

# The Missing Billion

How web analytics is wiping the  
emerging world off the map



---

2014

---

# Who are the missing billion, and **why** do they matter?

---

It's long been said that the internet is one of the most measurable mediums out there – a place where the behaviors of real-life consumers can be quantified, tracked and aggregated. We can analyze where people are, what they're doing online and, as a result, how they are likely to act in the future.

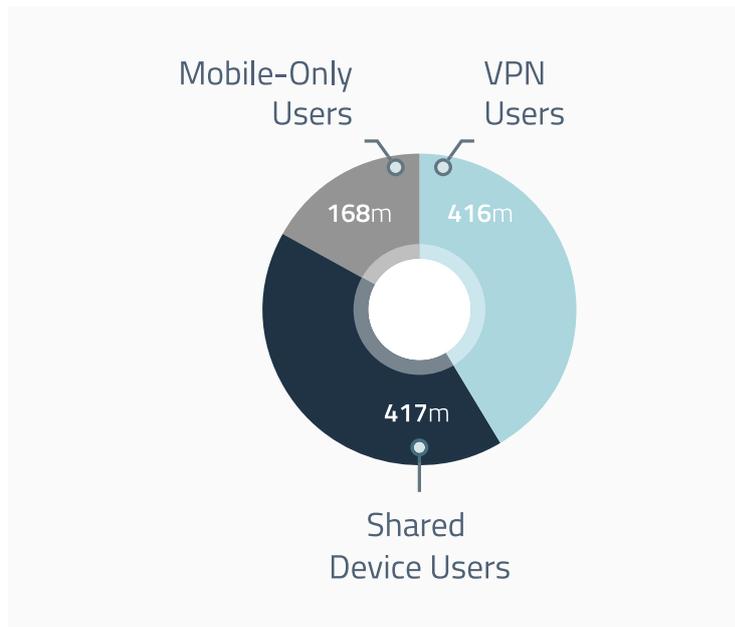
But the reality is a long way from this promise. As the internet landscape has grown increasingly sophisticated and diverse, it has become impossible to accurately measure or understand online behaviors through passive measurement techniques alone (whether in the form of analytics, passive panel-based measurement, third-party cookies or click-stream data).

Huge numbers of internet users have become invisible, absent as a result of:

- The **amount of web traffic that automatically flows through servers in the US**, as well as America's disproportionately high number of IP addresses.
- Tools such as **Private Browsers, as well as Virtual Private Networks and Proxy Servers** (which typically mean an internet user is incorrectly geo-allocated to the country of the server, rather than their actual location)
- **Multi-device** internet usage (which means people are connecting via multiple devices – and hence IP addresses – and are much more difficult to identify as unique users)
- High levels of **device sharing** (which forces us to discount the idea that 1 device = 1 user – the very foundation of analytics)
- Those who access the internet via a **mobile only** (a segment not typically included on measurement panels)

Individually, these factors have major impacts; in combination, they have created the Missing Billion – those internet users who are being **misunderstood or overlooked entirely by passive data collection techniques**. By our estimates, 416 million users are absent from passively collected data because they are using VPNs, another 417 million are missing because they are using shared devices and 168 million are being overlooked across GWI's 32 markets because they are accessing from a mobile only. The collision of these trends has generated a landscape of such complexity that we in fact need to question everything we think we know about digital data.

## THE MISSING BILLION IN NUMBERS



---

## Digital Imperialism

Consumers in fast-growth markets, where there are lower levels of investment in building measurement techniques and services, tend to be impacted much more heavily by these problems. In these nations – where online populations are growing at the fastest rates and where users are typically highly engaged in terms of how much they do online – our research shows that people are more likely to be using VPNs, more likely to be mobile-only users and more likely to be using multiple and/or shared devices. That means they can be dramatically under-represented in many passive datasets or measurement techniques, just as their counterparts in the most mature internet countries like the US are the most likely to be over-represented because of where major servers tend to be based.

In short, this skew in our understanding of the global internet population is creating a form of “digital imperialism” where consumers in the emerging / fast-growth markets are consistently losing out due to their invisibility. They’re receiving lower levels of commercial investment than they deserve, whether in terms of ad spend, the development of digital infrastructure or the provision of content. Quite simply, their overall internet experience is poorer – with fewer services and less content available to them – because money that could be spent on fast-growth markets is instead being funneled to mature internet nations where it is assumed (potentially incorrectly) that the biggest audiences are located.

All this should lead us to reevaluate the concept of analytics as well as the assumption that digital data automatically represents “actual” measurement. Indeed, too heavy a reliance on incomplete or inaccurate data means that many marketers, content-owners and brands are missing out on the biggest opportunity in a generation.

# Setting the Scene:

## Passive Analytics vs Active Data Collection

When website traffic to the biggest and most popular global websites is broken down by country of origin, **passively collected analytics will typically show very large spikes coming from the most developed and Westernized internet nations – and the US in particular.**

Clearly, global sites like Facebook, Yahoo, Twitter and eBay are very popular within countries like the US and the UK, and this goes some way to explaining their supposed audience breakdowns.

Very often, though, the sheer levels of traffic which appear to be originating from these nations are simply too big to be credible, especially in the case of America. This is a result of **analytics techniques failing to understand where users are located.**

Several factors contribute to this problem, including the **large number of servers which are based in the US and the large levels of traffic that are automatically routed through the country for reasons of cost or ease.** Indeed, as highlighted in the NSA Prism release, a significant amount of traffic flows via the US regardless of source and destination – a reflection of the legacy control that the US had in the foundation of the web, as well as the present reality in which American web companies are still dominant. For example, someone based in Brazil who is Skyping the person next door will be routed via the US. Much of this is happening behind-the-scenes, unknown to both the user and the analytics software used to allocate the user's location.

The over-representation of the US becomes particularly apparent if we look at the number of IP addresses per internet user (as calculated from the Internet Assigned Numbers Authority – IANA – which is the organization responsible for the allocation of IP addresses). This is a crucial data signifier, as an IP and its accompanying geo-location is typically used as the basis for identifying a user and their whereabouts.

---

### ACTIVE VS PASSIVE DATA

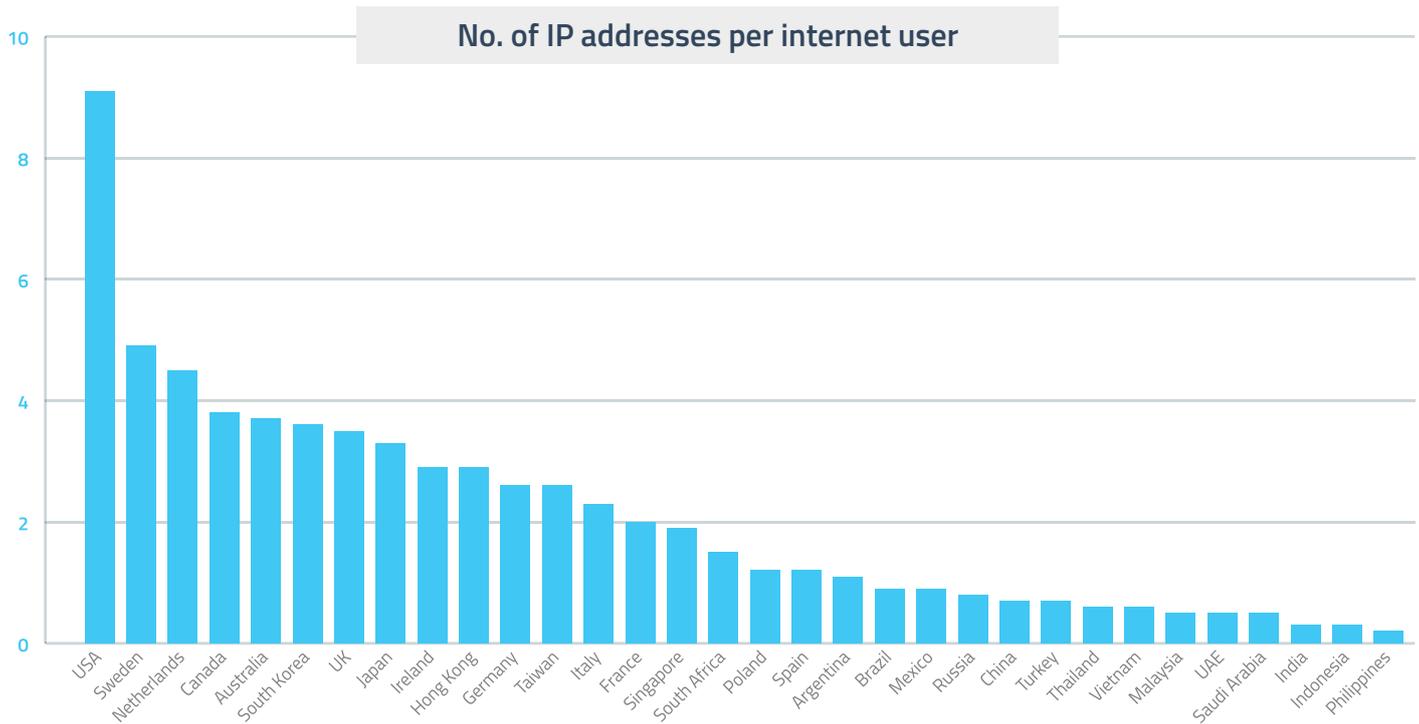
#### Active Data

Explicitly reported by the user. Active data is typically generated through form submissions, subscriptions, market research or other methodologies that ask for data from users themselves. It can also include data from Facebook and social media platforms.

#### Passive Data

Implicitly communicated by the user. Passive data is typically generated by web analytics such as Google Analytics, loyalty schemes such as Tesco Clubcard, or clickstream technologies that observe and record user actions.

As our chart shows, there are clear disparities in the numbers: in many countries, there are simply too many, or too few, IP addresses per internet user. In short, the figures do not, and cannot, reflect reality.



Source: Internet Assigned Numbers Authority (IANA) and GlobalWebIndex's 2014 Universe Estimates

Broadly speaking, countries fall into one of two groups here. In the most mature markets, there are typically more IP addresses than there are internet users. Often, the excess is massive – in no way can it correlate with actual on-the-ground internet usage in these countries.

This trend is most pronounced in the US, where there are more than 9 IP addresses per user. But there are also significant mismatches in places like Sweden, the Netherlands, Canada, Australia, South Korea and the UK (all markets in which large numbers of servers are based). The clear implication here is that users from outside of these countries are being incorrectly geo-allocated.

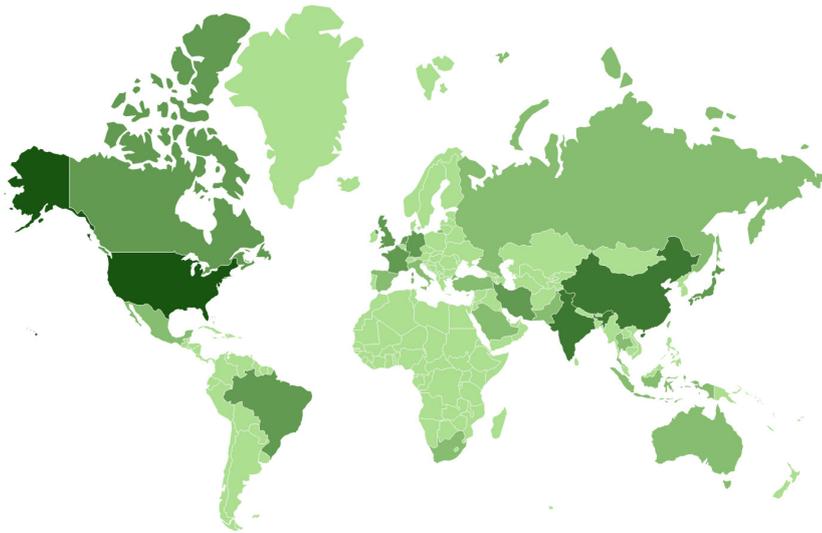
In contrast, fast-growth internet nations like Indonesia and China often have less than one IP address per user. That is, there are more people using the internet in these countries than there are registered IP addresses. Again, this simply does not reflect actual usage and strongly suggests that users are being under-represented because they are accessing via VPNs and Proxy servers.

Looking at some specific examples is perhaps the best way to illustrate these problems – and to highlight why actively collected data can bring a very different (and we would argue much more accurate) perspective.

**Alexa** is a service which offers 'intuitive analytics products'. According to its website, its 'traffic estimates are based on data from our global traffic panel, which is a sample of millions of internet users using one of over 25,000 different browser extensions. In addition, we gather much of our traffic data from direct sources in the form of sites that have chosen to install the Alexa script on their site and certify their metrics.'

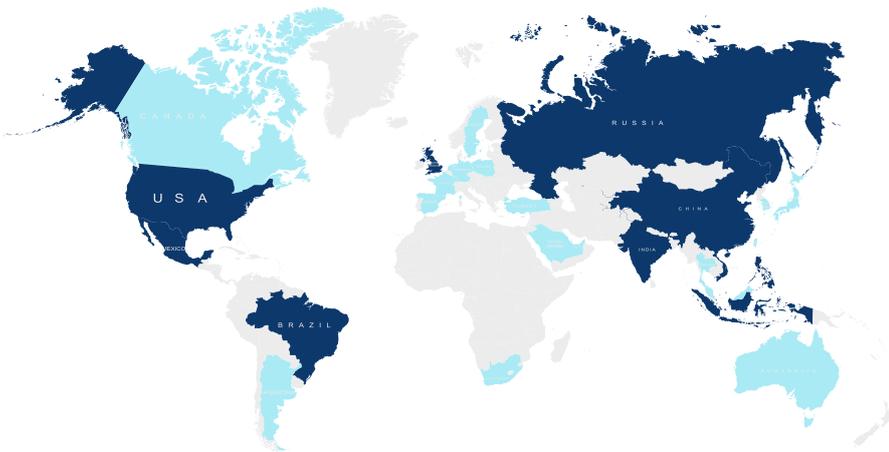
If we take **Bing.com's** visitor traffic as an example of the data that this methodology generates, we can see that the US is by far the biggest market; some 48.2% of Bing.com's visitors are identified as coming from America. All other markets within the top 10 account for very small shares only – with just 5.8% from India and 5.2% from China.

## Bing: Top 10 Markets for Traffic in Alexa's data



	USA	48%
	India	6%
	China	5%
	Japan	3%
	UK	3%
	Canada	2%
	Germany	2%
	Iran	2%
	Netherlands	2%
	France	2%

## Bing: Top 10 Markets for Traffic in GWI's cross-device data



	China	30%
	USA	17%
	India	14%
	Brazil	6%
	Indonesia	5%
	Philippines	3%
	Mexico	2%
	Russia	2%
	Vietnam	2%
	UK	2%

**Question:** Which of the following sites have you visited in the last month (via any device)?

**Source:** GlobalWebIndex Q1 2014 // **Base:** Internet users aged 16-64

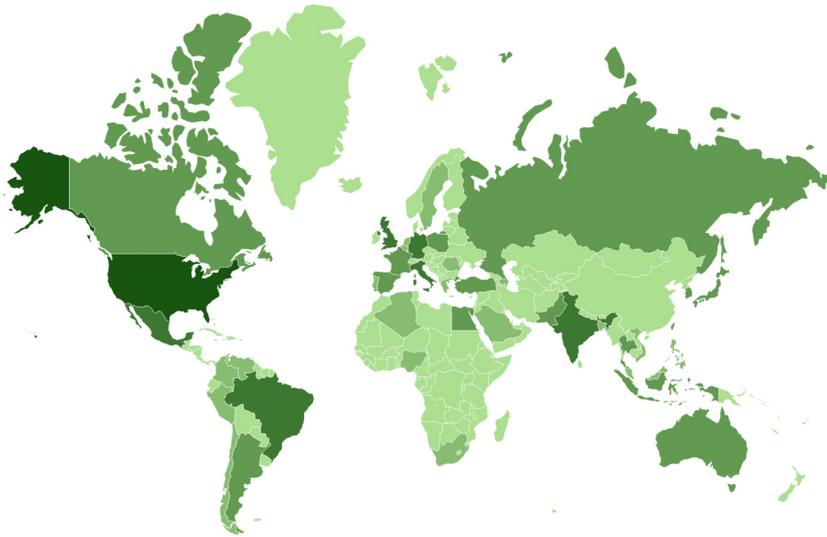
In GWI's actively collected data, which passes through extensive and robust accuracy checks and measures, we ask real-life internet users in 32 markets which sites they have been visiting. This produces strikingly different results.

Although the US is still a prominent market for Bing.com – just as we would expect for an American-founded service – it accounts for a much smaller (and more plausible) 17% of visitors.

Significantly, **it is overtaken by China** – which now represents 30% of the site's total traffic (far larger than the share identified in the passively collected data from Alexa). **India** has also risen strongly in GWI's data, reaching 14%.

Further down the list, there are countries not even featured in the top 10 by Alexa, including Brazil, Indonesia and the Philippines. Broadly speaking, then, while Alexa's list is dominated by mature internet markets, GWI's data gives a much more prominent role to fast-growth nations.

## Facebook: Top 10 Markets for Traffic in Alexa's data



	USA	22%
	India	8%
	Brazil	4%
	Germany	3%
	UK	3%
	Italy	3%
	Mexico	3%
	France	3%
	Spain	2%
	Japan	2%

## Facebook: Top 10 Markets for Traffic in GWI's data



	USA	15%
	India	13%
	China	13%
	Brazil	7%
	Indonesia	6%
	Russia	4%
	Mexico	3%
	Philippines	3%
	Germany	3%
	Vietnam	3%

**Question:** Which of the following sites have you visited in the last month (via any device)?

**Source:** GlobalWebIndex Q1 2014 // **Base:** Internet users aged 16-64

A similar pattern is evident if we look at **Facebook.com's** visitors. Analytics from Alexa shows that the US is the top market, accounting for 22.2% of all traffic. This is followed by India (8%), Brazil (4.4%), Germany (3.2%) and the United Kingdom (3.2%). Indonesia – well-known to be one of the world's most active Facebook markets – is absent altogether from Alexa's top 10.

GWI's actively collected figures once again give a different outlook: although the US is still the top market (15%), **India and China now represent much bigger shares of the total traffic going to the site.**

China is completely absent from Alexa's list, but our data shows that large numbers are using VPNs and other tools to bypass the official restrictions placed on the site by the Chinese government. And given that China and India are both

highly engaged markets for social networking, we would argue that these results give a much more realistic picture of what's actually happening on the ground. In short, emerging internet markets are once more much better represented in the data.

Similarly skewed results can be found for global services on other analytics sites, too. For Flickr, for example, **SimilarWeb** estimates that over a third of traffic is coming from the US – making it by far the biggest market for the photo-sharing site. In GWI's data, China and India are revealed to be much more important nations than passive data is able to recognize. In fact, more than three times as many Flickr visitors are coming from China as they are from the US.

For a service like LinkedIn, SimilarWeb figures once again put the US as the leading country, and by a huge margin. Relatively small markets like Canada and the Netherlands are also placed in the top 5. But GWI's data reveals a very different visitor landscape: only 12% of traffic is actually coming from the US, placing it some way behind India (21%) and China (19%). The Netherlands and Canada actually represent just 1% of visitors each.

Very consistently, what we're seeing here is **users in fast-growth nations actually form much bigger shares of traffic than is traditionally assumed**. It's passive analytics that is making them invisible, a problem being exacerbated by the trends we explore in the rest of this report.



### Flickr: Top 5 Markets for Traffic in SimilarWeb's data

 USA	37%
 United Kingdom	8%
 Canada	4%
 France	3%
 Germany	3%

### Flickr: Top 5 Markets for Traffic in GWI's data

 China	28%
 India	22%
 USA	8%
 Indonesia	7%
 Brazil	6%



### LinkedIn: Top 5 Markets for Traffic in SimilarWeb's data

 USA	40%
 United Kingdom	7%
 India	5%
 Canada	5%
 Netherlands	4%

### LinkedIn: Top 5 Markets for Traffic in GWI's data

 India	21%
 China	19%
 USA	12%
 Brazil	8%
 Indonesia	7%

# VPNs: A major internet gateway

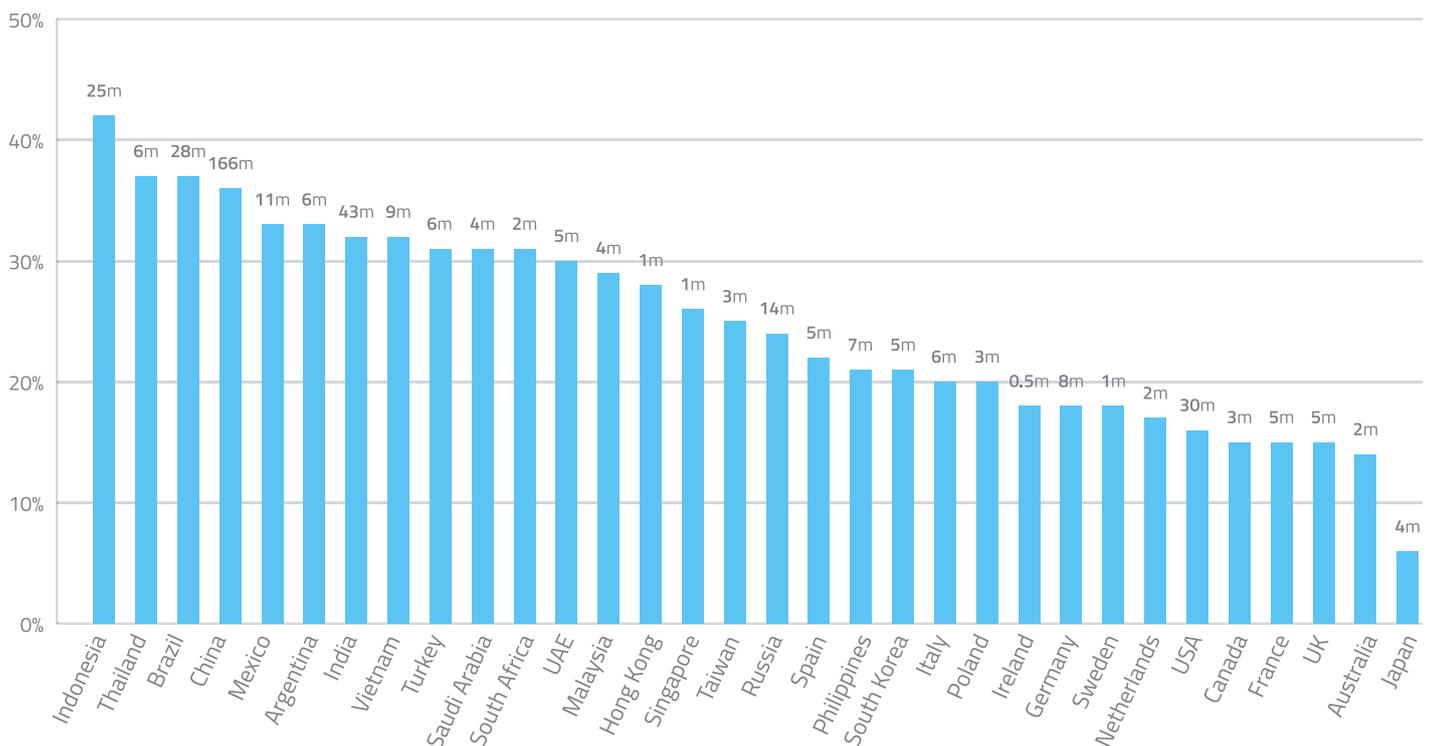
Although awareness of Virtual Private Networks (VPNs), Proxy Servers and Tor Servers is increasing, they're still seen as pretty niche – tools used principally by tech-geeks.

Our latest research shows that the real picture is quite different: VPN usage is already mainstream, with 28% saying that they have deployed one when using the internet. This equates to a staggering 419 million people in GWI's 32 markets.

But it's analyzing behaviors at a country level that shows us why this trend is causing millions to become invisible in traditional measurement techniques.

It's abundantly clear that VPN usage is most pronounced in fast-growth nations. Indonesia is the country in which people are the most likely of all to be using them, with 42% saying that they have done this. In terms of overall audience size, though, the importance of China is obvious: the 36% of its online population who report having used a VPN represents some 166 million people.

## VPN Usage by Market



**Question:** When you access the internet, do you ever do so using a VPN (virtual private network) or proxy server?

**Source:** GlobalWebIndex Q4 2013 – Q2 2014 // **Base:** Internet users aged 16-64

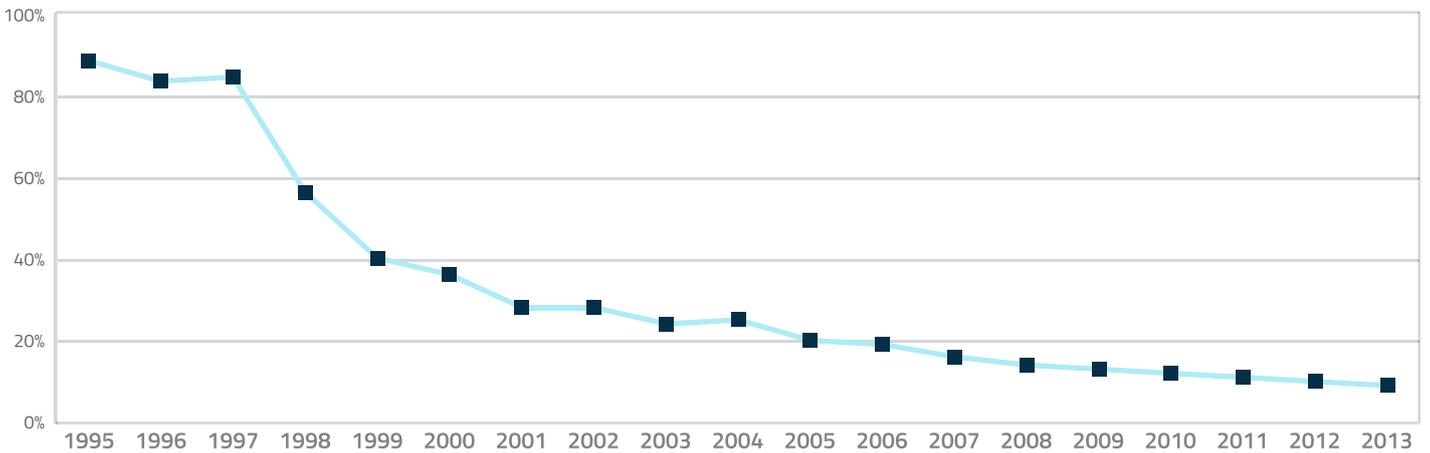
In contrast, usage is considerably less common in the most mature internet markets – with the USA, Canada, France, the UK, Australia and Japan coming at the very bottom of the table.

This geographical pattern has major implications for audience measurement. Most significantly, a user who connects via a VPN is highly likely to be geo-located to the country where the server is based, rather than being assigned to the nation in which they are actually located. This skews where website traffic appears to be originating from; some visitors recorded as coming from the most established markets like the US actually represent internet users in fast-growth nations.

As we explored above, the problem we need to overcome is that the US is still assumed to be a bigger force than it now is. Indeed, the composition of the global online audience has shifted dramatically over the last twenty years, eroding the historical dominance of America (and other mature nations). The late 1990s and early 2000s saw the biggest evolution of all, with growing numbers in other countries going online for

the first time. But this downward pattern for the US is still in evidence today; American internet users represented just 9% of the total worldwide internet population in 2013 (compared to a staggering 81% in 1995). It's now the fast-growth markets which dominate – and they will do so to an even greater extent in the future as their national internet penetration rates continue to increase sharply.

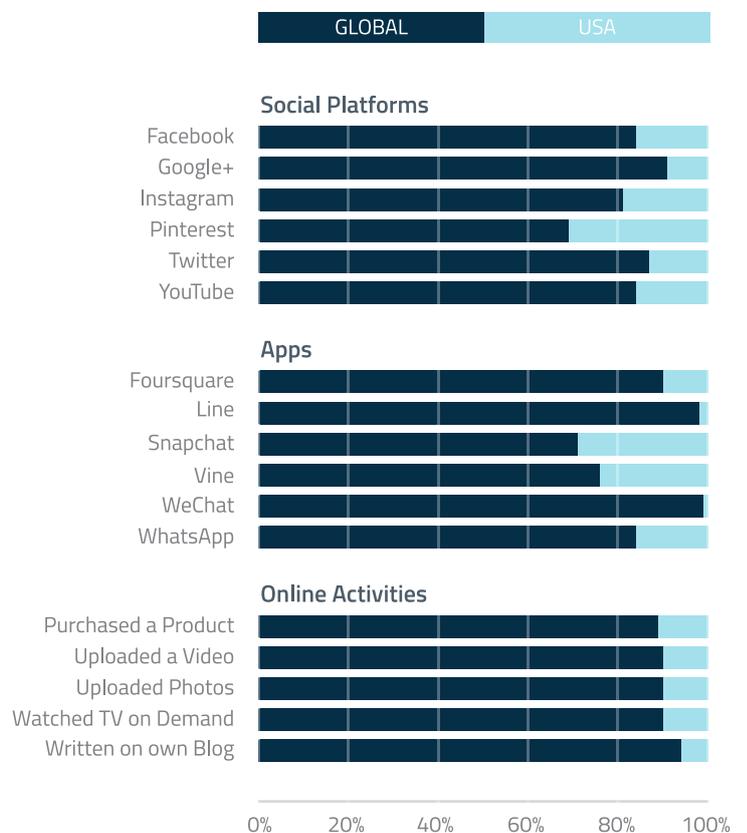
## US Share of Global Internet Population



Source: ITU

This shift is reflected in GWI's audience breakdowns for various social platforms and key web behaviors during Q2 2014. In each case, it's strikingly clear that American users now represent only a minority of the total audience; some 84% of Facebook's active users are non-US, for example, as are 87% of Twitter's and 99% of those on WeChat. Similarly, 90% of people who watch TV on-demand each month are located outside of America.

## The US vs Global Audience



Base: Internet users aged 16-64 who have actively used the following social networks or mobile apps in the last month, or who have carried out one of the following online activities // Source: GlobalWebIndex Q2 2014

# Motivations for using VPNs: Driven by content and networking

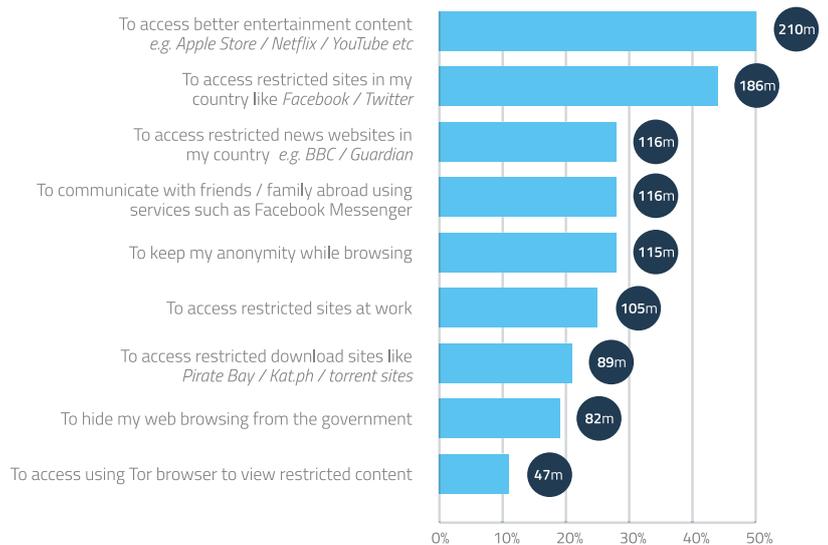
To understand why people are deploying VPNs to access the internet – and thus becoming invisible in passive datasets – we asked users about their motivations.

Clearly, gaining access to content and social networks which are not available in one's own country due to geo-restrictions is extremely important for VPN users. This helps to explain why usage is so high in fast-growth markets in particular and brings with it a number of implications.

Firstly, it demonstrates that users are willing to take control and use the internet as they wish – and that attempting to place national restrictions on particular sites is increasingly futile. A good illustration of this comes from looking at VPN usage in Turkey, a country in which access to Twitter was officially blocked earlier in 2014. Overall, VPNs are being used by 31% of Turkish internet users (representing 6.3m people). And some 55% of this group – corresponding to 3.48m people – say that they are using them to access sites like Facebook or Twitter. Quite simply, then, official attempts to block sites in the future will become ever less effective.

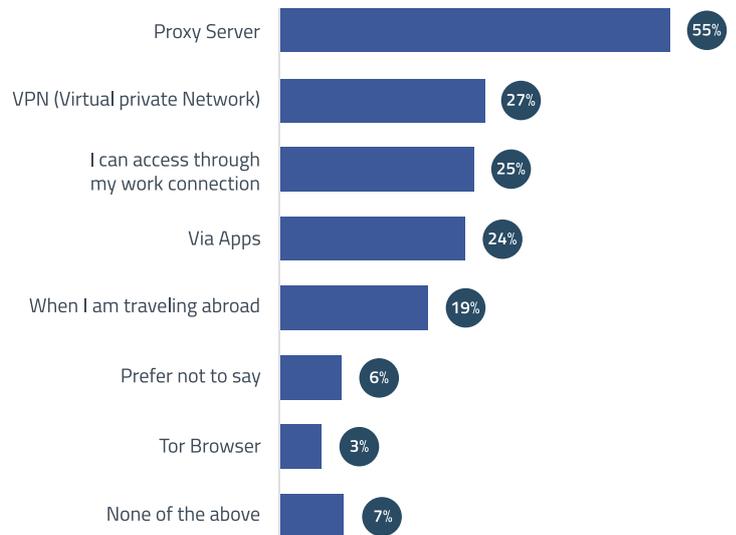
The second implication of VPN users wanting access to content and networks is that some national audiences for popular sites are likely to have been under-estimated. If we take Chinese users of Facebook as an example, our research shows that Proxy Servers and VPNs are the most popular way of accessing the site.

## Reasons for using VPNs



**Question:** Can you tell us why you use VPNs / Proxy Servers when browsing the internet?  
**Source:** GlobalWebIndex Q4 2013 – Q2 2014 // **Base:** VPN users aged 16-64

## How Chinese Facebook users connect to the site



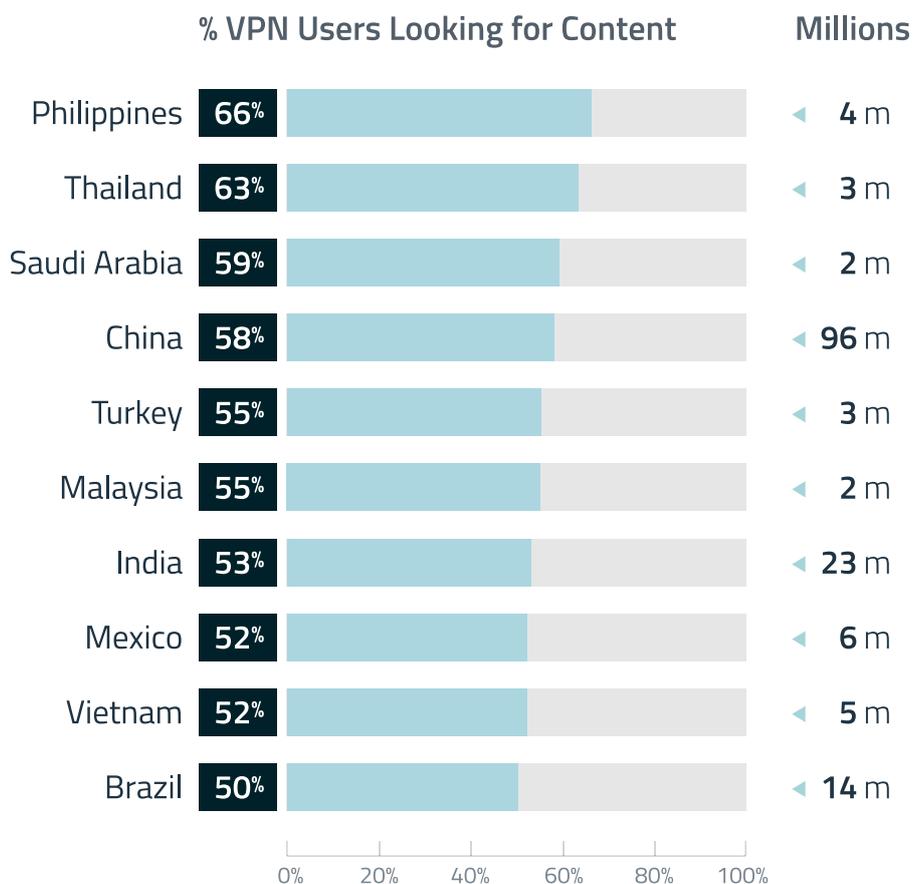
**Question:** Have you used any of the following methods to bypass government restrictions in order to use Facebook? **Source:** GlobalWebIndex Q3 2013 // **Base:** Active Facebook users in China aged 16-64

There's likely to be a degree of unintended proxy usage in evidence here – with those accessing at work or via their mobile not necessarily realizing that they are being routed by a server hosted in a different market (many companies proxy internet usage via head offices, while the majority of apps are hosted outside the country in question). Nevertheless, intentional avoidance is dominant: many Chinese internet users are deliberately using VPNs and Proxy Servers to bypass official restrictions. So, social audiences in China on international platforms such as Facebook and Twitter can be – and are – substantial even in spite of official bans.

A similar picture emerges in relation to the search for content. It's apparent from our table that there are significant numbers in emerging internet markets who are accessing content from major, global services like Netflix and iPlayer. This suggests that demand in these countries is not being sufficiently recognized or met, and that setting geo-IP limitations on access is not preventing people from obtaining the content they want.

Once again, it's easy for the US to become over-represented as a result of this trend: when we ask VPN users where the server they use is located, some 15% – or 62 million – say it's in America. It's also worth noting that a further 28% say that they simply don't know where their server is based, suggesting that the actual figure for the US is probably higher still.

### VPN Usage for Better Entertainment Content: Top 10 markets

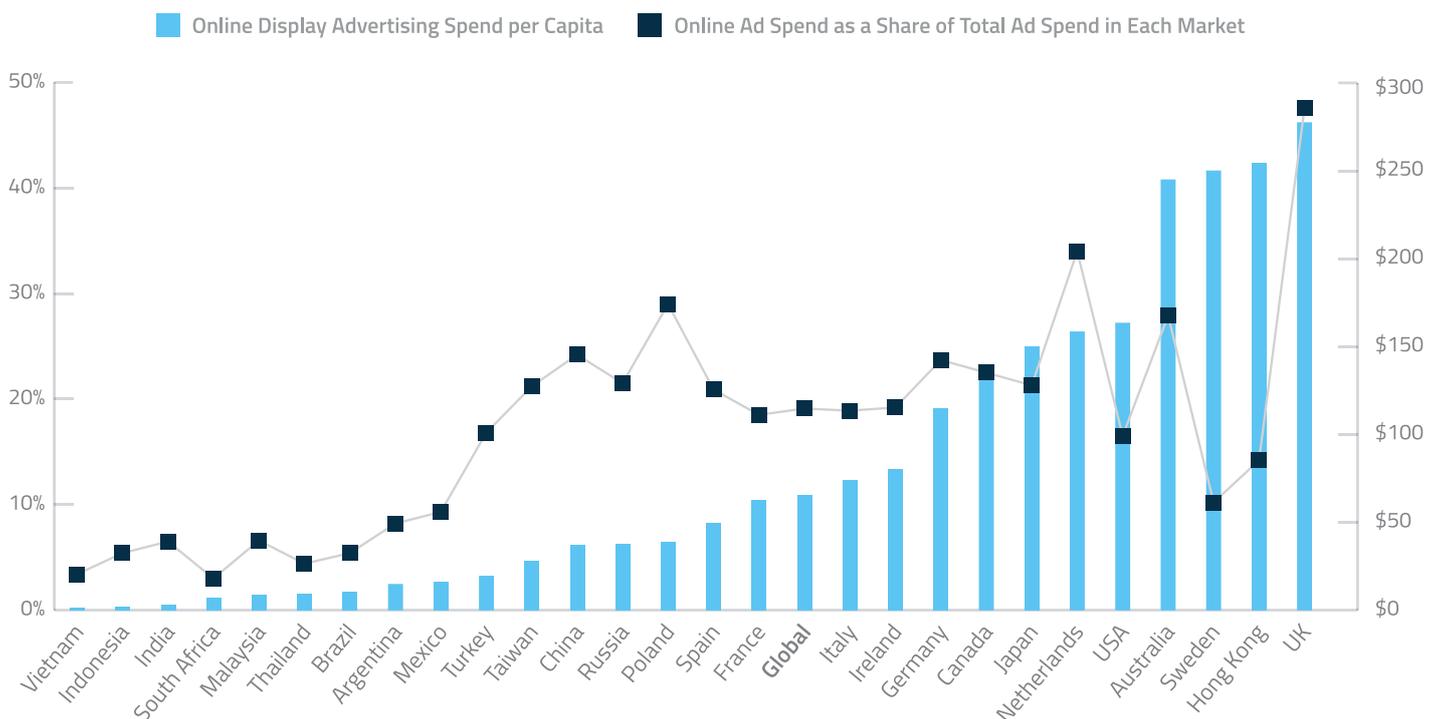


**Question:** Can you tell us why you use VPNs / Proxy Servers when browsing the internet?  
To access better entertainment content e.g. from Apple Store / Netflix / YouTube etc  
**Source:** GlobalWebIndex Q4 2013 – Q2 2014 // **Base:** VPN users aged 16-64

# Implications for marketing

Widespread VPN usage has major ramifications for marketing. Potentially, billions of geo-targeted ads are being mis-directed.

Most importantly, fast-growth markets might now account for a growing share of the global internet audience, but marketing investment has not followed suit. There remains a general lack of digital investment in these nations, with the opportunity to drive truly global marketing not being met.



This is reflected through annual levels of online ad spend per capita. As our chart demonstrates, a country like Indonesia might have the world's most active audiences on sites like Facebook and Twitter, but just \$2 per year is invested in each user. Compare this to the UK, where \$277 is spent annually on each person. Or to the US, which still commands more than 30% of the global digital ad spend despite accounting for just 9% of the global internet population. Although ad spend is not of course linked solely to user numbers – with the potential value of the return per user being a key metric – there's still an obvious skew to be seen here.

Just as significant is the incorrectly geo-targeted advertising that many of the 419 million VPN users identified above are likely to be seeing. If they are recorded as being in the US but are actually in markets like China or Indonesia, the relevance of the messages they are shown will be hugely diminished. This is having a major impact on the effectiveness of advertising; ultimately, some of the online ads currently being bought by US brands are being targeted at consumers in different countries – helping to explain why the Missing Billion represent such a major blackspot for advertising.

# Device Explosion: A significant challenge for audience measurement

The rapid emergence of smartphones and tablet devices has had a dramatic impact on how, when and where people go online. This has created major challenges in terms of tracking and understanding the online behaviors of individual users, something which the arrival of wearable tech and other connected devices will complicate still further.

Clearly, there is no longer a universally applicable model in terms of internet behaviors: some now access through a mobile only, others are using shared devices and hundreds of millions are connecting via a diverse combination of devices. All three of these trends present serious difficulties for passive tracking methods and – with their impact disproportionately high in fast-growth markets – they are significant contributors to the Missing Billion.

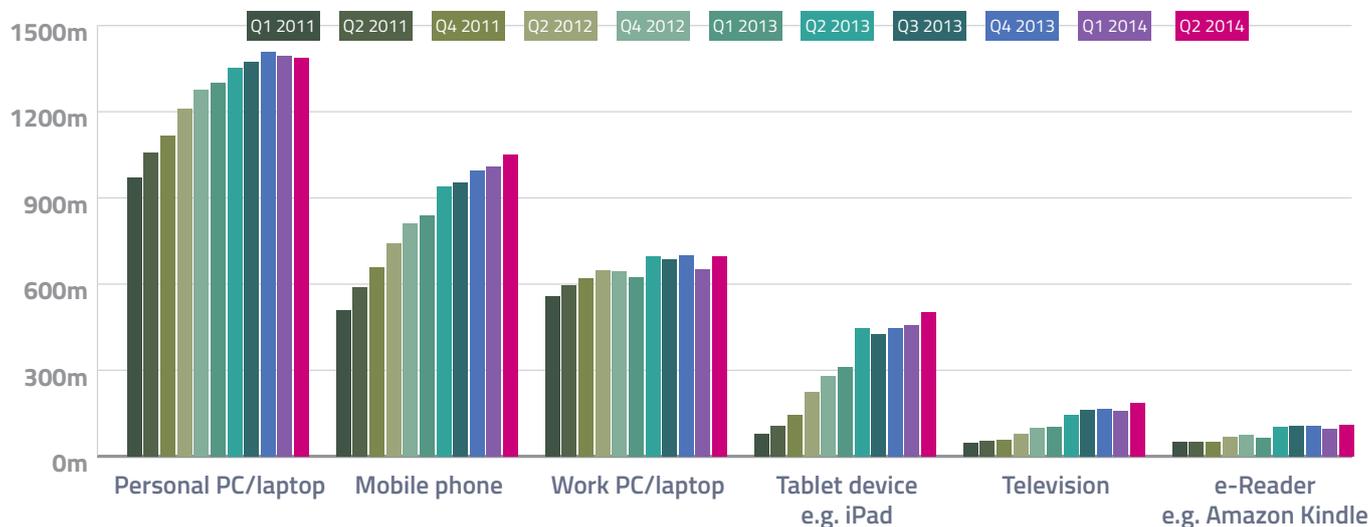
## Cross-Device Usage the Norm

While PCs and laptops still remain the dominant internet access points, mobiles and tablets are seeing (very) strong levels of growth.

Between Q1 2013 and Q1 2014, for example, the numbers using the mobile internet globally jumped by 20%. The equivalent figure for tablets was 48%.

It's not that some devices are being abandoned in favor of others. Rather, most internet users are now going online via multiple access points – turning to a PC, laptop, mobile or tablet as location and activity dictate. GWI's latest data in fact shows that we now use an average of 2.4 devices to go online each month, with only a relatively small proportion using *just* a PC.

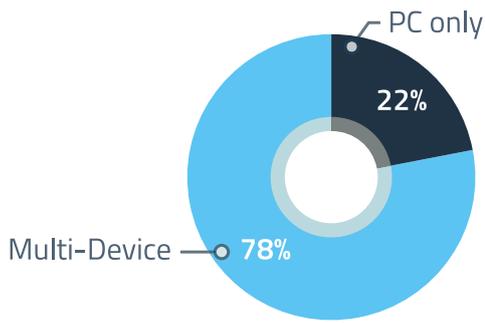
### Millions of Internet Users Going Online....



**Question:** In the past month, from which of the following devices have you accessed the internet either through a web browser or an application?

**Source:** GlobalWebIndex Q2 2014 // **Base:** Global internet users aged 16-64

## Multi-Device Internet Usage



Devices used to access the internet in the last month  
**Source:** GlobalWebIndex Q4 2013 - Q1 2014  
**Base:** Internet users aged 16-64

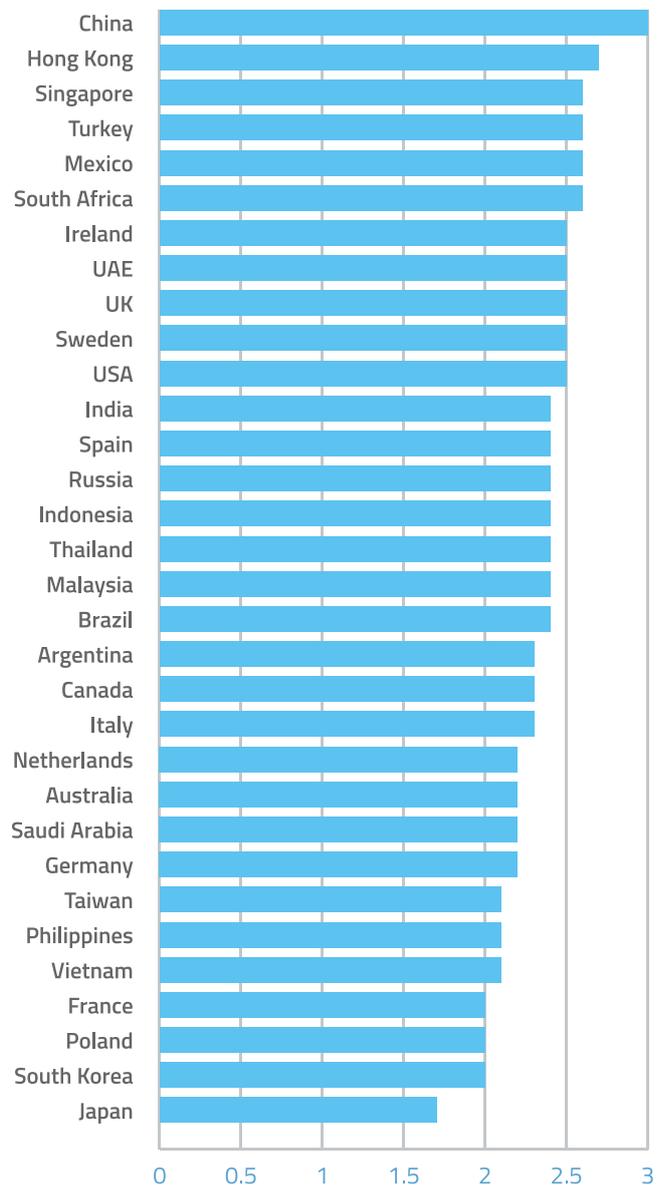
In many instances, this means that the same user will be visiting the same site via a combination of devices, a trend which makes it extremely difficult to track an individual's behavior (particularly if they are not logged in).

More significantly, though, the majority of passive measurement (particularly panel-based or cookie data) is focused on PC-sourced data. This problem becomes particularly pressing if we examine multi-device usage by country.

It is in China where people are typically using the highest number of devices to go online, followed by other fast-growth markets such as Hong Kong, Singapore and Turkey (something which reflects the less historically prominent role of PCs in emerging internet nations as well as the widespread availability of low-cost mobiles and tablets).

However, investment in non-PC measurement is virtually non-existent outside of the US and Western Europe. Once again, then, users in fast-growth markets are being overlooked.

## Average Number of Devices Used to Access the Internet Within the Last Month



**Source:** GlobalWebIndex Q4 2013 - Q2 2014  
**Base:** Internet users aged 16-64

# Device Sharing

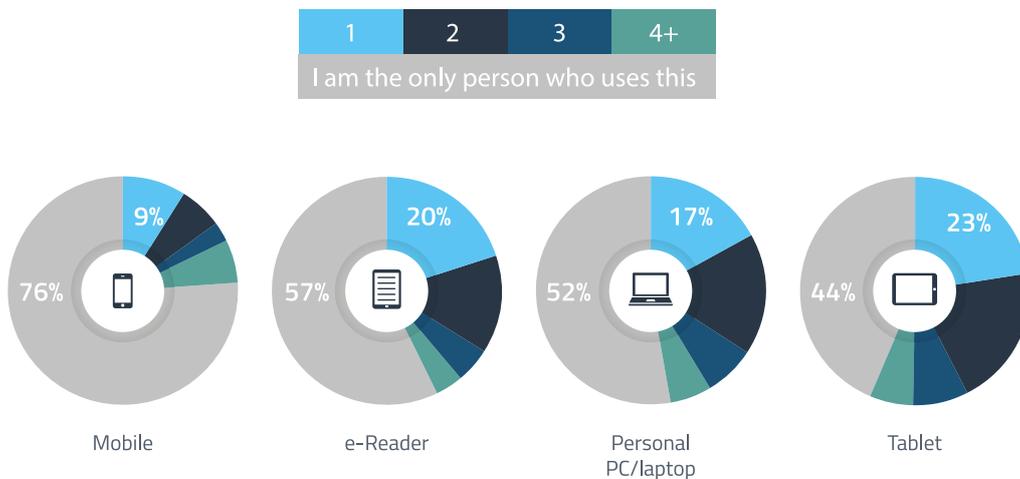
Further identification issues have been caused by the rise of device sharing – a trend which makes machine or browser identification increasingly unreliable.

As our figures show, this is a major trend across *all* device types. Certainly, tablets are the devices which are most likely to be shared with other people; globally, just 44% of tablet users say that no-one else has access to it. PCs and laptops are shared widely too; about half of these devices are used by at least one other person.

But even for mobiles – which are usually considered to be much more of a “personal” device – we find that about a quarter of users are sharing them with other people.

This creates major complications when attempting to track the behaviors of a single individual and shows why there can be no automatic assumption that one device represents just one user. In some cases, one device is in fact being used by 5+ people – a behavior which is most common of all in fast-growth nations.

## Device Sharing

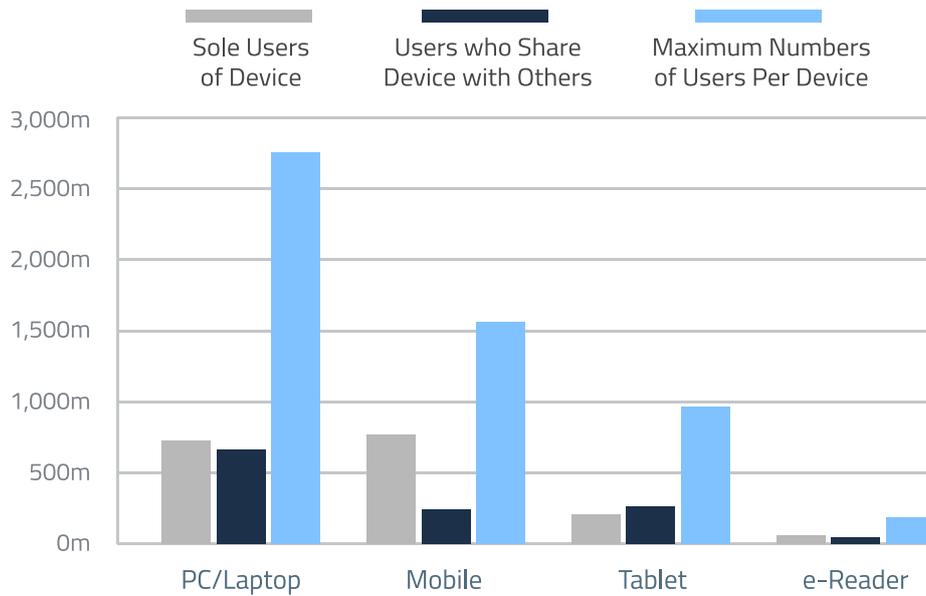


**Question:** Thinking about the devices you use to access the internet, can you please tell us how many, if any, other people use each one? // **Source:** GlobalWebIndex Q4 2013 – Q2 2014 // **Base:** Global internet users aged 16-64

The main implication here is that large audiences are going uncounted. For PCs/laptops, for example, more than 660 million people say they share them with other users. If we work out a total audience size which includes all of these shared users, the figure rises to a mighty 2.75 billion people.

Of course, these are not 2.75bn unique individuals as many will be utilizing other devices too (clearly, the multi-device nature of today's internet usage means that there will be degrees of duplication across audiences – especially for devices subject to the highest levels of sharing). However, this does highlight the perils of assuming that one browser or device equates with just one user. And perhaps most significantly of all, it reveals that the reach and audience of a particular website could be much larger than currently thought.

## Maximum Potential User Numbers

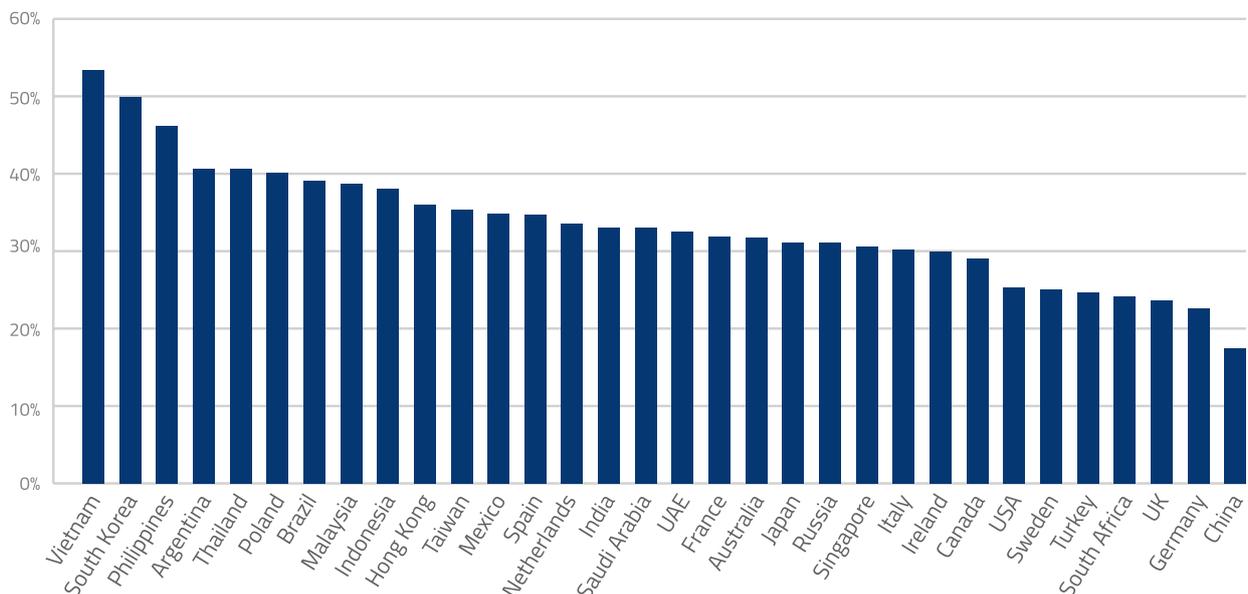


**Question:** Thinking about the devices you use to access the internet, can you please tell us how many, if any, other people use each one?  
**Source:** GlobalWebIndex Q4 2013 – Q2 2014 // **Base:** Global internet users aged 16-64

Our research also shows that there are large numbers who *only* use a shared device – that is, they do not have a single internet access point which is not shared with other people. As elsewhere, this tends to be most pronounced in fast-growth markets; in Vietnam, for example, 49% of the online population are using the internet via shared devices only. And even in China, the country where this trend is least common, there are still 17% who are sharers only.

Overall, there are in fact nearly 425 million people who are accessing the internet via shared devices only – with over 80 million in China alone. These individuals are highly likely to be mismatched or absent altogether from passively collected data and thus represent a huge segment of invisible users.

## Internet Users Connecting via Shared Device Only



**Question:** Thinking about the devices you use to access the internet, can you please tell us how many, if any, other people use each one?  
**Source:** GlobalWebIndex Q4 2013 - Q1 2014 // **Base:** Global internet users aged 16-64

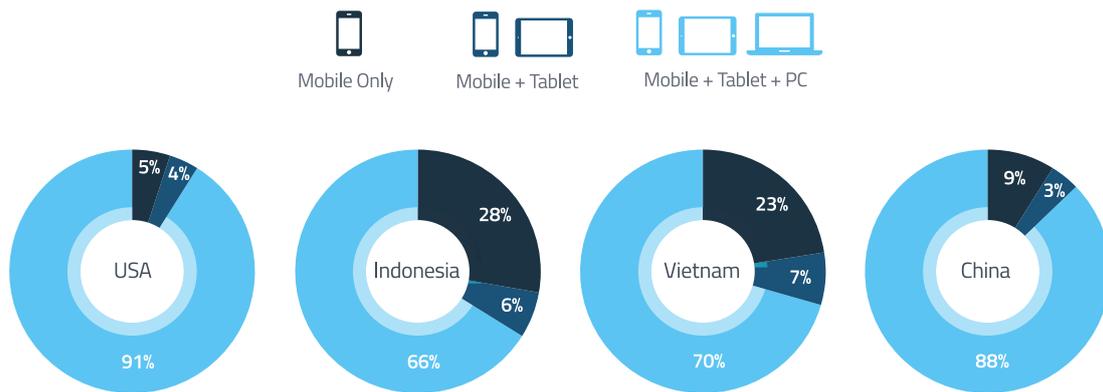
# Mobile-Only Internet Users

The rapid rise of mobile has created a new group within the total online population: mobile-only internet users.

These individuals are experiencing the internet via mobiles alone and represent a growing audience (which we have started to quantify and which, due to its importance, will be integrated inside our research program during 2014).

In our Mobile-Only pilot study (Jan 2014), we interviewed over 2000 mobile users in four markets of varying maturity (the US, China, Vietnam and Indonesia). The results showed that the mobile-only (or mobile- and tablet-only) audience is significant in size even in a particularly mature market like the US (9%). But in emerging internet markets – where there is the least investment and development in measurement as well as the highest levels of mobile sharing – it can grow to as much as 30% (Vietnam) and 34% (Indonesia).

## Mobile-Only Internet Users



Proportion of the mobile internet audience who are mobile only  
Source: GlobalWebIndex Q4 2013 // Base: Mobile internet users aged 16-64

We believe this audience could represent another 168m users across the 32 markets tracked by GWI, all of whom are not covered in audience measurement panels (and the majority of whom are to be found in the fast-growth nations). If we were to extend this to include other markets not currently surveyed by GWI – especially those with low overall internet penetrations but high levels of mobile internet usage – the mobile-only user numbers would climb further (and by an appreciable amount).

While it's long been known that mobiles are having a transformative effect on internet access, when we take these mobile-only users and add them to the sharers identified above, it's clear why this device is causing acute problems in terms of user identification; passive techniques simply haven't evolved to respond to mobile users and their behaviors. And that this is such a driver of the Missing Billion phenomenon, especially in fast-growth markets, is of particular concern given how quickly mobile internet usage is growing each year. In short, more and more people will become invisible if we continue to rely on passively collected data alone.

# IMPLICATIONS

## For Measurement

The perception that passively collected data provides accurate and indisputable “facts” is very far from the truth. There is no such thing as “actual” data: passive data is no more than an estimate of one aspect of an audience.

Significantly, **passive techniques do not reflect the complex, multi-device nature of internet usage today** and – by misrepresenting, incorrectly geo-allocating or overlooking people using shared devices, VPNs or mobiles only – they are causing more than a billion people to become invisible.

This has a **disproportionately strong impact in fast-growth markets**, despite these countries now dominating global internet usage. As a result, hundreds of millions of users in these nations are losing out on investment as well as access to content and other services. This form of “digital imperialism” is **reinforcing out-of-date views about the world being divided into developed vs developing markets**.

No company, brand or agency should be making investment decisions based solely on analytics or passive data; they should be used in tandem with other data sources.

## For Marketers

The global nature of the internet today provides an incredible opportunity to drive scale. But while it is possible to deliver globalized marketing and campaigns to a massive audience, this is being held back by inaccurate data.

**Budgets should not be set based on passive data metrics**, and nor should internet penetration be used as a figure for judging in which markets to invest. Accurate data reveals the dominance of growth markets for today’s most popular worldwide platforms (e.g. Facebook, YouTube, etc).

Accurate geo-targeting is essential. Currently, billions of dollars of ad spend are potentially being mis-directed.

## For Content Owners

Although the global marketplace for content streaming is huge, investment is still typically focused on the US or Western Europe.

The **number one reason for using a VPN is to access entertainment content**; this is clear evidence of significant yet untapped demand.

In the vast majority of markets, the best way to access content online is to source it illegally. With proper data representation, it’s clear that this audience can be quantified, valued, served and, ultimately, monetized.



[www.globalwebindex.net](http://www.globalwebindex.net)

T: +44 20 7731 1614

A: GlobalWebIndex, Bedford House, 69-79 Fulham High Street, London, SW6 3JW, England



**Tom Smith**  
**CEO**

E: [tom@globalwebindex.net](mailto:tom@globalwebindex.net)



**Jason Mander**  
**Head of Trends**

E: [jason@globalwebindex.net](mailto:jason@globalwebindex.net)