COLUMNS

Case for first-party data in a third-party world

October 12, 2015

Steve Bagdasarian is general manager of Liquid

By Steve Bagdasarian

In an ideal world, effective ad targeting would be easy: get some accurate and complete profile data, put it into segments and use it across channels and devices to reach the right person at the right time with relevant, personalized ads.

Most marketers would agree that the concept is simple enough. But knowing the theory behind data-powered ad targeting versus effectively putting it into practice can mean two very different things.

As Big Data gets bigger and increasingly commoditized, more confusion clouds the market about what constitutes accurate consumer data. Where should it come from? What are acceptable ways to collect and use it? How do we define accuracy?

The line between quality and quantity, first-party and third-party, becomes murkier by the day, and the integrity of first-party data is jeopardized in the process.

For the sake of targeted ad campaigns everywhere, it is time to clear the air about first-party data and what it is, what it is not, why it is so hard to come by, and most importantly why it is worth it.

Cracking the code to better data
No matter how you spin it, successful ad targeting is impossible without accurate information about your audience.

To be true to its name, accurate and complete profile data can only come from one place: a company that is collecting that data directly from its consumers, typically through registration forms, purchase behavior and other activities performed within that organization's owned and operated sites and applications.

This type of information called first-party data is obtained via legitimate collection methods and is therefore authentic, clean and true.

Everyone claims to have first-party data, but the current standard for accurate targeting methodology is often from the use of third-party data, information collected by a business entity (a "third-party") that does not have any direct correlation or relationship to the people from whom they got it.

Third-party data is gathered from a variety of sources and pieced together to form personas composed of hundreds, if not thousands, of demographic and behavioral data points.

Because it is simple to access, third-party data is cheap and widely available everybody is using it.

If you are Ford Motor Co. and you are buying third-party data, you are just doing everything that all the other

automotive companies are already doing. It is not unique. But it is easy.

Cracking the code to first-party data, on the other hand, demands that chief marketing officers and marketers adopt an entirely new way of thinking.

Planning an ad campaign around known customer data blurs the line between marketing and technology, requiring a people-based approach quite different from traditional retargeting or third-party data buying.

Scaling that data into an effective, ROI-boosting campaign is an entirely other hurdle, but it is well worth the effort.

When companies take on the challenge of harnessing first-party data, the results are incredible.

We have seen Facebook turn social media into the gold standard for ad products, Amazon become a retargeting powerhouse, and Twitter drastically scale its targeting capabilities.

These companies collect, scrub and map out their real consumer data, and they are confident that other companies will want to buy it spend billions of dollars on it, in fact.

Turns out, they are right.

Perks of first-party
Why are marketers clamoring to find clean, reliable first-party data for their ad targeting? There are three primary reasons, and they are game changers:

It is deep. When publishers collect information on their own customers, the sky is the limit.

Every click, every movement, every purchase can be tracked. Online and offline behavior is recorded, analyzed and categorized, woven together to tell a comprehensive story about each person who interacts with that company's sites and apps.

It is deterministic. When publishers require users to log in with a single sign-on across their devices, the customer's data story becomes even tighter.

That person's identity is linked to each device in which they are logged on, creating a unified device profile a cross-device targeting map for that user.

This type of first-party information called deterministic data is necessary to accurately identify and target the same person across their smartphone, tablet and desktop, thereby greatly increasing the chances for a conversion.

Facebook is a master at using deterministic data for ad targeting, and it is empowering advertisers with that data like never before.

Aside from the connected ad experience on their platform, Facebook partners with publishers to target and retarget consumers across various properties, on multiple screens.

It is the Holy Grail of ad targeting, providing tangible connected value to the advertiser and increased monetization for the publisher.

Expect to see more of this from Twitter's continued integration with MoPub, advertising consortiums such as Pangaea, and large programmatic platforms including AppNexus.

In addition to its cross-device targeting value, deterministic tracking is opt-in, so it is reliable and privacy-safe.

The same cannot be said for all third-party providers: some third-party data companies build profiles around users without their consent, such as demand-side platforms (DSPs) that listen and collect information without actually bidding, or user acquisition companies that grab other developers' data to build models for their own use.

The proliferation of data companies means that it is getting harder to regulate questionable data collection methods, and the industry has yet to really crack down on these types of operations.

Now more than ever, chief marketing officers need to be ever-diligent about screening potential partners to ensure that they are collecting data in an ethical, opt-in environment.

On the other side of the cross-device matching story is probabilistic data.

This type of data uses predictive algorithms to determine the likelihood that a specific device ID is tied to a user, typically resulting in match rates hovering around 40 percent to 70 percent accurate, at best.

Claims of higher probabilistic matching accuracy are oftentimes based on data sets pulled from a small sample of information or on probabilistic data layered with deterministic data making the probabilistic results look much better than they are.

In either case, those inflated accuracy rates are highly unlikely to hold up in a large-scale environment, so it is best practice for advertisers to make sure that data partners are separating their probabilistic and deterministic models when testing accuracy.

It is scalable. Many people argue that first-party data lacks scalability.

No matter how big their databases, companies have finite consumer data and therefore finite audience prospects, right?

Not quite.

A company's data can be scaled, and it does not even have to be that big to get started.

When an advertiser starts with first-party data, they can be confident that those profiles are scrubbed and legitimate. They are segmented into the very best targeting prospects the cream of the crop and you know those customers are going to convert.

Enter lookalike modeling. Creating lookalike audiences based on high quality, first-party data helps advertisers build incredible scale without affecting the integrity of their carefully curated insights. They can employ lookalikes from their own CRM databases or DMPs, or allow a publisher to create lookalikes on their behalf.

Caveat: Advertisers should approach this method carefully to make sure the publisher gives them the audience that they are asking for and not leaving out valuable data attributes.

Growing an audience based on a successful seed audience is the best way to ensure that the scaled pool of prospects will perform well.

Similarly, third-party data finds its niche with audience layering.

While employing solely third-party information in an ad campaign is less effective, large amounts of clean first-party data overlaid with specific third-party information can be a powerful combination. Layering on third-party insights can deepen data's relevance and greatly increase the number of usable targeting prospects.

A SUCCESSFUL AD targeting strategy begins with deep, deterministic, first-party data, but tapping into this valuable resource is not as easy as it seems.

Ideological challenges aside, today's data environment requires that CMOs and marketers be diligent about asking data providers the right questions to ensure that they are getting user information that is accurate and ethical.

If advertisers take the time to evaluate their partnerships, force transparency into data discussions, and be more selective with the partners they choose, the results will be staggering in a good way.

Steve Bagdasarian is general manager of Liquid, Portland, ME. Reach him at sbagdasarian@pch.com.