NARRATIIVE

Digital Platforms Inquiry

News media bargaining code – concepts paper response

June 2020

Response to Questions

Sharing of user data

22. Should the bargaining code include minimum data-sharing obligations for each of Google and Facebook? If so, what should these minimum data-sharing obligations require?

Yes, there should be a minimum data-sharing obligation in the code and it is essential this includes a consistent, privacy compliant visitor identifier shared with publishers that all other data sharing obligations revolve around.

Privacy compliance can be achieved by adopting the models already established with the Apple Advertising Identifier (IDFA) and Google Advertising ID (ADID).

27. Would it be appropriate for each of Google and Facebook to provide news media businesses with access to additional data associated with individual user (based on anonymised user IDs), such as whether a visit to a news media business's website follows previous interaction with this business's content on a digital platform? If so, what steps should be taken to ensure an individual's privacy is protected?

Yes, this would be appropriate however the key point put forth in this submission is that this needs to be identified consistently to individual visitors after they arrive on publisher's content (in all environments) and be provided in a way that is consistent with the publishers owned data assets so that this additional data is actionable by publishers.

To ensure privacy this should not include any Personally Identifiable Information (PII).

This can be modelled on best-practices already implemented in the Apple Advertising Identifier (IDFA) and Google Advertising ID (ADID).

Treatment of paywalled news content and alternative news media business models

41. How might any relevant mechanisms in the bargaining code ensure treatment of paywalled news content is fair, without interfering with the general operation of ranking algorithms or unreasonably limiting consumers' access to free news?

Due to the erosion of first-party tracking, existing paying subscribers to paywalled content face added friction accessing content via Google and Facebook that requires them to log in multiple times, even on the same device. This adds frustration for the visitor, and undermines the value of a paid subscription.

By implementing consistent, secure, privacy compliant visitor identifier as proposed in this response, publisher's will be better able to identify their paying subscribers and keep them logged in more consistently in different environments.

Display and presentation of news on digital platforms

44. Which specific digital platform policies and practices affecting the display of news have a negative impact on the business models of news media businesses and/or their ability to monetise content?

As detailed below the fragmentation of first-party tracking due to the development of Google Accelerated Mobile Pages (AMP), Facebook Instant Articles (FIA) and the use of embedded web browsers that affect the display of news content have all contributed to the erosion of first-party tracking that enable publishers to personalise, monetise and measure their websites.

Background on Narratiive

Narratiive was founded in Melbourne Australia in 2008 and is the official audience measurement currency for South Africa in partnership with the IAB South Africa, and the de-facto audience measurement currency in the Middle East & North Africa.

With over 10 years of experience measuring desktop and mobile audiences as a provider of market wide competitive intelligence and audience profiling services for publishers and advertising agencies, Narratiive has seen a significant evolution in publisher fortunes in the digital space and an ongoing erosion of key technologies used to personalise the visitor experience and monitor the effectiveness of their publishing assets.

When Narratiive was founded mobile traffic was around 5% of the audience and today it stands at over 95% for most publishers, highlighting the incredible shift to a mobile-first world as a consequence of the revolution kicked off by the release of the first Apple iPhone in 2007. On the same timeline Facebook, which had 145 million active visitors at the time of our founding, has seen unprecedented growth in referral traffic to publishers and now counts a global audience of over 2 billion people.

Narratiive can provide as part of the Digital Platforms Inquiry and mandatory bargaining code process a technical perspective regarding the changes imposed by the digital platforms that have impacted on publisher's ability to accurately understand their own audience and the additional impacts on targeting and personalisation solutions in a first-party context. These have resulted from an erosion of reliable visitor identity mostly brought about by changes and initiatives implemented by the digital platforms to further service their own business aims.

No suggestion is made that these technical developments have been specifically intended to erode publishers' businesses, however the unintended consequences of these changes and their impact on publishers need to be highlighted.

For any stakeholders that would like to contact Narratiive for further discussion on this response, we encourage you to email accc@narratiive.com.

The Problem

The erosion of consistent first-party tracking

The technologies required to enable tracking of visitors within a publisher's own website (specifically first-party Cookies and LocalStorage) support the most basic forms of value derived on the Internet – including visitors being able to "log in" to websites, apps and other services; the ability to track items in a "shopping cart" and "checkout"; and the ability to measure a reliable and consistent understanding of "unique audiences" visiting websites for analytical purposes.

It's important to contextualise the use of first-party Cookies and LocalStorage as fundamental building blocks of the World Wide Web given it is inherently *stateless*. This means from a basic network communication perspective every action on the internet is regarded as discrete and unrelated. It is only through the implementation of first-party Cookies and LocalStorage that a *stateful* experience is possible on the World Wide Web.

Narratiive has observed over the years a number of developments and initiatives by the digital platforms that have eroded the ability for publishers to reliably track their visitors in a first-party context. The distinction with third-party tracking (detailed below) is an important one, as none of the basic capabilities listed above that we take for granted as part of our daily experiences online are possible without first-party tracking.

The deprecation of all third-party tracking

A lot of emphasis recently has been placed on the tracking of visitors in a third-party context. Simply put, third-party Cookies and associated technologies such as Device Fingerprinting have been used for a long time in an opaque manner to track the same browsing device across different websites. These third-party Cookies have enabled an ecosystem of tracking, segmentation, personalisation and campaign effectiveness products that support publishers in monetising their audience.

Google plans to completely disable third-party Cookies and related tracking technologies in their Chrome browser within the next two years, which will have a significant impact on the advertising value chain. Other major browsers such as Apple Safari and Mozilla Firefox have already implemented extensive limitations on third-party tracking, such as Safari's Intelligent Tracking Protection (ITP).

This issue of third-party tracking will no doubt be raised in the context of the ACCC's 'Digital advertising services inquiry' by various industry participants.

Narratiive is not addressing these issues in this response and the problems we are highlighting, and the proposed solution, are related solely to first-party tracking.

Initiatives by the platforms that have eroded first-party tracking

Google Accelerated Mobile Pages (AMP)

The Accelerated Mobile Pages (AMP) Project is an initiative by Google to deliver web-pages faster and make content easier to crawl. However, the restrictions placed on the presentation of content and the lack of support for JavaScript in this environment restricts the ability for publishers to extract full value from their audience, handing a significant level of control to Google in the process.

AMP pages are cached by Google and served directly from https://www.google.com/amp, effectively taking over the hosting of publisher's content on Google's own domain.

Google heavily prioritises the 'Top Stories' section of search results from AMP enabled content, thus creating a strong incentive for publishers to implement this technology and have it served from the Google Cache. If they don't, they can miss out on a significant level of referral traffic from Google searches.

One of the consequences of caching and serving the publishers content from google.com instead of their own domain means that the publisher cannot set and retrieve first-party tracking, as all first-party tracking is associated with the domain name it is being served from.

After a visitor clicks into another article on the publisher's website it usually reverts to their own domain and the tracking in a first-party context is reset. The consequence of this is that the same visitor is at that point treated as a second and distinct visitor, inflating the publisher's audience and removing any ability for a publisher to personalise the visitor experience based on the previous article they viewed or enable additional tracking and segmentation to optimise revenue on advertising inventory on their own websites.

Google has implemented a 'Client ID' in their support for analytics on AMP, however a lot of additional work is required by publishers to reconcile this with their own visitor IDs by utilising a further module called 'AMP Linker' that does something called link "decoration" to pass the 'Client ID' through to the publisher so it is accessible when the visitors finally arrives on their own domain. However, conflicts are inevitable if the same device and browsing environment has already visited the publisher site directly prior to a page served from the AMP cache, as it is a one-way street, so this is only a partial solution to the problem.

Facebook Instant Articles (FIA)

Facebook Instant Articles (FIA) is a framework implemented by Facebook to deliver faster load times of the content being served to users within the Facebook mobile application.

All analytics and tracking code must be embedded inside something called an IFrame, which is essentially a webpage nested within a webpage, and this creates additional complexity and imposes a number of restrictions on what analytics and tracking technologies can do in this environment.

For example, Narratiive runs opt-in demographic surveys to profile audiences and it is not possible to run these constrained to the IFrame environment. As a result, a key pillar of publisher profiling is unavailable in this context.

Embedded web browsers in native mobile apps

On your mobile device there is a native browser installed with the operating system and the opportunity to download alternate browsers. Safari is the native browser on Apple devices however visitors can download alternatives such as Google Chrome or Mozilla Firefox.

This native browser has its own first-party cookie storage while pages loaded from within any mobile application that utilise an embedded web browser, including those from Facebook and Google, load web content from within the app and these have a separate first-party cookie database.

When you access an article shared within Facebook it does not trigger the launching of the native browser on your mobile device but loads the content in an embedded web browser which pops up within the Facebook application.

It is important to highlight the significant fragmentation of first-party tracking this causes as the unique identifiers of a visitor between their native browser, or a third-party alternative, and embedded web browsers are disconnected and thus from a measurement, personalisation and targeting perspective this same visitor is seen as two completely separate visitors - even when it is the same person accessing the same publisher on the same device.

NARRATIIVE

Consequences of eroded first-party tracking

Inflated unique audience measurements

Duplication caused by the erosion of first-party tracking means various analytics and market wide audience measurement tools overcount the unique audience of a publisher, often by a significant but variable margin depending on the breakdown of how a publisher's content is accessed by their visitors. This undermines the accuracy of the data they rely on to make business decisions.

Erosion of advertising yield through reduced segmentation & targeting

Duplication of unique visitors means the accumulated content consumption of their subsequent visits to a publisher website are split into multiple profiles. Effective segmentation and targeting of visitors support increased advertising revenue by making the individual ad impressions (known as inventory) more valuable as advertisers are willing to pay more for ads that are targeted and personalised to the audience they are attempting to reach.

One of the key pillars of both Google and Facebooks dominance of the advertising market is the incredible array of targeting data they can provide advertisers, so any further erosion of publishers' ability profile their audience as a consequence of the decisions made by the digital platforms further entrenches their market power and the disparity of the value proposition to advertisers.

In simple terms, while the digital platforms appeal to advertisers has grown stronger and stronger over time, the same initiatives used to lock publishers into a dependence on the digital platforms have further eroded the ability for publishers to monetise their own audiences making them weaker and weaker.

Inability to personalise visitor experiences

Segmentation and targeting are used, in addition to advertising, to personalise the visitor experience so that what is shown to the visitor can be tailored to their interests and improve the user experience.

Effective personalisation by publishers leads to deeper and longer engagement, and more advertising and related revenue as a result. Undermining the ability for publishers to personalise the user experience further compromises their ability to effectively monetise their content.

Poor experiences for paywalled visitors

The necessity of logging into a publisher's website across multiple contexts (native browsers, embedded browsers, via their desktop and mobile devices) is a point of frustration for paying subscribers and creates a substandard visitor experience. This adds friction to the process for the vital paying subscribers of publishers who have implemented a monetary paywall, or simply require their visitors to log in and identify themselves to access freely available content.

The Solution

Development of a consistent, secure, privacy compliant visitor identifier accessible to all publishers to counteract the erosion of first-party tracking

Narratiive believes the major digital platforms should develop, in consultation with the industry, a standard visitor data sharing protocol and framework with a focus on consistent visitor identification.

Each platform (and in particular Google and Facebook) possess a consistent visitor identifier in almost all instances, and a derivative of this visitor identifier should be passed to all publishers when their content is linked to from these platforms.

This could be a privacy compliant hashed derivative of their own visitor identifier, paired with the domain name of the publisher in question so it is scoped to a first-party only context and generally analogous with the existing technologies.

Publishers (and the various analytics, segmentation, personalisation and data management tools that support them) could access this hashed visitor identifier and associate it with the collected visitor data within a first-party context.

This can be modelled on similar solutions such as the Apple Advertising Identifier (IDFA) and Google Android Advertising ID (ADID) that are already accepted as best practice from a privacy and opt-out perspective.