

The Definitive Guide to Data Clean Rooms

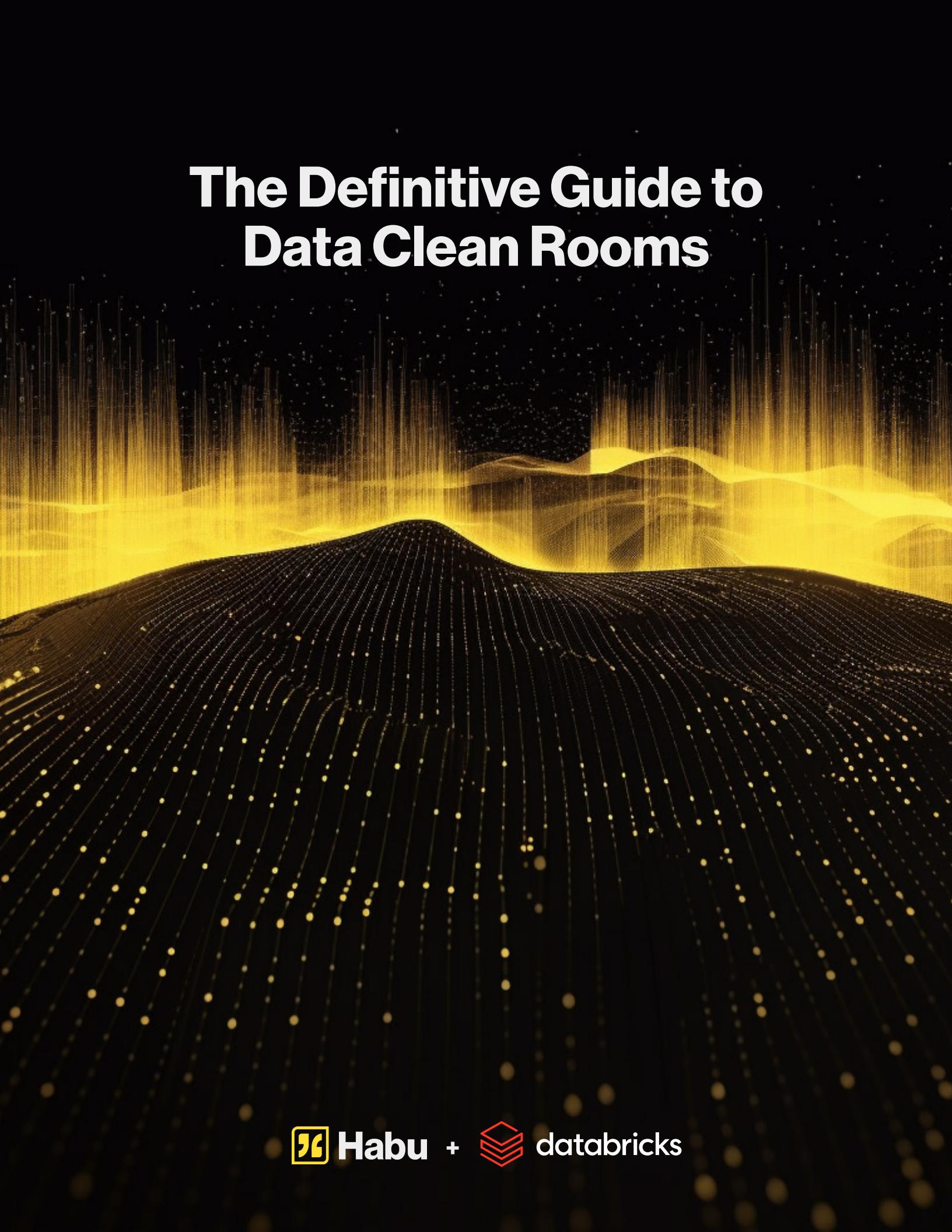


Table of Contents

I. Pages 1 - 5

**The Emergence of the
Data Clean Room**

II. Pages 6 - 9

**Essential Capabilities
and Diverse Use Cases**

III. Pages 10 - 14

**How Data Clean
Rooms Work**

IV. Pages 15 - 21

**Best Practices for
Data Clean Rooms**

V. Pages 22 - 26

Conclusion



I.

The Emergence of the Data Clean Room

In the world of digital marketing, we're witnessing profound changes to the data ecosystem. The past few years have seen the rapid deprecation of third-party cookies and other identifiers and the equally fast arrival of privacy-first mandates, all in the context of staggering growth in decentralized data. These trends have produced large amounts of siloed, hard-to-access data, much of it protected inside walled gardens, while marketing organizations urgently seek new ways to access the data they depend on.

At the same time, enterprises across industries have recognized the untapped potential of data analytics to drive insight and innovation while improving the customer experience. As with digital marketers, these enterprises understand the critical need for more data.

In response, smart brands have found a new way to access and analyze data in an increasingly fragmented ecosystem: the data clean room.



A Response to Significant New Challenges

In the past five to ten years, enterprises have confronted new and increasing restrictions on their use of data, specifically around the need to protect customer privacy. Efforts to legislate data privacy — with Europe's General Data Protection Regulation (GDPR) and then the California Consumer Privacy Act (CCPA) — have been followed by substantial privacy-enhancing changes in app-tracking technologies, such as Apple's ATT framework and Google's restriction of third-party cookies in Chrome. With access to customer data dramatically curtailed and controlled as a result, brands have turned to privacy-preserving data clean rooms to power data-driven marketing while also addressing new challenges:

Campaign Performance

Enterprises must find new ways to maintain and improve the performance of their data-driven campaigns without easy access to data.

Addressability and Personalization

Consumers expect a highly relevant and engaging shopping experience, which makes the need for accurate targeting and personalization more acute than ever, despite increasing privacy-preserving regulations.

Customer Sentiment

With consumers incorporating data privacy concerns into purchasing choices, brands need to maintain privacy-preserving operations to avoid reputational harm.

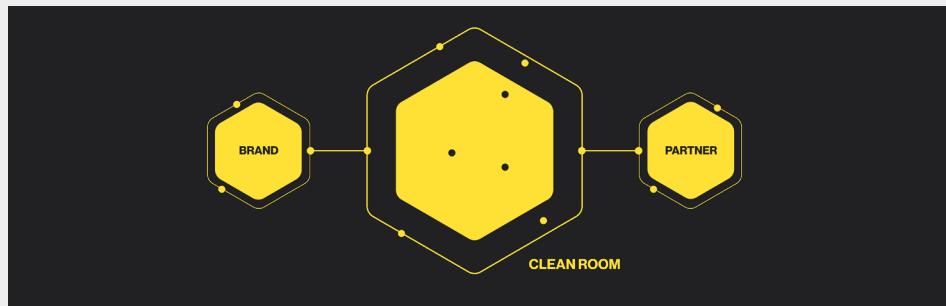
Cost Reduction

Cost is always an issue, and thus enterprises seek to reduce operating costs through automation and by eliminating advertising expenses that fail to drive suitable incremental revenue.



What is a Data Clean Room?

A relatively new technology, a data clean room is a privacy-first, closed-loop environment designed to enable brands to access second-party data from a partner or media platform and safely match it with their own first-party data. Data clean rooms offer brands a software solution that vastly expands access to data — and to data partners — via secure data collaboration that meets today's strict data governance and privacy regulations.



At a high level, we can view the benefits of a data clean room in five ways:

1 Unrivaled Access to Data and IP

A modern data clean room delivers direct access to a volume and variety of data that is simply unavailable by other means. Today, machine learning (ML) models and critical datasets — event-level data, user-level data, consumer data, transaction data, and more — are highly confidential and valuable, and a data clean room is the only means by which data owners will allow data collaborators to interact with sensitive data and IP.

Legacy tools from data warehouses are technical and can be hard to use for business users and others lacking data engineering skills. Modern data clean rooms, by contrast, offer highly-automated data collaboration for dramatic improvements in speed and privacy.

2 Expanded Access to Partners

In today's data environment, collaboration drives growth. Organizations with large and valuable datasets want a cost-effective, easy-to-use solution that will help attract partners to their platforms — and deliver rapid, unrestricted scale when needed. A modern data clean room provides seamless data collaboration, rapid scalability, and — via its attractiveness to partners — an additional product offering for data owners.

With a data clean room, an enterprise can expand its universe of data through a unified data collaboration solution for simple, consistent access to the unique datasets of multiple partners. A company advertising on major platforms such as Google, Facebook, and Amazon, for example, can radically streamline data operations by replicating analytics across platforms and avoid the need to field technical solutions and manual code queries for each new partner platform.



3 Use Case Enablement and Automation

When organizations gain access to new datasets, they're frequently enthusiastic about running new, ad hoc queries to answer new questions. That's understandable, but not where the real value of data collaboration lies. Instead, both for enterprises seeking new and easier access to data and businesses managing large, data-centric advertising ecosystems, it's the volume of recurring question runs, and their increasing scale, that truly unlock significant value. That's where data clean rooms shine. A modern clean room provides not only broad use case enablement via analytic templates, but the technical automation capabilities that enable organizations to run and scale those use cases faster.

4 Efficiency and Productivity Gains at Scale

With deep access to data and the means to simplify and automate recurring data analytics, a modern data clean room is an essential driver of business value. But it becomes an even more valuable tool the more you scale: For data-driven organizations, leveraging additional data collaboration partners by adding clean rooms dramatically increases the quality and resolution of customer journey understanding, which is critical to building loyal customers and expanding revenue. And for data ecosystem enterprises with a large and growing number of partners, easy scale and declining incremental cost leads directly to business value.

Beyond technical efficiency, the most sophisticated data clean rooms also enable an intuitive interaction with data that delivers productivity gains and additional resource efficiencies. Users can quickly query data in plain language to reduce time to insight, and there's no longer a need for specialists to code (and re-code) complex queries and interpret results.

5 Protection for Existing Investments

For data owners, a data clean room is the only means of simply and effectively leveraging your own distributed intellectual property — your machine learning (ML) models, containerized microservices, and other proprietary code — in the context of data collaboration. As your data analyses grow in detail and sophistication, this capability becomes more and more important.



Data Clean Rooms in the Real World

Adopting a data clean room for data collaboration can have a profound effect on a company's bottom line — and that's driving rapid clean room adoption. Analyst firm Gartner recently identified a 3x greater economic benefit for firms that share data externally,¹ and IDC predicts that by 2024, "65% of Global 2000 enterprises will form data-sharing partnerships with external stakeholders via data clean rooms."² So, there's plenty of incentive for organizations to seek access to external datasets. Across industries, data-driven enterprises understand the implications: nearly 80% of organizations want to collaborate with other businesses to share data in the next 12 months, and almost 70% want to expand their current data collaborations.³

3x

Economic benefit of sharing data externally

65%

Enterprises will form data sharing partnerships

80%

Enterprises want to collaborate in next 12 months

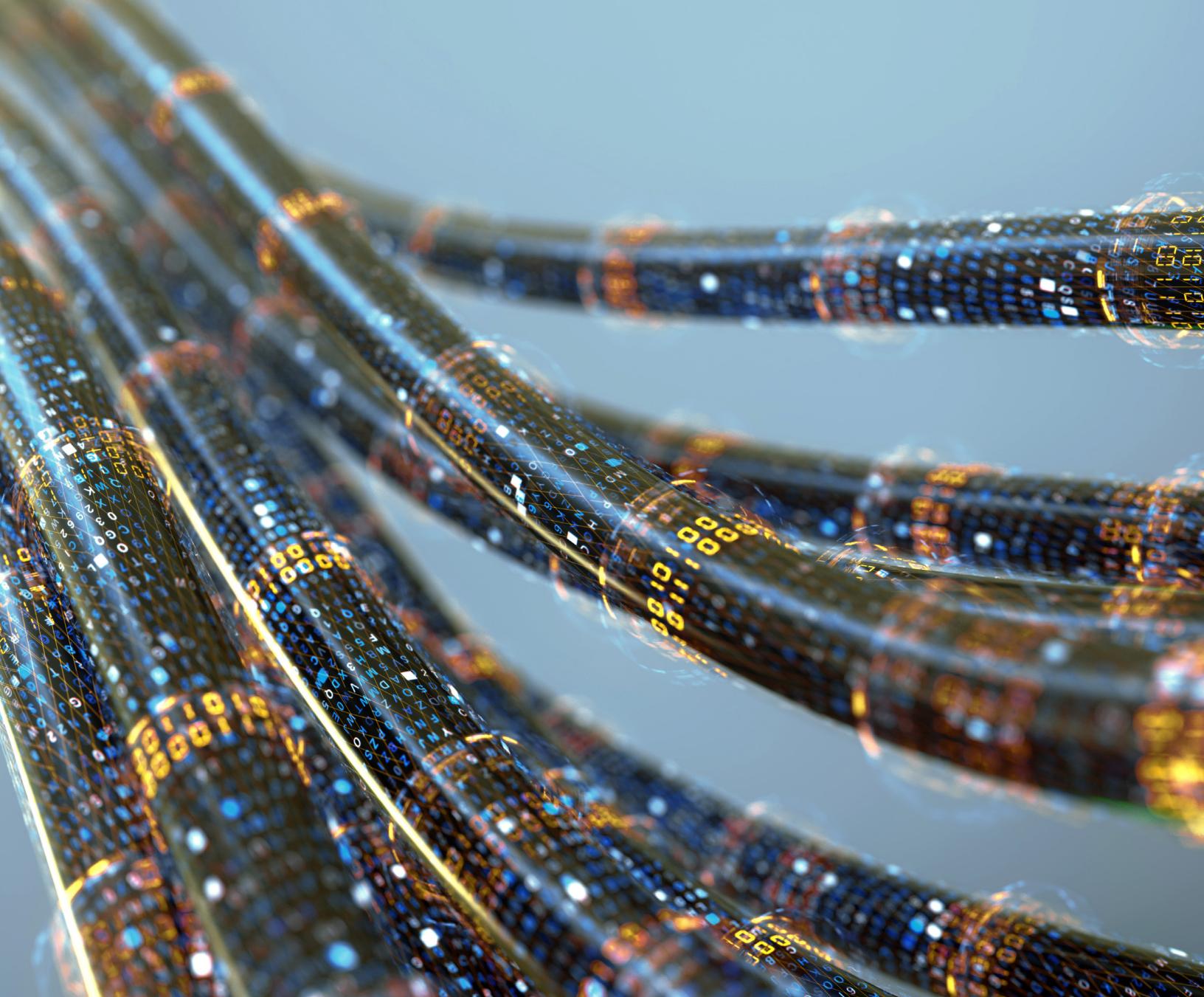
SOURCES

¹ [Gartner: Data Sharing Is a Business Necessity to Accelerate Digital Business](#)

² [IDC FutureScape: Worldwide Data and Content Technologies 2022 Predictions](#)

³ [Habu State of Data Collaboration report](#)





II.

Essential Capabilities and Diverse Use Cases

When thinking about implementing a data clean room, enterprises typically focus on use cases, and we'll get to those in a moment. But first, it's worth considering a set of higher-level capabilities that you'll need in your clean room solution — regardless of use case. This is especially true given that the data collaboration space is growing and evolving so rapidly. What's needed for your organization in terms of current use cases and data collaboration requirements is likely to expand substantially in the near future.



Look For These Capabilities in Every Clean Room

1 Full Interoperability with No Data Movement

Any modern clean room solution should offer the following five essentials:

Clean rooms need to be able to orchestrate workflows by connecting to data, models, and code across a plethora of cloud platforms. Your solution should securely connect to wherever data lives without moving any of that data. Besides being a requirement for many data owners, no data movement reduces latency and the risk of potential data leakage.

2 Future-Proof for Data Science

As mentioned, the use cases you're targeting today may not be those you'd like to leverage in the future. You may move further into machine learning tasks, such as training a model on a combination of first-, second-, or third-party data. To be future-proof, your clean room should provide a secure environment to run any containerized code — including SQL, Python, R, Spark, and other data science tools and libraries — without your partners being able to access your proprietary model or see your underlying data.

3 Enterprise-Grade Privacy and Governance

Whether you're the clean room owner or a data collaborator, you want full control over how your data, models, and code are used. That should include a positive opt-in for each collaborator in approving the use case and data assigned to a query. In addition, you want to have available all the enterprise privacy features you need to confirm privacy-preserving data sharing — and to adjust to evolving regulations.

4 User Experience for Any Use Case

Ensure you can speed time to insight with a clean room that offers pre-written analytics for common business use cases. Data science teams should be able to utilize APIs and code in multiple languages to develop advanced analytics and integrations. This flexibility means that as your organization's data needs grow, your clean room can accommodate different types of users and their unique needs without significant overhead or complexity.

5 Multi-Party Collaboration

Look for a clean room that supports templated analytics that allow you to efficiently work with data and service provider partners across multiple clean rooms. This time savings becomes even more powerful with a natural language framework for analytics, which makes it easy for business users to quickly find the queries they want and customize them as needed.



Use Cases Across Industries

Data clean rooms arose in part to fill the need for compliance with new privacy regulations. But secure and privacy-compliant data collaboration is only one facet of a clean room's usefulness. We can see this by looking at the wide diversity of use cases for clean rooms across industries. Let's examine a few:



Retail & CPG

Retailers and consumer packaged goods (CPG) companies can enhance customer engagement with personalized marketing strategies; improve forecasting accuracy for better demand planning; optimize inventory management to reduce stockouts and overstocking; and conduct product evaluations that inform development and optimization decisions.



Financial Services

Financial services companies, including real estate, consumer finance, banking, and insurance entities, can strengthen fraud detection and improve anti-money laundering tactics; gain deeper insights into portfolio risk and take appropriate mitigation measures; refine insurance rate calculations for more accurate pricing; and tailor adjacent product recommendations for individual clients.



Health & Life Sciences

Data clean rooms empower health and life sciences companies to develop a comprehensive, 360° understanding of patients, which enables personalized treatment plans; accelerate clinical drug development by identifying potential trial participants and collaborating securely with partners; and optimize manufacturing control by analyzing various data streams.



Media & Entertainment

Media and entertainment companies leverage data clean rooms to access valuable insights for planning, create precise audience segmentation, and measure the effectiveness of their campaigns. Furthermore, data clean rooms facilitate campaign optimization and enable seamless activation, resulting in a more effective and targeted approach to reaching audiences.



Use Cases Across Industries (continued)



Travel & Hospitality

Travel and hospitality companies can effectively predict customer churn, allowing for timely interventions and personalized offers to retain customers. Clean rooms also aid in accurate inventory forecasting, ensuring optimal resource allocation and minimizing waste, and contribute to efficient scheduling processes — helping to deliver a seamless customer experience and maximize operational efficiency.



And More!

Data clean rooms can play a valuable role in enabling companies to optimize their supply chains by identifying bottlenecks, improve supplier relationships, and increase efficiency. They can also enhance customer experiences by using data-driven insights to deliver personalized and timely promotions. Clean rooms also empower companies to accurately calculate customer lifetime value, allowing them to better allocate resources and tailor marketing strategies to maximize profitability.

Those are just a fraction of the many use cases that are ideal for data clean rooms. But now that we've defined what a clean room is, and seen some of the diverse use cases that are driving its adoption across industries, we need to pop the hood for a moment to understand how clean rooms actually work.





III.

Let's dig into the nuts and bolts of data clean rooms. Our goal here is to understand the technical details of how clean rooms are built, how they connect to and access data, and how they enforce privacy and security.

How Data Clean Rooms Work



How Data Clean Rooms are Constructed and Connected

First, a level set: There are all kinds of niche data clean rooms — among them, those for deriving insights, for training and inference on ML tasks, for enrichment of data, and for creating lists and pursuing activations. However, we're going to focus on the modern, hybrid data clean room, designed for use with multiple platforms, that empowers data collaborators to perform all of these use cases and with any partner — which is particularly valuable in the fast-evolving data collaboration space.

Back to the basics. To create a functional data clean room, we need four things:

1 Data Connections

The clean room owner's first step is to establish connections to the cloud accounts of all data partners. Note that these should simply be pointers to where the data, models, or code reside, so nothing needs to be moved or copied. A modern data clean room enables you to connect to data, models, and code wherever they live and ideally has no fixed schema — so your connections are not limited by data type.

2 Datasets

Contributed by the clean room owner and collaboration partners (depending on the particular use case), datasets are only accessed by the clean room at runtime. Only the rows and columns specified in the clean room use case policies are read into the analysis, and processed data is immediately deleted upon completion of the associated data join or query.

3 Use Cases

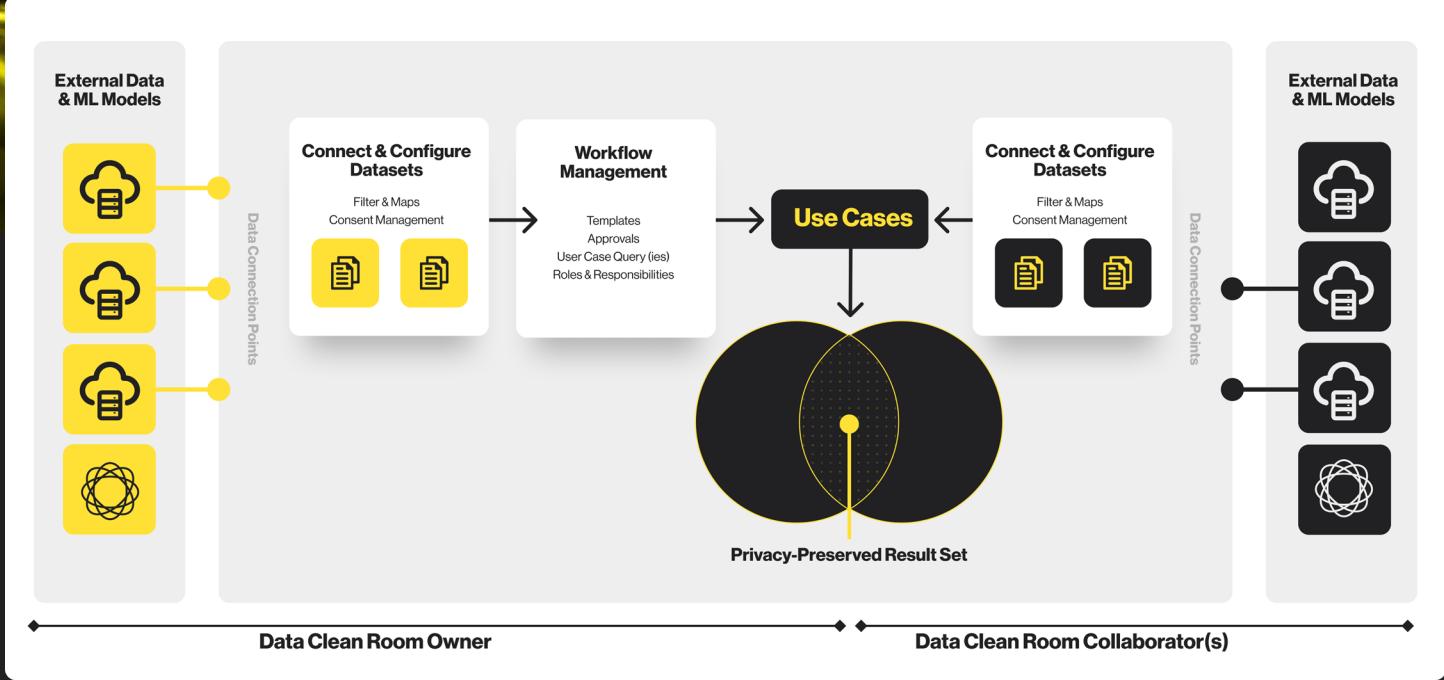
For each data collaboration use case defined in the clean room, the clean room owner and all partners must explicitly approve the query(ies) to be run, and each data contributor must also opt in to allowing their data to be used as specified. At runtime, only approved use cases can execute, only approved data can be accessed, and no party can see the underlying assets (data, models, and code).

Note that the best clean room vendors offer extreme use case flexibility: you should be free to run virtually any ML use case with privacy and governance as long as it's technically possible to do with the data. This broad clean room capability with ML use cases in particular is a significant advantage for data science teams.

4 Roles and Permissions

From managing datasets and viewing reports, to building and proposing questions and scheduling runs, the clean room owner decides the scope of each partner's available actions during onboarding, and these permissions are easily tracked and updated.





Data Clean Room



Solving for Interoperability

A major benefit of the best data clean rooms is their privacy-compliant interoperability. This is the crucial capability that enables complete freedom in your choice of collaboration partners.

Many data platforms offer the capability to freely share data across clouds. However, that data sharing often does not include privacy controls. A data clean room leverages platform data sharing protocols, but adds a crucial layer of policy-driven privacy control that protects against data leakage as you collaborate with diverse partners.

Ideally, with a modern data clean room any number of collaborators, working with any kind of data, code, or models on any number of environments, can come together to share data. Everything is interoperable and privacy-preserving. In a world in which everything is also multi-cloud, that's a crucial capability.

Delivering on Privacy

To see how clean rooms enable privacy-preserving data collaboration, we have to deep dive into the analytics process. To begin, clean room owners and collaborators agree to an analysis, assign data, and run one or more queries.

For each query, the clean room may offer differential privacy in the form of, for example, adjustable noise injection and crowd size minimums, so no specific user/consumer can be re-identified from the data. The clean room owner can specify these controls, as well as exercise governance over where and how data and code may be exported from the clean room.

Only once a query executes does the clean room solution open connections to each partner's data, and only to the necessary rows (e.g., Store #106) and columns (e.g., shopper ID, SKU, and price) to run the specified analysis. Within the clean room, a policy is created that says the current data connections can only be used for this specific query, with this specific data, right now.

Note that the data accessed for processing purposes is not shared; instead, it is read into a neutral and protected processing environment which is controlled by the clean room and which none of the collaborating parties has access to. After the report is run, the processing environment and associated data is deleted, and the output is sent to an agreed location, from which users can view results via embedded analytics in the clean room software or pipe the data to their own BI tools, such as Tableau or Looker.



Automation Dramatically Streamlines Analysis

We've just described a single instance of a query that required a specific set of inputs: data, models, and/or code. Without clean room automation, your technical teams would have to manually implement all the necessary primitives to prepare for this type of analysis. Now imagine doing that over dozens, hundreds, or thousands of queries. End-to-end automation dynamically implements all the necessary primitives, slashing the time required to prepare and execute queries. That's a huge time saver.

Caveat: Data Collaboration Still Requires Ownership and Trust

A data clean room is a powerful tool in the era of decentralized data, but it's important to mention two ways in which clean rooms assume a level of ownership and trust from their users.

First, keep in mind that a data clean room is not, by itself, a privacy enhancing technology (PET). Instead, the best data clean rooms provide access to a broad suite of PETs that data collaborators deploy in various combinations to ensure that each use case is completely compliant from a privacy and governance perspective. Among those PETs are:

- Encryption at rest, in transit, and in use
- Obfuscation
- Data minimization
- Differential privacy
- Injection of random noise

In other words, a clean room platform cannot ensure you are privacy compliant; rather, it empowers you with privacy tools and it's up to you to use them.

Second, data collaboration via a data clean room assumes a base level of trust between collaboration partners. Clean rooms enable a great deal of privacy: partners cannot access the data, code, or models of their counterparties. But all participants in the clean room must still trust that their partners' representations about the efficacy of code, or the quality of data, or the outputs of a model, are what they say they are. The clean room itself cannot vouch for these things.

Now that we've detailed the inner workings of clean rooms, we can examine the best practices that will enable you to make the most of them.





IV.

Best Practices for Data Clean Rooms

As enterprises consider expanding their data collaboration agendas with data clean rooms at their center, there's a need to standardize best practices for clean room use. In short, what's the optimal way to think about clean room implementation — and what are the key considerations for success?

We'll break it down into a four-step process:



Review Your Needs and Strategize with a Data Audit

01



In working with numerous customers and partners to enable clean room workflows, we've found that a vital first step is addressing a set of general questions around your data. You need to have a clear sense of:

What data you have. What's the extent of your data holdings, and is there a clear understanding of where the various "sources of truth" are? There may be gaps you need to fill.

What format the data is in. Is it structured data, with rows and columns, or is it unstructured or semi-structured, with images, audio files, and more?

What condition it's in. Is your data well-defined, with clear and accurate tags and headers, or do you need to do some clean-up before it's usable?

Where your data lives. If it's in multiple cloud locations, you'll need to ensure interoperability with each cloud vendor.

How often it's updated. For example, when capturing store transactions, is your data refreshed continuously, hourly, nightly, or even less frequently?

Who your data partners are. Who do you intend to partner with, and where is their data held?

What workload(s) you want to run. Given your use cases, what are your staffing, infrastructure, privacy, and data requirements?

Educate your stakeholders

At the outset, it's advisable to put together an operating model for your clean room. The clean room itself provides an innovative paradigm for data collaboration outside — and inside — your organization. But it's a different way of working that's going to require both education and evangelizing among your teams. Choose a core team of internal experts and empower them to bring the benefits and best practices of data clean rooms to the rest of the organization.



What's Your Clean Room Strategy?

With a core team in place, you're ready to formalize your data clean room strategy. This overall strategy should precede your work on learning agendas, use cases, and operational details. What is your overall objective? Increased revenue and/or efficiency? Better CPMs or ROAS? Who do you want to work with? Do they have a clean room set up? Answering these questions will help to establish the optimal clean room approach for your organization.

Distinguish Your Role

Your clean room strategy will necessarily vary depending on your organization's goals:

If you're a brand, you'll likely want to establish clean rooms for your top media investments, which may be a handful of large publishers and walled gardens. Others may want to enrich their first-party data with data from entities that are downstream from them, such as CPG companies accessing store transaction data, or parts suppliers accessing forecast information from automobile manufacturers. Build a learning agenda around a core set of business questions you'd like to answer, and be sure that you identify the key business drivers that you'll target with your data clean room initiative.

If you're a publisher, you should consider deploying clean rooms for your top 30-50 advertising partners and focus on creating a product roadmap for those partners. Think about how a data clean room fits into your product lineup, as well as the ways in which partners will access your data. What use cases do you intend to support? Will you offer templated queries or an open-ended framework? What changes to the organization of your data need to be made to support your intended offerings?

Keep in mind that, whether you're a brand or publisher, there will likely be a substantial amount of cross-functional alignment needed within your organization as you refine your strategy. Moreover, while a data clean room is a critical tool, it's also part of a broader marketing ecosystem. As you map roles and responsibilities, be sure to cover the complete picture of marketing workflows and account for tasks upstream and downstream of the clean room that will be required to accomplish your objectives. Understand where a data clean room fits within your strategy, rather than the other way around.



Implementation and Testing

02

Once you're satisfied with your data clean room strategy and you've clearly outlined the functional clean room model you intend to pursue, you're ready to implement your plans, deploy your data clean room, and test your new solution.

As you proceed through implementation and deployment of your data clean room, there are a few essential tips to keep in mind:



Automation is crucial. If you want to leverage clean rooms at scale, you're going to need to embed automation into the process — for scheduling, environmental configuration, portability, and much more. You should aim to automate your data collaboration process at every stage; assuming you can manually scale is highly impractical and invites complexity and additional cost.

Focus on portability. A modern data clean room offers you a huge potential in terms of the data and partnerships available to you. Thus, you don't want the automation you created for one platform to be incompatible with other platforms, forcing you to rebuild everything from scratch when you need to connect to a new platform. Ensure the cross-platform portability of your solution so you can take advantage of new data opportunities, quickly and efficiently.

Data normalization is a constant. As you collaborate with diverse partners and datasets, you'll frequently need to be able to normalize certain types of data so you can align cuts of the data across multiple clean rooms. As with other automation requirements, normalizing datasets one-by-one will be tedious and eventually a significant pain point. Be sure you have a solution for what will likely be fairly continuous data normalization needs.

Deliver a uniform experience. Over time, you will inevitably collaborate with partners whose data lives on a variety of different cloud platforms. Ensure efficiency and ease of use among your teams with a uniform experience across clean rooms. Your clean room vendor should be able to deliver an identical clean room experience across your clean room estate, regardless of cloud provider or partner clean room solution.



Centralize your outputs. Once data has been queried and outputs generated, the results are yours, and you'll want to be able to share insights within your organization. This enables you to look at all of your data queries and insights in one place, as opposed to leaving the results of isolated or episodic querying on sites such as Amazon Marketing Cloud (AMC) or Google Ads Data Hub (ADH). Instead, centralize the results data in your environment, run business intelligence on top of it, and build and scale this data over time. Then deliver your insights widely within your organization.

Maximize impact. You'll want to deliver insights widely within your organization. But insight is most effective when it's acted upon and its progress is quantified. For completeness, your clean room should support monitoring and alerting activities in conjunction with their analysis outputs. These features also make the technology easier to adopt and use.

Next, begin testing your clean room environment with a sandbox and synthetic data. See if everything works as expected, check that the privacy controls you want to apply are functioning, and then provide the solution to a few trusted customers and partners to react to, test use cases against, and generally understand whether the solution fits the needs you've identified. In the sandbox environment, customers and partners can safely prototype use cases and validate the process before working on live data.

Note that your initial partners — and all future partners, once you begin to scale — will need to have chosen a clean room solution for themselves in order to collaborate with you. Because they may be unfamiliar with the clean room model, your core team may need to educate your partners on deploying their own clean room software and aligning their workloads for data collaboration. Three cautions to implementation and testing: First, remember that testing with synthetic data is great for obtaining initial validation of your solution, but it cannot preview the full potential of the clean room. Only production data at scale can do that. Second, when bringing in early partners to help get your clean room operation off the ground, it's important to set expectations. Be clear with them that you are at the pilot stage, that not everything will go perfectly smoothly, and that you'd like their help and feedback in evolving the offering. Finally, be aware that your early partners will likely have their own data readiness tasks to ramp up — understanding what data they will provide, who will prepare it, and who will manage the data collaboration. If not done properly and efficiently, both data owner and data collaborator will face a delay in time to value.



Partner Onboarding

03



When you've fully validated your clean room solution with your initial set of partners, it's time to go live and begin broader onboarding. For each new partner, you will need to confirm their buy-in on key aspects of the data collaboration:

Business

Business teams need to align at the outset on the use cases you will pursue, as well as mutually agree on what data will be shared and how it will be used.

Technical

Technical teams must thoroughly review information security (InfoSec) protocols and establish a secure collaboration process that both parties can trust.

Legal

Departments from both sides should confirm data use via a detailed written agreement.



Managing Ongoing Clean Room Operations

04



As your data collaborations move forward, keep your key objectives top of mind while evaluating the effectiveness of your teams. The marketing team should be defining the broader workflows required to achieve your goals, while the legal department needs to ensure data operations stay within privacy guidelines as well as facilitate partnership agreements. Since data clean rooms emerged from the marketing world, if you have a data science team, they may need additional context from the marketing team to get up to speed with an unfamiliar tool.

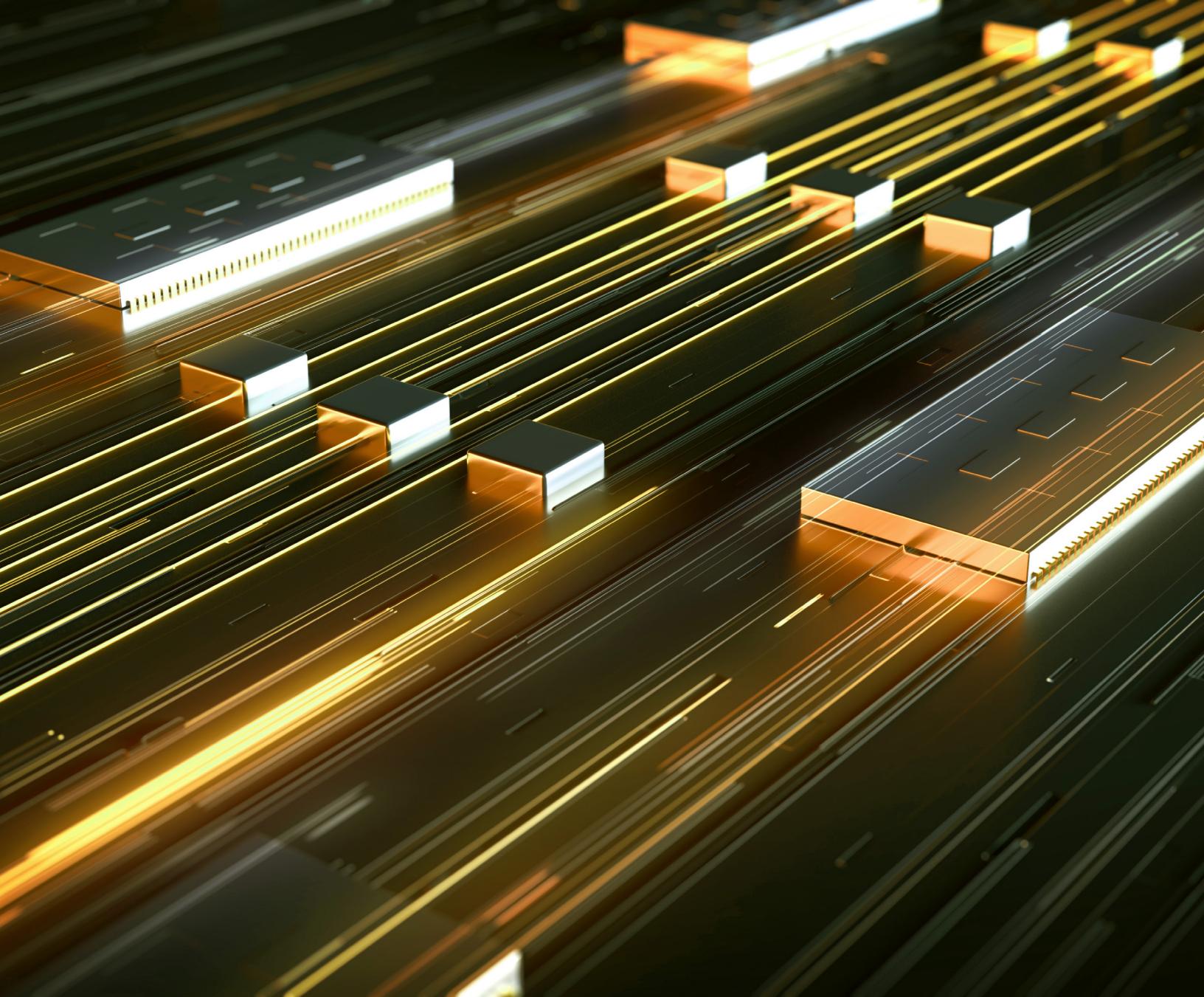
Prepare for Data Science

While you may not initially plan to pursue more advanced data science use cases, in a privacy-first world in which user-level data is hard to come by, leveraging data science for audience modeling is inevitable. Moreover, to scale data, or to compare data across environments, modeling is again required. So, you should plan for the fact that you're going to be doing some modeling — and build in a clean room data science capability from the beginning.

Spread Your Clean Room Insights Widely

A modern data clean room enables the entire organization to derive valuable insights from a much broader universe of data. In thinking about the data clean room model, keep in mind that a clean room is a mass tool for the insights it puts out, but not for conceiving and implementation that goes into it. That is, your core clean room team must be responsible for using best practices to optimize the processes through which you adopt and leverage your data clean room — so that you're prepared to share the powerful insights that result. Remember that a modern data clean room is an enabler of insight distribution. While data scientists can accomplish many tasks within the clean room, a simple interface means that they or business users can easily export data for use in business intelligence (BI) or other common tools that power data-driven organizations.





V.

Conclusion

In an age of decentralized data and privacy-first mandates, data-driven organizations understand that data collaboration is essential to gaining a competitive advantage and driving growth. It's for that reason that data clean rooms have rapidly become an essential component in enterprise transformation plans.





Data clean rooms deliver a secure, privacy-first environment for organizations to collaborate on sensitive data without compromising data privacy regulations or exposing proprietary information. Moreover, they provide an essential link to decentralized data that lives on diverse cloud platforms. By leveraging data clean rooms, businesses can effectively pool data resources, uncover deeper insights, and unlock new opportunities that would not have been possible otherwise.

Simply put, a modern digital enterprise needs a clean room solution in order to participate in the kind of data collaboration that underpins a growth strategy. Without a clean room, data owners and other partners simply will not provide access to their most valuable data. That's inherently limiting in an organization's drive for insight and innovation.

By streamlining data operations, enabling powerful use case automation, and driving increased efficiency, data clean rooms offer businesses a competitive edge in an increasingly fragmented data ecosystem. In addition, data clean rooms empower organizations to leverage their existing investments in machine learning models and proprietary code, enhancing the overall value derived from data collaboration.

Across industries, the fast growing adoption of data clean rooms is a testament to their efficacy and transformative potential. As more organizations recognize the value in data collaboration, data clean rooms will continue to play an essential role in unlocking insights and enabling growth in a privacy-first world.



The modern data clean room offers a powerful solution to the challenge of data collaboration in an ecosystem characterized by fragmented, hard to access data and growing privacy regulation. When thinking about adopting a data clean room strategy, enterprises should keep five key takeaways top of mind:

1 Best Practices for Data Collaboration

The modern data clean room emerged in the past few years in response to new challenges in the data ecosystem around data privacy and security. Recognizing the need for a new way to access and analyze data, data-driven marketers and enterprises across industries have turned to the data clean room for simple, efficient, and privacy-preserving data collaboration. Clean rooms provide unrivaled access to data and IP, far greater access to partners, and powerful use case enablement at lower cost.

2 Broad Applications Across Industries

From retail and CPG, to financial services and healthcare, to media, travel, and more, data clean rooms support diverse use cases. Regardless of vertical, a modern data clean room should deliver the following capabilities:

- Full interoperability with no data movement
- Flexibility for collaborative data science
- Enterprise-grade privacy and governance
- User experience for any use case
- Multi-party collaboration

3 Deep Interoperability and Automation

From establishing data connections to any cloud platform, to accessing datasets where they live, to powering a wide range of use cases, a modern data clean room delivers broad interoperability and high levels of automation that streamline and accelerate data operations for IT and lines of business alike.



4 Governance for Secure Data Sharing

To succeed with a data clean room initiative, organizations need to:

- Conduct a complete data audit and come up with a clean room strategy. Understand what the state of your data is, where it lives, who your partners are, and what workloads you want to run.
- Implement your clean room plans, deploy your clean room, and test the solution with a few trusted partners to confirm it works as expected. Ensure automation is embedded and portable across platforms, and that the user experience is consistent across clean rooms. Be sure to centralize your analysis results so you can collect and distribute your insights.
- Go live and expand your partner onboarding. Ensure business, technical, and legal teams from all collaborators stay closely involved as you scale.

5 Unique Competitive Advantages

Data collaboration is essential to gain a competitive advantage and drive growth — and that's driving rapid data clean room adoption. A data clean room dramatically streamlines data operations — delivering unprecedented access to decentralized data while enabling powerful use case automation and driving increased efficiency. Across industries, these capabilities are enabling businesses to uncover deeper insights and unlock opportunities that were previously out of reach.

Get Started Today

Request a demo to see the Habu and Databricks data clean room platform in action.





Databricks is the data and AI company. More than 9,000 organizations worldwide — including Comcast, Condé Nast, H&M, and over 50% of the Fortune 500 — rely on the Databricks Lakehouse Platform to unify their data, analytics and AI. Databricks is headquartered in San Francisco, with offices around the globe. Founded by the original creators of Apache Spark™, Delta Lake and MLflow, Databricks is on a mission to help data teams solve the world's toughest problems.

[DATABRICKS.COM](https://databricks.com)



Habu enables organizations to master decentralized data, uncovering insights that fuel better business decisions via data collaboration that's smart, safe, scalable, and simple. With flexible, multi-cloud deployment, high levels of automation, and an intuitive interface, a Habu Data Clean Room empowers data teams. Easily connect your data from any source and start accelerating growth today.

[HABU.COM](https://habu.com)