

October, 2020

The State of Identity Management: How Did We Get Here and What Comes Next?

614
GROUP

A's

CONTENTS

Acknowledgments 3

Background: The 4A’s Future of Programmatic and Automation Initiative.. 6

Introduction to Identity Management 7

Third-party Cookies/Privacy Sandbox. 7

Identity/Audience Co-Ops 8

Walled Gardens 9

First-Party Data, Third-Party Data, and Proxies 10

Anticipated Future Obstacles12

Conclusion12



Acknowledgments

The 4A's and The 614 Group gratefully acknowledge the contributions of the members of The 4A's Future of Programmatic and Automation Think Tank. Each person who served on the ThinkTank is a thought leader and each one brought knowledge, good will, and unique expertise to this initiative.

The Think Tank selected topics that will have an enduring influence on making the ecosystem move to a better future for programmatic and automation: identity management and workflow automation. Both topics are vital to the immediate future and to the long-term growth and health of the marketplace.

Think Tank members split into two working groups to develop, create and produce white papers on each of the topics. Each member of the Think Tank participated in one working group and each working group was led by two co-chairs and executives from The 4A's and The 614 Group. Without these working groups and the leadership of the co-chairs, the white papers would not have been possible. Below are the names of the working group participants and the respective co-chairs. Together they comprise The 4A's Future of Programmatic and Automation Think Tank 2020.

We are indebted to these individuals who contributed their time and insights:

Think Tank Participants

Josh Abneri - MatterKind
 Yvonne Abt - Sony
 Cliff Atkinson - RPA
 Josh Bock - UM Worldwide
 Daniel Bueckman - Hearts & Science
 Doug Chavez - Geometry
 Patrick Giblin - 451 Degrees
 Greg Johns - Canvas Worldwide
 Mike Margolin - RPA
 Meghan O'Keefe - Starcom
 Claire Russell - Fitzco
 Lauren Smith Ekstrand - Horizon Media
 John Stachulski - Laughlin Constable
 Allen Stern - AKQA
 Tom Swierczewski - Cramer-Krasselt
 Prerna Talreja - Crossmedia
 Kevin Villatoro - Amplifi USA
 Brittney Whisenant - MatterKind
 Grace Yeung - Horizon Media

Workflow Automation Working Group**Co-Chairs**

Yvonne Abt - Sony

Doug Chavez - Geometry

Workflow Automation Working Group Contributors

Josh Abneri - MatterKind

Greg Johns - Canvas Worldwide

Lauren Smith Ekstrand - Horizon Media

Mike Margolin - RPA

John Stachulski - Laughlin Constable

Allen Stern - AKQA

Brittney Whisenant - MatterKind

Identity Management Working Group**Co-Chairs**

Tom Swierczewski - Cramer-Krasselt

Grace Yeung - Horizon Media

Identity Management Working Group Contributors

Cliff Atkinson - RPA

Josh Bock - UM Worldwide

Daniel Bueckman - Hearts & Science

Patrick Giblin - 451 Degrees

Meghan O'Keefe - Starcom

Claire Russell - Fitzco

Perna Talreja - Crossmedia

Kevin Villatoro - Amplifi USA

With gratitude, The 4A's and The 614 Group acknowledge the generosity of the sponsor companies who underwrote the work of the Think Tank and made it possible for this group of thought leaders to convene. Because of our sponsor support, the Think Tank members were able to collaborate on creating a POV on two of the most pressing and complex issues for the future of programmatic and more broadly, the future of automated marketing and advertising.

Sponsor Participants

Basis
BY CENTRO

Tyler Kelly
Ryan Manchee

FACEBOOK

Erik Geisler

IBM Watson Advertising

Brendan Cleary
Jeremy Hlavacek

media.net

Larry Gelfand
Pranay Gupta
Peter Liao

Background: The 4A's Future of Programmatic and Automation Initiative

On March 5, 2020, The 4A's in partnership with The 614 Group launched The Future of Programmatic and Automation Think Tank. The formation of the Think Tank and creation of its intellectual property mark the second stage of a large-scale 4A's initiative that seeks to understand and influence the current and future state of programmatic and marketing automation.

Ultimately, the purpose of the initiative is to foster a better future for marketing automation. Marketing automation must have a foundation of respect for both consumers and practitioners. In the future, it should be easy to optimize the use of data and access interoperable technology in order to be smarter, work faster, and be efficient. Better technology, methodologies and tools will accelerate the process of identifying the right audiences for a campaign, delivering them and measuring the campaign's return. Practitioners will have accurate (where precision is determined through generally accepted standards), real-time inputs and be able to further optimize campaigns or readily course correct. In the envisioned future, data and creativity will come together to make compelling, effective marketing campaigns at scale.

The concept of The Future of Programmatic and Automation Think Tank originated as a real-world action step based on findings and recommendations of a qualitative and quantitative research program. The 4A's partnered with The 614 Group to investigate and understand how best to meet marketplace needs for the future of programmatic and automation. The comprehensive research program included one-on-one interviews with marketing, advertising, publishing and ad tech leaders as well as a quantitative survey. The research identified a number of issues that require cross ecosystem attention to ensure an optimized and efficient future of programmatic and marketing automation. The report, published in November, 2019, can be accessed [here](#).

The specific objectives of the Think Tank are:

- Review the research findings and recommendations to determine which key issues should receive immediate focus
- Put forth actionable POVs and recommendations within a relatively short timeline of six to eight months, once the topics are selected. (Note: this is a different goal and process than standards-setting and technical specification development undertaken by various industry groups)
- Develop ecosystem-wide guidance and POVs around the future optimization of programmatic and automation
 - Core elements of the POVs and guidance should express the agency perspective, while also fully recognizing a cross-ecosystem perspective.
 - Where possible, real world case studies exploring the selected topics should be used to describe successes that illustrate what to do and/or failures to exemplify what not to do.

The Think Tank determined that the two topics requiring immediate focus are: 1) identity management and 2) digital advertising workflow automation. To

develop POVs and recommendations for each topic, members of the Think Tank convened separate working groups for each subject. These groups consisted of members of the Think Tank who volunteered their expertise and time to the work product and are listed in the acknowledgment section above.

The pages that follow detail The 4A's Future of Programmatic and Automation Think Tank POV regarding Identity Management as developed by the working group that is a subset of the Think Tank.

Introduction to Identity Management

Identity management - the art and science of identifying ad recipients using probabilistic or deterministic methods - lies at the heart of digital advertising. Accurate and reliable identification enables relevant, one-to-one communication based on a host of recipient attributes and actions. Conversely, an inability to correctly identify ad recipients in many cases leads to a poor user experience and wasted ad dollars. Even a modest improvement in the reliability and accuracy of identity management and resolution could yield an enormous and immediate benefit--advertisers would be able to spend their dollars more efficiently and effectively; publishers would see increased revenue from their audiences; and consumers would enjoy a more relevant and personalized experience.

To that end, we have assembled senior experts from across the spectrum to lend their thoughts and expertise to this effort. It is our hope that by initiating an honest discussion about the pros and cons of identity management, we will be able to create a more positive and beneficial environment for advertisers, publishers, and consumers alike. (Note that for the purposes of this paper, we will be examining the connection between desktop and mobile only - the integration of offline [in-store, call center, IOT] and multi-channel activity is outside of our consideration.

Third-party Cookies/Privacy Sandbox

Just as personalization became more sophisticated for big advertisers and more widely adopted for small advertisers, Google joined Safari and Firefox in the fight to end tracking individuals across websites, and to limit the amount of information gathered from users on and off the site. As Google's Chrome browser makes up [69% of desktop browser usage](#) and [39% of mobile](#), according to Statista, this signals the beginning of the end for third-party cookies.

Their rollout plan is considerably longer (two years) and less jarring (no major changes out of the gates) than the efforts of Safari and Firefox. But why would Google do this knowing this too is how they make money? While Google uses third-party cookies for pieces of their ad ecosystem, many speculate it's a relatively small sacrifice to propel their dominance even further. Google (along with Facebook) has the unique position to create and keep a persistent identifier across devices as it's pretty difficult to get around the internet without having and being logged into a Google product somewhere - Gmail, Maps,

YouTube, etc.

Still, over the next two years, Google is working with several partners to explore a variety of ways to preserve the tracking needed to keep the Web free, but curb behavior that is considered overreaching by consumers.

One of the main solutions Google is exploring is a privacy sandbox. Google states the sandbox is intended to “[create a thriving web ecosystem that is respectful of users and private by default.](#)” There are three key components to the offering:

1. Replace functionality served by cross-site tracking, which includes solutions for audience targeting and conversion tracking
 - a. It’s worth noting that the conversion tracking proposals are mainly based on last touch attribution, which won’t account for view-thru
2. Remove third-party cookies carefully
3. Mitigate workarounds (e.g. fingerprinting) by pointing developers to new functionality

Other proposed privacy sandbox outcomes include [the introduction of “noise”](#) (artificial conversion data) for obfuscation purposes; aggregating individuals into [delimited cohorts](#) (grouping individuals by defining characteristics); limiting the amount of post-click/conversion data made available; and [aggregated reporting](#) (providing campaign results, but not event-level data), which restricts users to campaign-level (not site-specific) frequency capping.

Apple’s iOS 14 update means a similar situation in the mobile world. The IDFA (Identifier for Advertisers) is somewhat similar to an advertising cookie on the web in enabling advertisers to target users and track performance from mobile activity (including location-based activity, unique to mobile). Developers must receive permission from users in order to collect and share their data. While no one can predict what the opt-in rate will be, when iOS13 was released, millions of people opted to block applications’ ability to track their location when not in use, [resulting in 70% of the data disappearing](#). As such, the volume of IDFA-based measurement scale can be expected to drop significantly. Additionally, apps [cannot ask users to change the permission](#) after the first ask. Developers will have to adapt as Apple will also require them to report what will be tracked and shared in simple terms to consumers.

Identity/Audience Co-Ops

The extensions of products that rely on their own ID systems as well as individual publishers will allow for a rise of additional ID sets that can be used in a privacy compliant way.

For companies that work in hardware ID sets, such as OTT devices, Mobile Devices and gaming systems, the privacy of their users come to the forefront. The information of a transparent opt-in of tracking occurs at the device initiation. The data is generally used for optimization of those platforms, but can and has been extended out to the large digital ecosystem as a way to ensure that the device user gets the most relevant information that pertains to

their online behavior.

Similarly, publishers now have an opportunity to extend their own data networks outside of the purpose of fulfilling their own yield. By allowing a transparent exchange of content for consumer compliant information, they create a wealth of owned first-party data on readers. The opportunity is to start sharing this form of data sets with each other, and create a network of privacy compliant first-parties, that could be equally monetized by all contributors over time.

Additionally, the need for alternatives to third-party cookies is poised to shift the model for premium publishers who have their own wealth of information on their audiences. They can develop a strategy around user log-ins, along with the contextual value they can provide. The Washington Post, Vox Media and The New York Times, who all have their own subscriber data, are offering their own data solutions and are moving away from third-party data. This authenticated information can be translated into identifiers that can be used across the rest of the supply chain. Advertisers would shift to buying direct first-party data on publisher inventory, or rely on accessing third-party data via programmatic means. Publishers without first-party data at scale will need to rely on contextual opportunities. Largely, the sharing of publisher data sets between each other would circumvent some of those concerns.

Such data co-ops have traditionally been driven by ad tech in our industry. This usually led to several issues over time:

- Ownership of ID sets were never established, and monetization of them became cumbersome and problematic
- Taxonomy of the data would have limitations with the implementation
- Underpinning by 3rd party cookies
- Lack of transparency, with no account for user privacy
- Fragmentation of data across sources

Despite this, publisher-led co-ops already inherently have some of the necessary tools to succeed: They have a clear and fair exchange with the consumer, by representing content as a commodity. There is an inherent interest in monetizing data quickly and efficiently, but also allowing for their own content to grow based on learning, and they have the necessary site infrastructure. The ID must stay neutral, with equal share and opportunity across all participants to succeed.

Walled Gardens

Without third-party cookies, the walled gardens (Google, Facebook and Amazon) will benefit via their own identity solutions. Additionally, the need for alternatives to third-party cookies is poised to shift the model for premium publishers who have their own wealth of information on their audiences. They can develop a strategy around user log-ins, along with the contextual value they can provide. The Washington Post, Vox Media and The New York Times who all have their own subscriber data are offering their own data solutions and are moving away from third-party data. This authenticated information can be translated into identifiers that can be used across the rest of the supply chain. Ad tech vendors, such as Sovrn and Index Exchange, are exploring

replacements including first-party data solutions such as login identifiers.

Publishers without first-party data at scale, however, will need to rely on contextual opportunities.

Suggested responses to the deprecation of support for third-party cookies include:

- Invest in First-Party Data
 - A robust data collection, hygiene, and application apparatus is a minimum requirement.
 - Ensure data collection, storage, transmission and application is compliant with existing privacy laws.
- Assess partners' first-party data collection, availability, and strategy
 - Walled gardens will now have taller walls with these changes. Big platforms and publishers like Google, Facebook, Amazon and Verizon have their own first-party data that they collect and make available through the consumer's usage (and thus opt-in) of their services.
 - Focus on partners who have proactively moved toward cookieless buying and measurement solutions. For example, partners who have been using device ID or household-level data to care for user identity and privacy in combination with the partner's own first-party data.
- Explore contextual targeting
 - Consider loosening or broadening your contextual targeting parameters to reach a more upper-funnel audience. One of the paradoxes inherent in contextual targeting is that the more specific and restrictive your targeting parameters, the further down the funnel your audience will be.
 - Broaden ways to define context by going beyond site lists, categories, and keyword targeting to include additional levels of sophistication such as exploring sentiment (i.e. is the content on this page happy, sad, exciting).
 - Always apply brand-safety controls when applying semantic and/or basic keyword targeting.

First-Party Data, Third-Party Data, and Proxies

First-party data is the preferred underpinning of identity management; however, in the absence of robust and consistent first-party data collection, storage, and hygiene methods and apparatus, third-party data augmentation is an acceptable substitute or complement. Contextual proxies and online proxies for offline activity (coupon downloads, etc.) are also a viable opinion.

Contextual targeting has been primarily used in the industry to identify and avoid any brand safety concerns. However, many articles on the internet are created with a specific target audience in mind, and a page captures the latest signal of a visitor's interest. This suggests contextual data can be used for targeting the intended audience demographic. This raises one challenge for advertisers in that contextual targeting is inherently mid-funnel. It is unlikely to be a targeting method to bring someone from 'unaware' to 'aware'.

Still, the major benefit offered by using contextual data as a proxy is complete independence from cookie/device identifier-based systems. Another dimension of contextual data that many partners have been exploring is sentiment. Programmatic partners and YouTube alike have been testing the ability to serve ads that align with content that's happy, sad, exciting, etc.

There are several challenges in identifying accurate contextual data, including:

- **Ambiguity**
differentiating between The Notebook (a popular fictional book and movie), notebook (stationery) and Notebook (laptop). A simple keyword-based approach to identifying concepts would fail to capture this difference
- **Synonymy**
textually varying phrases which refer to the same concept e.g. Covid19 and Coronavirus, Meghan Markle and Duchess of Sussex etc.
- **Phrases**
ability to identify difference between 'Citi Field NY Mets Stadium' against Citibank, New York City, NY Mets baseball team
- **No standard classification process**
there are several grey areas without guidelines which make the classification process difficult and predicted results become subject to the algorithmic implementation, e.g. classifying a page about murder hornets as violent
- **Sentiment variation**
A mention on page needs to be complemented with sentiment associated to truly understand the context, but when only based on the words in the article, e.g. critical review of the latest smartphone, it can easily render an inaccurate categorization.

Ambiguity, Synonymy and Phrase related challenges can be solved by relying on a standard vocabulary of entities, instead of textual keywords. This vocabulary should encompass all relevant concepts in the Ad Tech world - people, places, companies, brands, movies, phenomena etc. and should be regularly updated with upcoming concepts. This would also help in standardizing the identification of concepts and defining sentiment on each concept.

Suggested criteria when evaluating Contextual Data Providers include: cookie dependence; GDPR/CCPA compliance; page-level content-based inference; phrase handling; ambiguity management; standardized classification; customization ability; sentiment analysis; brand safety and suitability management; app/video compatibility; and multi-language support capabilities.

The quality and accuracy of third-party audiences is notoriously uneven. Audience data is aggregated from multiple sources - some public, some private; some proprietary, some not. There are a number of industry initiatives to address this (most notably the IAB's [Data Transparency Label](#)), but all fail to address two simple but key questions:

1. Does this audience actually represent people it claims to represent?
For example, are Auto Intenders actually in-market for an automobile purchase?
2. What is the expected and demonstrated performance of this audience?
For example, what percentage of Auto Intenders have actually made a purchase after being exposed to an ad?

Anticipated Future Obstacles

Privacy regulations have transformed our online browsing environment, and proposed legislation such as the New York Privacy Act and amendments to CCPA may continue to add complexity to this evolving space. In efforts to create a new set of privacy focused on open web standards, Google announced plans in January 2020 to depreciate third-party cookies in its Chrome browser within two years. This joins earlier moves by Safari, Firefox, and new identity focused browsers such as Brave, which has amassed over 15 million monthly users since its release in 2019 by offering controls to block cookies and ads. Clearly there is a demand for sustainable identity management, and because [Chrome commands 69% of the browser market](#), Google's move has raised big questions about the future of cross-site tracking and how brands acquire and activate customer data.

Conclusion

In reviewing the information and insights above, several things become apparent:

- The pending deprecation of support for third-party cookies will increase the value of - and reliance on - first-party data. At the same time, legislation will make the collection and use of such data more restrictive. Advertisers must build a coherent and cohesive data collection and application structure, and include robust opt-in collection mechanisms.
- The increased use of multiple devices in a single household presents both challenges and opportunities. As devices within a household proliferate, audiences become more fragmented. However, this also provides multiple opportunities for data collection and analysis, and affords a more comprehensive view of the customer journey for those marketers willing to invest in robust data collection and identity resolution.
- At the same time, viable and sustainable proxies are available. Third-party privacy-compliant audiences - when used judiciously and optimized aggressively - can provide a suitable alternative to first-party identity management. Audience co-ops can be used to share and strengthen resources for those advertisers lacking robust first-party assets. Additionally, contextual advertising - when coupled with semantic analysis and robust brand-safety and suitability filters - can be used to create a semblance of personalized, 1:1 communication.