Clean Room Primer: A Practical Guide for Marketers, Retailers, Agencies, and Publishers
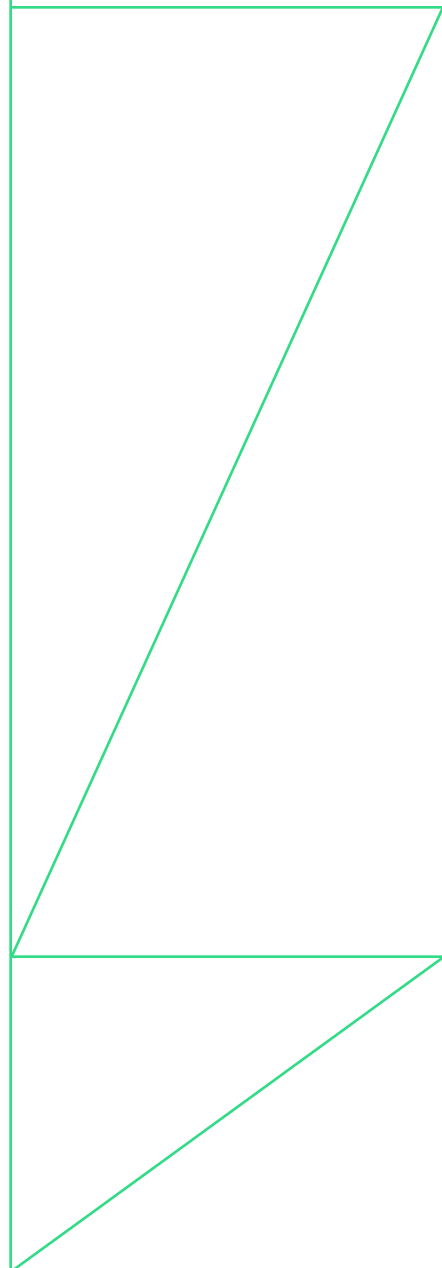
Brought to you by

/LiveRamp   The B2B Marketing Co.

Contents

Introduction

Now more than ever before, marketers are being asked to do more with less: tie more sales to ad spend, make do with one analyst instead of a team, and make a stronger case for every component of the tech stack or risk further budget cuts. It's a lot to take in when brands may need to spend up to 200% more to maintain healthy campaign metrics.

Underpinning each of the challenges mentioned above—and others that may be churning through your mind—is a lack of access to data that can help you deliver the business results you need, transforming marketing from cost centre to revenue driver. No company alone has all the data it needs to answer key questions on business results and shift the narrative when answering what marketing delivers. That's where collaboration through a data clean room comes in.

This primer shares everything you need to know about clean rooms—definitions, common use cases, and how to get started—so you can initiate new projects and move forward.

Common terms and definitions

To encourage productive discussions around clean rooms, the following is a list of foundational concepts, terms, and definitions.

1. Data clean room

A neutral and secure data environment that enables multiple parties to collaborate using proprietary data in compliance with privacy regulations, while allowing each participant to maintain privacy and control of its data. By matching (vs. sharing) data, a clean room enhances data value and usability, delivering an extra layer of intelligence not available from individual data sets in a safe and secure environment.

- **Data matching:** The process of comparing different datasets to find commonalities that allow participating parties to align the underlying data. The resulting dataset can be evaluated for duplicates and mined for insights, and/or resolve individual identities.

- **Customer data clean room:** A subset of data clean rooms designed to match the participating parties' customer data sets to strengthen their marketing and media efforts. In the marketing/advertising industry, data matching typically occurs between a brand, agency, and/or a publisher (content media or retail).

- **Enhanced data clean room:** This type of data clean room incorporates embedded people-based identity, measurement, and insight applications and is also connected with leading activation channels, which can help you build deeper partnerships and augment your sphere of influence.

Importantly, while clean rooms are designed to support privacy protections, each participant remains individually responsible for ensuring that its data collection and usage practices comply with applicable regulations.

2. Personally identifiable information (PII)

Information about an individual that, when used alone or with other relevant data, can reveal that person's identity. This data is generally provided by the end user. PII characterisations can vary by platform or governmental entity and evolve over time.

- PII may contain direct identifiers that can identify a person uniquely (name, address, passport information) or attributes shared with others (race, gender, date of birth) that can be combined to recognise an individual.
- Opt-in data refers to data in which individual customers are required to give permission before their data can be shared.
- Opt-out data refers to data in which individual customers withdraw consent for data processing.
- Non-PII first-party data is data that cannot be directly traced back to identify a person. It may include masked IP addresses, aggregated product or service users, device type, browser type, plug-in details, language preferences, time zone, and/or screen size.

3. Identity data

This is a term we will use later in the primer and wanted to provide a definition for reference, as follows:

- PII that a company owns, based on connecting two or more pieces of information that can be used to identify an individual.

4. Identity resolution

Building a view of an individual or household by matching records using PII and other identity data. Resolution can be built within a company's own dataset, with a platform's own set of data as a baseline, and/or with multiple companies' data-sets, using three distinct analytic approaches to build a view of an individual:

- **Deterministic matching:** Matching customer data on exact PII matches
- **Probabilistic matching:** Inferred customer data matching using both PII and non-PII data
- **Integrated matching:** Blending deterministic and probabilistic matching approaches to provide the broadest and deepest customer insights

Probabilistic matching and integrated matching are often accompanied by an accuracy rating to show the quality of the data match, usually expressed as a percentage. For example, Joe Smith and Joseph Smith residing at the same address with the same email address are a better match and would get a higher accuracy rating than two Joseph Smiths with different addresses.

5. Differential privacy

Processes and protocols that protect the underlying individuals or households from re-identification, reconstruction, or tracing. Those problematic practices are:

- **Re-identification:** Reversing the step of anonymisation or pseudonymisation to track an individual record back to a specific person
- **Reconstruction:** Uncovering identifying attributes for individuals in a specific data set using non-PII
- **Tracing:** Using data science to determine whether an individual is a member of a specific dataset

6. Publisher types

For the purposes of this primer, publishers will be referred to in two categories:

- **Content media networks:** Advertising-supported businesses that produce, deliver, or otherwise showcase original or archived content such as articles, videos, commentary, etc.

- **Retail media networks:** Retail or commerce businesses that sell advertising as a supplement to their commerce business, including advertising on their owned media as well as media purchased from other publishers.

7. Data monetisation

In the context of clean rooms, data monetisation refers to selling access to data to one or more partners that seek to drive customer behaviour or gain greater insights in a privacy-centric manner.
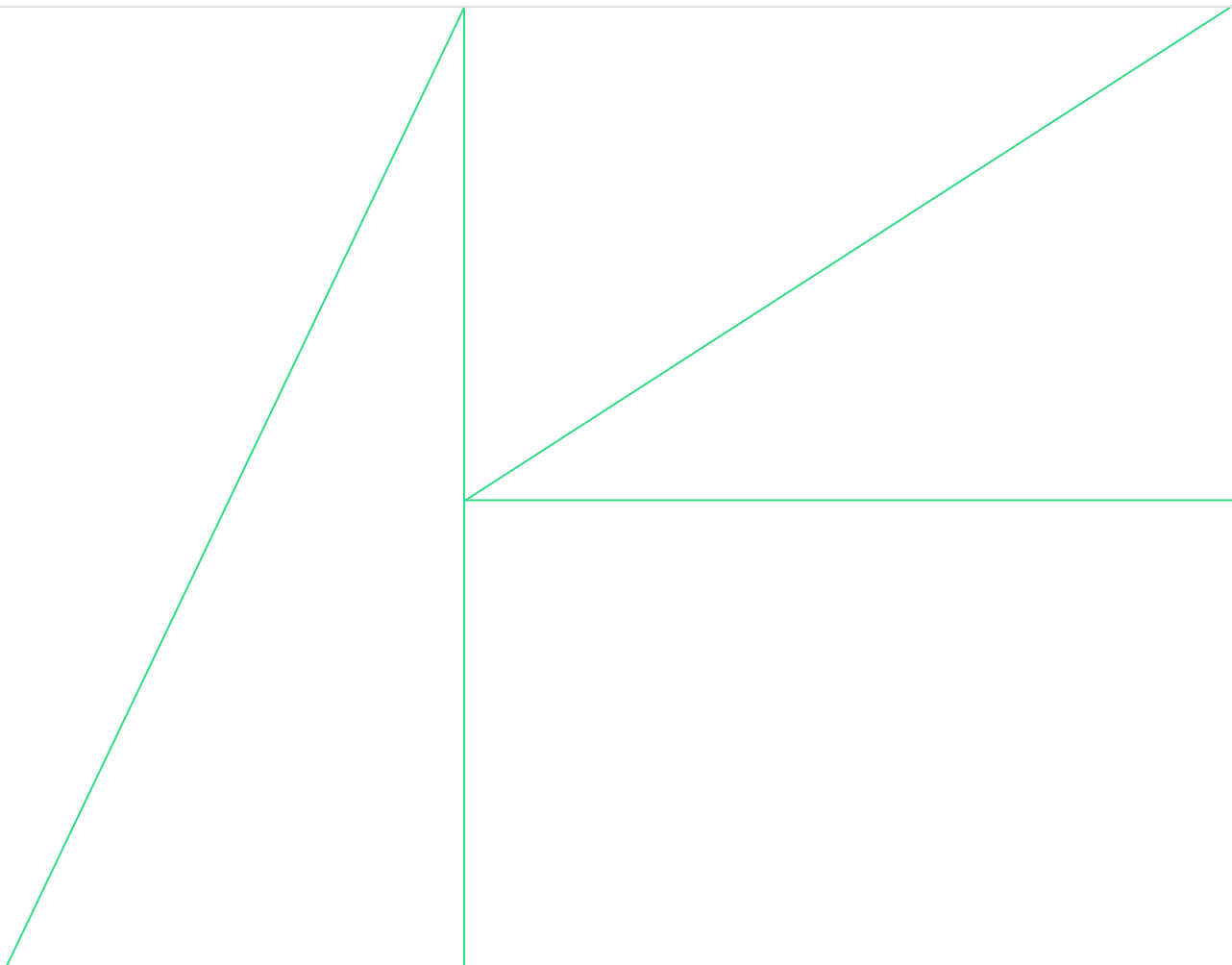
Use cases

While the core use of clean rooms in the marketing and advertising space is to match customer data across partners, the value of clean rooms is derived from specific use cases. Here are three of our industry's most common use cases:

1. Campaign measurement and attribution

Clean rooms must match customer and outcome data accurately and comprehensively to provide meaningful measurement and attribution. This includes the ability to:

- Effectively connect first-party data with data from walled gardens and publishers

- Connect campaign and conversion data in a privacy-centric environment

- Combine standard and experimental attribution and measurement methods

- Compare first-party data with available merchant data for more granular insights

As a result of new privacy standards, measurement and attribution approaches enabled by clean rooms must embrace a reality in which data remains

[continued]

distributed across multiple environments and identity remains fragmented, or even aggregated into micro-cohorts, while still offering these capabilities:

- **Channel-level attribution:** The ability to incorporate data from all marketing channels so that brands can adequately provision and configure their clean rooms to collaborate with strategic partners.

- **Cross-channel experimentation:** The means to automate A/B testing and cross-channel experiments so that marketers can accelerate their ability to test hypotheses and prove causation.

- **Media mix model enhancement:** The ability to train and calibrate models using the outputs of channel-level attribution analysis and cross-channel experiments, as well as incorporate "macro" datasets (e.g., econometric).

In addition, clean rooms should have the ability to automate data integration and allow for querying across all environments to facilitate real-time measurement and maximum flexibility.

2. Audience insights, planning, segmentation, activation, and analysis

Clean rooms permit the building of richer and deeper datasets by connecting first-party datasets that would otherwise be inaccessible due to changes in privacy laws and restricted data access. This improves the parties' ability to develop privacy-conscious, trafficable plans and define target segments for audience activation and insights before, during, and after campaigns run. For privacy reasons, the initiating party can only see its own data with the matched insights provisioned by its partner(s), not the underlying data itself. Clean room audience-based applications generally fall into the following categories:

- **Attribute-based:** Enriching consumer profiles/identity by leveraging another party's data, such as transactions, first- or second-party owned-and-operated media engagement, offline third-party data, customer preferences, and partner media interactions

- **Performance-based:** Using transaction data to segment audiences based on how and what drives views, clicks, conversions, and other actions

- **Lookalike-based:** Extending audiences via models that use attribute- or performance-based criteria to identify prospective customers who look and act like existing customers

Once audience segments are identified and built, they are assigned an anonymised audience identifier that can be deployed to other digital marketing tools for activation.
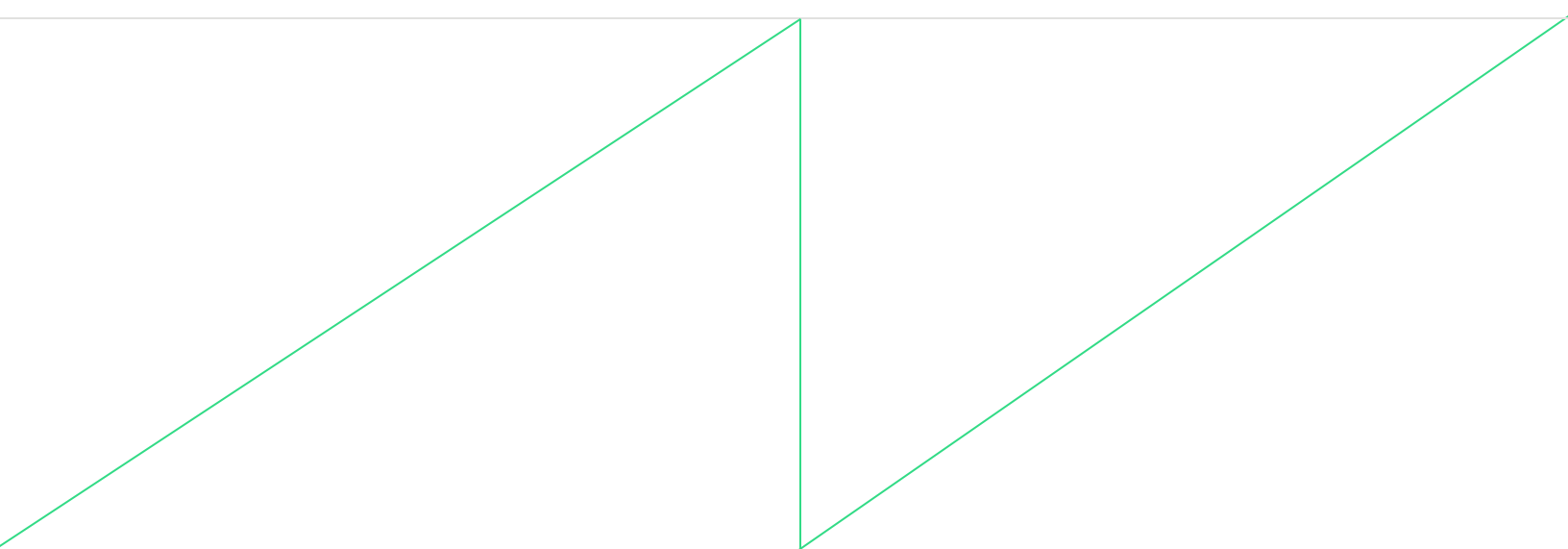
3. Safer monetisation of first-party data

Because of new privacy-focused regulations, business rules limiting the sale and use of third-party data, the rise of walled gardens, and the changing legislative landscape, finding safe new ways to monetise first-party data has become a core clean room use case for companies rich in first-party data. Clean rooms facilitate the monetisation of first-party data, since they are designed with security provisions that protect consumer data in accordance with privacy regulations and business rules in a safe and secure environment.

As in any marketplace, the ability to sell one's "merchandise" — in this case owned and managed data — is contingent on the amount and perceived quality of the product being sold. The value of a set of first-party data is contingent on its ability to:

- **Optimise personalisation** to drive customer engagement, interactions, loyalty, and conversions
- **Enhance audience segmentation and improve marketing performance** using supplemented data
- **Determine campaign return on investment (ROI)** for marketing and media plan optimisation and return on ad spend (ROAS) analysis via richer data sets

While investing in clean rooms can also require an investment in tools and time, the added revenue from sales or improvement to marketing performance can help offset the investment cost. Additionally, clean rooms provide adherence to any new data management rules or legislation.

Data management dynamics

Partnerships for matching data in a clean room will vary by use case:

- Enhanced audience insights may involve the following integrations:

    - A marketer's customers with a publisher's audience(s)

    - Two (or more) marketers' customer bases

    - A marketer with a data provider, DMPs, CDPs, and clean room platforms

- For campaign measurement and attribution, partnerships occur between a marketer or its agency* and a publisher to provide details on media delivery at the person level (impressions, reach, etc.) to understand what drove outcomes and at what cost without exposing underlying data.

- Monetisation of first-party data may include any of the above partnerships.

*A marketer may choose to have their own data lake or to have all or part of their data housed in an agency data lake.

How a data clean room works

- The diagram here illustrates components of a marketer and publisher partnership:

- The types of data, data flow, and interactions that may take place in a clean room environment

- What each party may see as a result of data matching

- Contractual needs. Each partner should have a contract with:
  - » The clean room platform to outline terms of usage
  - » Any other party sharing data to specify business terms, use cases, access, and outputs
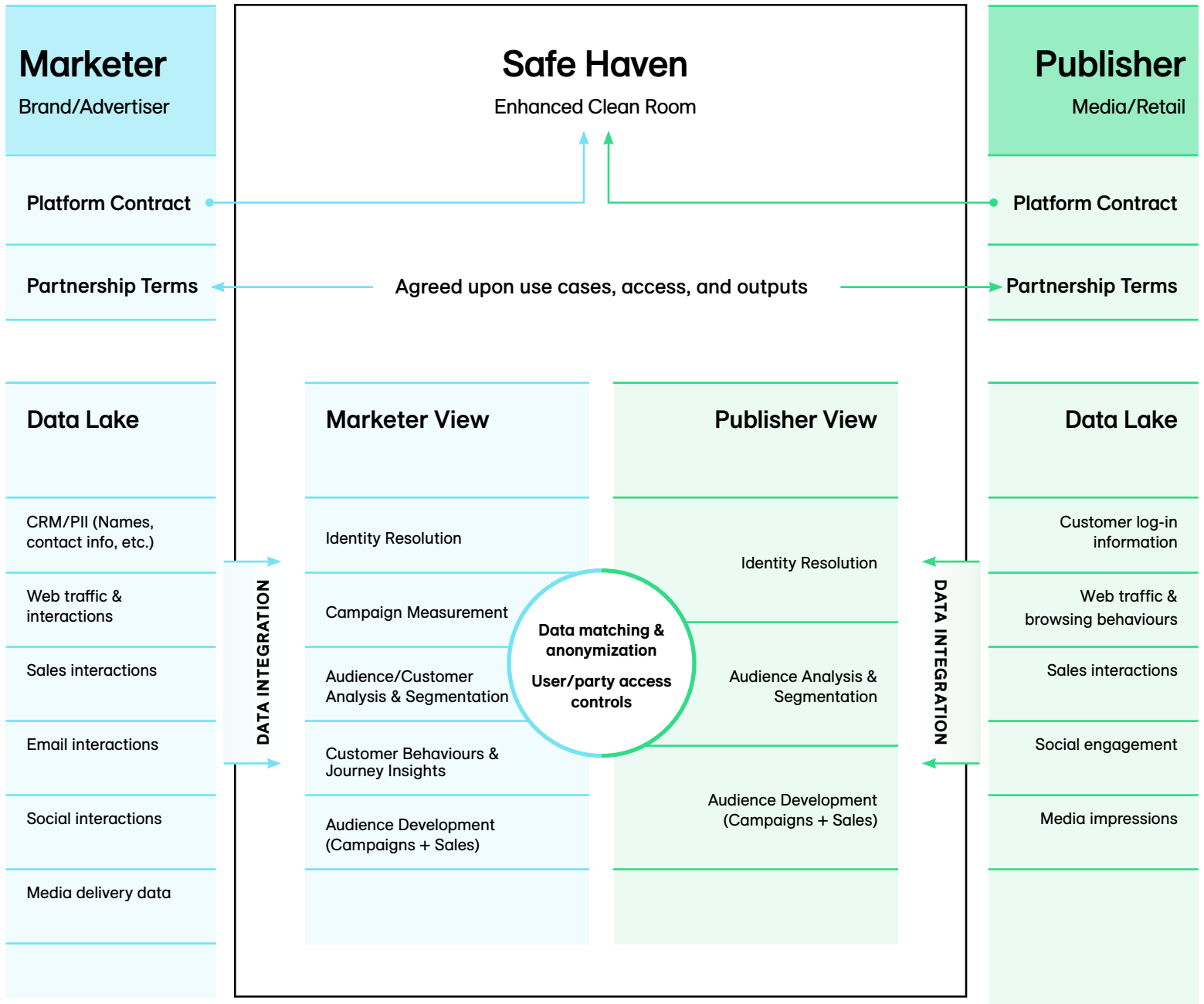
Similar dynamics will occur for other clean room partnerships, even though exact circumstances may vary.

"A clean room offers a privacy-conscious and cookieless way for partners to collaborate on valuable customer data in a safe and secure environment."

Veronica Luik
Director, Strategy and Market Intelligence, LiveRamp

How a Data Clean Room Works*

Marketer
Brand/Advertiser

Safe Haven
Enhanced Clean Room

Publisher
Media/Retail

Platform Contract

Platform Contract

Partnership Terms — Agreed upon use cases, access, and outputs → Partnership Terms

Data Lake

CRM/PII (Names, contact info, etc.)

Web traffic & interactions

Sales interactions

Email interactions

Social interactions

Media delivery data

DATA INTEGRATION

Marketer View

Identity Resolution

Campaign Measurement

Audience/Customer Analysis & Segmentation

Customer Behaviours & Journey Insights

Audience Development (Campaigns + Sales)

Data matching & anonymization

User/party access controls

Publisher View

Identity Resolution

Audience Analysis & Segmentation

Audience Development (Campaigns + Sales)

DATA INTEGRATION

Data Lake

Customer log-in information

Web traffic & browsing behaviours

Sales interactions

Social engagement

Media impressions

* What is included in a marketer's or publisher's data lake and what each party can view will vary in different situations

Data management principles

Clean room vs. data lake (vs. CDP vs. DMP)

In many cases, clean rooms are mistakenly conflated with data lakes.

Relative to a clean room, a data lake is:	Relative to a data lake, a clean room is:
• A data storage environment controlled by one company	• A data collaboration platform designed to operate with partners
• Focused on storage and hosting of that company's data sets (organising datasets across different sources occurs primarily for internal use)	• Focused on matching data sets across partner companies in a neutral environment that protects the privacy of each party's data

Data lakes and clean rooms are meant to work together. While some companies may offer both capabilities, they are separate products that do not need to live under the same platform to be successful and impactful.

Both customer data platforms (CDPs) and data management platforms (DMPs) offer collaboration as a capability. However, DMPs and CDPs are designed to serve the purposes of their owners, while clean rooms are designed to be neutral. Therefore, the independence of clean rooms protects data privacy for the purposes of data matching better than DMPs or CDPs.

Privacy vs. security

While privacy and security are often used interchangeably, they are not the same:

- **Data security** relates to keeping personal and sensitive data safe from unauthorised access. This is now common practice: companies are expected to employ updated technology and follow legal and ethical guidelines when handling data. This is the protection of the data itself.
- **Data privacy** involves protecting data while it's in use (vs. at rest or in transit), which includes data sharing, collecting, storage, and regulatory compliance. This is the protection of identity.

Data-matching protocols

Data-matching techniques provide the foundation for clean rooms to successfully join partners' data. Interoperability and scalability are the key criteria used to assess the likelihood of a successful data match:

- **Interoperability:** How well a platform interfaces with clients, partners, and different clean room platforms
- **Scalability:** The reach a platform offers based on the quantity and quality of identity data used in data matching across partners
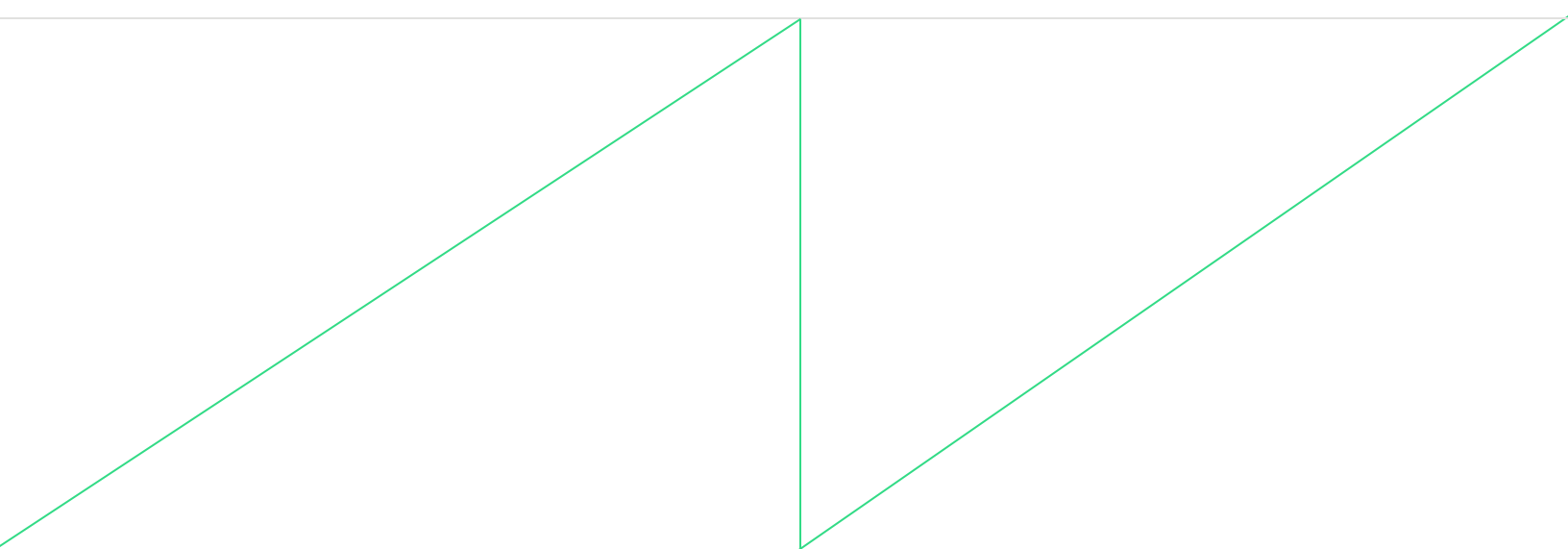
First-party data preparedness

Before investing in clean rooms, brands, agencies, and publishers should consider data quality and the specific use case:

- A high percentage of opt-in PII data within a customer base improves the odds for successfully matching data. Importantly, the opt-in PII data should include at least two deterministic characteristics (e.g., name and email address) to increase accuracy.
- Clean room participants should assess how their first-party and partners' data can collectively achieve their use case. For example, if a brand wants to attribute conversions to media spend effectively, it needs to be able to source a meaningful amount of transaction data from the various channels through which it sells, e.g., in-store, e-commerce, third-party e-commerce/distributors, and retail partners.

People and process preparedness

The opportunities offered by clean rooms can only be realised if all internal and partner team members share an understanding of basic terms, protocols, and desired use cases. As you invest in education and training, keep the following in mind:

- **Multi-department participation:** Data teams should include other departments that may assess and mine the insights from data matching throughout the clean room process (marketers, analysts, etc.). Even if they are not active in every step, these other teams should be available to weigh in as critical decisions are made, e.g., use cases, success metrics, segmentation, and reporting.
- **Clear communication:** Technical and business-side participants involved with the integration or use of the data should understand and employ consistent terms, standards, and processes.

A clean room road map for success

Moving forward in a changing landscape can be challenging. To facilitate clean room adoption in the face of an increasing number of third-party data usage limitations and privacy regulations, this section covers how marketers, agencies, and publishers can prepare for the time when they will need to rely heavily on first-party datasets to remain competitive, including:

- Guidance on selecting potential partners
- Addressing common misperceptions
- Steps to increase the likelihood of a positive outcome

What to look for in a clean room partner

Here are three things to consider when selecting enhanced data clean room technology:

1. Must be deterministic.

Enhanced privacy capabilities are often the first reason to gravitate toward a data clean room. Privacy-enhancing technologies (PETs) enable companies to analyse data without exposing it. However, the key to success for data clean rooms is to provide PETs which are modular so customers can leverage the required level of privacy thresholds when collaborating with their partners. In addition, having a deterministic approach to identity and matching is also key to ensuring better collaboration. Without deterministic IDs and matching, the outcomes of collaboration can lead to inaccurate outcomes.

In probabilistic matching, a profile of a person and their neighbour may be close enough that both receive the same advertisements. However, deterministic matching within an enhanced clean room environment can provide more exact targeting. Brands can see minute differences between the person and their neighbour, enabling one to be suppressed from targeting. This not only saves the brand time and money, but also provides a better customer experience.

An enhanced data clean room with deterministic matching powered by people-based identity also ensures you are able to accurately measure the impact of your marketing campaigns.

Leveraging an interoperable identity framework like LiveRamp's RampID can not only help you measure effectiveness more accurately but also can help drive personalized experiences across omnichannel touch points.

When it comes to measurement, the right enhanced data clean room provides similar capabilities as walled gardens, such as:

- Matching impression data to transactions to deliver conversion and sales lift reporting
- Enabling multiple partnerships for multi-touch attribution
- Incorporating TV partners to better plan omnichannel campaigns

2. Answers the right questions.

Enhanced or not, a data clean room would be moot if it didn't enable easier data collaboration within the privacy-conscious space. Once the data is accessible,

clean rooms should also provide additional capabilities to make collaboration easier and simpler.

At LiveRamp, we believe in helping brands collaborate better by:

- Helping build and reconcile their own first-party datasets to support first-party graphs
- Ensuring brands are bringing datasets into a unified taxonomy to support easy collaboration
- Pre-building queries across datasets to answer key business outcome questions without heavy lift from both parties
- Embedding analytics dashboards to ensure granular data can be easily leveraged to create better audiences at scale
- Embedding measurement protocols to automatically and continuously measure the effectiveness of collaborative campaigns

3. Must enhance personalisation.

Data is the key to great customer experiences, and an enhanced data clean room helps unlock those experiences.

Let's use frequent visits to your neighbourhood coffee shop as an example. By the end of the first or second week of daily visits, you would likely expect a more personalised experience from your regular barista, given they would know your specific preference for coffee. If nothing else, you would hope to be recognised. You never know, you might even get a pastry on the house as a thank you for being a loyal customer.

Another example would be retailers with transaction data that tell you what customers are buying, how much they are buying, and how often. That retailer could build audiences that are attractive for CPGs who may be first-party data poor. This is a win-win scenario for both parties, since the brand would achieve better targeting, reach the right audiences, and reduce their advertising waste. Meanwhile, the retailer (in this case acting as a media publisher) could improve yields and deliver better customer experiences.

Data, when leveraged properly in a privacy-safe manner, can do more to enhance the customer experience, ensuring that brands can recognise these preferences and deepen relationships with existing customers for increased brand loyalty and awareness. An enhanced clean room delivers on the ability to create exceptional customer experiences by helping activate these insights across marketing applications.

Common misconceptions

One-size-fits-all

Companies should anticipate using multiple clean rooms because separate data integrations may be needed with certain partners. For example:

- Walled gardens like Amazon, Google, and Meta require proprietary clean rooms for access to their customer data

- A partner may want a company to work with the clean room platform it already uses

Everyone sees everyone

In a truly decentralised clean room, no party should have access to any other party's underlying data sets. This includes the data clean room operator, who should remain neutral and restricted from accessing other parties' data for any reason other than technical support.

A clean room should offer specific measures to prevent inappropriate data access:

- Privacy and governance controls that enable every participant to manage data access and permissions with autonomy, including requesting and approving permissions regarding who can connect to which datasets for a specific purpose and over what time period

- End-to-end data encryption

- Contractual terms that specify privacy and governance controls among participants

- Differential privacy techniques to eliminate risk of re-identification or leakage

All consumer data is equal

Within clean rooms, one party will often have a richer level of first-party data than the other. This is not problematic as long as both parties:

- Review the types and volumes of their first-party data, the desired use cases, and potential audience applications before investing in a clean room integration

- Understand the likelihood of 100% deterministic matching and the implications of using an integrated model based on the accuracy of the underlying probabilistic matching; this will inform how each party can work with its view of the matched data

Readiness by use case

It's important to discuss these points up front to decrease frustration among partners. For each potential use case, a company should clearly understand the following seven elements to define its needs, establish guardrails, set expectations, and facilitate progress:

1. What type of data it has internally.

2. Desired data from partners.

3. Data that might be absent internally and across partners.

4. Budgeting needs and options.

5. Who needs to be involved in the process.

6. What, if any, training should be implemented.

7. Which standards need to be put in place internally.

In addition, before matching data, each company involved should establish an internal protocol for constructing and managing its identity data.

Campaign measurement and attribution

It's important that marketers, agencies, and publishers with a vested interest in campaign measurement and attribution understand that campaign results will differ from those seen in their prior analyses, because new datasets have been introduced. Specifically, matching data with a new partner will change the baseline data used for benchmarking.

As a result, before working with clean rooms, each participant should anticipate how the differences in the datasets that emerge from data matching will affect campaign measurement and attribution. This involves assessing:

- The data that will be used to build audiences going forward vs. what was used previously

- Which customer and/or audience-tracking methodologies will be used to measure results for KPIs, including the impact on pre- and post-campaign benchmarks

- How campaign reporting and insights will be impacted by new data and a different user interface

Audience insights, planning, segmentation, and activation

Each marketer, agency, and publisher using—or contemplating using—a clean room needs a firm understanding of how it applies to audience data and segments internally and externally, and which outcomes are desired from working with a data clean room. Each party will benefit from having:

- An existing library of audience segments using currently available datasets, which may include:
  - » Media consumption habits (types of content, engagement/attention, etc.)
  - » Actions (downloads, registration, sharing, etc.)
  - » Purchases (historical transactions on products and categories)
  - » Intent (specifically modeled behaviors that indicate a willingness or interest to purchase)
  - » Demographics (age, gender, presence of children, income, etc.)
- A list or examples of what additional data might be useful to apply to existing audience segments; desired new audience segments; and desired incremental insights, analytics, and reports
- A list of all technical and operational destinations where audience data is or might be used, such as advertising platforms, email service providers, CRM marketing tools, media planning models, etc.

Monetisation of first-party data

Monetising first-party data with minimal friction requires that data experts work with legal, IT, marketing, and sales teams to ensure there's internal alignment on how first-party data is sold, how it shows up in the marketplace, and how it's protected from data theft and privacy breaches. This includes establishing the following prior to going to market:

- A business case for selling first-party data that includes a list of prospective buyers and anticipated revenue
- Guidelines for how first-party data can be sold and applied by others to protect the value and privacy of that data

Marketing clean room use cases

1. Campaign measurement and attribution.

A common challenge for advertisers is accurately measuring the effectiveness of marketing campaigns to report on ROI. A clean room can act as a neutral environment to analyse both advertiser data and ad exposure data provided by the marketing platforms. Partnerships with various publishers, including walled gardens, through their clean rooms enable brands to measure all their media buys using neutral methodologies.

2. Collaboration.

Data clean rooms enable varying degrees of collaboration between providers and potential users of the data. Marketers can use clean rooms to set up collaboration efforts with their partners, for example between retailers, suppliers, hospitality groups, partner agencies and airlines.

3. Increased targeting and reach opportunities.

Through clean rooms, marketers can expand audiences, fill in customer details to build more complete profiles, and leverage first- and second-party data to identify and enhance target audiences.

4. Campaign optimisation.

With audience insight data, marketers can better segment audiences based on specific traits, habits, history, etc. for improved targeting and higher ROI.

5. Expanded insights.

A data clean room provides marketers with a safe and secure space to match first-party audience data associated with their advertising and marketing efforts with the transactional audience data from retail partners. Marketers gain insight into which campaigns drove purchases and can use that insight to refine targeting, build segmentation, and optimise ad spend.

How to get started:

1. Identify the types of insights from shared data that would most benefit your business.

2. Determine your potential data collaboration partners, either internally or externally.

3. Explore and select the data clean room solutions that best fit your needs.

Retailer clean room use cases

1. Expand audiences.

Through first-party data matching, brands can plan, activate, and measure advertising campaigns, both across the retailer's own audience and across other media audiences.

2. Collaborate without losing control of your data.

Through data clean rooms built with privacy in mind, retailers can safely and securely explore data collaboration opportunities and securely share transaction data all while maintaining full control and eliminating risk concerns.

3. Establish a retail media network.

Retail media networks (RMNs) allow for collaboration across partners and the advertising ecosystem. In order to get started, a retailer's first-party data must be uploaded, organised, and protected.

How to get started:

1. Set goals and priorities for developing a retail media business.

2. Ensure your first-party data is organised and protected.

3. Organise your current assets and capabilities.

4. Source partners or advertisers that will be a good fit for the network.

Agency clean room use cases

1. Increased targeting and reach opportunities.

Through clean rooms, agencies can help their clients expand audiences, fill in customer details to build more complete profiles, and leverage first- and second-party data to identify and enhance target audiences

2. Enhanced campaign planning and measurement.

Clean rooms allow advertisers and agencies to safely and securely combine formerly disparate data sets for more accurate measurement and sharper insights that can be used for planning or optimisation.

3. Collaboration.

Agencies can use clean rooms to set up collaboration efforts for their clients who wish to test and learn with a few use cases before going on their own clean room journeys.

How to get started:

1. Identify the business case and time-bound experiments you can run to achieve quick wins (crawl > walk > run methodology)

2. Find the right partner aligned with your business needs and overall approach.

3. Set the rules of engagement to ensure your partnership meets your requirements for provenance, governance, and permission.

Publisher clean room use cases

1. Keep data safe and secure.

Publishers need to keep their data safe so those interested in collaborating should consider doing so in a clean room where the rules of engagement for data use are clear and permissions are set for who can see and do what within this secure environment.

2. Increase engagement with advertisers.

Publishers have comprehensive knowledge of their readers and viewers through first-party data. Clean rooms can help make that data valuable for partners and advertisers, increasing ROI and trust in the quality of your audience.

3. Monetisation.

Clean rooms create an opportunity for publishers to connect data with adver-tisers for possible monetisation efforts—all without losing control of their data.

How to get started:

1.   Organise your first-party data.

2.   Determine potential data collaboration partners and advertisers.

3.   Match data to partners and/or advertisers.

4.   Set permissions to ensure control of data.

About LiveRamp

At LiveRamp, we know that simplicity is always welcome—no matter where you are or what you're responsible for. We live in a complex ecosystem where it feels easy to become disconnected from core audiences, and we help our partners gain the connection and control they're seeking by enabling them to gain permissioned access to data, wherever it resides.

If this sounds good to you, **we'd love to talk**.

www.liveramp.uk