

About CIMM

The Coalition for Innovative Media Measurement (CIMM) is a nonpartisan, pan-industry coalition of companies focused on cultivating and supporting improvements, best practices, and innovations in measurement and currency; data collaboration and enablement; and the use of new metrics and approaches to understanding the value of media. CIMM embraces the entire media and advertising ecosystem and prioritizes effective collaboration to deliver meaningful change.

About the Author

With more than two decades in advertising, Sable Mi has spent the last ten years at the forefront of data, identity, and measurement, helping organizations navigate a rapidly evolving ecosystem. Sable's expertise lies in elevating data integrity, uncovering meaningful audience insights, and advancing cross-platform measurement to drive smarter decisions and business growth. Widely recognized as an industry thought leader, Sable's work spans published research, advisory roles, and leadership positions across the industry's most influential analytics bodies.

A strategic analytics and research executive, Sable is known for challenging the status quo. She has spoken at leading conferences, authored multiple white papers and bylines, and led industry groups including CIMM's Identity & Data Collaboration Working Group, the ARF Analytics Council, the I-COM MMM & Attribution Council, and previously the IAB Measurement Committee and ARF Identity Resolution Group.

Sable's career spans full-service agencies, university-teaching, and leadership roles in top MarTech companies. She is a two-time IAB Service Excellence Award honoree. Sable holds an M.B.A. from Johns Hopkins University and an M.A. in Mass Communication from Towson University, and currently resides in San Francisco.

Acknowledgements

We are grateful for the participation of the following marketing professionals, who shared their experiences and insights with us.

Alysa Hutnik, Chair & Partner, Privacy, Advertising / Marketing Law, Kelley Drye & Warren LLP

Alyson Sprague, VP, Measurement Science, Samba TV

Brienna Pinnow, Co-founder, Product Strategy & Marketing Consulting for Data-Driven Companies, Blink

Debroop Basu, Principal Product Manager, Healthcare, LiveRamp

Dennis Buckkeim, Global Industry Lead, Snowflake

Don Black, Senior Director of Consulting Services, TransUnion

George Chien, SVP, Data Science Platform, Nielsen

Gerald Smith, (former) General Counsel & Chief Privacy Officer, Cuebiq

Greg Iocco, VP, Audience Measurement, Warner Bros. Discovery

Eli Heath, Principal, Product Management, Amazon

Francesco Guglielmino, Chief Executive Officer, Cuebiq

Ian Tattenham, VP of Sales & Strategic Accounts, Semcasting

Jennifer Yurko, VP, Data Product Strategy, Warner Bros Discovery

Jessica Hindlian, VP, Ecosystem Solutions & Enablement, Open AP

Jessica Hogue, Chief Data Officer, Consumer Media, Hearst

Jessica Simpson, (former) Global Solutions & Partnerships, AI, Choreograph

Josh Chasin, Principal, KnotSimpler

Kara Pellegrino, Managing Director, Crossmedia

Keith Camoosa, Chief Product & Technology Innovation Officer, Acxiom

Kelly Barrett, SVP, Product Management, Samba TV

Kym Frank, SVP, Research & Data, FOX

Lewis Abbey, Sr. Product Manager, Samba TV

Lindsey Woodland, VP, Data Science & Innovation, AMC

Lisa Hamilton, SVP, Data Science, Nielsen

Marc Sabatini, Managing Director, The CRO Collective

Margarita Lyadova, Product

Mathieu Roche, Co-founder & CEO, ID5

Melissa Grady Dias, (former) CMO, Cadillac

Michael Law, CEO, Carat North America

Noelle Huynh, (former) SVP, Measurement & Research, Warner Bros Discovery

Paul Chachko, CEO, Throttle

Paul Cimino, President, Cimino Collaborative

Peter Nummerdor, SVP, Product Management, VideoAmp

Rachel Cascisa, VP, Platform Adoption, Epsilon

Rachel Galvin, Client Advisor, ThinkMedium

Richy Glassberg, Co-founder / CEO, Safeguard Privacy

Scott Gordon, Founder, Epic Innovations

Steve Silvers, Chief Product Officer, Guideline

Suzi Palmer, Managing Director, Data Collaboration, Data Clean Room, LiveRamp

Venu Konda, EVP, Data & Media Solutions, Dentsu

Zora Senat, Chief Customer Officer, Cuebiq

Research Objectives and Approach

In October 2024, CIMM published [*Identity Resolution for Advanced TV and Video Advertising: A Case for Durability and Transparency*](#). The report provided an overview of the identity resolution (IDR) ecosystem, a diagnostic assessment of the marketplace, and a set of recommendations, including the need for a buyer's guide.

Building on that foundation, CIMM established the Identity & Data Collaboration Working Group, with this paper serving as one of its key initiatives. This paper aims to provide clarity and direction for brands and publishers, i.e., the “buyers” who own first-party data and maintain direct consumer relationships, helping them assess identity solutions that support both expansion and collapse to drive richer insights, more precise targeting, and more accountable marketing outcomes.

Because identity is an ever-evolving topic, ensuring relevance requires a mixed-method approach. In addition to reviewing dozens of industry references, the research incorporated extensive one-on-one interviews. To date, more than 40 industry leaders — some participating multiple times — have contributed their candid expertise and perspectives. This paper would not have been possible without their generosity and engagement.

This report is intended solely for educational purposes. Neither CIMM nor the author make any representations as to the accuracy or completeness of any information contained in this report or in any report or website linked to in this report, nor will either be liable for any errors or omissions in this information or for any losses, injuries, or damages incurred from the display or use of this information.

© 2026 ARF Innovation Studio, Inc. All rights reserved.

CONTENTS

Executive Summary	6
I. Identity Solutions Overview	7
II. Needs Assessment	10
III. Targeting Optimal Providers	11
IV. Key Questions for an RFP or RFI	12
V. Kicking the Tires	17
VI. Evaluation Criteria	20
VII. An Ongoing Process	25
Appendices	
Appendix A – Glossary	28
Appendix B – References	30

Executive Summary

Identity has become the connective tissue of modern marketing, yet many brands and publishers still struggle to cut through vendor noise, technical jargon, and shifting privacy rules to choose the right identity resolution partners. This guide is designed to help buyers who own first-party data turn identity from a confusing buzzword into a practical, durable advantage across targeting, measurement, and customer experience.

This paper provides a clear framework for understanding core identity concepts (including expansion, collapse, identity graphs, deterministic vs. probabilistic matching, interoperability), along with a structured needs assessment to determine whether, when, and how to engage external providers. The paper then walks through how to navigate the marketplace: segmenting provider types, asking the right RFI / RFP questions, “kicking the tires” with match tests and demos, and applying rigorous evaluation criteria across data quality, integration, privacy / compliance, service, and commercial models. It concludes with a peek into synthetic IDs and non-human identities that points the way to the near future.

Throughout, the guide translates hard-won lessons and input from more than 40 industry leaders into checklists, example scoring matrices, and practical prompts can be applied immediately in vendor reviews, internal strategy discussions, or board-level conversations.

Evaluating identity solutions can be confusing and exhausting, as in between expansion and collapse lies the real work: deciding which signals to trust, which partners to empower, and how to align identity with accountability. The purpose of this guide is not to crown winners and losers, but to equip buyers with a clear, structured way of thinking — so they can ask sharper questions, make more confident decisions, and build identity practices that are resilient, privacy-forward, and genuinely useful to the business for years to come.

Whether an organization is standing up identity for the first time or recalibrating an existing stack, this paper will help the buyers separate signal from noise and select partners that deliver accurate, privacy-conscious, and future-ready identity intelligence at scale.

I. Identity Resolution Overview

Research indicates that 70% of marketing leaders struggle to identify and reach audiences across multiple touchpoints due to media fragmentation, a complex network of marketing technology (MarTech) solutions, and identity challenges¹. And the term *Identity* can be vague under different context and adds more to the confusion. For the purposes of this paper, we will focus on identity in the MarTech space and define identity resolution (IDR) as the process of connecting disparate data points to create a unified, cohesive profile of a single “unit.” That unit could be a device, an account (e.g., an email), a person, a household, or a “micro-cohort” or “cluster.”

By linking multiple data sources to a unique identity “unit,” marketers can achieve a more complete view of their customers — enabling better personalization, more relevant engagement, and stronger long-term relationships.

There are a wide variety of solutions providers who help marketers craft these identity databases. While each provider offers a distinct methodology and scope, the core of identity resolution consists of two fundamental components — expansion and collapse.

Expansion

Also known as hydration or enrichment, expansion involves adding attributes to a known unit — essentially, adding new columns of data to each identifier. In today’s fragmented landscape, consumer-related data are scattered across online and offline environments like “digital bread crumbs.” Many of these signals — registration data, media consumption, purchase history, physical location, and more — can be connected to one or more identifiers to form a richer profile.

Collapse

Also referred to as deduplication or reconciliation, collapse links multiple identifiers (e.g., device IDs, email accounts, phone numbers, logins, names, addresses, IP addresses) and consolidates them into a single, persistent entity — effectively merging rows in a database to remove duplicates. This process enhances operational efficiency and provides a more holistic and accurate view of the identity “unit.”

The Role of Identity Graphs

By integrating multiple datasets and refining the data through processes of expansion and collapse, identity providers construct identity graphs that connect identifiers — such as cookies, mobile advertising IDs (MAIDs), logins, emails, names, and physical addresses — across devices and channels into a unified identity, typically representing a person or household.

These identity graphs form the foundation for a consistent, privacy-conscious view of consumers, enabling marketers to:

- Map and understand customer journeys across touchpoints to identify trends and gain holistic insights.
- Design a media plan, onboard and activate relevant and addressable audiences.
- Manage advertising reach and frequency across channels universally.
- Measure and optimize ad campaign performance holistically.
- Build closed loop measurement systems that can attribute ad exposures to business outcomes to measure return on ad spend (ROAS).

Broadly speaking, the identity ecosystem underpins nearly every aspect of modern marketing. Yet, with diverse data sources, evolving regulations, and varying business models, a “one-size-fits-all” IDR approach is neither practical nor effective.

This paper provides clarity and direction for brands and publishers, i.e., the “buyers” who own first-party data and maintain direct consumer relationships, helping them assess identity solutions that support both expansion and collapse to drive richer insights, more precise targeting, and more accountable marketing outcomes.

¹ Forrester Research, *Identity Resolution Survey* (Q4 2024)

I. Identity Resolution Overview

What Buyers Are Saying

“I assume they are telling me the truth”

Identity is complex and full of details. Some buy-side stakeholders take claims and the results at the face value without digging into data sources, definition, scale or validation.

“I don’t care” vs. “when does it matter”

While debates on methodology and accuracy continue, we heard brands note that unless it impacts the bottom-line, it doesn’t matter.

“Don’t tell me it’s perfect. Tell me where the bias / gaps are”

Contrasting to the “I don’t care” is the demand for transparency and realistic expectations, preferring clarity on bias and limitations over claims of perfection.

“I wish they would give me a clear pricing model that fits my purpose”

There is strong demand for more flexible, modular pricing. Buyers want the ability to select only what they need instead of being forced into all-or-nothing models.

“How can I be sure if they are doing the right thing?”

Even though vendors claim compliance, buyers find it difficult to validate those assurances.

“I asked a specific question and they gave me a fluffy answer without answering the question!”

Buyers complained that some of the responses lack specifics, and therefore did not help with the evaluation process.

Benefits of Deploying IDR Solutions

Identity has become the connective tissue of modern marketing. The industry’s intensified focus on identity resolution was triggered by Google’s 2021 announcement to deprecate third-party cookies. Even though Google has since reversed course and sunset its Privacy Sandbox initiatives, the anticipated loss of cookies accelerated widespread structural change. At the same time, the rapid proliferation of connected TV, digital audio, and all other IOT (internet or things) has further fragmented signals. Together, these forces have elevated identity from a technical consideration to a critical enabler of data-driven marketing, cross-channel measurement, and decision making.

While identity resolution, especially when related to digital identifiers, is often associated exclusively with media activation, its value extends far beyond ad delivery. The most common use cases include:

A. Enhancing Content Personalization

Identity solution enables marketers to identify and engage high-value audiences with tailored messages such as next-best offer or personalized content based on behavioral or transactional data. Leveraging identity can help build loyalty, drive cross-sell opportunities among existing customers, re-engage lapsed users, attract “moveable middle” audiences, or upsell less frequent buyers.

B. Improving Targeting and Media Efficiency

Identity empowers creative personalization, allows for sequential marketing messaging and frequency management across channels, ensuring the right message reaches the right audience at the right time, reducing waste and improving media ROI.

I. Identity Resolution Overview

C. Streamlining Customer Experience

Auth0 puts it succinctly: “Customers want frictionless, personalized, and instantaneous experiences when logging into apps and making purchases.”² Automated identity resolution enables marketers to build and share a unified customer view across brands, products, and business units. Consistently recognizing customers throughout their journey eliminates duplication, ensures singular user experience and preference, reduces friction, and enables personalized engagement in the most relevant channel at the optimal moment.

D. Strengthening Measurement and Attribution

Persistent identifiers that link known and anonymous customers across devices and channels support more accurate, closed-loop measurement and attribute credit accordingly, enhancing accountability and marketing effectiveness.

E. Compliance

As privacy regulations continue to evolve and consumer expectations for data transparency rise, maintaining a modern identity management practice is essential for compliance. Marketers must be able to authenticate users, manage consent and privacy preferences access or verify customer records, and delete personal data upon request. Implementing these capabilities not only ensures adherence to legal requirements but also strengthens consumer trust and supports responsible data stewardship.

² auth0, *Customer Identity Trends Report 2025: Securing Customer Trust in the Age of AI* (2025)



II. Needs Assessment

Buyers should explore a range of internal questions to assess their readiness for, and potential demand for, an identity resolution solution. The table below outlines the key questions organizations must evaluate to determine their need, and type of approach, for seeking IDR solutions support.

Figure 2.1: IDR Needs Assessment Approach

Critical Assessment Areas	Key Questions
Current Status	<ul style="list-style-type: none"> How is my current customer data collected, cleaned, organized, stored, and refreshed? Are there any gaps in my customer knowledge that can be fulfilled by another source? How is identity resolution currently used within our organization?
Why: Know Your Use Case(s)	<ul style="list-style-type: none"> Which use cases matter most? (e.g., personalization, media efficiency, measurement, cross-channel consistency) What limitations or concerns exist with current practices? How would we use identity resolution differently with a new solution? Are these short-term needs or part of a long-term data strategy? Which platforms or marketing channels will benefit most from an IDR program?
Who	<ul style="list-style-type: none"> Which departments will use the solution? (marketing, analytics, customer support, compliance, etc.) Who owns identity internally? (data team, marketing ops, IT) Do we need external partners / consultants for implementation? Do we need additional personnel to maximize the value of a new identity resolution solution?
When	<ul style="list-style-type: none"> What is the implementation timeline? (urgent vs. phased rollout) Do we need real-time resolution, or is batch sufficient? Are there seasonal or market-specific deadlines to consider?
Where	<ul style="list-style-type: none"> What systems will this integrate with? (CRM, CDP, data warehouse, media platforms) Where does the data live today? (cloud / on-prem, regional / global) Are there geographic / regulatory restrictions to consider? (e.g., EU or state-level compliance)
How Much	<ul style="list-style-type: none"> What is our budget? How do we finance it? Do we need flexible pricing? (modular vs. all-in-one) What ROI metrics or efficiency gains will justify spending? How would the benefit of new solutions offset the cost over time?
Other Considerations	<ul style="list-style-type: none"> How will we validate vendor claims? (audits, proofs of concept, match rate benchmarking) How will success be measured internally? (KPIs, adoption, cost savings, incremental lift) What risks are we willing / not willing to accept? (probabilistic matching, synthetic IDs, data leakage) Who controls my data and who is liable?

III. Targeting Optimal Providers

IDR platforms generally provide the following core services:

- Data onboarding (including online / offline matching),³ which allows clients to onboard their first-party data via secure file transfer (SFTP) or API, enabling the provider to match the individual ID within their ID graph, suppress unresolved IDs, pseudonymize, and validate the accuracy of the dataset.
- Proprietary identity graph.
- Client ownership of first-party data.
- Persistent individual and / or household ID.
- Compliance with privacy regulations.
- APIs for third-party system integration.

While providers may offer different service packages from one another, they can generally be segmented into three large buckets. To streamline the evaluation process, it would be easier for buyers to focus on the segment of solution provider appropriate for their specific needs, and avoid issuing RFI / RFPs to misaligned providers.

Figure 3.1: Identity Solutions Provider Segments

Segment	Description	Examples
A. Media Activation & Measurement Focused	Identity providers that connect attributes to digital identifiers for audience creation and activation. While they are able to leverage their identity data for measurement, they typically do not export individual-level profiles to marketers.	LiveRamp*, ID5, UID2.0 (TTD), Criteo; also walled gardens such as Meta, Google, Amazon, etc. *With limited exceptions
B. Comprehensive Identity Providers	Identity providers that connect both online and offline identifiers (ID graph) and attributes (data) and are able to export individual-level profiles to marketers beyond media activation.	Experian, TransUnion, Adstra, Verisk, Semcasting, LiveIntent; also agency holding companies' tech arms such as Epsilon, Acxiom, Merkurly, etc.
C. Tech Providers / Infrastructure (Identity Enablers)	Tech companies that provide software to enable identity matching without owning their own ID graph.	Snowflake, Databricks, AWS, Microsoft Azure, InfoSum (WPP), Aqfer, Tealium, Narrative, MadConnect, Optable, AppsFlyer, etc.

Of course, many data aggregators, providers, and retail media networks (RMNs) overlap across these groups.

³ Pamela Parker in MarTech, [What is identity resolution and how are platforms adapting to privacy changes?](#) (September 16, 2024)

IV. Key Questions for an RFI or RFP

Issuing an RFI (Request for Information) or RFP (Request for Proposal) is a common practice as the first step to evaluate new vendors. But the question lists can risk being too extensive, making the process exhausting for both brands and vendors. Eschewing long, complex RFPs is sensible, as the most effective requests are the ones that clearly focus on purpose, criteria, and relevant information.

While each company should customize its request to best suit its own purpose, below we list a set of key questions that should be included in the RFI or RFP. A simplified RFP template in Excel capturing these questions is provided as a tool at [this link](#).

Key RFI / RFP Questions

General Business

- What's your approach to identity?
- Please share examples of major clients using your identity graph.
- What differentiates your identity solution from others?
- What are the permissible use cases of your data?
- Do you have any success stories that demonstrate your identity solutions have worked in cookieless environments?

Data & Identity Graph

- How do you source your data? Please describe the approach and name the sources.
- How do you validate the quality of your data sources? Is there a threshold for you to accept any given source? If yes, please share the methodology.
- How often is your data refreshed? What is your validation and accuracy cadence (daily, weekly, monthly)?
- What is your coverage by market? (Ask for specific markets based on your business and roadmap). Which use cases are supported globally vs. specific locale?
- What types of identifiers does your platform support (e.g., name, email, device ID, home address, IP address, phone number, hashed PII)? How many records use IPV6 vs. IPV4?
- What is your methodology for combining different data sets to form a single source of truth?
- How many unique individuals does your data resolve to in the US (or another key market for your business)? How do you define a "person"?
- Does your identity graph support households (HHs)? How many HHs does your data resolve to in the US (or another key market for your business)? How do you define a "household"? If IP address is used: which ID signals directly connect to IP address, and at what frequency?
- How do you validate the accuracy of your data, and of your matches? Do you have any stats or case studies you can share that show high accuracy?
- Do you provide a confidence score? How do you assign a confidence score?
- Which channels / platforms are you currently integrated or interoperable with?

IV. Key Questions for an RFI or RFP

When discussing identity resolution (IDR), the terms “deterministic” and “probabilistic” often come up. These describe the methods used to match attributes to a unique identifier either by observation or algorithm. It is important to know that identifier itself, e.g., ID# 12345 is always deterministic, and many identity providers use a combination of both approaches to balance accuracy and scale.

Deterministic Matching

Deterministic matching relies on observing known identifiers such as cookies, logins, device IDs, hashed emails, phone numbers, names, or postal addresses that are persistently associated with a person or household. By linking these known identifiers through a direct match, IDR providers can enhance or collapse identity “units” across different sources. It is important to note that these known identifiers are not always personal identifiable information (PII).

Probabilistic Matching

Probabilistic matching, on the other hand, infers connections between attributes and identifiers based on relational patterns without requiring a direct identifier match. Common signals used in probabilistic matching include timestamps, user agents, IP addresses, location data, and browsing patterns. It is important to know that probabilistic matching can be applied to all signals including known identifiers such as email. Using these signals, statistical models estimate the likelihood that multiple attributes belong to the same identity, helping account for unknown or missing factors.

Complementary, Not Mutually Exclusive

Both deterministic and probabilistic matchings have their strengths and limitations, which often complement each other. Most IDR solutions use a hybrid model. For example, a user profile may be anchored with deterministic signals such as email logins and media consumption history. From there, statistical algorithms and predictive models extend the profile by connecting unauthenticated browsing data, e.g., user agents, timestamps, and geolocations, without relying on PII.

ID Bridging

ID bridging is a process of connecting all the different identifiers across platforms, devices, and data sources where traditional cookies or deterministic IDs are unavailable. ID bridging focuses on maintaining identity continuity across fragmented, privacy-restricted environments. In the context of programmatic advertising, ID bridging is primarily used by publishers and SSPs to generate identity signals in cookieless browsers (e.g., Safari). Many sell-side platforms leverage ID bridging solutions to increase bid responses and monetization with demand partners, often without transparency or permission.

IV. Key Questions for an RFI or RFP

Privacy, Consent & Compliance

- Is your platform compliant with GDPR, CCPA, and other relevant regulations?
- How do you maintain compliance with existing and emerging relevant privacy legislation?
- How do you capture, store, and transmit consent and opt-out status?
- How are access, deletion, and portability requests operationalized end-to-end?
- What identifiers are required for matching, and can specific fields be suppressed?
- What's your approach to hashing, salting, tokenization, and pseudonymization?
- Can you support regional data residency requirements?
- Can we restrict processing to specified regions and providers?
- How are sub-processors vetted, monitored, and audited?
- Can we license all of your data or just a portion? If the latter, what portion?
- How do you ensure transparency and auditability in identity resolution processes?
- What are your breach SLAs, notification protocols, and forensics capabilities?
- Do you conduct DPIAs and offer opt-outs for profiling / automated decisions?
- Can you produce logs for consent changes, matching events, and deletions?

Integration & Interoperability

- What integrations do you offer with CDPs, CRMs, media platforms, and analytics tools?
- Can your solution operate within our existing MarTech stack without major disruption?
- Do you support crosswalking with other data partners?
- Do you support real-time identity resolution and activation?
- What APIs or connectors are available for custom workflows?

Interoperability

Interoperability between IDs enables various parties in the digital ecosystem to communicate with each other and advertisers to ensure increased reach and scale for their audiences.

In the absence of a universal ID, interoperability becomes a cornerstone of identity solutions, essential for:

- **Advancing insights:** Facilitates collaboration with data partners to create a more granular and comprehensive view of consumer profiles.
- **Enhancing targeting:** Enables advertisers to activate the same audience across multiple channels and platforms, driving scale while minimizing media waste.
- **Improving measurement:** Allows measurement vendors to connect media touchpoints to business outcomes, strengthening evidence of campaign impact across consumer journeys.

Data Format & Delivery

- What types of data delivery do you support? Do you support S3 or SFTP delivery?
- What file formats do you support? (e.g., JSON, CSV, PSV, etc.)
- Do you support delivery of hashed PII? If so, what hashing methodologies do you support?
- What is your data normalization process? Do you work with a third party?
- Do you support the sharing of hashed PII to third parties to facilitate matching?
- Do you have the ability to send us raw identity data or does licensing your graph require us to receive data you have pre-processed?

IV. Key Questions for an RFI or RFP

Measurement & Attribution

- How does your platform support marketing mix modeling (MMM) or multi-touch attribution (MTA)?
- Can you link identity resolution to performance metrics across channels?
- How do you handle identity persistence over time for longitudinal analysis?

MMM and Identity

Marketing Mix Modeling (MMM) traditionally utilizes aggregated data to estimate the contribution of each channel to performance, and therefore does not rely on identity. However, identity can enhance MMM in many ways, including leveraging synthetic IDs to improve modeling, connecting online and offline data to better capture the full consumer journey across touchpoints, and enabling more granular reporting, e.g., at the audience level based on aggregated media exposure.

Transparency & Control

- As your client, can we access and control the identity graph directly?
- Do you offer visibility into match rates, confidence scores, and resolution logic?
- How customizable is your resolution methodology to fit our business rules?

Support, Scalability & Roadmap

- What onboarding and ongoing support do you provide?
- How scalable is your solution across geographies and business units?
- What innovations are planned in your product roadmap (e.g., AI, predictive identity)?



IV. Key Questions for an RFI or RFP

Pricing

- What is your business model and rate card?
- Are there other fees to be aware of?
- What are your payment terms?
- Do you offer any discounts? If so, what are the triggers for the discounts?

Pricing models are rarely the first discussion point when evaluating IDR solutions, yet often become the determining factor of final selections. There are three typical pricing models, though keep in mind that volume and commitment are always in the mix of negotiation. The higher the volume with the longer commitment, the more negotiation power a buyer would have.

Figure 4.1: IDR Solution Pricing Models

Pricing Model	Description
Volume-based	<ul style="list-style-type: none">• Records being managed and resolved.• Data returned to the client with appended identifiers, additional attributes such as demographics, propensity scores, etc.• Cadence of updates: daily, weekly, monthly, quarterly, or custom.
Media-only	<ul style="list-style-type: none">• Media fees.• Data fees.• Advanced reporting.
Flat Fee	<ul style="list-style-type: none">• A flat monthly fee depending on the scope of services, such as types of data required, amount of data, cadence of resolution, install or managed services.
Others	<ul style="list-style-type: none">• Additionally, it is possible to purchase a full data set / list and bring it into a client's environment. Financial institutions use this method where they outright purchase the Identity data and host it in their own environment.• Another method is pass-through technology fees, e.g., identity may come packaged along with a Clean Room or CDP solution.

V. Kicking the Tires

Match Test

Conducting a match test is a common practice for evaluating both the identity providers and the marketer's dataset. It should include two aspects: matching and reconciliation. While there are different ways to run a match test, most providers would ask for a sample from the marketer to provide a "match rate." Additional ways to conduct a match test include submitting an entire dataset to generate a diagnostic report, or for a client to do an analysis themselves by requesting a segment of the IDR provider's dataset (for example, a sample of 10% of their data for a given month, with a 30-day lookback). Generally speaking, the greater the sample size of data shared and evaluated, the more conclusive a match test can be. As an alternative to match tests, a client can simply request an estimated match rate given their vertical market and database size, but this approach is less conclusive than running an actual test.

It is important to remember that the result of a match test is more than just a number. Key criteria in evaluating the success of a match test include the following:

- What is the process — how do they determine a record is a "match."
- How many records you originally share vs. how many records they can "match" in their identity graph.
- Within the matched records, how many individuals and how many households are they able to identify.
- How many records they identify that are outdated or duplicated.
- Actionable match rate: of the matched records, how many are addressable across primary activation platforms.
- How many of the matched records do they have common identifiers that can be matching with the third party.
- Graph stability over time, e.g., a hashed email should match to the same person over time.
- Keep in mind that while many records can be matched / resolved, only a subset user records can be deployed through personalized marketing channels.

Match Rate

Match rate refers to the percentage of records from your dataset that an identity resolution provider can successfully match to known identities in their database. For example, if you submit 10,000 customer records and 7,000 are matched, the match rate is 70%. Match rate is often used as a signal of dataset compatibility between you and the identity providers.

However, match rate definitions, methodologies, and reporting practices vary widely across providers, making it difficult to compare results. A high match rate may not necessarily translate to better performance or business outcomes.

For instance, if you submit 10,000 records and the provider identifies that 2,000 of them are duplicates, the true customer count becomes 8,000. In this scenario, the reported 7,000 "matched records" could reflect somewhere between 5,000 and 7,000 unique individuals — a meaningful difference depending on your use case.

Additionally, match rate does not necessarily account for *recency*. Some providers use "all-time" records to match. Nevertheless, what matched three months ago may not be true today or three months later.

Other considerations include:

- Accuracy and precision (e.g., false positives, or overly permissive matching rules).
- Use case-specific addressability such as digital activation vs. analytics only.
- Sample representativeness — since match tests are often run on subsets of data rather than full files.

V. Kicking the Tires

Ask for a Demo

Many platforms will offer to demonstrate their user interfaces as part of the pitching process, which gives buyers an opportunity to get a look-and-feel of the services provided. Skipping this step, or a declining a trial period with the data, can lead to surprises after commitment. Demos and trials allow you to evaluate the data's quality, relevance, and integration capabilities firsthand, ensuring it meets your specific needs.

However, keep in mind that the demo is a sales event, and can showcase the “shiny objects” of the product solution while skipping less attractive features. If possible, ask for a sandbox or trial instance as well.

Evaluating the strength of a product demo can be buttressed by asking the following questions⁴:

Platform, Data, and Onboarding

- Does the platform support first-party data onboarding? Can we incorporate any of our private customer IDs into the platform? If so, what's the onboarding process and how long does it take?
- Do you own or license your referential identity data?
- What are your identity data sources? Ask for the sources of specific data used in the demo to gain clarity.
- How do you validate the quality of your identity graph?
- How much of your data is addressable? Ask for specific platforms and identity units that are strategic in your marketing plan.
- How is your identity graph linked to offline PII? Ask for specific examples, e.g., how do you connect app usage to name and address.
- If the demo is on the U.S. market and your business covers other regions, ask if identity capabilities apply to non-U.S. markets, and what's the difference in capacity.
- How does the platform integrate with other MarTech platforms (e.g., CRMs, DSPs, CDPs)? Ask for a few examples.
- Does the platform feature any built-in data activation capabilities (e.g., personalized email or ad campaign execution)?
- Are APIs available for data import / export?
- What analytic support and reporting do you provide that will document ROI for our identity investment?
- Offboarding portability and data migration (upon “churn” events).

⁴ Pamela Parker in Martech, [24 questions to ask identity resolution vendors during a demo](#) (March 27th 2024)

V. Kicking the Tires

Customer Support

- What kind of customer support is included — can we pick up the phone to report problems?
- Will we have a dedicated account manager and technical support?
- Do you offer a proof-of-concept trial to measure potential performance and scale?
- Do you provide a self-service option in which we can manage identity data?
- What kind of professional services are available, and how much do they cost?
- What is your SLA for critical issues, and how do you measure adherence?
- How do you handle escalations, and what's the average resolution time?
- Can you share examples of proactive service (e.g., quarterly business reviews, optimization plans)?
- How do you handle requests for product modifications?
- What new features are you considering? What is the long-term roadmap and launch dates?

While some of these questions may seem redundant with those in the RFP, they provide more clarity during the demo process as you get to see how the vendor's data is arranged and applied to specific use cases. While the questions listed above are applicable in most cases, the best way is to build upon them and probe issues specific to your company's needs and circumstances, e.g., if I give you X, can you provide Y?



VI: Evaluation Criteria

Accuracy and Quality of Data⁵

Not all data is created equal, and the sources and methodologies used by vendors to collect, verify, and update their data can greatly influence its reliability. High-quality data can be defined by several key characteristics: accuracy, timeliness, relevance, precision, refresh rates, last validated timestamps, and completeness.

Before examining the data itself, it's essential to assess the credibility and expertise of the data sources. Consider the reputation, credentials, and expertise of the entity responsible for collecting, cleaning, and preparing the data. Review whether the data collecting methodology adheres to established standards and look for potential biases, flaws, or limitations in the process.

It's important to distinguish between **data accuracy** vs. **data completeness**. Data accuracy refers to the correctness of values within a dataset, whereas data completeness refers to the extent of data coverage. While data accuracy is an important component of data quality, these two are not synonymous. For example, data can be accurate yet incomplete, outdated, or inconsistent. True data quality depends on balancing both attributes.

Figure 6.1: Precision vs. Recall

	Matched	Not matched
Agree with truth set	True Positives	True Negatives
Disagree with true set	False Positives	False Negatives

To evaluate data matching accuracy, you will need to establish a truth set against the data you wish to evaluate. Common evaluation metrics include precision, recall, and accuracy⁶:

- **Precision** — measures the trustworthiness of the matches.

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}}$$

- **Recall** — measures the completeness of the matches.

$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}}$$

- **Accuracy** — measures the overall correctness of the match test.

$$\text{Accuracy} = \frac{\text{True Positives} + \text{True Negatives}}{\text{True Positives} + \text{True Negatives} + \text{False Positives} + \text{False Negatives}}$$

In most matching systems, there is an inverse relationship between precision and recall. Together, these metrics help quantify how effectively a matching system identifies true matches while minimizing false ones, whether using rules-based or machine learning-based resolution methods.

Data Quality

A common mistake is prioritizing the volume of data over its quality. While having a vast amount of data sounds appealing, irrelevant or outdated data can clutter your systems and lead to misguided decisions. It's vital to focus on the accuracy, relevancy, and freshness of data rather than just the quantity. We call this the data.com fallacy.

⁵ LeadGenius, [7 Critical Factors in Evaluating a Data Provider](#) (May 23, 2024)

⁶ CVAT, [Is Your Training Data Trustworthy? How to Use Precision & Recall for Annotation QA](#) (September 12, 2025)

VI: Evaluation Criteria

Integration and Compatibility

Seamless integration of external identity data into your Customer Relationship Management (CRM) and marketing automation platforms is essential. The ability to ingest and activate data without manual manipulation safeguards operational efficiency, while integration challenges can lead to costly disruptions and inefficiencies.

Your identity provider should support flexible data structures that align with your internal taxonomy and data dictionary, enabling straightforward mapping to existing fields. The tool should cover your standard CRM / MAS platforms and have an Open API for more sophisticated data refreshes. This ensures that identity data can be operationalized quickly and consistently across systems.

Compatibility should be assessed not only within your internal data architecture but also in terms of interoperability across your broader ecosystem. Providers that offer established crosswalks with strategic partners and / or that support high-fidelity mapping of widely used identifiers such as email or hashed email (HEM) enable greater data fluidity. This level of interoperability is critical for maximizing the utility of identity data across platforms, partners, and use cases.

Integration, compatibility, and interoperability are the keys to operational scale. When evaluating scalability, consider the following factors:

- Does the solution have enough coverage in your relevant channels and markets?
- Does the solution allow you to seamlessly combine your first-party data with data from other sources?
- Is the ID solution widely adopted by your strategic partners?
- How well the ID graph is populated with common identifiers, e.g., HEMs, IP address (Ipv6 vs. Ipv4)?
- Can you easily transfer your first-party data across different platforms using this ID solution?
- Does the solution enable you to expand your audience targeting to reach more potential customers? How much does it expand your reach beyond current levels?

Compliance and Privacy

Compliance and privacy are not just a standard requirement; they are the foundation to maintaining customer trust. While every provider asserts they are fully compliant, the burden is on customers to apply scrutiny and ongoing diligence to this complex and constantly evolving territory.

When evaluating IDR providers in compliance and privacy, ask specific questions, request documentation, and watch out for red flags.

Documentation to request:

- Data Processing Addendum (DPA), SCCs, and regional hosting options.
- Records of Processing Activities (ROPA) and recent DPIAs.
- Sub-processor list with services, regions, and DPAs.
- Security reports (SOC 2 Type II, ISO 27001) and penetration testing summaries.
- Consent / opt-out propagation design and sample logs.
- Retention / deletion policy and operational playbooks.
- Clean-room enforcement documentation and export auditing.

What to watch out for:

- Vague role definitions: Lack of clear delineation between controller vs. processor responsibilities by use case.
- Opaque data sourcing: Inability to demonstrate consent provenance for third-party data.
- Unsalted hashing: Use of plain or shared salts for HEM.
- Retention ambiguity: Lack of configurable retention policies or deletion audit trails.
- Sub-processor opacity: No published sub-processor list, or reliance on broad “blanket approvals”.
- Bundled profiling: Default profiling without opt-out or DPIA documentation.

VI: Evaluation Criteria

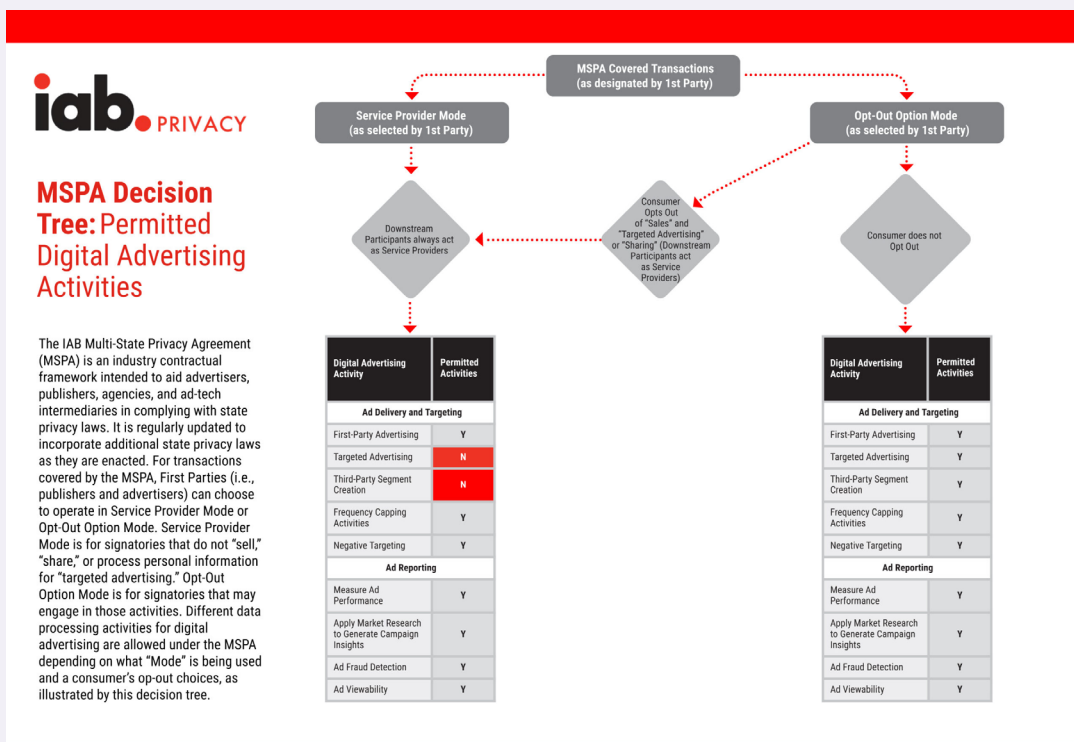
IAB MSPA

The IAB’s Multi-State Privacy Agreement (MSPA) is an industry contractual framework intended to aid advertisers, publishers, agencies, and ad tech intermediaries in complying with five state privacy laws that became effective in 2023 in California, Virginia, Colorado, Connecticut, Utah. The MSPA is not a “model contract” or a template agreement; instead, it is a set of privacy-protective terms that spring into place among a network of signatories and that follow the data as it flows through the digital ad supply chain.⁷

The MSPA provides a single, standardized framework to help businesses comply with multiple state privacy laws simultaneously rather than developing separate contracts for each state. It reduces the complexity of creating numerous custom agreements and provides a clearer path for compliance.

The figure below shows the IAB’s depiction of how this framework can be applied to digital advertising decision scenarios.

Figure 6.2: IAB’s MSPA Decision Tree



Source: IAB

IAB Diligence Platform

The IAB Diligence Platform, powered by [SafeGuard Privacy](#), is a data privacy platform that contains a set of standardized privacy diligence questions that are specially designed for participants in the digital advertising industry. It includes a vendor compliance hub that allows each company to complete the diligence materials once and share it with other IAB member and non-member companies within the Platform. Importantly, participating companies choose who they share their privacy diligence responses with when engaging in a digital ad transaction.⁸

NAI Guidance, Standards & Best Practices⁹

While the NAI discontinued its cookie-based and email-based opt-out tools as of September 15, 2025, it released a primer of Privacy Enhancing Technologies (PETs) on October 6, 2025 which explains the four methods in detail: Trusted Execution Environments (TEEs), Multiparty Computation (MPC), Differential Privacy (DP), and Zero-Knowledge Proofs (ZKPs), and discusses the trade-off between accuracy and utility with limitations.

⁷ IAB, [How the IAB Multi-State Privacy Agreement Can Help Industry Participants Meet their 2023 Privacy Challenges](#) (March 2023)

⁸ IAB, [IAB Diligence Platform](#) (Jan 29, 2024)

⁹ NAI, [A Primer: Privacy-Enhancing Technologies in Digital Advertising](#) (October 6, 2025)

VI: Evaluation Criteria

Support and Service

The level of support and service offered by an identity vendor is often underestimated. After the sale, you may need ongoing support for integration, troubleshooting, or extracting insights from the data. A vendor that offers strong customer service and meaningful consultative support can greatly enhance the value of the solution.

Keep in mind that the high-touch support comes at a cost, and the right level will depend on how your team operates. Useful criteria for evaluating support capabilities of an IDR solution include:

Onboarding & Implementation

- **Dedicated support team:** Do they provide a named account manager or implementation specialist?
- **Training resources:** Availability of documentation, workshops, or self-service portals.
- **Integration assistance:** Hands-on support for mapping data structures, APIs, and CRM / marketing automation connections.

Ongoing Customer Support

- **Availability:** 24/7 vs. business hours; global vs. regional coverage.
- **Channels:** Email, phone, live chat, ticketing system.
- **Responsiveness:** SLAs for response and resolution times.
- **Escalation paths:** Clear process for urgent issues.

Service Quality & Expertise

- **Industry knowledge:** Do support teams understand identity resolution, compliance, and marketing use cases?
- **Proactive guidance:** Regular check-ins, best practice sharing, and optimization recommendations. For example, ask how they stay up-to-date and compliant with consumer data privacy regulations around the world.
- **Customization:** Ability to tailor support to your workflows and taxonomy.

Partnership & Relationship Management

- **Strategic alignment**
 - Do they act as a partner, not just as a vendor?
 - Do they have the domain knowledge and experience of your business?
 - Do they specialize in certain areas / verticals and have unique datasets related to that market?
 - **References:** Vertical-specific case studies or client testimonials.
- **Executive engagement:** Access to leadership for roadmap discussions.
 - Review the credentials of the most relevant executive team such as Chief Product Officer or CTO.
- **Community & ecosystem:** User groups, partner networks, and forums for peer learning.
 - **Proof of Reliability.**
 - **References:** Case studies or client testimonials on service quality.
 - **Metrics:** Published support KPIs such as average response or resolution time.

VI: Evaluation Criteria

What to watch out for:

- Vague or generic support commitments (“we’ll help as needed”).
- No published SLAs or performance metrics.
- Limited communication channels (e.g., email-only).
- Lack of industry-specific expertise in identity and compliance.

Cost Effectiveness

Cost is often a heavily weighted factor in final vendor selection, but focusing solely on price can be shortsighted and ultimately more expensive. While budget constraints are real and all deals are negotiable, keep in mind that the lowest-cost option may compromise on data quality, service reliability, and long-term support. True cost-effectiveness comes from evaluating the total value delivered: high quality identity data, operational efficiency, and measurable business outcomes.

When evaluating cost, avoid focusing only on headline pricing. Request itemized pricing for maximum flexibility, especially if you are working with multiple providers across different capacities. This allows you to model total cost across various configurations, including implementation, ongoing support, incremental usage, and payment terms.

What to watch out for:

- “Free identity” claims — identity is not cheap. If there is no explicit cost for identity, the cost is baked in somewhere else, often less transparently.
- Opaque pricing models with unclear usage thresholds or unpredictable overage charges.
- High upfront integration fees without guaranteed or clear ROI milestones.
- Bundled pricing with no ability to customize modules or scale selectively.
- Contracts that lock you into long terms without flexibility.
- “Black box” resolution methods that make efficiency impossible to measure.
- Potential switching cost: high switching cost can create vendor lock-in, limiting flexibility and reducing negotiating leverage over time.



VII. An Ongoing Process

The landscape and demand for identity resolution continues to evolve, and there is no one-size-fits-all solution. Evaluating providers can feel complex and overwhelming, especially when differentiators span beyond features and pricing. This guide is designed to offer a structured approach and a way of thinking when choosing IDR partners — calling attention to both foundational requirements and often-overlooked considerations across data sources, methodology, integration, compliance, support, and commercial models.

As businesses grow, their data and identity needs evolve. Selecting an identity provider without considering scalability, adaptability, and global applicability can constrain future growth and create costly transitions later on. The ideal partner should support expansion across regions, channels, and segments, including SMBs and emerging markets, while offering flexible solutions that grow with your data maturity and business strategy.

Finally, to ensure alignment with go-to-market objectives and measurable business outcomes, it is essential to evaluate providers against a transparent, measurable, and repeatable framework. The goal is not simply to choose a vendor, but to identify a trusted partner capable of delivering accurate, actionable, and scalable identity intelligence that drives meaningful business results.

The figure below shows an example of an evaluation scoring matrix when considering multiple IDR vendors.

Figure 7.1: Evaluation Scoring Matrix

Dimension	Weight (%)	Criteria Examples	Scoring Scale: 1–5 (1 = poor, 5 = excellent)
Strategic Fit & Vision	20	Roadmap, industry expertise	1–5
Data & Technology	25	Accuracy, scalability, integration	1–5
Privacy & Compliance	20	GDPR, CCPA, security	1–5
Support & Service	15	Onboarding, training, success	1–5
Cost & Commercial Model	10	Transparency, flexibility	1–5
Trust & Partnership	10	Reputation, cultural fit	1–5

What's Next

As the identity landscape continues to evolve rapidly, several emerging technology developments reshaping how organizations will approach identity, measurement, and activation in the years ahead. Two areas in particular — **Synthetic Identity** and **Non-Human Identity (NHI)** — are becoming increasingly important for brands, publishers, and technology buyers to understand.

Synthetic Identity

Even though Google has reversed its third-party cookie deprecation plans for now, the industry's long-term reliance on cookies remains unsustainable. As consumer privacy expectations rise and identifiers continue to disappear, the AdTech ecosystem has accelerated the search for alternative, privacy-preserving ways to unify audiences and measure performance across platforms.

One emerging approach is the use of **Virtual IDs (VIDs)** — a privacy-safe, modeled approach, designed for cross-media measurement.

The concept of synthetic data itself is not new. Its origins date back to the 1940s with early Monte Carlo simulation techniques¹⁰. Synthetic data is created through mathematical models or algorithms to mirror the structure and patterns of real-world data, without relying on actual individual records. Because it is generated rather than collected, synthetic data allows organizations to control how much sensitive information is revealed and how closely it resembles real behavior.

¹⁰ The Alan Turing Institute and The Royal Society, [Synthetic Data — what, why and how?](#)

VII. An Ongoing Process

VIDs extend this principle to identity. They are non-existent digital personas that mimic the statistical patterns, characteristics, and behaviors of real people without corresponding to actual individuals. While VIDs are not part of a “real-world” identity graph, they aim to support many of the same use cases, particularly in cross-platform measurement, audience analysis, and privacy-friendly activation.

Adopting VID does not replace your existing identity graph. Instead, they serve as an additional tool, or an alternative solution, to cover the areas where your ID graph may have limited reach, e.g., walled gardens, enabling more complete measurement and addressability in a privacy-first environment.

Non-Human Identity (NHI)

Not all identities in the digital world belong to people. Increasingly, they belong to machines, systems, and software — and these non-human identities are multiplying far faster than human identities. Frost Radar reports that large enterprises now have over 17 times more machine identities than human identities, while CyberArk cites an even more dramatic ratio of 82:1¹¹. The numbers may differ, but the conclusion is the same: NHI is scaling exponentially, fueled by rapid growth in AI, automation, cloud infrastructure, APIs, and microservices. Each system requires its own credentials, permissions, and verification processes, creating an identity volume and operational complexity that traditional human-centric identity tools were never designed to handle.

The rapid growth is reshaping the market. The global NHI solutions category is expected to grow from \$5.0 billion in 2024 to \$11.1 billion in 2030, a 14.2% CAGR. North America remains the largest and most mature region, while Asia-Pacific is the fastest-growing with a projected 15.9% CAGR¹².

Yet most legacy identity and access management (IAM) tools still treat identity as fundamentally human-centric. As a result, organizations face significant and costly security risks such as identity spoofing and impersonation of machine agents, privilege creep and excessive, unmonitored permissions, compromised keys or tokens granting unauthorized access, and untracked service accounts that persist long after they should be retired.¹³

These risks are already materializing in the real world. For example, The Hacker News reported that 12,000+ API keys and passwords were discovered in public datasets used for LLM training¹⁴, illustrating how easily machine credentials can lead and proliferate beyond an organization’s control.

For brands and publishers that are new to this space, Frost Radar highlights several best-practice principles¹⁵:

1. *Give machine identities only the access they truly need*

Machine identities often accumulate permissions over time, creating unnecessary risk. Limiting each identity to the minimum required access — and only at the moment it’s needed — helps prevent misuse or exploitation.

2. *Automate management and tracking*

With thousands of tokens, keys, and service accounts operating across cloud environments, manual oversight is no longer viable. AI-driven tools can automate rotation, classification, and de-provisioning to ensure nothing is forgotten or left exposed.

3. *Embed identity security into digital workflows*

Machine identities are created inside CI / CD pipelines, Kubernetes clusters, and other automation environments. Security shouldn’t slow teams down — it should be built directly into these workflows so protection happens by default.

¹¹ CyberArk, [2025 Identity Security Landscape](#) (2025)

¹² [Non-human Identity Solutions 2024-2030](#), Frost & Sullivan (November 14, 2025)

¹³ Christian Simko for Token, [The top 10 identity-centric security risks of autonomous AI agents](#) (October 7, 2025)

¹⁴ Ravie Lakshmanan for The Hack News, [12,000+ API Keys and Passwords Found in Public Datasets Used for LLM Training](#) (February 28, 2025)

¹⁵ See footnote 12

VII. An Ongoing Process

There's little doubt that NHIs are going to expand dramatically, as AI agents, whether they are autonomous or semi-autonomous, begin to interact with websites, apps, and APIs on behalf of users. These agents behave like users: shopping, searching, requesting info, generating behavioral signals, producing device fingerprints, initiating sessions, and triggering events. As such, they introduce a new class of identity entities that behave like users but are not human. Over time, these entities will need to be recognized, validated, managed, and incorporated into identity graphs, much like any other participant in the digital ecosystem. In practical terms, identity graphs must evolve to recognize and model these NHIs, distinguishing them from humans while understanding their roles, relationships, and interactions across the digital ecosystem.

Identity Resolution as a Strategic Differentiator

Selecting the right identity partner is not about chasing the lowest cost or the flashiest claims — it's about aligning with a provider that delivers accuracy, interoperability, and trust at scale. As the ecosystem evolves toward privacy-safe solutions such as VIDs, cross-media measurement frameworks, and AI agents acting autonomously across digital environments, marketers must prioritize partners who can safeguard long-term investments while driving measurable business impact today.

Evaluating partners through a lens of total value: data quality, activation capabilities, compliance rigor, and long-term adaptability, ensures that identity resolution becomes more than a tactical tool. It becomes a strategic differentiator.

Ultimately, identity resolution is no longer just a data exercise, it's the foundation of trust, performance, and future-ready growth in the modern AdTech landscape.



Appendix A — Glossary

Activation

The process of deploying audience data to deliver targeted media across channels, connecting the right message to the right person at the right time.

CCPA (California Consumer Privacy Act)

A state-wide data privacy law that regulates how entities handle the personal information of California residents.

Data Clean Rooms

Secure, controlled environments that allow multiple parties to collaborate and analyze sensitive data in a privacy centric way without sharing or exploring the raw data.

DPIA (Data Protection Impact Assessment)

A systematic process used by organizations to identify and minimize the privacy risks associated with processing personal data. It involves describing the processing, assessing its necessity and proportionality, evaluating potential risks, and determining measures to mitigate them. DPIAs are required under regulations like the EU's [GDPR](#) and help ensure compliance, accountability, and the protection of individual rights.

GDPR (General Data Protection Regulation)

The privacy regulation that governs how the personal data in the European Union may be processed and transferred.

HEM

Hashed email, which translates a known email address (PII) through an irreversible hashing algorithm, e.g., SHA-1 or SHA-2, to an anonymous customer identifier.

Identifier

A unique value assigned to a unit, which could be a device, an account, a person, a household, or other micro cohort that is defined by the identity provider. The common examples include but not limited to device IDs, mobile ad IDs (MAIDs), emails, hashed emails (HEMs), user names, names, phone numbers, postal or IP addresses.

Identity Graph

An identity graph maps various identifiers such as devices, logins, emails, names and addresses to the same unit, creating a cluster that enables advertisers and publishers to effectively plan and target relevant audiences, and measure the impact of the marketing efforts.

Interoperability

The ability of different platforms, tools, and systems to work together seamlessly — enabling data exchange, audience activation, and campaign orchestration across the ecosystem.

IPV4 (Internet Protocol Version 4)

A fundamental Internet Protocol in data communication that uses 32-bit addresses to identify and route devices on a network. IPV4 provides the logical connection between network devices by identifying each device. There are many ways to configure IPV4 with all kinds of devices, including manual and automatic configurations, depending on the network type.

Appendix A – Glossary

IPV6 (Internet Protocol Version 6)

The latest version of the Internet Protocol, which allows computers to uniquely identify and locate other computers and devices on the Internet. Using 128-bit addresses vs. IPV4's 32-bit, IPV6 was created by the Internet Engineering Task Force (IETF) to replace IPV4, supporting 340 undecillion unique addresses vs. IPV4's 4.3 billion. Apart from the larger address space, IPV6 provides a more efficient packet handling mechanism, better security and improved performance.

MAID

Mobile Ad ID.

MMP

Mobile Measurement Partners.

Onboarding

Process to connect offline or CRM data to online identifiers such as MAIDs and cookies to enable digital activation, strengthens customer databases. Onboarding typically involves three steps:

- Uploading data: anonymize and upload your first-party data to your onboarding partner's system.
- Matching data: match the data you uploaded with specific online identifiers.
- Activate data: Create addressable audience segments for targeting digitally.

Salt

In hashing, salt is the “extra entropy” (the extra pinch of salt) added to something that is already entropic (a fancier word for random). A “salt” is a unique, randomly generated piece of data added to a password before it's hashed. This process makes each password's hash unique, even if the original passwords are the same. The salt is stored alongside the resulting hash.

Universal ID

A privacy-compliant digital identifier that persists across websites and devices, e.g., UID 2.0, ID5 ID, RampID, Panorama ID, Yahoo ConnectID, Criteo ID, Fabrick ID, Prebid SharedID.

Appendix B – References

- Acxiom, *See Your Customers as Real People: Your Guide to Choosing an Enterprise-Wide Identity Solution*, (2023).
- Aleman, Dolores, Frost Radar: [Non-Human Identity Solutions, 2025](#), (November 2025).
- Amazon Web Services (AWS), [Measuring the Accuracy of Rule- or ML-Based Matching in AWS Entity Resolution](#), (September 29, 2025).
- Auth0, [Customer Identity Trends Report 2025](#), (2025).
- Avatier, [Selecting the Right Identity Vendor: 12 Critical Evaluation Criteria to Future-Proof Your IAM Strategy](#), (July 8, 2025).
- Byrd, Andrew, [ID Bridging Explained: Benefits, Controversies, and the Battle for Transparency in Digital Advertising](#), (June 13, 2024).
- CVAT, [Is Your Training Data Trustworthy? How to Use Precision & Recall for Annotation QA](#), (September 12, 2025).
- CyberArk, [2025 Identity Security Landscape](#), (2025).
- Experian, [Identity Resolution Guide](#), (2023).
- GDPR.eu, [Data Protection Impact Assessment \(DPIA\)](#), (2024).
- Heath, Eli for Epsilon / Lotame, [Top Questions to Ask When Evaluating Identity Resolution Vendors](#), (June 12, 2024).
- IAB, [How the IAB Multi-State Privacy Agreement Can Help Industry Participants Meet Their 2023 Privacy Challenges](#), (2023).
- IAB Tech Lab, [Identity Solutions Guidance](#), (2023).
- IAB Tech Lab, [Online Advertising and Ad Tech Glossary](#), (2024).
- Lakshmanan, Ravie, “[12,000+ API Keys and Passwords Found in Public Datasets Used for LLM Training](#),” The Hacker News (February 28, 2025).
- LiveIntent, [Identity Resolution: Everything You Need to Know](#), (2024).
- LiveRamp, [Identity Explainer](#), (2024).
- Lead Genius, [7 Critical Factors in Evaluating a Data Provider](#), (May 23, 2024).
- Monte Carlo, [What Is Data Accuracy? Definition, Examples and KPIs](#), (July 11, 2023).
- Network Advertising Initiative (NAI), [A Primer: Privacy-Enhancing Technologies \(PETs\) in Digital Advertising](#), (October 6, 2025).
- Non-Human Identity Management Group, [Non-Human Identity Threats in Cybersecurity](#), (2024).
- Parker, Pamela, “[24 Questions to Ask Identity Resolution Vendors During a Demo](#),” Martech (March 27, 2024).
- PwC, [PwC’s 2024 Trust Survey: 8 Key Findings](#), (March 12, 2024).
- Roqad, [Identity Resolution Buying Guide](#), (2023).
- RTInsights, [Why Managing Non-Human Identities \(NHIs\) Must Be a Central Concern for Identity Access Management](#), (August 29, 2025).
- Salesforce, [Data Cloud Bootcamp: Identity Resolution](#), (2024).
- Simko, Christian, [The Top 10 Identity-Centric Security Risks of Autonomous AI Agents](#), (October 7, 2025).
- The Alan Turing Institute & The Royal Society, [Synthetic Data – What, Why and How?](#), (2024).

Appendix B – References

- Throtle, [*Your Ultimate Checklist for Identity Partnership Success*](#), (July 31, 2024).
- TransUnion, [*New TransUnion Research Reveals Seven in 10 Marketing Leaders Struggle to Connect with Audiences Due to Complex Network of Martech Solutions and Identity Challenges*](#), (January 8, 2025).
- Verisk Marketing Solutions, *Identity & Attribute RFI Starters*, (2024).
- Winterberry Group, [*Outlook for Identity in Advanced Television: Challenges and Opportunities*](#), (June 2022).



cimm

Coalition for Innovative
Media Measurement

