

Untangling Terminology Within Data, Identity & Privacy

Why do we need to simplify common 'data, identity & privacy' terminology:

Due to the many challenges that the advertising industry is facing today – measurement for modern times, the deprecation of third-party cookies, increased desire for consumer privacy - data, in all its forms, has become one of the most heavily discussed topics in marketing.

Therefore, as marketers, it's important to understand the related verbiage and terminology being used in the industry as you navigate through this complex ecosystem that includes a multitude of data sources, targeting approaches, platforms, solutions and regulations.

Below, we seek to simplify the frequently used terms and provide a comprehensive glossary:

Data Sources

- **Automatic Content Recognition (ACR):** *Captures and identifies content (both ads and programming) across internet connected devices via audio or video fingerprints.*
- **Behavioral Data:** *Data generated from an individual's engagement with a company's owned media or other marketing channels (e.g. web pages/mobile apps, CRM) - or recorded as a response to that engagement—such as page views, email list opt-ins, etc.*
- **Cookies:** *Unique identifiers that can assign a given Internet browser or device to an individual which allows a website to recognize a specific user and their shopping behaviors as well as remember information that the user may have previously entered.*
- **Psychographic Data:** *Information about a consumers' values, attitudes, interests and personality traits, collected by analyzing online behavior (including viewing, website history and social media activity) as well as conducting surveys or traditional focus groups.*
 - **Related Term:** *Attitudinal data*
- **Set-Top Box Data:** *Viewership data that is collected directly from the set-top-box in wired cable homes via MVPDs.*
- **Transactional Data:** *Purchase history associated with audiences/segments, it is often provided in aggregate and used to determine interests and needs in support of target segmentation.*

What are Census- and Panel-based approaches?

Census-based Approach

Digital devices and some cable or satellite set-top boxes report back viewing behavior for every single ad opportunity in near real time. As a result, everyone on the platform is included and more granular targets (e.g., audiences or geographies) will be included. However, because the data is collected from a device, rather than a panelist, it can be challenging to gather information about viewers.

Panel-based Approach

Measurement providers create a representative sample based on the total population of households, or households with access to platforms (Internet households, TV viewing households or CTV households). Participants in the panel report back on their individual viewing or engagement, which is then extrapolated out to the broader population.

- **Related Term:** Sample-based approach

Targeting

- **Behavioral Targeting:** Uses data from visitor browsing habits such as search terms, sites visited or purchases to display relevant ads and offers.
- **Cohort Targeting:** Clustering large groups of people with similar interests. This approach effectively hides individuals “in the crowd” and uses on-device processing to keep a person’s web history private on the browser.
 - **Related Term:** Federated Learning of Cohorts (FLoC)
- **Contextual Targeting:** Places ads based on content on screen or on the webpage. Advertisers can use contextual advertising to target consumers based on the environment of content they are in, for example, this could be ads for running shoes on a news article about running, or it could be ads for laptops on a tech ecommerce site.
- **Demographic Targeting:** Demographic categories like age and gender provide marketers categories that directly align to measurements in the US census.
- **Device-Level Targeting:** Granular targeting that can identify a specific device and deliver tailored ads based on user behaviors.
 - **Related Term:** Identifier for Advertising (IDFA)
- **Geographical-Targeting:** Based on consumers’ real-time location and enables delivery of advertisements based on where they are, such as a promotion for a local car dealership

- **Index-Based Targeting:** *Allows marketers to target based on indexed viewership data against traits like income, hobbies or food preferences. Programs with the right audience composition can then be targeted based on the index of a given characteristic.*
- **List Targeting:** *Shows ads to a list of anonymized users that share a common characteristic.*

Platforms & Solutions

- **Authenticated Audience:** *A group of audiences that has gone through a validation process. For example, this might be a user sign-in via a website or streaming service. For linear TV, this must be supplied by the MVPD that has the identifying information.*
- **Data Management Platform (DMP):** *A platform that unifies and centralizes collecting, organizing, and activating large sets of data from disparate sources. Any audience built within the DMP can be defined and analyzed using audience profile reporting.*
- **Demand-Side Platform (DSP):** *A technology platform that provides centralized & aggregated media buying from multiple sources including ad exchanges, ad networks and sell-side platforms, which often leverages the real-time bidding capabilities of these sources (e.g., Amobee, VideoAmp).*
 - **Related Term:** *Buy-Side Platform*
- **Open AP:** *An open platform unifying digital and linear television for cross-publisher audience targeting and independent posting. Data agnostic, with relevant data sets used in audience buying to support data-driven marketing through transparency.*
- **Open ID:** *a unique identifier that enables resolution of both linear and digital viewership to a single identity framework.*
- **Supply-Side Platform (SSP):** *A technology platform that provides outsourced media selling and ad network management services for publishers. The business model resembles that of an ad network in that it aggregates ad inventory, however they serve publishers exclusively and do not provide services for advertisers (e.g., FreeWheel, SpotX).*
 - **Related Term:** *Sell-Side Platform*
- **Unified ID 2.0 (UID 2.0):** *An alternative approach to “cookies” that aims to solve for identity across the open internet, increasing privacy while supporting targeted advertising. UID 2.0 utilizes individuals’ anonymized email addresses; this is gathered by a user logging into an app or website.*

Identity

- **Cross-Device Identity Graph:** *A database of devices that are deterministically or probabilistically linked based on available identifiers. Consent to use device information is provided through the website, app or manufacturer, and licensees are assigned specific rights by the company assembling the cross-device graph.*
- **Data Consent:** *The ability of a data subject to decline or consent to the collection and processing of their personal data. Consent can be explicit, such as opting-in via a form, or implied, such as agreeing to an End-User License Agreement, or not opting out.*
- **Deterministic ID:** *Uses personally identifiable information (PII) like email addresses, names and phone numbers to associate a specific device with a user and attribute both digital and offline interactions to them.*
 - **Related Term:** *Deterministic Matching*
- **First-Party Identity Graph:** *A database of profiles consisting of deterministic, first-party identifiers and attributes such as email addresses, physical addresses, mobile numbers and usernames. The data collected is consented and aligned with both browser restrictions and regulatory changes, and first-party identity graphs provide marketers with the opportunity for personalization within their own domain.*
- **Identity:** *The effort to recognize and understand individual audience members (including customers, prospects and other visitors) across channels and devices such that brands can interact with these individuals in ways that are relevant, meaningful and supportive of business objectives*
- **Identity (ID) Graph:** *A database that can establish a profile on an individual consumer by creating links between identifiers associated with them - such as demographic, geographic, behavioral or purchase data - to deliver a more tailored ad experience.*
- **Identity Solutions:** *The coordinated activation of platforms, data and supporting services (provided by third parties and sourced from marketers) that support persistent recognition of audiences / unique individuals across devices and other touchpoints*
 - **Related Term:** *Identity Resolution*
- **Internet Protocol (IP) Address:** *A numerical label assigned to each device connected to a computer network that uses the Internet Protocol for communication. An IP address serves two main functions: host or network interface identification and location addressing.*
- **Personally Identifiable Information (PII):** *Any data that could potentially identify a specific individual. Information that can be used to distinguish one person from another, and can be used for deanonymizing previously anonymous data, can be considered PII, such as name,*

email address, physical address, date of birth, etc.

- **Private Data Clean Room:** *A secure and privacy-compliant data environment that facilitates individual-level insights, measurement and targeting at an enterprise level across all media, marketing / CRM, sales and services.*
- **Probabilistic ID:** *Uses predictive algorithms based on anonymized data points like IP address, device type, browser or OS or locational data and behavioral patterns like timing and patterns of web browsing to identify a customer via online touchpoints.*
 - **Related Term:** *Probabilistic Matching*
- **Third-Party Identity Graph:** *A database of profiles based on third-party sourced identifiers and attributes. Data must be permissioned and auditable for third-party usage, and marketers licensing this type of ID graph may have less control over the data.*
- **Virtual ID (VID):** *An identity graph approach utilized to assign record level data instances to demographic or behavioral groups.*
- **Walled Garden:** *A closed ecosystem in which all the operations, technology, information, and user data are controlled by the ecosystem operator. In the advertising technology space, companies like Facebook and Google own each step of their advertising process and have come under scrutiny for a lack of transparency.*
- **Walled Garden Clean Room:** *A secure and privacy-compliant data environment that facilitates individual-level insights, measurement and targeting only for audiences targeted across walled garden media ecosystems.*

What is first-, second- and third-party data?

First-Party Data



Data that a brand or media owner has collected directly with permission from the consumer.

Second-Party Data



Data that comes from another company's own first party data but with a defined set of rights usage and permissions.

Third-Party Data



Data that is purchased or obtained from a variety of sources by entities that do not have a direct relationship with the end user.

Related Term: Syndicated Data

Data Regulations

- **California Consumer Privacy Act (CCPA):** *A California law that provides residents to know what data is being collected about them, where/if it is sold, the right of refusal, personal data access, data deletion and anti-discrimination. Any business that conducts business with California residents must adhere to this law.*
 - **Related Term:** *Virginia Consumer Data Protection Act (CDPA)*
- **California Privacy Rights Act of 2020 (CPRA):** *A California law that allows consumers to prevent businesses from sharing their personal data, correct inaccurate personal data, and limit businesses' usage of 'sensitive personal information,' which includes precise geolocation, race, ethnicity, religion, genetic data, private communications, sexual orientation, and specified health information. It also removes the set time period in which businesses can correct violations without penalty and prohibits businesses from holding onto personal data for longer than necessary.*
 - **Related Terms:** *California Proposition 24*
- **General Data Protection Regulation (GDPR):** *A regulation on data protection and privacy in the European Union and the transfer of personal data outside of the EU. The GDPR's primary aim is to enhance individuals' control and rights over their personal data and to simplify the regulatory environment for international business.*

What are the 3 things to know about data, identity & privacy:

1. *Just a few years ago, only **15%** of marketers said they could identify their audiences accurately and consistently, while the majority (69%) reported only partial success with their identity efforts.¹*
2. ***49%** of marketers view 'third-party cookie deprecation' as the top *digital media challenge* in 2021 as its phase-out across browsers and mobile devices is due to be complete in 2023, meaning marketers have a need to develop more sophisticated marketing technologies such as identity solutions and tech stacks.²*
3. *With the need to understand their audiences more thoroughly within an environment of elevated consumer privacy coupled with the deprecation of cookies, U.S. marketers are expected to invest an estimated **\$2.6 billion** in services and solutions focused exclusively on identity in 2022, almost 3x the amount invested in 2018 (~\$900 million).³*

Industry Perspectives:

1. *“Creating a responsible digital ecosystem continues to be a priority for Unilever and adding an ethical overlay for use of data in media is a key next step. Consumers are growing increasingly distrustful of advertising and lack clarity in how their data is being collected and used in profiling tactics... We must all hold ourselves, our partners and agencies to the highest standards, truly putting consumers first.” - Jennifer Gardner, Senior Director of Media, Unilever* ⁴
2. *“In the long term, (desire for increased privacy) offers a unique opportunity to (re)build trust between brands and consumers around the data issue. Success will hinge on increasing efforts to educate audiences (consumers) and defining a right value exchange that works for all.” – Dentsu ‘The Cookieless World: A Guide for the New Era of Digital Marketing’* ⁵
3. *“Leaders across the media and technology landscape have a unique opportunity to reimagine the role of data and identity in day-to-day business operations, and to create a simple, interoperable future that yields impact for partners and improves the consumer experience, all while upholding the privacy standards and trust the world demands.” – Chris Crayner, Chief Digital Officer, Universal Parks and Resorts* ⁶

Future Outlook:

Marketing is experiencing enormous change and the next era of media and advertising will be defined by data and identity as consumers, who have become more distrusting of intrusive marketing tactics, demand more control over their own information and, in response, marketers search for solutions that strengthens their relationship with the consumer while respecting their privacy.

By truly putting the consumer first, both marketers and media companies will need to evolve, adapt, build and optimize privacy-compliant, data-driven approaches to drive future business growth. Staying informed on the latest terminology can help marketers better understand this ever-evolving industry and make more informed decisions.

Want to learn more?

Click on the images below for the content.



About VAB

VAB is an insights-driven organization that inspires marketers to reimagine their media strategies resulting in fully informed decisions.

Drawing on our marketing expertise, we **simplify** the complexities in our industry and **discover** new insights that **transform** the way marketers look at their media strategy.

We are committed to your business growth and proud to offer VAB members, brand marketers and agencies **complimentary access** to our continuously-growing Insights library. **Get immediate access at theVAB.com.**

Sources

1. Merkle / Merkury, in association with Winterberry Research, 'Identity' webinar presentation, 2020.
2. Integral Ad Science, The 2021 Industry Pulse Report, also available through VAB's [Data-Driven Opportunities in a Cookieless World](#), 8/11/2021.
3. Merkle / Merkury, in association with Winterberry Research, 'Identity' webinar presentation, 2020.
4. AdWeek, [New Data Ethics Compass Looks Beyond What's Simply Legal](#), 2/1/2021.
5. Dentsu, [The Cookieless World: A Guide for the New Era of Digital Marketing](#), 7/8/2021.
6. NBCUniversal, [John Lee Joins NBCUniversal as Chief Data Officer, Spearheading New Enterprise-wide Data Unit](#), 7/22/2021.