



Understanding Internet Audiences

Andrew Green

Warc Best Practice

July 2014



Title: **Understanding Internet Audiences**
Author(s): **Andrew Green**
Source: **Warc Best Practice**
Issue: **July 2014**

Understanding Internet Audiences

[Andrew Green](#)
Ipsos MediaCT

[EXECUTIVE SUMMARY](#)

[ESSENTIALS](#)

[WHERE TO START](#)

[FUTURE DEVELOPMENTS](#)

[CONCLUSION & CHECKLIST](#)

[FURTHER READING](#)

[REFERENCES](#)

EXECUTIVE SUMMARY

According to ZenithOptimedia's April 2014 advertising expenditure forecasts, advertising on the internet will surpass \$121 billion this year (2014), a twelve-fold increase since 2002. This represents a 22.9% share of major media adspend globally. In 2004, Internet's share was just 3.9%; five years later it had reached 13.5%.

Advertising spending in the medium surpassed the amount spent in Out of Home media globally in 2006, overtook radio in 2007, exceeded magazine advertising levels in 2009 and caught up with newspapers in 2012 – overtaking them in 2013. Only television remains larger, yet internet advertising continues to power ahead, with forecast growth of 16% in 2015 and 14% in 2016. The UK in 2009 and Canada and Australia last year are the first three of the major advertising markets where internet spending has already surpassed the television total.

Generally speaking, spending is categorised as 'Display' (including banners, advertising in social media etc.), 'Classified' (small ads, mainly taken from newspapers in the early years of the medium's development) and 'Paid Search' (e.g. Google's AdWords, but also Search Engine Marketing (SEM), Search Engine Optimisation (SEO), pay-per-click (PPC), cost-per-click (CPC), cost-per-impression (CPM) search engine advertising, sponsored listings and paid for placement).

Paid Search is currently the largest category worldwide, with 46% of total spending in 2014, but is being rapidly caught up by Display (43%).

Downloaded from warc.com

The internet – however accessed – is already taking up a greater part of consumers' media time than any other individual media in advanced industrial countries. An eMarketer report on media usage in the United States¹, for example, suggests that time spent with all forms of digital media (including mobile) represents 47% of all time spent with media (although this is counted even when people are surfing at the same time as, for example, watching television). TV represents a further 36.5% of time, while reading printed newspapers and magazines accounts for just 3.5% of the total.

The same source puts the digital share of time in the UK at 43%, compared to television's 38%². Other sources have put the number even higher, although there are questions about the sort of people polled for these studies (generally heavy internet users) and the value of 'claimed' versus 'measured' behaviour³.

There is no doubt the internet has revolutionised how people access news and information, how they shop and how they interact with others. There are more than a billion websites available to users (a number that is growing by the minute⁴)

The web can be accessed from multiple devices, including PCs and laptops, tablet computers, Smartphones, Games Consoles and 'Smart' television sets. These will make the internet even more omnipresent in peoples' lives. According to NPD Group, the number of streaming devices in the United States will surpass 200 million by 2017, double the current level and twice the number of households connected to the internet⁵. Gartner has forecast that global Tablet shipments will overtake those of PCs in 2015, eventually leading to PCs becoming a much smaller player in overall internet usage. All need to be measured.

The medium (insofar as it can be called a medium) has massively impacted the economics of newspaper and magazine publishers, television and radio networks, music and film producers and retail establishments.

It is said to be inherently 'accountable', with advertisers able to reach the people they want, when they want, paying only for those people they want to talk to. Of course it is not that simple.

ESSENTIALS: The Power of Internet Advertising

Advertising on the internet is marketed as offering many advantages over advertising in other more 'traditional' advertising media. The case for internet advertising is widely documented, notably by Internet Advertising Bureaux in many countries. These include:

- *Global reach* – everybody in the world can access the same sites and information
- *Mass coverage* in every major country. An estimated 3 billion people can now access the internet – 40% of the world's population⁶. It continues to grow.
- *Precise targeting* – tools exist that enable advertisers to target users where they live, at particular times and to do so amongst content that is relevant to them.
- *Cost efficient* – companies selling advertising on the internet argue that advertisers need pay only for what is useful to them. This may be measured by the number who 'view' a page or who 'click' on a link.
- *Measurable* – every visit to a web page is recorded, offering the ability to obtain precise counts of the number of times a banner ad or other internet page is opened.
- *Trackable* – individual machines can be tracked over a period of time, allowing advertisers to paint a picture of the route users travel to reach the page where their message appears and also to see which pages are viewed after an ad has been served.

-
- *Fast and flexible* – buys can be rapidly altered or updated in order to respond to changing usage patterns. This has spawned developments such as near real time optimisation and programmatic buying of space.
 - *Engaging* – people use the internet to search for information or to connect with other people. As such, it is an inherently engaging experience, which may benefit advertising appearing within and around the content being actively looked at.
 - *Versatile* – advertisers can make use of a whole range of creative executions on the internet, including text, video, photography and the use of sophisticated graphics and other imagery.

But, equally, there are areas where the medium can fall short of these ideals:

- *Advertising clutter*. Many web pages are chock full of advertising banners, some of which pop up in front of the page a user wants to look at. This can cause what is known as 'banner blindness'.
- *Invisibility*. Ads are often barely visible to web page viewers, hidden 'below the fold' or playing on tiny video players. Yet they are still counted as 'views.'
- *Click fraud*. This occurs when a person, automated script or computer programme imitates a legitimate user of a web browser by clicking on an ad or link. For those advertisers who pay for every click their ad receives, this raises the cost of their campaigns without generating any real impact on the people they are trying to reach. Click fraud is very hard to detect and, according to some sources, very significant.
- *Privacy considerations*. There is a growing realisation by many internet users that everything they do online can be recorded and then used to 'target' them or simply to spy on them. Capturing and using this intelligence through techniques such as the use of cookies (a small piece of data sent by a web site and stored on a user's browser so that the browser is recognised every time it returns to the site) is commonplace. Because there may be many different advertisers appearing on a website, all of whom set cookies on the user's browser, it is not a simple matter of one cookie being set by the website being visited. Sometimes one website will place several hundred cookies on the browser, each of which can enable tracking of user behaviour over a period of time and over multiple sites. In some countries, legislators are seeking to ensure users are made aware of these tracking cookies in advance of them being set. They can also be deleted by users.
- *We still live offline*. Not all buying behaviour can be directly linked to internet usage. People often research products online and then purchase them in traditional retail establishments.

WHERE TO START: Measuring internet audiences

Every internet browser contains an Internet Protocol or IP address in order for it to send and receive data, just as an individual needs a postal address to send and receive mail.

Websites which send cookie files to a user's computer browser can then recognise that browser on future visits. Third parties such as advertising networks, which serve ads across multiple sites, also install cookies on the devices used to access websites on which their ads feature (called 'Third Party' cookies).

Server logs record all transactions on a website and the IP addresses from where the websites were viewed. They are collected by the website server and can be accessed for analysis by the website's owner.

Many websites also add special 'tags' to their webpages, enabling them to track when the page is visited. These are essentially small image files which send a message to a server every time the page in which they are embedded is accessed. Counting is therefore activated by opening the page, not by requesting it from the server. The tag can capture information about the computer retrieving it (e.g. the IP address, browser type etc.) as well as when it was seen, although it does not allow the same depth of analysis as server logs.

Cookies, server log files and tags are all different ways of tracking what happens at the website and web page level. Each approach has a role to play in helping website owners understand their audiences, although each also has weaknesses and shortcomings. They represent the foundation of what is called 'site-centric' website measurement and also the basis for saying that 'everything on the web is measured.'

Metrics such as the number of Page Views (the number of times a page is viewed over a period of time) and the number of Unique Visitors (the number of different individual browsers visiting a page or site in a given period at least once) can be generated from these sources.

Analysis of site-centric data is known as web analytics. It describes ways in which visitors to a given set of websites behave – for example isolating the particular landing pages they click on and gauging how effective these pages are in leading them to make purchases or provide information. The performance of any given page can then be compared with others and used to help marketers optimise their campaigns – often very rapidly (unlike other audience data, these data can be made available in almost real time).

But browsers are not, of course, the same as people (although they are often treated as such). If people use a different computer, open a new web browser on the same computer or delete their cookies, the website will treat them as first time users and set a new cookie. If three people use the same computer and browser, they may well be treated as one person.

Many browsers are not people at all, but the activities of automated search engines, indexers, robots and spiders. Or they may simply be people outside the territory where an advertiser is marketing his products.

As noted above, browser counts may contain a certain amount of fraudulent click activity. All this needs, ideally, to be isolated and excluded from the audience counts.

Cookies have a shelf life. They are often deleted by users or by their anti-virus software either rapidly or on a regular basis.

As if all this wasn't enough, cookies do not work in quite the same way on Smartphones and Tablets. Most mobile browsers accept First Party cookies (from the websites being viewed), but many do not accept Third Party cookies (e.g. from advertising networks). Mobile apps do not allow sharing of information between apps or with the device's web browser.

So, as mobile internet usage continues to grow and, in the near future, will probably surpass usage levels on PCs (it currently represents around a quarter of internet traffic), the usefulness of cookies will decline. In short: site-centric audience measurement is not, in itself, an accurate measure of the number of people (or, indeed, of the type of people) viewing web pages.

Other methods of collecting data are being developed, but no one of them has yet been adopted as a standard.

User-centric Measurement

User-centric measurement is based on recruiting panels of internet users and placing a kind of meter onto the devices they use to access the internet. These meters can record everything the user does on the monitored devices including the sites they visit, the time they spend on each site and page and the 'journeys' they make from site to site.

Providers of user-centric audience measurement services begin by recruiting a panel of internet users (using various approaches, but all with the goal of somehow representing the internet population in a country).

Panellists are asked to provide some basic demographic information (such as location, household size, age and gender) which can then be used to weight the panel back to the characteristics of the internet population in a country. A software meter is downloaded by panellists which monitors internet traffic going to and from the users device, as well as tracking and capturing certain other information about how it is used.

A challenge in any multi-person household is how to know which household member is using the computer at any particular point in time. In some services, this is asked via a pop-up or log-in screen at the start of a session; it can also be inferred from such data as the sites visited by a user, their keystroke patterns (knowledge built up over time about how each person in the household types) and certain clues in the web pages being visited and in the data being entered (e.g. names) during browsing sessions. comScore, once of the largest services, calls this process Session Assignment Technology (SAT)⁷.

Hybrid Audience Measurement

Most major advertising markets now have the ability to use 'hybrid' audience estimates created out of a mix of site-centric and user-centric methods. This essentially involves first filtering the site-centric measure of audiences, with those originating outside the country, as well as the best estimate that can be made of non-human and other invalid traffic being subtracted from the totals. Data from the panel being used to measure audiences is then integrated with these audience estimates, adding information (such as the demographics of users and patterns of usage between different websites) not picked up by the site-centric method.

In this case, the site-centric data is usually derived from asking companies that use the measurement service to tag all their web content including web pages, videos, apps and advertisements. When tagged content is accessed by any device, it is picked up by monitoring servers, essentially giving a 'census' estimate of audiences. Panel owners will also be able to identify when members of their panels access these same pages and to combine the two sources of information.

For many years now, internet audiences from PCs and laptops have been the only data reported. Of course people also access the internet from other devices such as tablet computers and Smartphones.

Services measuring all these devices have been launched in a handful of countries (notably in the USA and Canada) in recent years, using separate panels to capture data from each device and then integrating it statistically. It remains a complex and emerging area.

There is much more to internet audience measurement than simple reporting of Page Views, User Sessions, Unique Users and Page Duration. Using any or all of the above methods, alongside other specialist techniques, users of internet audience data can drill down into the content being viewed to get deeper insight. Examples include text analytics (analysing words and phrases used in emails and social media postings, for example) and social listening (a related approach for monitoring what people are saying about brands on social media).

FUTURE DEVELOPMENTS: Towards Perfect Internet Measurement

It is clear that internet audience measurement is far from perfect. There are growing challenges for cookie-based measurement, issues with 'viewability' (do people even see the ads served to them?), problems with clutter (according to DoubleClick, the average click through rate for banner ads is 0.1%; according to comScore, only 8% of users account for 85% of display ad clicks!) – and with click fraud.

Advertisers often find, on auditing a campaign, that their content can appear in what they would deem to be 'inappropriate' sites. And studies have shown that, in many cases, Paid Search is no more effective than 'natural' search at generating impact.

Towards the end of 2012, the Media Ratings Council in the US announced that they were looking to introduce a new set of standards for defining 'viewability' of advertising impressions. This was to address the challenge of internet advertising sales organisations basing their pricing on impressions served (i.e. sent), rather than impressions *seen*.

Industry consultations had determined that the definition of 'viewable' should be that at least half of the pixels (parts of a visual image) of a particular screen should be in view for at least a second. This, it was believed, would be the closest representation of 'audience' for online advertising to the 'Opportunity to See' definition used for other media.

There were plenty of complications in how this would actually be measured and managed, but the definition has been widely adopted in the US and overseas and looks to be a standard towards which every web vendor can work. A similar standard has now also been adopted for video. In this case, a minimum of 50% of advertising pixels must be viewable for at least two continuous seconds during any part of the video content being served to be counted.

As will now be clear, internet audience measurement is very complex. However, moves to tighten the definitions of how audiences are defined and reported continue and are constantly improving. A key challenge now is to include measurement of internet access across all of the devices used. This is slowly being achieved with the addition of Tablet and Smartphone user-centric measurement in a handful of countries. The size and quality of the panels being used for the user-centric component of the measurement needs comparable attention.

CONCLUSION & CHECKLIST

The internet generates Big Data like no other medium. It is vital that those investing in the medium for advertising purposes can be confident that a Page View or an impression is seen and noted. The danger of trends like the move to Programmatic Buying, where ads are served automatically in near-real time to those considered most likely to respond, is that the audience data upon which machine decisions are made may not be fit for purpose.

Preparing for the multi-platform future

As noted, the numbers of devices used to access the internet continue to grow and multiply. Site-centric measurement is moving towards greater accuracy – though the one and two-second duration definitions may ultimately need to be expanded so that advertisers can move beyond these very short minimums and towards time periods where people have more time to take things in. Some of the things advertisers should be asking their media agencies to advise on include:

- How many viewable impressions will my ad achieve when viewability is defined as those having had the content open for

5 seconds, 10 seconds or longer?

- Given the clutter of different messages on each screen, how likely are people to really notice an ad? In other words, how does 'banner blindness' affect performance?
- How can click fraud and other fake traffic be identified better and eliminated?
- Should TV online (familiar TV shows accessed online) and online TV (e.g. YouTube) be treated differently?
- What is the relative value of an internet ad impression versus one in other media?
- Should measurement of direct response ads be different to branded ads and if so, how?

FURTHER READING

The following is a list of some of the main books and articles about internet audience measurement.

Books

Green, Andrew (2009). *From Primetime to My Time: Audience Measurement in the Digital Age* (Warc)

Websites

<http://www.iab.net/guidelines/508676/guidelines/audiencemeasurement>

<http://www.iabuk.net/>

<http://www.iabeurope.eu/>

<http://www.iabaustralia.com.au/>

REFERENCES

1. eMarketer, April 2014
2. <http://www.marketingweek.co.uk/disciplines/mobile/news/consumers-to-spend-more-time-on-digital-than-tv-in-2014/4009797.article>
3. <http://blog.globalwebindex.net/online-time-now-exceeds-offline-media-consumption-globally/>
4. <http://www.internetlivestats.com/total-number-of-websites/>
5. <https://www.npd.com/wps/portal/npd/us/news/press-releases/the-number-of-connected-tv-devices-will-be-more-than-double-the-number-of-us-internet-households-by-2017/>
6. <http://www.internetlivestats.com/internet-users/>
7. <http://www.journalism.org/files/2014/03/comScore-Media-Matrix-Description-of-Methodology.pdf>

About the author

Andrew Green is Chief Marketing Officer at Ipsos MediaCT and the author of the book, *From Prime to My Time: Audience Measurement in the Digital Age*. He has held senior research positions at Zenith Optimedia, Billetts, OMD, Nielsen and Carat.

© Copyright Warc 2014

Warc Ltd.
85 Newman Street, London, United Kingdom, W1T 3EU
Tel: +44 (0)20 7467 8100, Fax: +(0)20 7467 8101

www.warc.com

All rights reserved including database rights. This electronic file is for the personal use of authorised users based at the subscribing company's office location. It may not be reproduced, posted on intranets, extranets or the internet, e-mailed, archived or shared electronically either within the purchaser's organisation or externally without express written permission from Warc.

Downloaded from warc.com