Who Ate All Our Cookies? Investigating Publishers’ Challenges Caused by Changes in Third-party Cookie Tracking

Valerio Stallone a, Aline Gägauf and Tania Kaya
Institute of Marketing Management, ZHAW School of Management and Law, Winterthur, Switzerland

Keywords: Third-party Cookies, First-party Cookies, Digital Advertising.

Abstract: This paper investigates the potential reactions of Swiss publishers, as actors within the digital advertising ecosystem, to the forthcoming fundamental changes to user tracking in the world wide web. The results of this mixed methodical study initiate the discussion on the future of cookie tracking by setting and then answering to four hypotheses regarding first-party tracking, shared ID solutions, Google’s Privacy Sandbox, and a national walled garden system. The results show a clear inclination of Swiss publishers towards first-party tracking and shared ID solutions, neutral standing towards Google’s efforts to undo their harm provoking with their upcoming change, and an aversion towards a nation-wide walled garden. These findings intend to increase the volume of the discussion on the effects of BigTech’s changes on the digital advertising ecosystem as a whole and therefore stimulate further research on the effects on single actors within this ecosystem – beyond the publishers themselves.

1 INTRODUCTION

The online advertising industry is currently facing significant challenges. With Apple’s Safari and Mozilla’s Firefox disabling third-party cookies in 2020 and Google Chrome’s announcement that third-party cookies will be disabled by 2023 (Sparkes, 2022; Szabocsik, 2021), lately postponed to 2024 (Love, 2022), many applications of relevant stakeholders of the online advertising ecosystem will be eliminated (Szabocsik, 2021). The decision to phase out tracking cookies was made to protect the privacy of users. In this research, the aim is to initiate an open discussion lead by the clarification of the standpoint of publishers as suggested by the Marketing Science Institute (2020) in their research priorities. It is yet unclear, how disabling third-party cookie tracking will affect publishers and their business models.

A cookie is a text string that is stored in the Internet user's browser when the user accesses a particular website. The web cookie was invented in 1994 with the intention of maintaining status between clients and servers (Cahn et al., 2016) and is used to store and read different data. For example, products that have been placed in a shopping cart on a website still appear in the shopping cart when the website is accessed later thanks to the cookies. Furthermore, log-in data, details on personal information and other data can also be stored by a cookie (Github, 2019).

If a cookie is set by the website on which an Internet user is staying, it is called a first-party cookie. These cookies are generally used to identify users, remember the user's settings, or save the shopping cart (Cookie-Script, 2021). A company that owns first-party cookies can enter into a partnership with another company. If the first-party cookies are forwarded to the partner company, then the partner company subsequently owns second-party cookies (Cookie-Script, 2021). Cookies that are set for the Internet user by domains other than the one that appears in the URL line of the browser are called third-party cookies.

Cookies are domain-related, meaning that a third-party provider has a different ID stored for the Internet user than another third-party provider. Cookie synchronization provides a channel for information exchange between different third-party providers in the background, in order to obtain information about Internet users and, for example, to serve user-specific advertisements (Papadopoulos et al., 2019).

Worldwide, different approaches to replace use of cookies are already being developed. However, it is
unclear whether one of these solution approaches is sufficient for a publisher or whether a combination of different solutions must be used to avoid a potential drop in revenue. Also unclear is, whether these changes will affect national publishers differently. The solution approaches need to be investigated so that a publisher can properly plan its strategy for selling programmatic inventory. The goal of this work is to present the different solutions and prepare for discussion. The advantages and disadvantages of these approaches should be explored, and its implications investigated.

The focus of this work is only on publishers, in order to being able to deliver specific recommendations for action for this essential participant of the digital advertising ecosystem (Gusic & Stallone, 2020). Effects on other participants of the online advertising industry ecosystem (such as advertisers, intermediaries, and users) are not analyzed. Although being a worldwide phenomenon, we geographically limit our work to Switzerland. Mobile applications are not considered. We start this paper by highlighting the online advertising ecosystem narrowing down the focus of the potential revenue loss due to the disablement of Third-Party-Cookies. We then move on to our mixed method procedure: We present (1) our hypotheses deduced from interviews with experts and (2) show the survey with Swiss publishers. After showing the results of the survey, we discuss them and interpret them, in order to derive a differentiated set of implications.

2 RELATED WORK

2.1 Publishers and Tracking

McKinsey already examined the implications for US publishers of disabling third-party cookies (Brodherson et al., 2021). For this purpose, the authors conducted 28 expert interviews in the US and came to the conclusion that 80% of online advertising activities on non-premium publishers is placed with third-party cookies. For premium publishers, this percentage is lower, because many of these publishers already use first-party cookies. The aim of another study was to find out the actual use of ad choices and financial impact (Johnson et al., 2020). The results showed that only a small part (0.26%) of advertising inventory sold to publishers in Europe, had been shown to users who opted out of behavioral targeting. Further on, the authors found out that these advertising content generated 52% less revenue than advertising content with behavioral targeting (Johnson et al., 2020). Similar arguments are made in two other studies in this area way back in 2014 and 2013 respectively. 2014, authors concluded that the revenue loss for impressions without cookies ranges from 37.5% to 66% (Beales & Eisenach, 2014). 2013, scholars noted in their publication a 30% loss of revenue if the top 5% users do not allow web tracking (Gill et al., 2013). Earlier on, authors concluded in their study that a publisher can make twice as much revenue when buying via programmatic with behavioral targeting than compared to revenue via the traditional sales channel without behavioral targeting (Chen & Stallaert, 2010). In contrast to the studies listed so far, which all predict a significant loss of revenue for the publisher, scholars concluded that a publisher should expect only a small loss, when avoiding behavioral targeting as an offering (Marotta et al., 2019).

2.2 Third-party Tracking Alternatives

2.2.1 First-party Data

First-party cookies enable a publisher to create its own walled garden by aggregating first-party data. An example of how first-party data can be linked is shown by the solution from Meta, which offers advertisers a solution on how target audiences can be found again within their platforms with the help of the advertiser's first-party data. With Meta's conversions API, the advertiser has the possibility to share his first-party data with Meta's server. Meta promises that in combination of the Meta Pixel and the Conversions API will improve the performance and measurement of the advertising campaign.

The study of Diener et al. (2020) goes on to state that first-party data can be used as a basis for personalization and measurement if publishers and advertisers collect the data of their users and process it. Diener et al. (2020) have concerns about the fact that the open Internet is becoming a more and more proprietary, leading to smaller, national publishers having no chance to successfully use their first-party data when compared to the data held by the big walled gardens Alphabet, Meta, Apple, Microsoft, and Amazon (Diener et al., 2020).

2.2.2 Shared ID

With the elimination of third-party cookies, however, this solution is gaining more importance. The shared ID solution is based on first-party cookies. When a website is visited, different first-party data is shared with it. In the process we can talk about declared and derived information (ID5, 2022; Papadopoulos et al.,
2019). **Declared information** is the e-mail address that is voluntarily entered by the user on the website, also known as first-party data according to Hassler (2021). When visiting a website, passive identification signals such as the IP address or the user agent of the device are exchanged via the http-protocol, also called **derived information**. It can be processed by algorithms to deduce the uniqueness of the user. This information is considered personal data and requires the consent of the user due to legal regulations (e.g. GDPR and CCPA), also known as zero-party data according to Hassler (2021).

In order for the collected first-party data to be further used, a publisher must work with an ID provider. Via an API, the collected first-party cookies are forwarded to the ID provider. If the required consent of the user is available, it is the task of the ID provider to create an ID for this user. This ID consists of random sequences of numbers and letters. The API then sends this ID back to the publisher so that the publisher can store the ID in the user's first-party cookie, which can be read by SSPs and DSPs (Davies, 2019). Thus, advertisers can recognize their target audience on the publisher's website when using this approach (Wakefield & Mussard, 2021).

### 2.2.3 Privacy Sandbox

In addition to the information that Google will block third-party cookies in their browser, readers of the post were also informed that Google has launched a new initiative called Google Privacy Sandbox (Cooper et al., 2022; Geradin et al., 2021). One component of this solution is interest-based advertising technology. Federated Learning of Cohorts (FLoC) is an API that can be used to target groups of web users with similar interests into clusters. The FLoC API relies on a cohort assignment algorithm, assigning web users to cohort ID based on their browsing history. The browser updates the cohort ID when web users are active. In order to ensure privacy, the browser requires that this cohort ID be shared by at least n different users (Turati, 2022), whereas n has to be above a certain predefined threshold.

A cohort assignment algorithm is a trade-off between privacy and utility: The more users share a cohort ID, the harder it is to use this signal to infer an individual user's behavior from the entire Web. On the other hand, a large cohort is more likely to consist of a large number of users, making it harder to use this information to personalize ads (Turati, 2022). According to Bindra (2021) and Diemert et al. (2022), initial tests with simulations based on FLoC suggestions from Chrome were successful. The result was that advertisers can expect at least 95% of the conversions per dollar spent when compared to cookie-based advertising.

### 3 HYPOTHESES BUILDING

#### 3.1 Interviews

In order to build our hypotheses, we relied on online advertising experts. We planned to find experts within the online advertising ecosystem of the online advertising industry (see Gusic & Stallone, 2020 for a thorough presentation of them). Individuals were contacted via email or LinkedIn. We ended up with six experts willing to support our research.

In order to be able to evaluate the six interviews, we decided for this work to use content analysis according to Mayring (2020). This method is used for the systematic processing of texts and should help to gain new insights. We evaluated the transcripts of the interviews summarizing the content analysis. The deductive approach pursues the goal of classifying and utilizing the statements from the interviews based on predefined categories.

#### 3.2 Hypotheses

The experts see various advantages in first-party data. Since the data comes directly from the publisher, they can decide for themselves which data is shared with whom. However, first-party data also requires the publisher to install a login wall in order to access the data. All experts recognize similar side effects with first-party data. A lot of resources must be made available by the company so that first-party data can be built up. The experts assume that only larger publishers will be able to provide these technical and human resources at the beginning. It is not only the resources that the experts consider problematic, but also the density of the data. A large publisher will be able to provide much more data than a small publisher with generic content. A login wall could have a deterrent effect on an Internet user, which could lead to a publisher losing a website visitor because an attempt is made to build up first-party data. From an advertiser's perspective, the problem is that first-party data cannot be used across publishers. Advertisers would therefore have to plan a strategy per publisher. Based on this evaluation, the following hypothesis was made:

**H1**: Large publishers will invest resources to strengthen their first-party data.

With regard to the shared ID, the experts see the advantage of standardization. The ID created via a
shared ID. The ID can be understood by all parties affiliated with this provider, although login data being required. The experts strongly advise against working with shared ID providers who still rely on fingerprinting. Different shared ID providers give rise to the problem that the IDs are not understood if different providers are used by the parties. Another concern expressed by the experts is the long-term nature of this solution. This is because individual user data is shared between the parties. Some experts fear that for this reason the solution will be restricted again by law in a few years. The evaluation therefore raises a second hypothesis:

**H2:** Smaller publishers who market themselves will work with shared ID providers.

Products such as contextual or geo-targeting not only offer new ways to track a user in the industry, but could also focus more on the environment a user is in. These targeting methods have existed for some time and are likely to be used forever. With contextual targeting, the advertiser buys specifically on an inventory. Despite the side effects, experts argue that contextual targeting is essential for a publisher. This is also because this targeting method represents a long-term solution.

**H3:** All publishers already know about contextual targeting.

The experts were asked whether they considered a Swiss walled garden to be possible. The experts were of very similar opinion on this question. A great potential is seen in this approach. It would make it easier for advertisers to plan their advertising buying, as it would be possible to buy from several publishers with the same strategy. But it would also bring advantages for Swiss publishers, as the publishers would then be more competitive against the big companies Meta, Alphabet, Apple, Microsoft and Amazon. Publishers also benefit from shared data. For example, Publisher 1 only knows the age and Publisher 2 only the gender of the Internet user. Through shared data, this results in an Internet user where age and gender are known. However, a Swiss walled garden requires that all publishers participate. Some experts express the fear that certain publishers would not participate at the beginning, but that these publishers will join at a later stage. What would need to be considered in a Swiss Walled Garden is a way to share the data between the parties. After this evaluation, a fourth hypothesis is made.

**H4:** All publishers are ready to build a Swiss Walled Garden.

Google's Privacy Sandbox was also mentioned as another approach to solving the problem of eliminating third-party cookies. Various advantages are noted here. On the one hand, this approach allows access to data, even to parties who otherwise do not have any data. Another important aspect of this solution is the fact that an Internet user is no longer tracked individually. Rather, the Internet user is hidden in a group with similar interests. The reach that can be achieved with this solution is also listed as an advantage. From the Internet user's point of view, it is further assumed that they feel less tracked with this solution. Especially because Google builds the solution on the W3C-standard, it would be optimal for the online advertising industry if other browsers would also implement this approach. However, there are also some concerns expressed about Google's solution. Why an Internet user is assigned to a cohort is no longer comprehensible. The solution is therefore a black box for Internet users, but also for online advertising industry participants. Experts see another side effects in the even greater market power that Google will gain as a result. Two experts expressed the wish that such a solution should have come from an independent organization such as the IAB.

### 4 ASKING THE PUBLISHERS

#### 4.1 Methodological Approach

A total of 12 questions on various topics were asked in the survey. The questions were mainly closed questions, which respondents could answer with single or multiple choice. The operationalization of the survey is shown in Table 1. In some cases, hybrid questions were asked. For these questions, respondents had the option of choosing between either predefined answers or their own answer under "other". Four open-ended questions were also asked. These questions always involved a statement of reasons for an answer previously given. At the beginning of the questionnaire, six questions were asked about the company. They were served to find out differences between different attributes. If it was stated that no programmatic inventory was offered, then the respondents had to answer whether there were plans to offer the programmatic sales channel in the future and via which channel in order to be forwarded to the questions for publishers. The condition for this diversion was also that the inventory is marketed by the publisher.
In order to check whether the survey functioned correctly, a pretest was first carried out. People from the authors’ environment filled out the survey and checked whether the branchings were correct and whether there were any spelling mistakes. The respondents could fill out the survey in German or English. The survey took about five minutes to be completed. The survey with the Swiss publishers was conducted between May 10 and May 23, 2021, one year after the first announcement of phasing out third-party cookies by 2022, one month before the announcement of delaying it to 2023 and almost a year before the second postponement to 2024 (Love, 2022). The participants were deliberately written to. Only people who work in the programmatic area for Swiss publishers and advertisers were contacted. We contacted a total of 20 Swiss publishers.

4.2 Results

We excluded survey that were not completed resulting in ten response sets to be considered. All of the respondents confirmed, they were selling programmatic inventory. We depict the results of all the questions in Table 2.

5 FINDINGS

The analysis of the individual hypotheses is based on the size of the companies. They are subdivided as follows: Large companies (250 or more employees),...
medium companies (50-249 employees) and small companies (1-49 employees).

**H1**: Large publishers will invest resources to strengthen their first-party data.

To evaluate the first hypothesis, the responses on the login wall for first-party data, as well as the information on the probability of building up first-party data, were compared with the size of the companies. Out of the three large companies overall, all would install login walls at websites to collect first-party data (100%). In comparison, five of the respondents who do not work for a large company say they would not build a login wall, while three of the respondents in this category would consider a login wall. Regardless of whether a login wall is installed, the inclusion of first-party data in the strategy of both large and small companies is considered very likely. Based on the results of this analysis, the first hypothesis is confirmed.

**H2**: Smaller publishers who market themselves will work with shared ID providers.

For the evaluation of the second hypothesis, the company type Publisher was taken into account. Larger collaborations consider cooperation with a shared ID provider. A collaboration with a shared ID provider is indicated as "very probable" or "probable" by three respondents. All these three individuals work for medium or large companies. Respondents from smaller companies, indicate that collaboration with a shared ID provider is "not probable" or "not probable at all". According to this analysis, the second hypothesis must be falsified. Currently, it looks more like medium-sized and large companies will cooperate with a shared ID provider. This result is hardly surprising insofar as this solution approach has generally received little approval.

**H3**: All publishers already know about contextual targeting.

The third hypothesis’ statement cannot be confirmed: Only 7 out of 10 said they were familiar with this potential alternative. For this reason, the third hypothesis must be falsified.

**H4**: All publishers are ready to build a Swiss Walled Garden.

In order to be able to answer the fourth hypothesis, the question about the Swiss walled garden was evaluated. Since an initial analysis already showed that less than the half of the respondents would be willing to join a Swiss walled garden, the question was examined in more detail to identify any patterns. We encountered that willingness is highest among medium-sized companies, at 60%. In a second comparison, where the willingness was put in relation to the type of company, a higher agreement can be seen among publishers (50%). Not all publishers and marketers are willing to set up a Swiss walled garden. For this reason, the fourth hypothesis must be falsified. Although the establishment of a joint Swiss walled garden cannot yet be given any real chance, half of the publishers are open to such a project. The fourth hypothesis has therefore not been confirmed. It will take time and experience to win over more companies for a joint Swiss walled garden.

The result of the analysis shows that there is currently no consensus among Swiss publishers regarding the right solution. The survey was able to confirm only one of the four hypotheses. Agreement can be seen with the approaches of first-party data and the Google Privacy Sandbox.

### 6 DISCUSSION

The aim was to find out, from the publishers' point of view, which approaches could represent a possible solution. The results of the survey clearly show that publishers are generally very familiar with the various solutions. The solution approach of building up first-party data itself achieves the highest probability of being included in the strategy of the companies. There is also a high level of agreement with the Google Privacy Sandbox. Shared IDs were exclusively indicated by large companies as a probable solution. What does not represent a solution approach for publishers is the establishment of their own walled garden. The participants in the survey also do not see a Switzerland-wide walled garden as an optimal approach. At least the importance of the topic was recognized. This is because the programmatic sales channel is relevant for many of the publishers, which consequently means that all of these publishers must expect a drop in revenue if third-party cookies are deactivated.

#### 6.1 Implications

Many publishers see the creation of first-party data as a possible solution. In order to implement this solution, the publisher must first clarify important strategic questions. Decisions must be made about how to access the data, how to interpret the data, what data and also how and with whom the data will be shared. Some experts have expressed fears of a waste
of resources on this point. In this work, it was assumed that only the large companies would build up first-party data due to the resources required and the high effort involved. However, regardless of the company size, it has been shown that this approach is followed in all the companies surveyed. The actual effort of each individual company should not be underestimated. After all, SSPs and DSPs cannot read first-party data from publishers without additional effort. If all publishers collect their own first-party data and offer it as targeting, this means that an advertiser must increasingly buy from individual publishers and can no longer apply a unified buying strategy.

A smaller effort exists for the publisher in a cooperation with a shared ID provider. The assumption that smaller publishers use this approach is related to the fact that these publishers have fewer human and financial resources. However, the survey showed that large publishers are considering such collaboration. Shared ID providers have been around for a while, but the awareness of this solution approach is the lowest. The fact that there are already many different providers could make it even more difficult to understand the solution approach. It could be that due to the lack of understanding, there is a reluctant attitude towards this solution approach. The acceptance of the Google Privacy Sandbox by Swiss publishers should also be viewed with caution when it comes to weakening Alphabet's monopoly position: This could intensify the mechanism of "digging one's own grave".

Not all types of targeting are affected by the deactivation of third-party cookies. Contextual targeting is already used in campaigns. The solution approach is very well known among publishers, but not as well as we originally thought. Some experts also pointed out the advantage that this type of targeting will never be affected by data protection laws and can always be used. This fact led to the assumption that this solution approach is followed by all publishers. However, based on the survey results, this assumption had to be falsified. From the author's point of view, this approach represents a sustainable solution. However, it was probably rejected because the question was formulated imprecisely. The survey should have been supplemented with the probability of cookieless solutions. The fact that this type of targeting is already in use perhaps means that contextual targeting is not perceived as a solution to this problem, but more as simply targeting that still works. Disabling third-party cookies might lead to the risk of shifting the competitive advantage to the big foreign companies such as Alphabet, Meta, Apple, Amazon, and Microsoft. This is because the solutions from these companies also work without third-party cookies. To counteract this, a Switzerland-wide solution seems sensible and appropriate. However, the participants disagreed on the establishment of a Swiss walled garden. Such a project requires technical know-how as well as the firm will of all participants in the online advertising industry. The experts considered a nationwide solution across all market participants to be the best way to achieve this goal. To be able to do this, the participation of the large media houses and marketers is particularly needed. However, the survey shows that approval from these companies is rather low. Various points are cited as reasons for their rejection. For example, there are fears that Switzerland is not technically ready for such a project, that closed systems limit innovation and that data protection could suffer as a result. In particular, little attention is paid to walled gardens because they are closed systems. Consequently, solutions such as the development of open systems would increase the interest of market participants.

6.2 Limitations and Future Research

In this paper, we undertook an investigation into the world of the online advertising ecosystem and the perspective of publishers on it. Since the number of publishers relevant to this ecosystem in Switzerland is not large, we ended up with a reduced number of respondents. Because the topic is also relevant for other countries, the same study should be extended to the surrounding German-speaking countries like Germany and Austria. With this further research, differences and similarities between the countries could be identified and it could be found out whether, for example, other data protection laws influence the assessment of the solution approaches. Another recommendation is to repeat the study when the solutions are in use.

REFERENCES


