The ANA
Trust Consortium

Restoring Trust through Transparency, Integrity, and Growth

DATA SOURCES FOR MEDIA: A BUYER'S GUIDE

MAY 2020

INTRODUCTION

Now representing over 80 percent[1] of all United States digital ad spending, programmatic media buying, or the automated buying of ad impressions enhanced by data, has substantially changed both the media landscape and the way advertisers operate.

One of those changes has been the creation of a new ecosystem of companies selling data to use in media buying, promising the ability to deliver a more on-target message more efficiently. Advertisers can easily buy and deploy data allowing them to target consumers based on a wide range of criteria, ranging from simple demographic attributes all the way to detailed purchase intent data. According to the IAB, in 2018 advertisers spent $19.2 billion[2] on third-party audience data and related solutions.

This new data ecosystem is deeply complex, and often opaque. Without transparency, a solid definition of data quality, and tools to evaluate the data quality, advertisers are at a significant disadvantage in navigating this new ecosystem. This leaves them at risk of spending on data that isn't suitable for their desired purpose.

To make these evaluations, advertisers need transparency. They should be able to get detailed answers to questions regarding the sources, methodologies, and processes used by their data vendors. To that end, the ANA Trust Consortium has developed this buyer's guide, aimed at educating marketers about the data ecosystem for media buying with suggestions for the questions they should be asking when acquiring or using a dataset.

This guide may be useful for advertisers who are:

- Evaluating data to use for an upcoming campaign/initiative
- Conducting a post-mortem to understand why targeting an audience did or did not work
- Conducting a review or audit of vendors
- In the process of evaluating or signing a contract with a new data vendor

---

[1] **U.S. Programmatic Digital Display Ad Spending**, eMarketer, November 2019
[2] **The State of Data 2018**, IAB, December 2018

Trust Consortium ANA

ACKNOWLEDGEMENTS

Trust Consortium ANA

EXECUTIVE SUMMARY

This paper focuses on the use of external (third-party) data for digital media programmatic buying. Third-party data is data that is purchased from or traded for something of value with the party which had the consumer relationship. Overall, when acquiring data from an external source, it's likely that the data is third-party.

This data most often takes the form of "audiences," which are lists of consumers who fall into a particular category or exhibit a specific behavior. Advertisers can then target those lists (or combinations/overlaps of those lists) in their programmatic media buys. Some common examples include demographic/census, behavior, location, and identity.

In addition to buying audiences, advertisers can also use data related to a given piece of advertising inventory. This is often included by the publisher in the bid request and sometimes acquired from another third party. This data is most commonly related to verification, viewability, and context.

Third-party data can come from many different sources, and multiple methodologies can be used in its collection, structuring, and marketing. This often makes it difficult for advertisers to understand exactly what they're buying. This leaves advertisers at a disadvantage and at risk of purchasing data that's unsuitable for their purpose. To address the problem, advertisers need to evaluate the quality of a dataset before purchasing it. There are five criteria to focus on when evaluating data:

1. **Data Accuracy:** Does the data actually mean what I think it does, e.g., are visitors to an auto website actually more likely to buy cars?

2. **Data Precision:** Are the data collection and modeling procedures sufficiently precise to avoid a large number of false positives, e.g., does the vendor use a lookalike model that assigns people to the audience who shouldn't actually be included?

3. **Data Recency:** How regularly is the data refreshed? When was it last refreshed?

4. **Data Coverage:** Does the dataset cover enough of my intended campaign audience to provide necessary scale for my client's campaign?

5. **Data Deployability:** Can I use the data with my chosen tech partners?

The process of selecting a data provider should be treated as an important vendor selection procedure. Advertisers considering a new data partner, or a new offering from their existing partner, should follow this checklist:

- Read the Data Label.

- Ask additional questions of the provider/data source. (See **Appendix** for some suggested questions.)

- Consider engaging a third party to verify the data.

TABLE OF CONTENTS

Trust Consortium

ANA

I. HOW IS DATA USED IN DIGITAL MARKETING?

DATA USE CASES

Data permeates the digital advertising ecosystem. It is used in nearly every step of the media value chain, from pre-campaign planning to post-campaign measurement and reporting. Data is commonly applied for:

- **Media planning:** Data is used in media planning to generate insights that can help construct the media plan and profile high-value segments.

- **Programmatic media buying:** Data is used to improve media efficiency by suggesting whether to purchase a given impression. Advertisers can purchase data to help them identify high-value media audiences or determine whether a specific page context is appropriate for their ad.

- **Personalization:** Advertisers may also employ data about users or about a given piece of inventory to decide which creative version to display. For example, an advertiser may use a third-party audience to identify users who are currently in the market and show those users an offer-focused message, while showing other users who aren't on the list a prospecting message.

- **Reporting and Measurement/Attribution:** In certain industry verticals, external data can form a key part of the measurement strategy as well. When there is no direct relationship with the consumer (e.g., most CPGs), marketers frequently turn to sales data from third parties to measure performance.

- **Accountability and Insights:** Data is often used to match across different segments to learn about which consumers are receiving and responding to advertising. For example, mobile data might reveal if a person went to an auto website while at a dealership or might find overlap between people in Southern states and those who clicked on a skin care message.

Data features heavily across all digital marketing functions. However, this white paper focuses on the use case that is top of mind (and wallet) for most advertisers: external (third-party) data for programmatic buying.

Trust Consortium
ANA

I. HOW IS DATA USED IN DIGITAL MARKETING?

TYPES OF DATA

The primary types of data are audience data and inventory and verification metadata.

Audience Data: This data most often takes the form of "audiences," which are lists of consumers who fall into a particular category or exhibit a specific behavior. Advertisers can then target those lists (or combinations/overlaps of those lists) in their programmatic media buys. Some common examples are:

- Demographic/Census: Data related to the attributes of an individual or household
  - » Examples: HHI (household income brackets), financial status (e.g., net worth brackets, creditworthiness), gender, age bracket

- Behavior: Data related in some manner to a user's behavior — either having previously taken some action/made some purchase in the past, or having performed some activity that expresses an interest in or intent to take some action in the future
  - » Examples: Affinity for a specific product category, intent to purchase a particular category

- Location: Data related to a user's geographic location at a point in time or over time

- Identity: Data which allows advertisers to link an offline identity (such as name, address, or email address) to users' anonymous behavior across online environments. Unlike other categories of data, identity graphs don't describe attributes of the user; they provide a way to unify multiple different profiles that refer to the same user.
  - » Example: An advertiser might have a list of opted-in email addresses that they would like to target on the web; an identity graph would allow them to do so.

Inventory and Verification Metadata: In addition to audience lists, advertisers can also use data related to a given piece of ad inventory (often included by the publisher in the **bid request**; sometimes acquired from another third party).

- Verification: Data on whether a piece of inventory is likely to be fraudulent (pre-bid) or was fraudulent

- Viewability: Data on whether the ad is likely to be in view (pre-bid) or was actually in view

- Contextual: Data on the context of a page (e.g., is it sports content? news?)
  - » In addition to targeting page contexts that are likely to reach the targeted audience, contextual data is also frequently used as a part of marketers' brand safety procedures. It allows them to avoid content that could have negative associations for their brand.

Trust Consortium ANA

I. HOW IS DATA USED IN DIGITAL MARKETING?

SOURCES OF DATA

There are nearly as many different sources of data as there are types. Some common examples are:

- Website owners (publishers)
- Retailers/loyalty card administrators
- Credit card companies
- Mobile apps
- Public records
- Consumer panels

Please note that data is often classified as first-, second-, or third-party. First-party data is data that is collected by the party with whom the consumer has a relationship. Third-party data is data that is purchased from or traded for something of value with the party which had the consumer relationship. When an advertiser shares a partner's first-party data, it is sometimes referred to as second-party data. Overall, when acquiring data from an external source, it's likely that the data is third-party. For more detailed definitions of first-, second-, and third-party data, see the **Glossary**.

DISTRIBUTION OF DATA

Data has a supply chain in which raw data is collected from a source, and then cleaned, processed, and made ready for distribution. In many cases, data is collected and aggregated by data brokers which package the data into audiences that can be used by advertisers. However, not every data transaction is so straightforward. Advertisers can also acquire data:

- Directly from a data source or broker
- Included as part of a media buy (either from a publisher or DSP)
- Via their agency as part of the agency's media buying process
- Via their agency from a data vendor owned by the agency holding company
- Via a data co-op with other advertisers

Depending on the nature of the transaction, advertisers may not see a separate line item for data on their invoice or necessarily even be aware that data from external sources is being used. Therefore, it is important that advertisers be aware of what their partners, both agency and technology, are doing on their behalf.

Trust Consortium ANA

II. EVALUATING DATA

Third-party data can come from different sources, and many methodologies can be used in its collection, structuring, and marketing. This makes it difficult for advertisers to understand exactly what they're buying. In addition, this leaves advertisers at a disadvantage and at risk of purchasing data that is unsuitable for their purpose.

Unfortunately, most advertisers have yet to establish a standard approach for evaluating datasets or audiences in general. Without clear criteria to evaluate a dataset before buying it, they often take a "let's see what sticks" approach — starting off with an array of datasets that seem to have potential, and then winnowing them down over time based on performance metrics. Some advertisers may also decide to simply ignore data quality altogether, and just go with whichever audience is cheapest.

DATA EVALUATION FRAMEWORK

To address the problem, advertisers need to evaluate the quality of a dataset before purchasing it. They need to know what questions to ask of data vendors, and what they should be looking for in the vendor's response. There are five criteria to focus on when considering a data source.

1. Data Accuracy: Does the data actually mean what I think it does, e.g., are visitors to an auto website actually more likely to buy cars?

When data is used for media targeting, it almost always involves some sort of inference that an observed behavior, such as visiting a certain category of website, is indicative of a mindset or intention. For example, a vendor may compile a list of users who visited sites related to travel destinations and market it as a "Travel Intenders" audience, making the inference that visiting a destination website is an indicator that the user is looking to book a trip.

These inferences are an essential part of data-driven marketing. Without making them, it would be impossible to have audiences who scale — surveying the entire internet population to see who is in the market for a vacation would be impractical. However, advertisers need to be aware of the inferences that are being made in the construction of their audiences. For example, those users who visited travel websites could just as easily have been fantasizing about a trip they have no intention of ever actually taking. Understanding the inferences that are being made is essential to understanding whether the data is likely to add value. The degree to which these inferences are likely to be accurate is a key factor in whether the dataset is likely to perform as expected.

II. EVALUATING DATA

2. Data Precision: Are the data collection and modeling procedures sufficiently precise to avoid a large number of false positives, e.g., does the vendor use a lookalike model that assigns people to the audience who shouldn't actually be included?

Another factor is how the dataset is modeled. Customer lists can be modeled as either deterministic, meaning that every individual on the list was actually observed performing the target behavior, or probabilistic, meaning the list includes individuals with demographic/behavioral profiles similar to the ones who were actually observed (also known as lookalikes). Deterministic methods are generally considered to be more accurate, but often come with either reduced scale or increased cost, while probabilistic methods trade off some degree of accuracy in exchange for increased scale and/or potentially lower cost.

It is also important to consider the underlying provenance or source of the data. Even if a dataset is constructed deterministically, it is possible for the data to be imprecise. For example, a segment of auto owners built off DMV registrations is much more likely to reflect actual car ownership than a segment developed using data from a Facebook survey, but a segment of luxury auto intenders built from observing visits to auto endemic sites may be a weaker indicator of intent than a dataset of visitors to auto dealerships based on location data. Note that some data suppliers refuse to disclose their data sources, which is likely a red flag from a compliance or a risk management perspective.

3. Data Recency: How regularly is the data refreshed? When was it last refreshed?

Recency, or "data freshness," is the frequency with which the data is updated and old users are expunged. Even if the underlying inferences used to construct the dataset are correct, and if any statistical methods applied are valid, it is still possible for a dataset to contain users who should not be on the list. For targets who are time-sensitive (e.g., many intent-based audiences), if the list is not refreshed on a frequent basis and old users are not expunged, the data can contain a high number of users who should have been removed from the segment either because they decided not to buy or because they already made their purchase.

An example from the consumer electronics category:

» An advertiser looking to sell televisions is very likely to buy media using intent or affinity audiences. The typical path to purchase for a television is about 11 weeks[3]. If the advertiser's intended audience is only updated to expunge old records every six weeks, then at any given point it is conceivable that half of the users in the segment have already made their purchase and in fact are entirely out of the market for televisions.

Before buying a dataset, advertisers should understand how often the data is refreshed as an indicator of the percentage of the audience that is still valid for their campaign. The method and process of the refresh also has an impact on how refresh translates to attribute concentration. For example, "continuous" refresh might allow data over a year old to be deleted daily, which would not improve a television segment.

---

[3]**Navigating the New Path to Purchase**, MillwardBrown Digital

Trust
Consortium

ANA

II. EVALUATING DATA

4. Data Coverage: Does the dataset cover enough of my intended campaign audience to provide necessary scale for my client's campaign?

Buyers often stress the importance of "scale" in their digital media buys. This is with good reason — even with automated tools, managing large numbers of media buys against small audiences can become very labor-intensive. It is therefore important for advertisers to consider the ability of datasets to provide the necessary scale for their media buys, but not necessarily at the expense of the other data quality considerations. The largest dataset may not always be the best. The data provider may have made overly broad inferences or used overly loose statistical modeling, resulting in users being included who shouldn't have been in the dataset. It is not uncommon to hear stories of claimed audience sizes that clearly exceed the actual population that would fit the target.

This situation can be relatively benign. A dataset constructed with cookies or other device-level identifiers would have individuals who appear multiple times in the same dataset because they own multiple devices. However, it is also possible that the data provider may be making unrealistic assumptions or using an overly loose statistical modeling process. Data coverage claims that appear too good to be true should be scrutinized closely.

Advertisers should carefully consider the degree to which scale is a requirement based on use case. An upper-funnel branding campaign might knowingly trade some fidelity in exchange for scale. But if the goal is a more targeted lower-funnel campaign, then a smaller, higher-precision dataset may be more suitable.

5. Data Deployability: Can I use the data with my chosen tech partners?

Closely related to the question of coverage/scale is the ability to activate the data. When using a dataset for media targeting, it is common for multiple tech vendors to be involved. For example, if an advertiser wants to target a list of users it has acquired from a data aggregator, that list of users would need to be shared with a media partner, usually a Demand Side Platform (DSP). Sharing data between data sources and different media/technology partners (called onboarding) can come with significant data loss. Not all the users in the source data may be matchable to corresponding profiles in the media partner's database, meaning that the partner will be unable to target them. Additionally, limitations on the ability to consistently identify users can result in difficulties in matching profiles. For example, if a user clears his/her browser cookies, the user's identifier in the segment will no longer match to anything that a media partner can target.

A certain level of data loss is normal when audiences are transferred from one partner to another. However, the data partner and activation platform should set reasonable expectations to avoid disappointing results. Other characteristics that might be important to evaluate for insight into data quality include the mix of first-party and third-party data in the set, the provenance (i.e., where the data originated), and how the data is cleansed and processed.

Trust Consortium   ANA

DATA COMPLIANCE CONSIDERATIONS

Apart from how effective a dataset is at driving value and performance for the advertiser, it is equally important that the dataset be compliant with all relevant/local laws and best practices for gathering user consent, and for protecting the privacy of users. Rules around data privacy are currently in flux. Both the EU's General Data Protection Regulation (GDPR) and the California Consumer Privacy Act (CCPA) represent evolving legal guidelines. Enforcement can vary significantly by location. Additionally, certain industry verticals (e.g., financial services) are subject to additional regulations governing their use of data for marketing purposes.

Providing a robust set of data compliance guidelines is not within the scope of this paper, but it is strongly recommended that advertisers question their data vendors on the policies and procedures used to ensure compliance and engage their own independent data privacy experts/counsel to validate what they are told by the vendor(s). In particular, it is important to note that the concept of user consent (i.e., what defines consent, how should it be properly acquired and tracked) is an area of special focus within compliance legislation. Most data suppliers either utilize a specific consent management platform (CMP) or at a minimum have clear policies and examples that data buyers can and should review.

Trust Consortium

ANA

III. WHAT SHOULD ADVERTISERS DO?

Large advertisers make many data purchase decisions. Some are tactical, specific to a given campaign or sub-audience within a campaign. Some can be of significant consequence. For major advertisers, data budgets can easily run into the millions or tens of millions of dollars. As a result, some large advertisers are beginning to initiate RFP processes for preferred data providers. Similar to how they run an RFP to identify an agency of record, they select a set of potential data partners and have them pitch the details of their offering, including data quality and associated services. As part of this process, advertisers are requesting that the data provider participate in live validations or offer a trial run of the data at a free or discounted rate so that they can test it.

While not all advertisers have the resources to dedicate to a full RFP process for data, the process of selecting a data provider should be treated as an important vendor selection procedure, not as an afterthought. Advertisers considering a new data partner, or new offerings from their existing partner, should follow this checklist:

MEDIA DATA BUYER'S CHECKLIST

Read the Data Label.

In 2018, the IAB introduced the **Data Label**: an industry standard for third-party data providers, aimed at giving advertisers a better understanding of the contents of their audiences. The Data Label gives advertisers a standard set of information about the contents of an audience, such as size, refresh date, and lookback window (i.e., how long a record can remain in the dataset before it is expunged), and can provide helpful information for evaluating both the fidelity and freshness of the dataset. If vendors do not provide the Data Label, ask them for their equivalent and the corresponding information.

Ask additional questions of the provider/data source.

In addition to the information provided by the Data Label, advertisers should ask the data provider questions to help determine the suitability of the dataset for their purposes. These questions should address the five data quality dimensions and should also address any data compliance/privacy requirements that your experts identify. For a set of sample questions, see **Questions to Ask Data Vendors** in the appendix.

Trust Consortium

ANA

III. WHAT SHOULD ADVERTISERS DO?

When asking questions of a vendor, advertisers should keep a few things in mind:

- Before asking, you should determine which of the data quality dimensions are most important for your use case. Often, there are tradeoffs that need to be made between the different factors (e.g., coverage may come at the expense of precision, and vice versa). Knowing what you're looking for will help you evaluate whether a vendor's responses are enough.

- It is natural for vendors to want to withhold some information, including proprietary sources and methods on how they compete. However, this should not come at the expense of advertisers' ability to understand what they are buying. The vendor should be able to answer questions such as whether the source was online or offline and whether lookalike modeling was applied without revealing any competitively sensitive information. Don't be afraid to ask probing questions, and be skeptical of a vendor that appears overly evasive. Remember, a vendor that is selling a truly differentiated and higher-quality dataset should want to talk about it.

- If you do not understand a vendor's response, ask for clarification. Jargon is common in the industry, and many advertisers may find themselves confronted with unfamiliar terms and concepts. These terms and concepts all have straightforward explanations. Don't be afraid to ask for them.

CONSIDER ENGAGING A THIRD PARTY

Independent data-quality verification services are a relatively nascent industry but have the potential to significantly improve an advertiser's ability to evaluate data quality. Vendors have different approaches to data quality verification. Interested advertisers should discuss methodology with prospective vendors and select the one that best suits their purpose. The most common approaches include:

- Survey-based: The vendor can survey a sample set of audience members and measure the degree to which the audience matched the intention and attributes the advertiser was looking for.

- Audit-based: Other third-party vendors provide data certification solutions, offering an independent, comprehensive evaluation of the dataset and data provider that is akin to a financial audit and covering all aspects of data quality and compliance.

For a sample of companies active in data quality, see **Data Resources**.

Trust Consortium

ANA

QUESTIONS TO ASK DATA VENDORS

Data Accuracy

- What is the ultimate source of the data (is it from aggregated web visits, public records, etc.)?
- What is the unique lens this data provider gives our campaign?
- What are the behavioral qualifiers for the target? Is it purchase-based behavior? If so, what is the data source (Oracle, Nielsen Catalina Solutions, etc.) and which channels are the data from (e.g., retail)?

Data Precision

- How much of the data set is actual owned/resolved/deterministic data vs. modeled?
- If the data is modeled, what procedures were used to validate the model's accuracy?

Data Recency

- What are the procedures for removing outdated information from the data? How often is this performed?
- What is the age of the data? How often is it refreshed, when was it last refreshed, and how often is it updated?

Data Coverage

- How many users/households/devices are included in the segment?
- How many individuals in the actual population meet the criteria (e.g., how many individuals in the U.S. are female)?
- What is the reason for discrepancy?
  - » If the number of identifiers in the segment is smaller than the actual population, who is being excluded and how much of the population are you losing?
  - » If the number of identifiers exceeds the population size, why? Does each individual use multiple devices? Has the data vendor used lookalike modeling to expand the size of the audience?

Trust Consortium   ANA

RESOURCES AND APPENDIX

Data Deployability

- For large deals, ask the provider to run tests to demonstrate match rates with your activation platform. For smaller deals, ask the provider to provide client references who can attest to match rates.

- How does my first-party data match with the agency's data source recommendations? What is the deduplication process? How is personally identifiable information (PII) stripped so it is usable and not still PII?

- How are audiences used across media platforms? Are the same audience data sets used consistently across multiple platforms/publishers? How often is the data refreshed, cleaned, and/or updated? Is this data typically device-agnostic?

- What form is the data to be shared (e.g., third-party cookies, mobile user IDs)?

- Can this data be used to reach consumers across desktop and mobile devices and on most internet browsers? Mobile apps?

- Is there a deliverability assurance/guarantee with buying the dataset?

Data Compliance

- What is the average quality/brand standard this data provider or audience provides?

- Is the data CCPA- and GDPR-compliant?

- Is the data provider a member of the **NAI** and/or **DAA**? Does it certify adherence with the **NAI Code of Conduct**?

- Was the data collected passively (websites visited, supermarket scanner data) or actively (consumer provides data and opt-in consent for use in exchange for coupons, discount loyalty card, etc.)?

- What permissions were granted for contacting those individuals? (i.e., audit trail for permission)

Trust Consortium ANA

AGENCY VS. ADVERTISER RESPONSIBILITIES

Doing all of this requires work. A basic but often overlooked question in this process is: Whose job is it to vet data quality? Currently, there is no industry standard. When polled, approximately 50 percent of media decision makers said that the agency was responsible, while 33 percent said it was the marketer's responsibility, and 17 percent either had no procedure for vetting accuracy or was not sure[4].

Agencies certainly have an important role to play in vetting data. However, it is clear that advertisers need to play an active role in the process, and of course, advertisers which have brought all or part of their media operation in-house will need to shoulder the responsibility themselves. Additionally, there are special considerations that need to be accounted for when the agency is also acting as a data vendor and aggregating or selling white-labeled or agency-owned datasets. For additional details on this scenario, see the following section, "Recommendations for Agency-Owned or White-Labeled Datasets."

While agencies can be responsible for doing much of the heavy lifting, clients can and should assign resources to oversee the process, similar to how they would oversee the process of creating the media plan. Advertisers are used to providing oversight for their agencies in the development of a media plan (i.e., having checkpoints throughout the strategy and planning process, and expecting to provide review and feedback on deliverables at those checkpoints). Advertisers should apply the same rigor that they use in the process of building the media plan to the process of planning which data vendors and which audience segments to use.

Of course, oversight requires a commitment of resources. The exact structure and level of commitment will depend on the advertiser's specific organization. Some common approaches are:

- Assigning responsibility to a member of the media team
- Creating a dedicated role (or team, if there is enough work to require it) that holds responsibility for evaluating data and operates separately from (but works closely with) the media team

There is no single right answer. Advertisers will need to evaluate their options and pick the right one for themselves, based on their current team structure and bandwidth.

---

[4]Segment & Campaign Validation Score Study, Advertiser Perceptions, April 2019, proprietary research on behalf of Lucid

Trust Consortium

ANA

RECOMMENDATIONS FOR AGENCY-OWNED OR WHITE-LABELED DATASETS

Frequently, advertisers may find that their agency is in fact the one selling data. In the last few years, many agency holding companies have made multi-billion-dollar acquisitions to build significant data capabilities, including proprietary data sets for media buying as part of their overall offering. For example, Dentsu Aegis now owns Merkle, IPG owns Acxiom, Publicis owns Epsilon, and WPP owns KBM/iBehavior and has an ownership stake in Kantar.

Agency ownership of data assets can come with advantages for advertisers. Integrated data sources can streamline the media buying process for the advertiser and the agency can tailor the offering to the marketer's needs better than some other third-party data sources. However, these advantages should not come at the cost of transparency. When your data is coming from an agency, you should ask all the same questions regarding data quality that you would of another vendor, plus:

- What is the ultimate source of the data?
- What data is proprietary to the agency and how was it obtained, cleaned, and refreshed?
- What is the relationship and preferred cost of the data vendor with the agency?
- How am I being charged for the data? Is there a separate line item, or is it baked into my media cost?
  - » How does the fee compare to other vendors of similar data if I were to purchase the data directly from its source?
- Can I use the data with my choice of buying platforms or am I limited to using agency tech?

Trust Consortium

ANA

CONTRACTUAL CONSIDERATIONS

Obtaining data from a third party generally comes with a license and use restrictions. Most data providers have their own set of terms they use when licensing their data. Understanding the limitations of your right to use the data is equally as important as evaluating the data you are licensing. Once you have decided that you indeed want to license certain data from a third party, consider the following additional questions when negotiating your data license agreement:

- What is the term of the license?

- What is the manner and frequency of the data that you will be provided?

- Are there storage and location restrictions to the extent that such data is given directly to you?

- Is there a limit on the number of users who can access and use the data? Are your affiliates and customers allowed to access and use the data? Can you share the data with third parties (i.e., your agency or a DSP) for use on your behalf?

- Is the data being licensed for a specific purpose (e.g., attribution, targeting)? Or can it be used for any purpose?

- What warranties/representations or disclaimer of warranties/representations are present in the agreement from the provider? Are there disclaimers as to accuracy and completeness? What about compliance with applicable data/privacy laws?

- Does the agreement address derived or resultant data (i.e., any new data that is created from the licensed data)? Who owns such data? Are there limitations on the use of such data?

Trust Consortium

ANA

GLOSSARY OF KEY TERMS

First-Party Data
First-party data refers to any data that is collected by the advertiser from its own direct interactions with the consumer. For websites, the "first party" is the party which had the domain in the address bar, and a third party is anyone else. Often, this data is collected via web interactions (i.e., retargeting audiences built from visitors to an advertiser's website). However, it can also include other interactions, such as offline sales data or CRM records. First-party data is an increasingly important focus for advertisers as it can be one of the richest sources of data for their customer base, and they can exercise complete control over its collection, analysis, and activation.

Second-Party Data
Second-party data is data that comes from another party which collects that data directly from its own audience/customer base. This is different from third-party data in that the advertiser is acquiring the data directly from its source rather than through a vendor which has aggregated it from multiple different sources or which is acting as an intermediary. A common example of second-party data is when a publisher includes data/audience segments derived from its own data as part of an ad buy.

Third-Party Data
Third-party data is data that comes from another party, most typically through a data vendor which has itself acquired the data from its original source. Third-party data has the advantage of giving the advertiser insight into users/behaviors that it would not be able to identify from its own data. However, because a third party is sourcing the data, the advertiser must carefully evaluate the vendor and its data to make sure the data is suitable for its purpose.

Difference Between Individual, Household, and Cookie Data
Audience data comes in the form of lists of users who fall into a specific category with associated identifiers. It is important to note that there are several different types of identifiers. Most commonly, the identifier used is a cookie. Cookies are unique identifiers that identify a given browser/device. However, sometimes the data provider may use a different identifier. It may provide a "person" ID that attempts to unify all the devices/browsers a person may be reachable on or it may provide a household ID that groups multiple individuals as part of the same household unit.

Trust Consortium   ANA

DATA RESOURCES

The following data resources are provided for reference purposes only. Their suitability for any project should be determined independently. The Trust Consortium has not vetted or verified these companies.

Data Vendors/Brokers – Demographic/Behavioral
- Acxiom (IPG)
- Epsilon (Publicis)
- Experian
- Merkle/M1 (Dentsu/Aegis)
- Nielsen
- Oracle
- Transunion

Data Vendors/Brokers – Identity
- LiveRamp
- Neustar
- Signal
- TapAd
- Throtl

Data Vendors/Brokers – Location
- Factual
- Foursquare/Placed
- Groundtruth
- PlaceIQ
- Safegraph

Inventory Quality/Verification
- DoubleVerify
- Integral Ad Science (IAS)
- Moat

Data Quality Validation/Audit
- Lucid
- Neutronian
- Truthset

Trust Consortium ANA

The ANA
Trust Consortium
Restoring Trust through Transparency, Integrity, and Growth

DATA SOURCES
FOR MEDIA:
A BUYER'S GUIDE