

ID-Less Solutions Guidance

Version 1.0

Released November 20, 2024

Please email support@iabtechlab.com for questions and feedback.

This document is available online at <https://iabtechlab.com/idlessolutions>

About this document

With the many challenges in advertising that come with the loss of identity signals there is a need to solve core advertising use-cases when identity signals are not available. Contexts in which identity signals are not available are called “ID-Less” environments.

The target audience are entities wishing to learn about advertising privacy and developments in the ad ecosystem, regulators interested in the direction of advertising technologies and privacy enhancing technologies, and product managers who want more exposure to ID-Less solutions and how they compare to ID-Based solutions. Additionally, sellers and buyers who are concerned about changes in ID and cookie policies and wish to find innovative solutions to address those changes.

This document seeks to explain what ID-Less solutions are, how they differ from traditional ID-Based solutions, their advantages and disadvantages, describe scale and availability, and also describe how these solutions can be used to solve common advertising use cases.

This document is developed by the IAB Tech Lab [Addressability & PETs Working Group](#).

Note: The use of words or phrases ‘Privacy’, ‘Private’, ‘Security’, ‘Control’, ‘Processing’, ‘Personal Data’, ‘PII’ in this document is generic and does not refer to definitions in any specific regulation e.g. GDPR or CCPA.

License

ID-Less Solutions Guidance document is licensed under a [Creative Commons Attribution 3.0 License](#). To view a copy of this license, visit creativecommons.org/licenses/by/3.0/ or write to Creative Commons, 171 Second Street, Suite 300, San Francisco, CA 94105, USA.

Significant Contributors

Chris Watts, *NumberEight*; Arthur Coleman, *ThinkMedium*; Brian May, *Individual*; Albert Thompson, *ID Privacy*; Dan Pike, *Covatic*; Airey Baringer, *TripleLift*; Brooks Dobbs, *The Trade Desk*; Jonathan Caines, *Anonymised*

IAB Tech Lab Leads

Miguel Morales, Director Addressability & Privacy Enhancing Technologies (PETs)

Shailley Singh, EVP Product & COO, IAB Tech Lab

About IAB Tech Lab

The IAB Technology Laboratory is a nonprofit research and development consortium charged with producing and helping companies implement global industry technical standards and solutions. The goal of the Tech Lab is to reduce friction associated with the digital advertising and marketing supply chain while contributing to the safe growth of an industry.

The IAB Tech Lab spearheads the development of technical standards, creates and maintains a code library to assist in rapid, cost-effective implementation of IAB standards, and establishes a test platform for companies to evaluate the compatibility of their technology solutions with IAB standards, which for 18 years have been the foundation for interoperability and profitable growth in the digital advertising supply chain. Further details about the IAB Technology Lab can be found at <https://iabtechlab.com>.

Disclaimer

THE STANDARDS, THE SPECIFICATIONS, THE MEASUREMENT GUIDELINES, AND ANY OTHER MATERIALS OR SERVICES PROVIDED TO OR USED BY YOU HEREUNDER (THE “PRODUCTS AND SERVICES”) ARE PROVIDED “AS IS” AND “AS AVAILABLE,” AND IAB TECHNOLOGY LABORATORY, INC. (“TECH LAB”) MAKES NO WARRANTY WITH RESPECT TO THE SAME AND HEREBY DISCLAIMS ANY AND ALL EXPRESS, IMPLIED, OR STATUTORY WARRANTIES, INCLUDING, WITHOUT LIMITATION, ANY WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE, AVAILABILITY, ERROR-FREE OR UNINTERRUPTED OPERATION, AND ANY WARRANTIES ARISING FROM A COURSE OF DEALING, COURSE OF PERFORMANCE, OR USAGE OF TRADE. TO THE EXTENT THAT TECH LAB MAY NOT AS A MATTER OF APPLICABLE LAW DISCLAIM ANY IMPLIED WARRANTY, THE SCOPE AND DURATION OF SUCH WARRANTY WILL BE THE MINIMUM PERMITTED UNDER SUCH LAW. THE PRODUCTS AND SERVICES DO NOT

CONSTITUTE BUSINESS OR LEGAL ADVICE. TECH LAB DOES NOT WARRANT THAT THE PRODUCTS AND SERVICES PROVIDED TO OR USED BY YOU HEREUNDER SHALL CAUSE YOU AND/OR YOUR PRODUCTS OR SERVICES TO BE IN COMPLIANCE WITH ANY APPLICABLE LAWS, REGULATIONS, OR SELF-REGULATORY FRAMEWORKS, AND YOU ARE SOLELY RESPONSIBLE FOR COMPLIANCE WITH THE SAME, INCLUDING, BUT NOT LIMITED TO, DATA PROTECTION LAWS, SUCH AS THE PERSONAL INFORMATION PROTECTION AND ELECTRONIC DOCUMENTS ACT (CANADA), THE DATA PROTECTION DIRECTIVE (EU), THE E-PRIVACY DIRECTIVE (EU), THE GENERAL DATA PROTECTION REGULATION (EU), AND THE E-PRIVACY REGULATION (EU) AS AND WHEN THEY BECOME EFFECTIVE.

Glossary

Term	Description
<i>Addressability</i>	Ability or extent of capability to uniquely identify an individual or a device between data sets of two or more parties in a given context e.g. targeting individuals with advertisements
<i>Anonymous</i>	Where the identity of a natural person is unknown, but some attributes about the user (such as saved preferences) may be known.
<i>Attribution</i>	The process of identifying and assigning credit to the specific exposure that contributed to a conversion.
<i>Cohort</i>	A group of Users. Also known as "interest groups".
<i>Contextual Advertising</i>	Ad placement based on page content rather than user identity
<i>Creative</i>	The content of an ad.
<i>Demand Side Platform (DSP)</i>	Entity servicing advertisers which bids on advertising opportunities presented by an SSP or (sometimes) a header bidding solution.
<i>Deterministic Identifiers</i>	Unique, fact-based identifiers like email or phone numbers
<i>Differential Privacy</i>	A method to add noise to data to preserve individual anonymity
<i>DSP</i>	Demand-Side Platform: a technology platform that enables advertisers to purchase and manage digital ad inventory through automated, real-time bidding processes.
<i>Frequency Capping</i>	Limiting the number of times an ad is shown to a single user
<i>Identifier</i>	A mechanism to assign a unique value to a device or a browser or a user for associating with a unique person or to identify a group of devices and browsers or users and assign a unique value for associating with a unique

Term	Description
	household
<i>K-Anonymity</i>	Ensuring that any given data point is indistinguishable from at least 'k' others
<i>Key</i>	Identifiers which do not refer to a specific household, device or user.
<i>Multi-touch attribution (MTA)</i>	MTA is a method of attributing credit to different touch points in a customer's interaction (for e.g different media channels where the customer viewed or engaged with an advertisement) with the advertiser that resulted in a customer action (for e.g. purchase of goods or services).
<i>Noise</i>	Random data that is added to the output of a query or computation to obscure the influence of individual data entries.
<i>OS</i>	Operating System: the software running on the User's device.
<i>Personally Identifying Information (PII)</i>	Any data that can, independently or in combination, be used to identify a person, either directly or indirectly.
<i>PETs</i>	Privacy enhancing technologies (PETs) are technology solutions that use one or more of the privacy technologies (differential privacy, secure multi party compute and on device learning) to accomplish complex data processing functions in digital advertising without revealing the individual, household or device level personal information to parties that do not already have them
<i>Privacy Sandbox</i>	A Google initiative proposing privacy-preserving advertising solutions
<i>Private Aggregation API</i>	Tool for generating aggregate reports while preserving privacy
<i>Publisher</i>	An entity that controls a website, app, or service designed for Users.
<i>ROAS</i>	Return on Ad Spend (ROAS) is a measure of total returns from an ad campaign arrived at by calculating the total revenue earned and direct expenses. It does not include other expenses and does not tell if a paid

Term	Description
	campaign is profitable for the advertiser.
<i>ROI</i>	Return on Investment (ROI) is a measure of overall return on investment arrived at by calculating total profit and all expenses- both direct spend on an ad campaign as well as other expenses. ROI determines how profitable an ad campaign is.
<i>SSP</i>	Supply-Side Platform: a technology platform that enables Publishers to manage, sell, and optimize their ad inventory through automated auctions to maximize revenue.
<i>Supply Side Platform (SSP)</i>	Entity servicing publishers, responsible for receiving ad requests from publishers or publisher header bidding systems, requesting bids from DSPs and running an auction to determine the ad to show, or respond with a bid to the header bidding system.
<i>Third-party Cookie</i>	Browser cookies used across multiple sites to track user activity
<i>Token</i>	A synonym for Key, more often used for authentication.
<i>Trust Token Framework</i>	Mechanism to verify the legitimacy of users without revealing identity
<i>User</i>	A natural person as the user of a website, app, or service.
<i>User Agent</i>	The software used by a User to access a website, app, or service, e.g. browsers such as Mozilla Firefox.

Table of Contents

- About this document..... 3**
 - License..... 3
 - Significant Contributors..... 3
 - IAB Tech Lab Leads..... 4
 - About IAB Tech Lab..... 4
 - Disclaimer..... 4
- Glossary..... 6
- Table of Contents..... 7**
- Introduction: Why ID-Less Solutions..... 9**
- What are ID-Less Solutions?..... 10**
 - What is an ID?..... 10
 - What are ID-Less Solutions?..... 12
 - What are not ID-Less Solutions?..... 13
 - Benefits and Disadvantages of ID-Less Approaches..... 14
- Overview of ID-Less Solutions..... 23**
 - Area: Attribution/Measurement..... 24
 - Challenge: Campaign Reporting..... 24
 - Propagated keywords..... 24
 - Promotional Codes..... 26
 - Aggregated Attribution Reporting..... 26
 - Probabilistic Cohorts..... 28
 - Cross-Context Deterministic Cohorts..... 29
 - Challenge: Multi-Touch Journey Mapping..... 30
 - Media Mix Modeling (MMM)..... 30
 - Brand Lift Studies..... 31
 - AdAttributionKit (Mobile)..... 31
 - Challenge: Attention..... 32
 - Rewarded Ads..... 32
 - Open Measurement SDK..... 33
 - Attention Scores..... 33
 - Area: Targeting & Prospecting..... 34
 - Challenge: Audience Prospecting..... 34
 - Contextual Data..... 34
 - Probabilistic Cohorts..... 34

- Cross-Context Deterministic Cohorts..... 35
- Seller-Defined Audiences (SDA)..... 35
- Private Marketplace (PMP Deals)..... 36
- Private Aggregation, Reach Estimation (Privacy Sandbox)..... 37
- Challenge: Audience Enrichment..... 38
 - Cohort Lookalikes..... 38
- Area: Retargeting..... 39
 - Challenge: Bring the Customer Back..... 39
 - On-Device Auctions..... 39
- Area: Frequency & Recency Capping..... 42
 - Challenge: Preventing Oversaturation..... 42
 - On-Device Frequency Capping..... 42
 - Challenge: Creative Sequencing..... 44
 - Device-Side Creative Sequencing..... 44
 - Challenge: Ad Pacing..... 44
 - Device-Side Pacing..... 44
- Area: Fraud..... 44
 - Challenge: Automated bot detection..... 45
 - Device Attestation..... 45
 - Private State Tokens..... 46
 - Statistical Determination..... 47
 - Challenge: Human bot detection..... 48
 - Statistical Determination..... 48
 - Challenge: Signal spoofing..... 48
 - Challenge: Impression spoofing..... 49
 - Impression Tokens..... 49
- Examples of How ID-Based and ID-Less Technologies Differ..... 50**
 - Retargeting..... 50
 - ID-Based Retargeting..... 50
 - ID-Less Retargeting..... 51
 - Frequency Capping..... 55
 - Fraud Detection..... 59

Introduction: Why ID-Less Solutions

The advertising industry is undergoing a profound shift in how it identifies and reaches specific audiences with relevant messages. This shift is being driven by a change in market power between consumers and marketers. On one side, there are consumers and their advocates - regulators, browser platforms, device manufacturers, and other privacy-sensitive technology companies. On the other side are companies in the ad tech value chain - advertisers, publishers, DSPs, SSPs, third-party data providers and various intermediaries. Consumers, directly and through advocates, are increasingly asserting data rights through increased control of their personally identifiable information and the ability to identify them online.

A very conspicuous example of this shift is the deprecation of third-party cookies, first in Safari and Firefox, and planned for Chrome for 2025 before being retired in favor of an as yet undefined user choice mechanism assumed to be similar to Apple's ATT. Until recently, third-party cookies have been a major mechanism used by advertisers and publishers to support browser identity use cases across the open web. The ability to identify browsers allowed companies to deliver relevant ads in the right setting, at a pace and volume that optimized return on their marketing dollars. The IAB has estimated that the loss of cookie-based identifiers and similar privacy-driven modifications to the digital ad supply chain will drive up costs to maintain campaign ROAS/CAC/CPMs from 29% to as much as 200%¹

The critical ability to consistently identify and reach specific audience members once third-party cookies are no longer viable has spurred intense innovation in technologies which will allow companies to reliably maintain relationships with consumers. These technologies, while not dependent on 3rd-party cookies, are not generally "ID-Less". Instead, they use deterministic data like email addresses, combined with probabilistic methods to create persistent identifiers. For more information about these solutions, please refer to our guidance on [ID Solutions](#).

At the same time, the privacy concerns motivating the decisions to deprecate third-party cookies have caused browser implementers like Google and Mozilla to revisit approaches that were rejected as being too difficult/expensive to implement in a cookie-driven world. These approaches don't rely on identifiers. Instead, they keep all user-related data in tightly isolated environments within a browser or secure server which also support the capabilities needed for advertising use cases. This collocation of user data with ad buying capabilities allows for privacy-preserving ad targeting without identifiers.

This document provides guidance for supporting addressability and measurement goals using ID-Less solutions and for reducing the reliance on traditional cross-context identifiers as part of a holistic approach to advertising in a post-cookie environment.

¹ <https://www.iab.com/insights/2024-state-of-data-report/>

What are ID-Less Solutions?

What is an ID?

The term identifier, or ID, in this document refers to a data value that is persistent and consistent across contexts and can be used to resolve the identity of a household, device or user. To be useful as an identifier, a value must have the following characteristics:

- Unique - it has a high probability of identifying a single entity from a group.
- Persistent - it is available across a number of transactions.
- Consistent - it reliably identifies the same entity across transactions to all parties to the transactions.

This definition excludes attributes which may be used as private first party keys that are readable solely by first-parties and their data partners with whom they have a direct relationship, so long as the attributes cannot be used by unconnected businesses to identify the same household, device, or user.

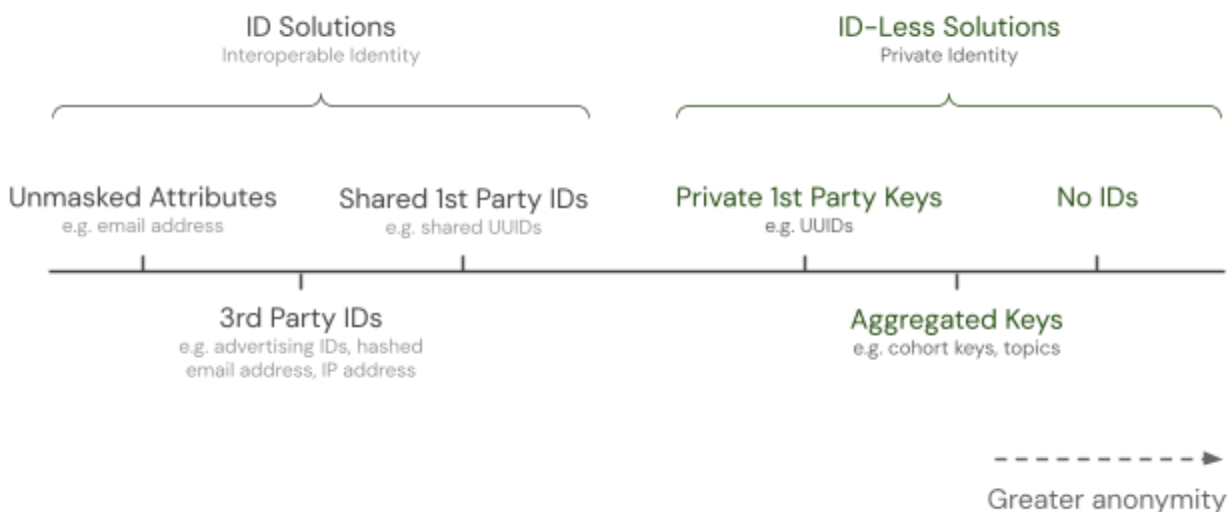


Figure 1: A Spectrum of ID Types from Least to Greatest Anonymity.

Identifiers may be either deterministic or probabilistic:

- A deterministic identifier is based on factual data and derived in a way that always returns the same value. Examples of deterministic identifiers are a person's name, date of birth, phone number or email address, as well as values generated by functions which take these values as inputs. A hashed email address is an example of the latter.
- A probabilistic identifier is derived from best-guess inferences based on sets of data which may contain values that independently do not distinguish or reidentify entities with a high level of coincidence, but in combination do. Statistical models are used to determine the likelihood that given sets of inputs identify the same entity. An example would be associating an identifier with a web site visitor based on the confidence that the combination of their IP address, user agent string and geolocation accurately distinguishes them from other visitors as indicated by historical data. The rarer the combination of these values, the higher the confidence would be that when seen they identified the same browser.

It is important to acknowledge that the terms "identifier" and "ID" are also used generically to denote a wide range of entities unrelated to advertising and addressability. In the most general terms, an identifier is simply a data value that consistently refers to other data values. To clarify: the use of "ID-Less" in this document means "without identifiers capable of distinguishing individuals, devices, or households". To avoid confusion, we will use the term "Key" with a capital K to refer to identifiers which do not refer to a specific household, device or user.

An example of a Key might be a cohort which identifies a (suitably large) group of users so as to be considered "ID-Less".

What are ID-Less Solutions?

ID-Less solutions are methods for targeting ads and measuring advertising campaign performance without revealing information that could allow someone to learn who an ad was delivered to. They take a variety of approaches from using contextual information related to the ad placement to determining general audience attributes like those provided by [Seller-Defined Audiences](#)².

ID-Less solutions will typically:

1. Use first-party user-level data to identify and share general audience categories rather than sharing per-user details.
2. Use information about the context in which ads are shown.
3. Share general and aggregated data about groups of users that cannot be used to identify a specific individual.

ID-Less solutions will not:

1. Share data that can reliably identify an individual.
2. Share data that can be linked to an individual and track their activities across contexts, such as between websites.

What are not ID-Less Solutions?

Technologies that require "match keys" are out-of-scope for this guidance. For example, the following would be classified as ID Solutions:

- Technologies that use an email address or phone number, whether in original or modified (e.g. hashed or encrypted) form, to identify the user.
- Technologies that use an IP address, alone or in combination with other values, to identify the user.

² See IAB Tech Lab's [Seller Defined Audiences Specification](#) for a broader discussion of this type of ID-Less audience based on first-part data.

- Technologies that generate a probabilistic ID by applying statistical models to constellations of device attributes such as installed RAM, storage space, screen size, and user agent data.
- Any other technology that is able to convey the identity of a specific user, device or household between multiple unconnected parties (e.g. two publishers can independently use it to identify the same user with reasonable certainty).

Note that whether or not solutions are "cookie-less" has no bearing on whether they are ID-Less. Many cookie-less solutions depend on commonly available user information that is consistent across contexts, such as email addresses or phone numbers, to identify activity related to the same user. Conversely, there are a number of cookie-based solutions that use cookies only to maintain non-identifying information between sessions such as shopping cart contents.

Furthermore, some Privacy-Enhancing Technologies (PETs) that rely on the use of match keys are out-of-scope, such as:

- Data Clean Room facilitated interactions that use a match key.
- Trusted Execution Environments where cross-party IDs are used to identify user records.

For more information on ID solutions, refer to our [ID Solutions Guidance](#).

Benefits and Challenges of ID-Less Approaches

Benefits

Some benefits of ID-Less solutions are described in Table 1 below.

Benefit	Beneficiaries	Details
Greater coverage.	Publishers	Publishers are able to associate data with more ad requests to enable better decisioning by bidders.
Simpler compliance with privacy legislation.	Publishers, Advertisers	Due to the absence of identifiers, the level of measures needed to keep data safe is reduced and proportionality of processing is easier to justify.
Personalized interactions for users without knowing their identity.	Consumers	Personalized experiences can be created based on insights gleaned from content and first-party data.
Improved consumer opinion of online advertising as being privacy-respecting while still delivering relevant content.	Publishers, Advertisers, Consumers	Anonymity becomes financially viable to more publishers and advertisers, leading to less regulatory circumvention.
The ability to complement ID-Based solutions by delivering relevant ads to unidentified, anonymous consumers based on what is learned from consumers that have an ID.	Publishers, Advertisers, Consumers	Cross-domain IDs are valuable when they are available, but coverage is low. Advertisers and publishers both benefit by leveraging data from interactions which includes identifiers to effectively target ad impressions when

Benefit	Beneficiaries	Details
		cross-domain IDs aren't available.
The ability to measure ID-Less ad campaign performance for anonymous traffic.	Publishers, Advertisers	Measurement is possible based on first-party data, such as ad engagement metrics and attention metrics as well as by modeling outcomes based on impressions that do have identifiers.
Improved ROAS for advertisers on traffic where an ID-Based solution is not available.	Advertisers	Performance is improved for impressions that supply ID-Less data in the absence of IDs.
Improved revenue for publishers as a result of delivering higher value to the advertiser.	Publishers	Publishers are able charge more for ID-Less impressions as a result of improved ad targeting capabilities that deliver better outcomes for advertisers.

Table 1 - Benefits of ID-Less Approaches

Challenges

On the other hand, ID-Less approaches also have a number of challenges (Table 2).

Challenges	Challenged	Details
Some ID-Less technologies are still in their infancy.	Publishers, Advertisers	Some of the most promising new ID-Less technologies are still at an early stage of development and not widely adopted, so their performance on a number of dimensions, including scalability, measurability and transparency, among others, is still unknown.
Use case coverage is incomplete.	Advertisers	Some use cases do not yet have known ID-Less solutions, or have unresolved limitations preventing their use. See Table 3 for examples.
For use cases which are supported, ID-Less may not be as effective as ID-Based solutions.	Advertisers	
Support for ID-Less solutions is limited by the current ecosystem.	Publishers, Advertisers	The digital advertising ecosystem has entrenched dependencies on identifiers. It will take time for a critical mass of industry participants to adapt their systems to support ID-Less alternatives.
Measurement of ID-Less solutions requires the industry to think differently.	Publishers, Advertisers	ID-Less solutions, by their nature, remove the ability to attribute conversions to specific impressions.

Challenges	Challenged	Details
		Innovation in modeled conversions, aggregated measurement, browser standards, and updated media mix modeling (MMM) capabilities are showing promise as alternatives for measuring ID-Less solutions, but metrics may be fundamentally different.
Statistical bias introduced by privacy-focused practices.	Publishers, Advertisers	As with ID-Solutions which employ Privacy Enhancing Technologies, the data reported to participants in an ad interaction may differ due to inconsistent addition of noise or other treatments employed to hide user identities. This is of particular concern for use cases that impact billing and payment. Parties will need to be aware and accommodate potential discrepancies.
Costs to implement can be substantial.	Publishers, Advertisers	Adoption of ID-Less solutions will impose significant financial burdens on publishers, advertisers, and other participants in the adtech value chain.

Table 2 - Challenges of ID-Less Approaches

Table 3 shows how the benefits and challenges of ID-Less solutions recast themselves across the major use cases advertisers and publishers depend on today in an ID-Based world. The rows indicate which use cases are well supported and which become more difficult in an ID-Less environment.

Use Case	ID-Based Solutions	ID-Less Solutions
Insights and Campaign Planning		
Universe-to-Universe Matching	✓	✗
Audience Discovery & Creation	✓	✓
Audience Identification with Historical Reach	✓	✓
Pre-campaign Insights	✓	✓
Media Mix Modeling (MMM)	✓	✓
Targeting and Activation		
Audience Activation	✓	✓
Bidstream Augmentation	✓	✓
Campaign Optimization	✓	✓
Retargeting	✓	✓
User-Agent Frequency Capping	✓	✓ ³
Global Frequency Capping	✓	✗
Fraud		

³ User-Agent frequency-capping can be achieved by the user's device to limit an ad being shown more than a set number of times within the context of a browser, website, app, or device.

Use Case	ID-Based Solutions	ID-Less Solutions
Automated Bot Detection	✓	✓
Human Bot Detection	✓	✓
Other Fraud Detection	✓	✗
Reporting and Attribution		
Aggregated Campaign Reporting	✓	✓ (Limited ⁴)
Event-Level Campaign Reporting	✓	✓ (Delayed / imprecise ⁵)
Aggregated Audience Insights & Trending	✓	✓
User-Level Insights & Journey Mapping	✓	✗
Attribution - Campaign	✓	~ ⁶
Attribution - Conversions	✓	~
Attribution - App Installs	✓	✓ ⁷

⁴ Campaign reporting is often limited in ID-Less contexts to the ad provider's own ecosystem, such as Apple's AdAttributionKit.

⁵ Event-Level campaign reporting is delayed and/or imprecise in ID-Less contexts for privacy.

⁶ Campaign and conversion attribution is under active investigation with cross-context deterministic cohorts.

⁷ App installs through AdAttributionKit and similar approaches.

Use Case	ID-Based Solutions	ID-Less Solutions
Attribution - Multi-touch (MTA)	✓	✓

Table 3 - Comparison of Use Cases Between ID-Based and ID-Less Solutions

Overview of ID-Less Solutions

The following analysis describes common adtech challenges and the extent to which a variety of available ID-Less implementations can address these challenges.

This list does not aim to be exhaustive. It is intended to illustrate various approaches to the challenges, especially as the state of ID-Less solutions is rapidly advancing. Each solution has pros, cons, and proposed improvements based on the current state of the industry, and the extent to which the solution has been adopted according to the following categories:

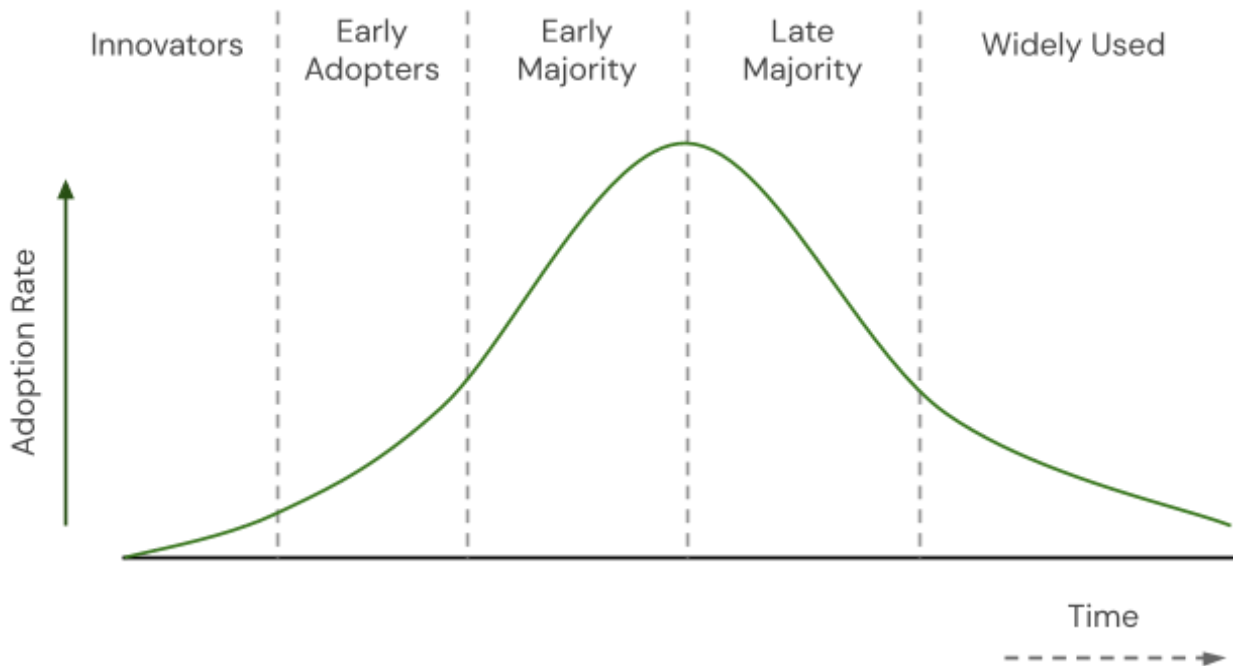


Figure 2: The Typical Technology Adoption Curve Showing the Rate of Adoption of a Technology at a Given Point in its Lifecycle.

Area: Attribution/Measurement

Challenge: Campaign Reporting

Also known as Click-Through or View-Through Impressions

“ As an advertiser, I want to know which of my customers found me using my ad campaign because it will inform how effective my ad spend is and help me to calculate ROAS.

Promotional Codes

Adoption: Widely Used

Distributing discount codes in an ad creative to attribute a sale to a particular creative. Used extensively in audio advertising, social media marketing and influencer marketing.

Example: "Use code TRAVEL1 at checkout for a 10% discount"

Pros:

- Allows campaign tracking at a coarsely granular level
- No limitations on the number of discount codes, just needs to be memorable
- Works without click-through attribution
- Works across any channel: display, audio, out-of-home, social media

Cons:

- Advertisers have a financial disincentive - running promotions increases cost of acquisition
- Only works when a purchase or conversion event is observed (unless used in combination with [Propagated Keywords](#))
- Scale of users redeeming promotional codes may be too small for effective attribution

Improvements:

- Dynamic discount codes per cohort or per creative for more granular reporting

Aggregated Attribution Reporting

Adoption: Early Majority

Browser and app providers (e.g. Apple's AdAttributionKit and Google's Attribution Reporting API) can provide aggregated reports with noise-induced data and limited reporting frequency to preserve individual user privacy while still providing some useful insights.

Other implementations provide a similar service using Multi-Party Compute (MPC), such as Interoperable Private Attribution (IPA). IPA entered the standards track at the W3C in late 2024 and will become the primary attribution reporting mechanism for Mozilla and Meta.

Pros:

- Privacy by design - limited risk of abuse
- Highly scalable
- High level of privacy
- Multi-touch reporting is possible

Cons:

- Delays in reporting are limiting
- With AdAttributionKit, app developers need to whitelist all parties that can send attribution reports to them, creating friction for implementation
 - Makes sense if there are only "ad networks" in the ecosystem, but becomes more complex when SSPs, DSPs, and other partners are considered
- Implementers have to make careful choices about using summary-level reporting (more expensive) and event-level reporting (limited frequency)
- There are a variety of solutions which do not have a common framework
- Implementers must typically use a Trusted Execution Environment (TEE)
 - TEEs have an additional latency
- The mechanics of privacy budgeting cause bias in reporting with the result that sales can be attributed to the wrong impressions
- Utility of the reports is dependent on the level of noise

Improvements:

- Currently implementers cannot host their own TEE server; only AWS and Google Cloud is available as of November 2024

Probabilistic Cohorts

Adoption: Early Adopters

Measure digital events, for example, campaign impressions or website visits based on shared behaviors or characteristics of groups of users. Includes simple cases such as measurement by country, and more advanced cases such as measurement by browsing behavior. The creation of cohorts may be implemented using on-device technology and federated learning for greater privacy.

Pros:

- Arbitrary level of conciseness depending on cohort membership rules
- Encourages use of first-party data
- Some implementations need no personal data
- Match rates between cohorts measured and cohorts targeted is typically high if the same cohort methodology is in use
- Semantic relationships between cohort behaviors and buying patterns is common, assisting the campaign planning process (e.g. early morning commuters may tend to buy more coffee)

Cons:

- Probabilistic cohorts tend to be less specific than ID-Based solutions as data is scarcer
- Solutions may still require access to user data that is subject to consent
- Attribution is only possible where both publishers and advertisers use the same technology (limited scale)
- Only provides aggregate reporting data

Improvements:

- K-anonymity or differential privacy can be used to limit re-identification risks

Cross-Context Deterministic Cohorts

Adoption: Early Adopters

By storing observed events (such as ad views, newsletter signups, and content viewed) in shared storage on the user's device, campaign outcomes and reports can be measured by grouping the events into cohorts. When integrated by both advertisers and publishers sites, it offers attribution capabilities.

Pros:

- Allows for post-view campaign measurement
- Encourages use of first-party data
- Deterministic - guarantees that all consented users will be present in the relevant cohort
- Operates similarly to third-party cookies, but with additional privacy guarantees

Cons:

- Does not work across multiple devices or multiple channel technology (limited scale)
- Consent is required in many jurisdictions in order to store data
- Attribution is only possible where both publishers and advertisers use the same technology (limited scale)
- Only provides aggregate reporting data

Improvements:

- K-anonymity or differential privacy can be used to limit re-identification risks
- Modeling the pathways that cohort users take to reach an advertiser can be recorded to replicate some elements of multi-touch journeys

Challenge: Multi-Touch Journey Mapping

“ As an advertiser, I want to know how many times a user saw my advert before making a purchase (especially if there was no click-through event) because it will inform how effective my ad spend is and help me to calculate ROAS.

Propagated keywords

Adoption: Widely Used

Metadata including the source, campaign, medium, and content, for example, are sent as parameters to the advertiser's website when the ad is clicked or interacted with. This data is typically presented as URL parameters and is commonly implemented as UTM tracking.

Example: https://travel.com/book-holiday?utm_source=tiktok&utm_term=travel_enthusiasts

Pros:

- Any information known by a publisher can be sent to an advertiser, within reason
- Suitable for attribution at a cohort-level
- When combined with first-party cookies, the advertiser is able to save the information contained in the parameters in the user's browser session to track multi-touch journeys

Cons:

- Privacy depends on the mitigations taken by publishers and advertisers and may enable re-identification attacks if poorly implemented, leading to regulatory risk
- Openness of networks - bad actors can build databases of users over time if the data can be attributed to an identifier
- The parameters may be subject to truncating by browser privacy controls and length limitations

Improvements:

- Differential privacy, when enforced, can mitigate some reidentification risks

Media Mix Modeling (MMM)

Adoption: Widely Used (*in channels such as Out-of-Home and Linear TV*)

Statistical models are used to attribute upticks in purchases and other metrics with campaign activity, typically using A/B tests to validate hypotheses and causation.

Pros:

- Privacy by design - limited risk of abuse
- Links directly to ROAS or any business-level metrics
- Can attribute across any channel including Out-of-Home

Cons:

- Perception that modeling is imprecise
- Can only reliably attribute one advertising channel at a time, although workarounds exist to attribute more channels concurrently
- Measurement is post-campaign and not real-time; typically delayed by 6-8 weeks
- It's time consuming to do the data processing

Improvements:

- Existing MMM services are costly and time consuming. Better technology decreases the time taken per report, which in turn allows more frequent A/B tests
- By carefully choosing the frequency, duration, and geographic location of campaigns, multiple campaigns can be measured simultaneously

Brand Lift Studies

Adoption: Widely Used

Typically involving surveys, advertising campaign goals are measured against "exposed" and "control" groups to form scientific conclusions about the efficacy of the campaign. For example, two groups of people – one who saw the ad campaign and one who did not – may be asked if they recall the product advertised or have made a purchase some time after the campaign has finished.

Pros:

- Uses a more scientific approach to gauge whether campaigns have had a significant impact
- Works for a variety of media channels
- Can measure business-level goals including brand awareness and purchase intent

Cons:

- It may be difficult to get users to respond to the survey at a large enough scale
- They are post-performance and not real-time
- It is time consuming to do the data processing
- Users who have been exposed to the campaign need to be remembered in some way (such as using first-party cookies)

AdAttributionKit (Mobile)

Adoption: Late Majority

AdAttributionKit (formerly SKAdNetwork) measures the effectiveness of app install campaigns while preserving user privacy. It provides aggregated reports and delayed event data to advertisers without revealing individual user details.

Pros:

- Privacy by design - limited risk of abuse
- Officially supported by Apple for iOS app install conversions

Cons:

- Limited reporting data
- 24-hour delay
- Only covers app install events
- Limited number of event types

Challenge: Attention

“ As an advertiser, I want to make sure my ad has been actively considered by the user because I don't want to pay for ad slots that aren't viewed.

Rewarded Ads

Adoption: Widely Used

Users can opt to watch unskippable ads in exchange for virtual rewards or access to content.

Pros:

- Shows the advertiser that their ad is actively considered for a given amount of time
- Consensual value exchange between publisher and user
- Well-suited for gaming

Cons:

- Scale is limited by the number of users willing to watch rewarded ads
- Users are at risk of being oversaturated with the same ad if appropriate frequency capping measures are not in place and the number of rewarded campaigns is low

Attention Scores

Adoption: Early Adopters

By measuring how much attention users have given to ads, campaigns can measure effectiveness by total attention received. Furthermore, advertisers can bid on the highest attention inventory.

Pros:

- Equal opportunity for all publishers
- Offers both measurement and targeting opportunities
- Well-suited to awareness campaigns
- Can be used directly as a campaign goal

Cons:

- Can be perceived as invasive, especially if eye tracking is in use

- Measurement methodology must be trusted to prevent fraud

Improvements:

- Encrypted attention signals by trusted vendors to assure trustworthy data

Area: Targeting & Prospecting

Challenge: Audience Prospecting

“ As an advertiser, I want to see the scale of various target audiences because I want to find the best scale of the right potential customers that suits my budget.

Contextual Data

Adoption: Widely Used (in channels such as Linear TV, audio and news)

Advertising by considering the likely audience for the content rather than user attributes. This method targets consumers who read, watch, or listen to certain content and can be combined with panel and census data to determine other attributes of the audience such as demographic and socioeconomic factors.

Pros:

- Proven model that works for awareness campaigns - has been used in radio and traditional media for a long time
- On the rise due to increasing investment in CTV content
- Content consumption may be tracked over time to form [Probabilistic Cohorts](#)

Cons:

- Assumes that consumers of the content are a single audience indistinguishable from each other

Probabilistic Cohorts

Adoption: Early Adopters

See the section on [Probabilistic Cohorts](#) above.

Cross-Context Deterministic Cohorts

Adoption: Early Adopters

See the section on [Cross-Context Deterministic Cohorts](#) above.

Seller-Defined Audiences (SDA)

Adoption: Innovators

[Seller-Defined Audiences](#) represents a way to transact on data that cannot be attributed to IDs programmatically without providing the data itself. For example, publishers may insert their first-party data into bid requests to allow the targeting of user-level and content-level labels by buyers. Often used with various cohort generation techniques, as listed above.

Pros:

- Supports standardized and custom audience segments
- Encourages use of first-party data
- Potential to use high quality, directly observed, and correctly consented data
- Cheaper to transact upon - the data is directly available in bid requests
- Scalable

Cons:

- Uncertainty around trusting the various models used by publishers
- Publishers may use inconsistent logic despite using the same standardized taxonomy as other publishers
- SSP traffic shaping rules may discard the data
- Larger publishers with lots of user data have an advantage
- Bad actors can scrape user data from RTB requests to build profiles
- DSPs tend to prefer data with IDs and browser signals directly - some DSPs may fail to bid when there is no ID regardless of the data available

Improvements:

- Encryption of signals prevents data being scraped while also verifying the source of the data
- Trusted data providers provide consistency in data signals
- Validation, transparency, and certification processes such as datalabel.org can be used to improve consistency of signals
- Differential Privacy can be used to mitigate re-identification attacks

Private Marketplace (PMP) Deals

Adoption: Late Majority

PMP deals, like SDA, is a method of transacting upon ID-Less data programmatically. They comprise invitation-only auctions where publishers or SSPs agree on targeting criteria with DSPs or advertisers directly.

Pros:

- Similar benefits to SDA
- Activation mechanism is widespread
- Prevents data leakage from publishers - the user groups within a PMP deal remain private (compared to SDA where they are publicly available)
- Theoretically offers the ability to bid on any first-party data with no format restrictions
- Allows more control over ad placement and reduces the chances of targeting Made-for-Advertising sites

Cons:

- Relies on curating deals for each desired targeting option
- There is no centralized database of PMP deals
- Lack of control over various ad-ops settings on the DSP side
 - Frequency capping and similar use cases may still rely on cookie technologies
- DSPs tend to prefer data with IDs and browser signals directly. Some DSPs may fail to bid when there is no ID regardless of the data available

Improvements:

- Similarly to SDA, validation, transparency, and certification processes such as datalabel.org can be used to foster a level playing field between publishers

Private Aggregation, Reach Estimation (Privacy Sandbox)

Adoption: Innovators

Using technologies such as Privacy Sandbox's Shared Storage API and Private Aggregation API, it is possible to count the number of ad campaign views per target audience across publishers.

The Shared Storage API allows storing data across multiple domains and publishers. However, the Shared Storage API only allows exporting of that stored data through output gates. Advertisers can specifically use the Private Aggregation API to generate aggregate reports of unique counts. So while the individualized view events do not leave the device, advertisers are able to receive privatized aggregated reports of unique views and the site in which they occurred.

For more information see the “Unique Reach Measurement” demo [here](#).

Pros:

- Privacy by design - limited risk of abuse
- Data can be shared between multiple publishers
- Offers some functionality of third-party cookies
- The tools are generalized and are suitable for many different use cases

Cons:

- Output data must be sent to the Private Aggregation API for noisy reporting

Challenge: Audience Enrichment

“ As an advertiser who has identified a target audience, I want to know what other attributes my audience has because it gives me the ability to personalize the messaging to that audience and bring more conversions.

Cohort Lookalikes

Adoption: Widely Used

Where users belong to more than one cohort, correlations can be mapped to build relationship mappings between cohorts and discover similar attributes for campaign planning.

Pros:

- Lookalikes can be discovered whether the user is known or anonymous
- The relevance of each lookalike can be quantified with data

Cons:

- The number of cohorts recorded together in any one observation needs to be controlled to prevent accidental user reidentification (e.g., limiting to 3 cohorts per observation)

Improvements:

- More cohorts may be simultaneously considered using Multi-Party Compute or Federated Learning

Area: Retargeting

Challenge: Bring the Customer Back

“ As an advertiser, I want to bring a potential customer back to the point-of-sale to convert (which may be a sign-up, a purchase, or some other call-to-action) because on average it takes 7 impressions to "break through the noise"⁸.

Related to [multi-touch attribution](#), [audience prospecting](#).

On-Device Auctions

Adoption: Early Adopters

As users browse the web or use apps⁹, they can be assigned to custom cohorts (known as interest groups) which reside on the user's device. Instead of these cohorts being sent to ad servers, the device itself runs a local auction and uses the bidding logic of the interest group to decide whether to bid or not.

Pros:

- Privacy by design - limited risk of abuse
- The interest groups are arbitrary and can be defined by the buyer depending on the needs of their campaigns
- No sharing of the user cohorts themselves for additional privacy
- More reliable retargeting than other tracking methods: the user is assigned the exact criteria that is relevant to the advertiser
- There are protections in place to prevent overly specific interest groups to mitigate the risk of re-identification attacks

Cons:

- Cross-device isn't supported
- Interest groups can't be combined (it's not possible to see which other interest groups a user was in when an auction bid was won)

⁸ For a basic discussion of “the rule of seven” see [The rule of 7: The power of social media](#).

⁹ As of November 2024, the Protected Audience API is not yet available in mobile apps.

- There's a lag between creating an interest group and being able to target it, and not all website/app visitors who get assigned to the interest group will run an auction, resulting in wasted effort
- Up-front effort to create interest groups might not result in successful auction bids if enough publishers don't support the Protected Audience API
- Reach estimation is more difficult and may rely on modeling
- On-device auctions have limitations for the number of bidders that may participate in that auction
- As of June 2024, the latency can be in the magnitude of seconds

Improvements:

- Can be combined with other cohort mechanisms to add retargeting capabilities to other addressability solutions
- Buyers (i.e. DSPs) currently own the interest groups - in the future this could be extended to publisher-owned or vendor-owned interest groups

 **Example: Retargeting in Both ID-Based and ID-Less Approaches**

There are subtle differences between retargeting in an ID-Based versus an ID-Less approach that are important to understand. A

You can find a [deep dive into these differences](#) in the examples below.

Area: Frequency & Recency Capping

Challenge: Preventing Oversaturation

“ As an advertiser, I want to prevent my ad from being seen by the same user too many times no matter where they are because it's wasted ad spend and it becomes annoying to the user.

On-Device Frequency Capping

Adoption: Innovators

When a device receives an ad creative with a frequency cap, it keeps a count of how many times that creative is seen over a period of time. The device can then signal to an ad server whether the frequency cap has been reached or not, which ads have been received, which creatives should be delivered next, and more.

Pros:


- No delay between reaching the frequency cap and preventing further ads from being shown.
- Some user agents and operating systems allow multiple contexts (e.g. different websites) to share the same data storage, e.g. Shared Storage (Privacy Sandbox)

Cons:

- May result in large bid request lengths if many creatives have frequency caps.
- Each solution vendor will maintain its own frequency records.

Improvements:

- OS-level or browser-level support can apply frequency caps to multiple apps and websites, such as when using Shared Storage (Privacy Sandbox)
- If sending the count of ad views per creative, consider adding Differential Privacy noise.
- There's an argument of "different site, different context" which implies it's less important to have a shared frequency cap between sites.
- Modeled data rather than deterministic to resolve issues between vendors not being able to share their observed frequencies.
- Devices can determine what "same" means, whether that is a creative ID or a hash of the pixel values for example.

 **Example: Frequency Capping in ID-Based and ID-Less Approaches**

The [frequency capping example](#) below builds on the prior retargeting example and explains how the Shared Storage API can be used in the browser to undertake frequency capping in an ID-Less approach.

It also introduces the concept of a privacy budget and how that can cause an underreporting of impression totals.

Challenge: Creative Sequencing

“ As an advertiser, I want the user to see several creatives in order because my campaign has multiple messages that are displayed in a storytelling manner.

Similar to "ad break management" for TV content.

Device-Side Creative Sequencing

Adoption: Innovators

Refer to [On-Device Frequency Capping](#) above.

Challenge: Ad Pacing

“ As an advertiser, I want to ensure that my ad is neither displayed too frequently nor too sparsely over a given time period because it may reduce the effectiveness of my campaign's message.

Related Use Case: [Retargeting](#)

Device-Side Pacing

Adoption: Innovators

Refer to [On-Device Frequency Capping](#) above.

Area: Fraud

Challenge: Automated bot detection

“ As an advertiser, I want to know if my campaign was delivered to real audiences or not as I don't want to pay for ads served to robots.

Device Attestation

Adoption: Late Majority (*used at OS-level, not typically used in advertising*)

A form of Zero-Knowledge Proof, a publisher's application is able to generate an attestation message which is cryptographically signed by the device itself. These frameworks can provide different types of attestations, including where the impression:

- Was generated by a legitimate device
- Was generated by a legitimate binary of the publisher's application
- Was generated by a binary installed through the device's app store

Pros:

- Privacy by design - limited risk of abuse
- Included in most Android, Apple, and Roku devices

Cons:

- Must be implemented manually by each publisher into their applications
- Low limit on the number of attestations that can be generated per day (for Apple and Android)
- Attestations can be copied and reused by fake devices and such attacks must be mitigated by attestation validators

Private State Tokens

Adoption: Innovators

A user visits a token issuer's website, and the issuer believes that they are a real human based on their behavior. A token is stored by the issuer on the user's browser. When the user visits another website, that website is able to check the trust token with the issuer to verify that they are a real human. A token issuer may be a reCAPTCHA provider for example.

Pros:

- Privacy by design - limited risk of abuse
- Anonymity is preserved
- Arbitrary and updatable criteria for "human-ness"
- Token recipients can guarantee that it was issued by the issuer
- Similar trust model to HTTPS (using trusted certificate authorities)

Cons:

- It relies on an issuer being trusted and uncompromised
- Relies on the existence of scaled issuers
- If users do not have a trust token for a legitimate reason, it's unclear how this may affect the value of ad requests
- Tokens are limited to within a single browser
- Mozilla has stated that they will not support this, and other browsers are pending as of June 2024
- Different issuers have different definitions of "human-ness"
- Low limit on the number of tokens that can be generated per browser per day, which may impact their usage in programmatic advertising

Improvements:

- More browser support
- More issuers

 **Example: Private State Tokens and Ad Fraud in an ID-Less Approach**

Private State Tokens, one implementation of the Privacy Pass API, are a new mechanic designed specifically to help distinguish “real” viewers from bots.

You can find a [detailed explanation of Private State Tokens](#) in the examples below.

Statistical Determination

Adoption: Innovators

Using sensor and content interaction patterns to detect anomalies that indicate non-human behavior.

While it may be possible to circumvent these measures on a small scale, in order for the fraud to be “worthwhile”, it needs to be on a large enough scale which makes it more susceptible to detection.

Pros:

- Fraud signatures can be continuously updated
- High frequency signals are difficult to forge without introducing repeating or predictable patterns

Cons:

- Must have access to sensor data or raw user interaction data (e.g. screen taps or mouse movement)
- Adversarial AI can train systems to avoid detection until fraud signatures are updated, similarly to the cat-and-mouse chase of virus detection

Challenge: Human bot detection

“ As an advertiser, I want to know if my campaign was delivered to people being employed just to view ads because I don't want to pay for these ad impressions.

DRAFT FOR PUBLIC COMMENT

Statistical Determination

Adoption: Innovators

See [Statistical Determination](#) above.

Examples of How ID-Based and ID-Less Technologies Differ

Retargeting

One use case causing significant concern in the evolution to ID-Less solutions is retargeting. This is because it takes multiple impressions to break through the noisy media environment to have a viewer take some specific action, such as a click-to-purchase. The inability to repeatedly identify a well-defined group of viewers would degrade the performance of advertising campaigns, as it would be impossible to reliably and consistently deliver a series of specific creatives to them.

This particular example explains the common behavior of some Chromium-based browsers (e.g. Chrome and Edge) and Android apps, but some browsers and other technologies may differ in their approach. These differences will be expanded upon in a later example which describes a mechanic using shared storage for retargeting with frequency capping. But for now we'll keep the example relatively simple to highlight a few of the major differences.

ID-Based Retargeting

Figure 3 shows how retargeting differs between ID-Based and ID-Less approaches. In Figure 3a, an advertiser shares a well-defined group of individuals with a publisher site which serves them an ad. While the group is served in aggregate in a pseudonymous manner to meet privacy requirements, each individual in the group has a specific unique user ID (UUID). These UUIDs can be matched through some mechanic to corresponding UUIDs on other publisher sites. So when the user agent visits a second or third publisher's site that has one of those matched UUIDs, that viewer can be retargeted through those publishers with a high degree of certainty.

However, the IDs used to create the matches do have a lifetime. In the case of third-party cookies, the average life of a desktop-based third-party cookie is approximately 30 days and for a mobile third-party cookie it is approximately 7 days¹⁰. Retargeting campaigns can take anywhere from two weeks to two months depending on the industry/product being advertised. So without some other mechanic, the limited lifetime of an identifier limits the ability to retarget in a single campaign with a duration longer than that lifetime. Equally of interest, what if the advertiser wants to reach members of that group in a second campaign several months from now. How does that work?

¹⁰ For desktop-based cookie lifetimes, see as one example [“What is the Real Lifetime of Online Analytics Visitor Cookie?” by Varpu Rantala \(Medium, June 2022\)](#). For mobile cookies, see [“Safari ITP update limits cookies to 7 days for responses from 3rd party IPs.” \(Stape.io, September, 2024\)](#).

Figure 3 - Comparison of Retargeting with ID-Based and ID-Less Solutions

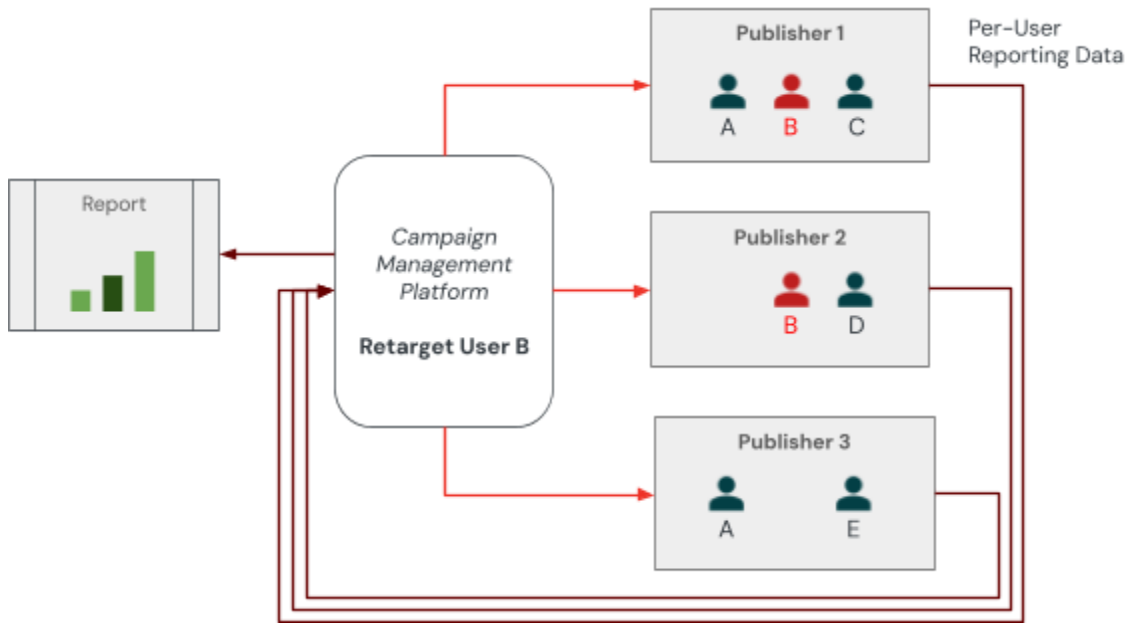


Figure 3a: Retargeting with an ID-Based Solution. The user IDs (A-E) are the match keys between the campaign management platform and the publishers. The specific individual identified as "B" can be retargeted. Some implementations may use browser characteristics as the match key if a deterministic ID is not known.



Figure 3b: ID-Less Retargeting. In this case, the advertiser can reach members of the cohort "Y", but they cannot know which specific members have been retargeted. Reporting information is delayed and noise is added to maintain anonymity.

The good news for the advertiser is that as cookies expire, identity providers or others who maintain identity graphs can reassociate the new third-party cookie (or other identifier) in that user agent with the existing identity graph. As a result, retargeting is possible on desktop and mobile devices across both long-running single campaigns or multiple campaigns over extended periods.

ID-Less Retargeting

Figure 3b shows the comparable ID-Less approach where an ID is not available but first-party attributes, such as cohorts, are available. In this case, as in the first, the advertiser can deliver an ad to viewers assigned to a group via their publisher relationships. The individuals in the group are assigned a unique and temporary Key. However:

- That Key is assigned at the time an auction occurs and disappears once the ad is delivered and the transaction recorded.
- What is maintained in the user's browser or application (user agent) is a cohort name, to which the user agent belongs. Under the Protected Audiences API (Privacy Sandbox), each advertiser can place up to 2,000 interest groups in a specific user agent. Cohort solutions typically limit the number of groups that can be assigned to the user agent, with the Topics API (Privacy Sandbox) allowing for five groups at any given time.
- Cohorts expire after a predetermined length of time. Advertisers can then recreate/update the group using the same name.
- The revised group may share some of the same members, perhaps even the exact same members, although that is unlikely. This is true whether the group is built from an advertiser's deterministic, first-party identifiers or from behavioral signals stored only in the browser.

Cohorts can be retargeted consistently within the expiry window because the cohort Keys reside in the user agent. If an advertiser's campaign on any publisher targets "baby monitor buyers", then the specific user agent instances will be considered for and will potentially receive an ad if it is part of that cohort. However, unlike in Figure 3a, an advertiser cannot know which members of the group have been reached on the second and third publisher's website because

- a. Publishers have no reusable ID to match across sites or auctions, only a one-time, per-site, per auction Key.

- b. Event level reporting is only available to publishers but with limited signals, including winning bidder, price paid, and participating interest groups. This data is intended strictly for the publisher to optimize their ad placements.
- c. Advertiser reporting is only done in aggregate across all websites on which the ads for that specific campaign were shown.

All the advertiser can know is that some members of the group received an impression. In this case they can retarget and can know on average how many impressions have been received by individuals, but they cannot know as specifically as in the ID-Based solution. This is true for both a short-term campaign that runs within the expiry window, campaigns longer than the expiry window, and for retargeting across multiple campaigns.

Duration of campaign does matter, however, if a campaign runs longer than expiry or if the campaign starts on day 25 of a group's 30-day lifetime in a specific browser, the group disappears. Now a new group may have been created with the same conceptual description "baby monitor buyers", but there is no guarantee that this new group will contain the same members. Similarly if the advertiser wants to target baby monitor buyers in a second or third campaign outside the expiry window, there is no guarantee that there would be any overlap between the similarly named groups created in different periods. So consistent retargeting is not guaranteed in these cases, in the short-term or long-term.

These are only a few of the subtleties when shifting this use case from an ID-Based to ID-Less solution. The example aims to give a sense of the differences that should impact the decisions of advertising entities considering ID-Less solutions.

Frequency Capping

Frequency capping is a variant of retargeting, with the addition of limits of how many times an ad can be shown. Thus in the ID-Less serving case, it is subject to the same challenges that were described in the retargeting example. We are going to drill a tad further in this example into some other important differences between the ID-Based and ID-Less solutions, especially around reporting and something called the privacy budget. Once again we are going to focus on Chrome and Android, but some of the same restrictions described here apply equally to Firefox, Safari and iOS, even though the specifics can differ significantly between platforms.

Figure 4a shows frequency capping as it is done today. Assume a frequency cap of 3 impressions for a specific campaign ad. As with retargeting, an individual browser or device, when seen by a publisher for the first time, is identified by a first-party cookie or mobile identifier. At the same time, there is most likely a tag on the publisher's page from a provider of identity services. This takes the identifier for that device, attempts to match it in the provider's identity graph, and stores it (blue lines). Behind the scenes the advertiser (or its representative) has a platform tracking impressions served across all partners (e.g. DMPs) and publishers that the advertiser uses to deliver its ads. That platform captures the serving of the first impression to the various devices. When the device goes to a second site it is either recognized by the new publisher based on its own first-party cookie/mobile ID or, if not, a request is made to the identity graph provider to find this ID in their identity graph and match it to that previously seen ID. So when the ad is served again, it is counted by the advertiser against the original browser/device.

Figure 4 - Frequency Capping

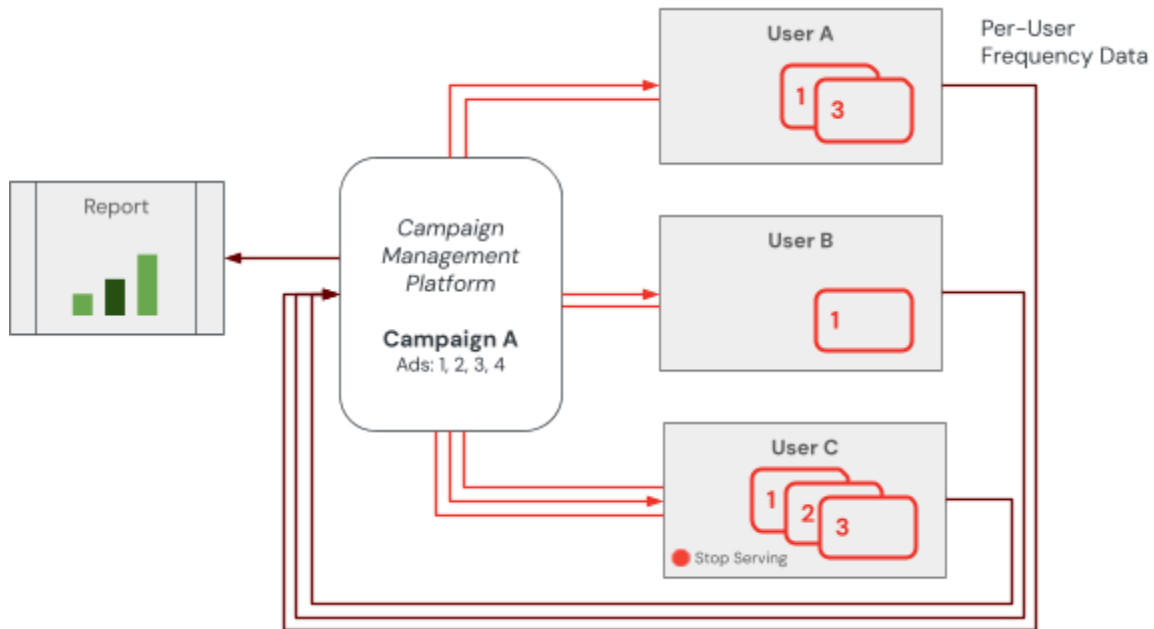


Figure 4a: Frequency Capping in an ID-Based Scenario where the campaign management platform receives data indicating which user IDs have seen which ads. The platform then chooses which users should stop seeing the ads.

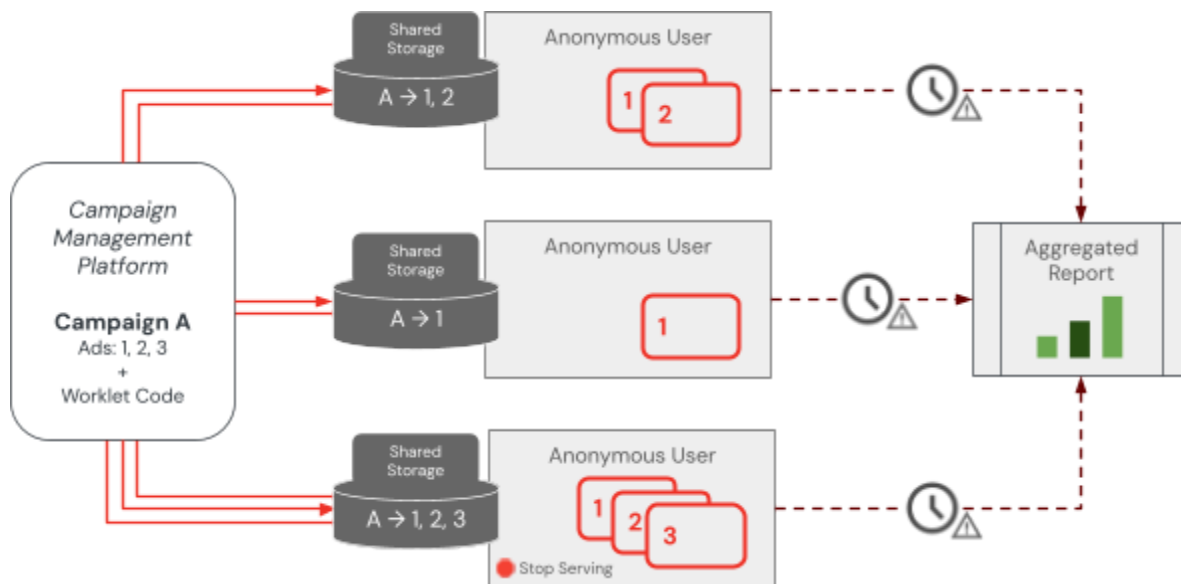


Figure 4b: Frequency Capping in an ID-Less Scenario, where shared storage in each user's browser maintains which ads have been shown for each campaign and worklet code controls the ad serving logic in the user's browser. Reporting information from the browser is delayed and noise is added for anonymity.

This process repeats itself across the entire programmatic ad ecosystem until the browser/device has reached the frequency cap on that specific ad campaign for that advertiser (different ad campaigns can have different frequency caps, obviously). At that point, no further ads are served from that campaign to that viewer (e.g. ID2 in the diagram has reached the frequency cap of 3 so it will no longer be served ads).

The advertiser or DSP can now report on an individualized basis for where and how many impressions were seen. This allows for predictive models to be built on an individual-by-individual basis to inform what ads, what creative, etc. to serve to that device in a future campaign. With a high-performing data pipeline and real-time modeling, the data can be used for real-time campaign optimization, media-mix modeling, or yield-optimization for this campaign.

Figure 4b shows one method for frequency capping in an ID-Less world using Chrome and Edge.¹¹ This approach uses a new standard, the Shared Storage API and something called “seeds”. The advertiser wishes to serve an ad up to a maximum of 3 impressions per device. When the first ad is served to a device, a seed is placed in the advertiser’s shared storage for that campaign. The seed contains a field for the impression cap, and a field for a counter. When the user agent goes to a second site, the publisher sends a “worklet” of Javascript code that accesses the seed from the advertiser’s shared storage.¹² When the ad is served, the counter is incremented by one. This process repeats itself until the frequency cap is reached. After that, no further impressions are served for that advertiser for that campaign for that device.

While frequency capping has been achieved on an individual browser/device basis, there are substantial differences between the two cases.

1. The advertiser (or their DSP) cannot make adjustments to frequency capping in real time.
2. The advertiser in the ID-Based case can manage the frequency cap globally across multiple browsers and channels. In the ID-Less case, the advertiser has no such control. All capping is local and any global result is an aggregation of the individual browser results after the fact.
3. Once again, reporting. Whether in a reporting worklet on the browser, in a SDK on a mobile device, or in a Trusted Execution Environment, the data is aggregated across all

¹¹ Android uses a slightly different mechanic, called [Ad Filtering](#). The underlying concepts are the same but the implementations reflect the different nature of storage on desktop versus mobile devices.

¹² The advertiser provides the publisher the key needed to access the advertiser’s shared storage.

user agents where that ad for that campaign is shown. There is no individualized data for the advertiser to model, only cohort level data.¹³ That data can be both noised and time-delayed on any platform. The advertiser can know on average how many ads were served to any individual in the cohort, with a standard deviation showing the distribution. Modelling can only occur at the cohort level, and cannot be used in real-time audience targeting due to the time delay. When it comes to attribution, the average and standard deviation can be used to determine, on average, the ROAS for the campaign. But the advertiser cannot match an individual purchase to ads served to an individual.

4. There is another subtlety that applies to reporting in both retargeting and frequency capping that has to do with privacy budgeting. Privacy budgeting is a concept from information theory. Information theory quantifies how much information is contained in some data set - in this case a single data export.¹⁴ Releasing too much information through too many exports would allow a malicious actor to potentially reconstruct individual identities. So a privacy budget is put in place by browser and OS owners that limits how much data can be shared. When a user agent reaches its privacy budget, no data can be exported from that browser/device. Thus any aggregation of advertiser data for a specific campaign may not contain all the impressions served. There will thus be a bias towards under-reporting the frequency of ad serving. Moreover, the advertiser will only have a limited ability to understand just how “biased” the data may be. They can know something is not right if the average impressions served shown in the reports are lower than the frequency cap. Lacking that, only the standard deviation might provide some clues that something is incorrect.

¹³ As in retargeting, publishers do receive event-level data for optimizing their ad placements, but we ignore that aspect for now.

¹⁴ For a good introduction to information theory see Stone, James V. [Information Theory: A Tutorial Introduction](#). (Sebtel Press, 2015)

Fraud Detection

There is a saying in adtech that perfect privacy opens the way for perfect fraud because lacking any signals it is not possible to tell a real viewer from a “fake” or invalid one. ID-Less solutions remove a significant amount of signal for identifying an individual user agent. As a result, they create an opportunity for substantial increases in ad fraud in comparison to the current mechanics, which are often based on third-party cookies. Fortunately the Internet Engineering Task Force (IETF) has developed new technology, the Privacy Pass API, that actually improves the ability to track fraudulent traffic. It does not cover all fraud use cases (see table below), but it certainly helps reduce the impact of click and impression fraud that occurs in an ID-Less environment, which is the focus of this discussion. There are two major instantiations of this API/standard: Private Access Tokens on iOS and Safari, and Private State Tokens in Chrome and Firefox.

Fraud Type	Description
Click Fraud	Fake clicks generated by bots or malware instead of real users. Can be automated or involve incentivized clicks.
Impression Fraud	Fake ad views recorded through hidden ads, pixel stuffing, or ad stacking.
Domain Spoofing	Fraudsters misrepresent low-quality websites as premium sites to get higher payouts for ad views.
Cookie Stuffing	Bombarding a user's browser with cookies to create a fake user profile for targeted advertising.
Pixel Stuffing	Hiding a full-sized ad within a tiny 1x1 pixel to register a view without the user noticing.
Ad Stacking	Layering multiple ads on top of each other, where only the top ad is visible but all register views.

Ad Injection	Injecting malicious code into legitimate websites to display ads from the fraudster's network.
Geo-masking	Hiding the true location of the user clicking on the ad to appear like a more desirable demographic.
Bot Traffic	Using automated software (bots) to mimic real users and generate clicks or impressions.
Malvertising	Deceptive ads that redirect users to malware-infected sites or install malware on their devices.
Mobile App Fraud	Fraudulent activity targeting mobile apps, including click injection, click spamming, and install hijacking.

Third-party cookies and identifiers allow fraud detection algorithms to recognize a specific browser/device as the first step in identifying suspected impression or click fraud.¹⁵ These algorithms then look for patterns of behavior from that specific browser/device that can help identify whether this is a “real” person to whom an ad should be served versus a bot or other malicious actor (e.g. manual click fraud) who should be blocked from receiving ads. Warning signals that ad fraud is occurring:

- An unusually high number of impressions served or clicks from the device in a short period of time or at unusual times (e.g., the middle of the night).
- An unusually high number of impressions or clicks coming from a device in an unusual location, such as an obscure country.
- A significant number of very rapid clicks from the browser/device.
- User behavior on the page from a specific browser/device is unusual (e.g., time on page is very low).

¹⁵ IP addresses can also be used. However, while IP addresses will no doubt be phased out as PII over time, right now they are still available so we will exclude them from this discussion.

- Conversion rates on the device are significantly below normal (are clicks leading to actual expected next actions, such as opening a landing page or buying a product?).

Even with these precautions, it is not always easy to detect invalid or fraudulent traffic. And the loss of third-party cookies takes away the basic mechanic that currently exists.

Fortunately, in the world of ID-Less ad serving there is a new, token-based approach to identifying invalid traffic. Private state tokens are a form of encrypted token designed to enable trust in a user's authenticity without allowing tracking. Private state tokens were designed to allow one website or app to validate that a user is "real" and place a series of tokens confirming that fact in the user's browser or app. Later a second website can use that act of validation, contained in those tokens, to verify the user or user agent represents a real person without having to do their own validations and token issuance procedure (Figure 5).



Figure 5: Using a Trust Issuer to validate the authenticity of a user and retrieve a Redemption Token as proof of the user's validity.

Tokens are issued by trusted third parties that provide the tokens to websites. There can be as many of these as the market has room for. A trusted issuer is likely to be a PKI certificate authority of some kind, although nothing in the specification requires that. One of the very unique, but hugely important features of private state tokens is that the issuer is unable to correlate its issuances on one site with redemptions on a different site. As a result, private state tokens are protected from a malicious issuer reidentifying a user and their behavior across websites.

Private state tokens are actually a stronger method for detecting invalid traffic than systems based on third-party cookies. So this is one of the cases where ID-Less approaches can be superior to current, ID-Based approaches.