The Cookiepocalypse: Why first-party data is going to matter to your newsroom

By Paul Herman

The end of third-party cookie tracking could be a huge issue for media companies that rely on digital ad spend, but it could also be an opportunity, *Paul Herman* writes.

Cookie Monster and Elmo from Sesame Street stand in Times Square. REUTERS/Shannon Stapleton

In digital newsrooms around the world, there is a clock counting down to the second half of 2024 and the demise of the third-party cookie.

The question of first-party data, and what it's going to mean for our business, is at the heart of what we will be able to do as an industry going forward – both commercially and, increasingly, editorially.

A solid registrations strategy will be gold for any news publisher wanting to offset the effects of the 'crumbling Cookie'. But this change is also an opportunity to improve both the ad and subscription revenue models, and improve analytics reporting. To understand this, we need to start with a short recap.

What is the cookiepocalypse?

A First-party cookie is a tracker created and stored by the website you are visiting directly. It helps the site remember your login, or what articles you read. This cookie is safe from the cookiepocalypse. A Third-party cookie is a tracker placed on a website by someone other than the owner – most frequently, advertisers – that collects user data for that third party. This cookie is crumbling.

In 2020, Google announced it will phase out the use of third-party cookie trackers on Chrome, and replace it with technology stemming from their workgroup, the [Privacy Sandbox](#).[1] Apple began making similar moves even earlier, phasing out third-party cookies with their ITP initiatives starting from 2017. Other browsers like Firefox and Brave now also block third-party cookies.

All of these moves were, in the main, brought on by increasing legislation around data privacy, including the General Data Privacy Regulation Act (GDPR) and [ePrivacy Regulation](#) in the EU.[2]

The GDPR is generally regarded as a win for privacy for the consumer, and among one of its many implications has been the greater protection of users' online data, how it is processed, stored and shared in the absence of consent.

Google's Privacy Sandbox, set to roll-out in the second half of 2024, promises to replace the old dominant way of tracking people through third-party cookies, and bring in a "secure environment for personalisation" that "still protects user privacy". This has not stopped uncertainty in the digital advertising market however, and has been dubbed the looming "Cookiepocalypse".

What does the end of anonymous, consentless tracking across platforms and websites through third-party cookies mean? Going forward, the currency for ethical tracking will be upfront consent (think of those "Allow Cookies" pop-ups).

Google has a plan to roll-out a new technology called Fledge to replace this old model, but it is still in the trial phase. There is also Google Consent Mode, which can make

---

[1] Google Blog: https://www.blog.google/products/chrome/building-a-more-private-web/

[2] https://www.lexology.com/library/detail.aspx?g=125b3c74-c34a-4939-aa7a-0ee6904de6c5

your Google Analytics platform GDPR-compliant, but it will lead to less detail in reporting if users opt out of consent.[3]

While there is still a lot of uncertainty, there are two definite implications: firstly, a fundamental change for the digital ad revenue model, specifically programmatic advertising that many digital news publishers benefit from. Secondly, it will have minor implications for analytics platforms that measure audiences, like Google Analytics (GA). GA uses first-party cookies and so its primary function will remain intact, but without an ability to count users across devices or browser logins, 'unique' numbers are not always reliable. Let's address the digital ad revenue side first.

Anonymous audiences are going to become less valuable

The current ad revenue model is valuable because advertisers can track, target and retarget specific demographics of users across various websites and serve them with ads across multiple devices. Using tracking systems like third-party cookies, advertisers can buy up inventory to serve these users with relevant ads based on their preferences or history. This inventory is usually purchased through what is called programmatic advertising, and likely includes the ads displayed on our news websites. But aspects of this system are now going to change upon the end of third-party cookies.

For instance, being able to retarget users across websites through what is called the open-auction or remnant inventory will be the biggest casualty, because this will now be in breach of the new privacy laws. While there are other tracking technologies available that take into account consent and allow for tracking, the third-party cookie is the most famous and the most widely adopted, so its fall in 2024 will be the most widely felt. Crucially, advertising clients will find it difficult to measure the effectiveness of their ads on your site (i.e. "conversions": when someone clicks on a link and leaves your website to go to theirs to buy something). This won't be measurable anymore under the new system.

The results of this, is best put by George Montagu, FT Strategies Subscriptions Lead: "When we think of the term 'unique visitors', a lot of that data that relies on the Google platforms is going to change. You won't be able to target repeat visitors because you won't know who they are. A lot of browsers are going to look the same, and therefore you won't be able to understand what is and what isn't 'unique'," said Montagu. "The

---

[3] https://www.cookiebot.com/en/google-analytics-gdpr/

main concern is you won't be able to tell advertisers exactly how many humans these ads on your site are reaching. The way people have sold this in the past is impressions. The more sophisticated advertisers could say you've reached 600,000 people, and they are typically from these demographic groups.

"You can still continue doing that, but that is now going to be done at a much more aggregate level [with less detail], which marketers do not like. And what advertisers won't be able to say is that these people clicked through and then purchased something [on another site]. So there's likely to be no or limited attribution on spend or [attributable] return on investment."

Gareth Lloyd, UX researcher at Meta and former head of data at 24.com, added: "This really is the biggest issue. Publishers can still build audiences using first-party cookies, but then when advertisers actually use these audiences they can't properly track performance. Performance advertising, or conversion tracking is not going to be possible. It's already being felt across the industry with advertisers having to pay more per conversion, or simply not being able to track conversions."

It bears repeating: this uncertainty doesn't mean that you won't be able to track users within your own ecosystem or website. The ability to track and collect your users' first-party data will still be possible provided you have upfront consent, and you can still build audiences. Media houses traditionally have always been the first stop for advertisers looking for an audience, and this change is an opportunity in disguise to become so again, but it will require publishers "coming to the party".

Lisa McLeod, Publishing Lead at the FT Strategies, explained: "In a nutshell, these enormous anonymous audiences are going to become less valuable. For large news publishers of largely unidentified users, the opportunity therefore is that advertisers are still going to need access to those audiences, and media houses hold the keys in being able to do that, which puts you in a position of power.

"Digital media companies will be one of the first buys for every single media planner, because they will ask who has the biggest audience and the biggest heft and how can I reach the widest variety of people. But, in order to do that, you have to play the game on your side, too."

What does playing the game mean? Simply put: you need to harness your first-party data in a way to show advertisers you have an audience that is reliable.

How will the third party cookie issue affect your analytics?

From an editorial analytics point of view, the effects of the cookie changes are not as world-changing. You will still be able to measure visits and engagement on your stories through first-party cookies if you use a tool like Google Analytics. You will still be able to measure performance of stories if you are tracking consistently. But there will be changes regarding what data will be allowed to be displayed, if it was previously collected via a third-party cookie, or if that data input is deemed not 'necessary' for the functioning of your website.

Secondly, an existing problem that pre-dates the cookie issue is the double counting of users who visit across different devices. For example, if I log onto a website on my desktop browser, a first-party cookie logs me with a unique ID. If I then log onto the same website on my phone, I'm counted with another unique ID. It's a duplicate of the same person. A third-party cookie can help here in 'de-duplicating' some of that double count across browsers or devices, but that will no longer be possible.

The numbers you read on analytics platforms therefore are better described as an indication of the number of page views or 'visits', but not the number of unique users, (which is what advertisers really want). But editorially speaking, you'll still be able to tell which stories were 'visited' the most, which performed okay, and which performed poorly, relative to one another.

Having said that, this form of anonymous analytics measurement has another weakness when compared to first-party data where the users are known. Currently, most analytics tools use sophisticated data models to infer the demographics of a particular user based on their patterns. There are different types of cookies which track different types of inputs, like repeat visits. But, the users are still largely anonymous – you can infer their gender and age bracket and preferences, but we don't know who they are, and the data is only calculated inferences.

The inability to track across multiple sites now will also lower the ability to include as many data points as possible to make better inferences. The fall of the third-party cookie therefore will likely "remove a lot of the granular detail" involved in understanding who's reading your stories, said Montagu. This is especially true if your publication is listed across various domains.

Ruby Norris, an audience editor for the *Guardian*'s GLabs, explains why her commercial content team will be affected by the cookie changes. They have an internal analytics tool called Ophan which helps them measure the performance of their stories internally, but when performance of ads and commercial content is reported to clients, they use Google.

"We're heavily reliant on Google Analytics for our analytics reporting. Agencies represent clients and they use standardised Google Analytics reports when we set our key performance targets."

Why registrations and first party data will be gold for your newsroom

When a user has created an account on your site, you will be able to tell if Joe Black has logged in from his desktop or his phone, and you can get closer to not counting him as a duplicate. Provided Joe stays within your ecosystem, you will be able – in theory – to accurately track that logged-in user as a unique visitor across your ecosystem on different devices or browsers.

If you're a news publisher, registrations and account logins are great ways to ethically collect first-party data. According to Montagu and McLeod, the value of registrations and logins are:

- Newsrooms get first-party data in which to track their audience from an analytics point of view. Consent must still apply here. Readers must have the free choice to opt in or out.

- That can in turn assist in building audiences for advertising clients with confidence and in more detail.

- It is great for getting readers into the subscriptions funnel.

On that last point, media houses running some form of reader revenue model have a strong focus on their subscriber or paying base. But typically, these paying readers account for only a small percentage of their total audience. If roughly 10% of your readership are paying subscribers, then that means 90% of your audience below that is being tracked in the 'anonymous bucket'.

By pulling your anonymous browsers into a known bucket and being able to track them directly and more carefully, you will have the three benefits mentioned above. Let's discuss these in turn.

Advertising targeting

The third-party cookie issue is going to stop advertisers from targeting audiences across different sites and networks, and it's going to mean that retargeting is going to fall away. "It's going to completely undercut the programmatic businesses. This represents the opportunity for publishers to get back to direct sales, deeper relationships with advertisers, stronger sponsorship opportunities, and better reporting," said McLeod.

This is because first-party data means you can say with confidence that you know who your audience is, which marketers will like. "For those that have a lot of first-party data, they can measure their demographic data against their users' email addresses, and for the first time will be able to say with confidence that 'you have 40% women reading your articles' or 'people within this age bracket', whereas previously it's calculated guesses with data overlays built into these tools," said Montagu.

The *New York Times,* in its [Investor Day](#) presentation, said it grew by $100 million in the last five years.[4] One of two key drivers of growth cited in their report was first-party data (the other being audio). First-party data helped them create advertising products that service their 135 million registered users and 9.1 million subscribers – a huge, known audience to target ads to.

Subscriptions funnel

Secondly, registered users are great for getting readers into the "subscriptions funnel". There are three stages in the funnel: anonymous, registered and subscribed. The goal is to move readers along each stage from first to last, until they've gone from anonymous or casual users, to involved and registered, to finally happy to pay for your journalism.

One of the core strategies to achieve that movement is through high quality content. But being able to target your readers with this content is also key to moving them along the funnel. Media houses are better able to do this when anonymous readers become registered or logged in.

---

[4] https://nytco-assets.nytimes.com/2022/06/NYT-Investor-Day-2022-Presentation-mC05z.pdf

At the *FT*, they found that a registered reader was 260% more likely to become a subscriber in their ecosystem, said Lisa McLeod. According to the *[New York Times](#)*, registered users, on average, convert to subscribers at rates more than 40x higher than anonymous users.[5] Since 2016, the *NYT* has grown its U.S. registrations four-fold and international registrations eight-fold, which in its own words "created a strong prospect pool for subscriptions". A reminder: that prospect is a whopping 135 million registered accounts on top of 9.1 million paying subscribers. A survey of these registered users showed that more than 50% were open to paying for at least one NYT product, and just over 33% were open to paying for two or more products. (And they were able to survey a sample of these users precisely because they were registered.)

Analytics

Lastly, registered users and subscribers give media houses first-party data that can be filtered into custom-built data analytics tools – which can then be used by both editorial and commercial teams.

Hannah Sarney, Head of Audience Engagement at the *FT*, said her team doesn't use Google Analytics at all thanks to their custom-built tool, Lantern. The *FT* has a hard subscriptions model, requiring most readers to be paying subscribers in order to access their content. A free trial offer acts as their primary registration offering. This gives them a trove of first-party data every time any subscriber or registered reader visits their site, because the reader has already logged in directly to their website.

"We're lucky that the tracking is all set up [and filters into our ecosystem on Lantern]. The third party issue will affect us as well, but it is more in the marketing wheelhouse, so the work is underway to register more users. When it comes to Lantern, registered users is one of the filters we brought in a couple of years ago, because we knew the reader behaviour and habits were so different for subscribers versus anonymous browsers. We can filter every page in Lantern now by subscribers, registered and anonymous users. So now we can go into a chart and see what our registered readers are dabbling with, what their interests are and what their patterns are."

Lantern therefore helps tell them the behaviour of their subscribed and registered audience in a measure of detail that the post-cookie anonymous browser system may not be able to.

---

[5] https://nytco-assets.nytimes.com/2022/06/NYT-Investor-Day-2022-Presentation-mC05z.pdf

It's important that newsrooms understand their own brand, their audience and their reach before deciding on a registrations strategy that suits them. A brand with high visibility and credibility like the *New York Times* can implement a very strict registrations wall. Currently, the *New York Times* allows readers to read one article for free, and issues a registration prompt on the second. News titles with less visibility may still need to attract ordinary readers for the ads model. Knowing your position therefore is important.

There are two streams of login your website might belong to:

1) Voluntary account logins - Your website can offer the ability to register a free account with your website, either through normal email signup, or linking it to existing accounts like Google or Facebook. It is crucial that the signup process is easy and efficient, and that the account only requires a single sign-on across devices, for a smoother user experience, but also to keep them logged in.

2) Mandatory logins - At Facebook and Google, you and I have to have accounts to login there, also known as 'walled gardens'. No account means no access. For media houses that have a hard subscriptions model, paying readers have to have an account in order to access the content. No account usually means only having access to a homepage, but no content. Every reader has some form of account, which can include a free trial account.

There are a number of methods one can employ to register users, and I list a few here:

- Close-out boxes - this is a simple popup on a webpage which asks readers to register for free. It usually includes a message of why this will be beneficial to them as a reader. This popup can be closed if the reader is not interested, and they can continue browsing.
- Metered access - This method measures how many stories a reader has clicked on in a given day. After a set amount, a popup will appear asking the reader to register to continue reading, and usually also includes a message as to why this is important for them. Depending on your strategy, this can also be a close-out box, or mandatory registration wall.
- Exclusive content - Stories are a great way to get readers to register, and arguably the greatest value proposition we offer our readers. It's at the heart of

what we do as news publishers. High-value stories created specifically with a registered or subscribed audience in mind can be locked behind a registration wall, and users can access it by providing their details.

- Added value offerings - Newsletters, curated opinion weeklies or the ability to comment on articles are examples of value-added products for a reader when visiting your website. There is usually an offer of value here for the reader in exchange for their details. The *New York Times* offers all of its newsletters for free in exchange for registration.

In theory, having as many logged in users on your website as possible can unlock opportunities for publishers' and their different growth strategies. At the *FT*, which has over a million paying subscribers, to navigate anywhere on the site you have to be logged in, something McLeod described as "the keys to the kingdom".

"If you can get your readers to log in and associate with your content as logged in users, it means better personalisation, it means better targeting for advertisers, it means better tracking of what the audiences are doing on your side, it means better opportunities with actually sharing data and enabling retargeting in a more structured way with trusted partners."

Here is Gareth Lloyd with the final word: "Identity is one of the most valuable things on an internet built off anonymous browsing. If you have the means to know who someone is, and tie their behaviour data-wise to them, you're building a much better data-set. 'Walled gardens' that ask us to login, like for Facebook and Google, were years ahead of the trend, and have been a huge part of the success of their business models."

Setting up a cross functional team

The final recommendation is for publishers to set up a cross-functional team to determine its first-party data goals across the business. It is easy for different units to prioritise what's important to them, but an overview across units allows data to become part of the design as a media business, with all teams aware of the goals rather than it being centralised in one unit. At the *FT*, they set up a team across marketing, product and tech, data, advertising and editorial units.

This approach allowed all the different units of the business to decide which first-party data goals were important to them, and prioritise which ones overlapped with one

another. "A collaborative approach allows you to achieve both advertising and subscription goals," said Montagu of FT Strategies.

As a result, you can create a strategic vision with a view to actioning through testing and experimenting. This vision can include how you collect first-party data, how it will be measured and maximised, how you protect your audience's data and their interests, and ensure the strategy is designed throughout the business and not centralised by one team. If your news business has not set up such a team, today is a good day to start.