST etc.)
  - IP Address (or name) of the **server**
  - Bytes transferred

ii) For App traffic you must be able to identify app data (e.g. via URL, domain, separate tag or account ID). These metrics are all subject to appropriate auditable data being provided that typically needs to contain the following fields:

- Device identifier
- Event Date
- Event Time
- Application Name
- Application Version
- Event Description (or Ad URL for Mobile App Ad Impression)
- IP address
- User-Agent

iii) Note that the W3C CLF (common log format) does not include fields essential to the accurate counting and auditing of web site activity, such as User-Agent, and as such is NOT suitable audit evidence. Also, some proxy log types (e.g. Squid) may not by default contain the fields typically required.

iv) You must not change the format of your log files during the Reporting Period.

- You should ensure that any third parties managing your log files for you are also aware of these requirements.

- If you operate a mixed logging format (e.g. some domains on Apache servers, others on IIS), you may fail to exclude robots, and miscount Unique Browsers and also Visits, unless you ensure that the User-Agent strings are normalised to the same format throughout before the claimed metrics are calculated.

Appendix 1. Key Unique Browser/Impression/Visit metrics – Detailed Definitions

Metric Name	Explanatory Notes
Unique Browser	This metric measures each browser on a given device; it does not measure a person. Counting of Unique Browsers may overstate or understate the real number of individual devices concerned due to factors such as dynamic IP address allocation, significant levels of uniformity in IP address and browser configurations operating through a proxy, cookie blocking and cookie deletion. Other device identifiers may be allowed as Unique Browser identifiers when they can be proved in an auditable manner to be persistent and consistent across the domains being measured. Note: App Unique Browsers may use consistent, persistent identifiers of an application that are wholly or partially based on identifiers (e.g. UID) passed to the App by the device. The media owner should be aware that using the unmodified device identifier may lead to privacy issues

APPENDIX 1: ABC DETAILED METRIC DEFINITIONS

Metric Name	Explanatory Notes
Page Impression	In effect, one request by a valid browser should result in one Page Impression being claimed. The counted Page Impression may not necessarily be in focus and all content may not be fully visible in the browser window. In most cases, a single request from a browser causes the server to send several files to satisfy the request. For example, the server may send an HTML file followed by several associated graphic images, audio files and other files such as stylesheets. A single request from a browser may also cause the server to send several additional HTML files to build a frameset. The site must ensure that all additional files are excluded when counting the claimed number of Page Impressions. Generally, subject to the guidance principles issued by the auditor, directly attributable user-initiated logged events* for content (typically mouse clicks) can be used to count Page Impressions, whether served in HTML, Ajax, Flash or other environments. Please note that files that contain specific types of advertising creative, such as banners or skyscrapers, and files that represent Streams are not valid for the counting of Page Impressions but should be used separately to identify Ad Impressions or AV Plays. Page Impressions must contain textual content beyond simple advertising. *Such events are typically captured by browser-side measurement. Examples include mouse-overs, link views, menu selections or filling out of form fields. The use of such events allows more granularity in the measurement of Visit and Duration metrics.