



DATA CLEAN ROOMS DE SNOWFLAKE: CÓMO AYUDAN A EDITORES Y RESPONSABLES DE MARKETING A MEJORAR LA EFICACIA DE LOS ANUNCIOS

POR QUÉ UN MUNDO SIN COOKIES REQUIERE UN NUEVO MÉTODO DE FOCALIZACIÓN Y MEDICIÓN

En los últimos años, los consumidores se han ido preocupando cada vez más por la privacidad en línea. Según un estudio de KPMG de 2020, el 54 % de las personas no confía en que las empresas utilicen sus datos personales de forma ética, y el 53 % no confía en que las empresas recopilen datos personales de forma ética.¹

Mientras tanto, según un estudio de la empresa de selección de ejecutivos Spencer Stuart, los directores de marketing cada vez duran menos en el cargo² y los responsables sénior de marketing están sometidos a escrutinio para demostrar el retorno de la inversión (ROI) de cada dólar invertido en publicidad. Recurren a modelos de atribución para cuantificar el impacto de cada campaña de medios y garantizar la optimización del gasto. Por lo general, estos modelos se han basado en cookies de terceros que realizan un seguimiento de la actividad de los usuarios en Internet, al igual que los métodos de focalización para mostrar anuncios y contenido personalizados.

Sin embargo, las medidas de privacidad que han implementado algunos de los principales actores del ecosistema publicitario han provocado que los métodos tradicionales de focalización y medición de los anuncios sean cada vez más precarios. En concreto, Apple exige ahora a los usuarios que autoricen el seguimiento de su actividad en todas las aplicaciones³, y la eliminación de las cookies de terceros en Chrome en 2023 es inminente⁴. Por ello, las data clean rooms han surgido como una estrategia para que los responsables de marketing realicen sofisticadas tareas (segmentación del público, focalización, atribución y medición de la actividad, y análisis de públicos similares basados en el aprendizaje automático) y, al tiempo, preservar la privacidad.

Las data clean rooms son plataformas seguras que permiten a dos o más partes combinar y consultar conjuntos de datos, sin que ninguna pueda ver los datos de la otra.

Con las data clean rooms de Snowflake, las empresas pueden compartir vistas de datos de forma segura y permitir que los datos se combinen con otros conjuntos de datos para elaborar modelos y análisis, sin que otras partes puedan ver información confidencial de los usuarios. Esto también puede beneficiar a los editores, que están sometidos a una enorme presión para proteger la privacidad de los usuarios. Las data clean rooms de Snowflake pueden ayudarles a mejorar la eficacia de los anuncios y, en última instancia, a reducir el coste por cada mil impresiones (cost per mille, CPM) y a aumentar el retorno de la inversión en publicidad (return on ad spend, ROAS) sin tener que compartir la información ni los datos de actividad que los usuarios hayan proporcionado a los responsables de marketing.

POR QUÉ LOS RESPONSABLES DE MARKETING Y LOS EDITORES USAN LAS DATA CLEAN ROOMS DE SNOWFLAKE

Dada la inminente desaparición de las cookies de terceros en el sector publicitario⁵, tanto los responsables de marketing como los editores se concentran en fortalecer las relaciones directas de datos de los consumidores. Un pilar fundamental de esa estrategia es gestionar los datos que recopilan directamente de los usuarios y los clientes (denominados “datos propios”) de forma que se preserve la privacidad.

Las data clean rooms de Snowflake permiten utilizar datos confidenciales derivados de identificadores únicos (como correos electrónicos, correos cifrados con hash, nombres, ID de dispositivos y direcciones IP) y, a la vez, proteger la privacidad. A partir de ahí, los responsables de marketing pueden segmentar y orientar la publicidad a los clientes existentes buscando solapamientos con el público de un editor sin recurrir a un proceso de extracción, transformación y carga (extract, transform, and load; ETL); es decir, sin tener que mover ni copiar datos.

La focalización en un entorno de clean room contrasta notablemente con la compra de anuncios programáticos, que a menudo carece de transparencia. Cuando los responsables de marketing compran anuncios en una plataforma de demanda (demand-side platform, DSP) o a través de un intercambio de anuncios, no tienen modo de saber si han llegado a clientes reales o si los usuarios a los que han llegado realmente han completado una conversión comprando algo, especialmente cuando la conversión ocurre de una manera no lineal. Imaginemos que un consumidor ve un anuncio en Internet o lee un correo electrónico, pero luego compra un producto en una tienda física. Para empezar, lo habitual era que los enfoques basados en cookies no pudiesen gestionar ese flujo.

Con las data clean rooms de Snowflake, la focalización y la activación se pueden ajustar en función de una amplia variedad de parámetros, como el consumo del contenido

(por parte de los editores) y el historial de compras o los datos demográficos (por parte de los responsables de marketing). Por ejemplo, los minoristas pueden orientar anuncios sobre equipamiento deportivo a los segmentos de clientes existentes que leen mucho contenido sobre *running* en el sitio web de un editor.

Los responsables de marketing también pueden aprovechar el modelado de públicos similares basado en el aprendizaje automático en un entorno de clean room. Así, se identifican cohortes en el público del editor que tengan atributos similares a los que tienen los clientes existentes o, incluso, los clientes de alto valor.

Todo esto ocurre sin que el anunciante pueda ver los datos del editor, sin que el editor pueda ver los datos del responsable de marketing y sin que ningún otro tercero pueda ver los datos de ninguna de las partes, ni siquiera saber que están realizando este tipo de análisis.

Las clean rooms también ayudan a los responsables de marketing que analizan el ROI a atar cabos. Solo tienen que realizar una consulta en el entorno de clean room para ver cuáles de sus segmentos de clientes que compraron un producto específico estuvieron expuestos a impresiones de anuncios del editor. Así, pueden acceder a la información sobre las conversiones de su campaña e, incluso, de campañas que se han ejecutado en varios canales de medios. Esto elimina la necesidad de recopilar y analizar datos de comportamiento a través de cookies de terceros para conocer los recorridos de clientes valiosos.

Un triunfo para los responsables de marketing también es un triunfo para los editores, ya que unas funciones de segmentación más potentes y una atribución más precisa pueden ayudar a justificar un mayor gasto publicitario. Por su parte, el sofisticado modelado de públicos similares ayuda a los editores a vender más anuncios.

NBCUniversal se prepara para el futuro de la colaboración de datos multinuble con data clean rooms de Snowflake

En 2021, NBCUniversal anunció sus planes de lanzamiento de NBCU Audience Insights Hub, basado en un entorno de data clean rooms multinube con tecnología de Snowflake, que facilitará la interoperabilidad de datos entre NBCUniversal y sus partners del ecosistema publicitario. La nueva solución permitirá a NBCUniversal incorporar datos propios de audiencia, que los partners publicitarios podrán combinar de forma segura con sus respectivos conjuntos de datos, sin trasladar, copiar ni exponer información de identificación personal (personally identifiable information, PII) subyacente.

El marco de Snowflake permitirá a NBCUniversal y sus partners controlar qué datos se alojan en la clean room, cómo se pueden combinar los datos, qué tipos de análisis puede realizar cada parte con los datos y qué datos pueden sacarse de la clean room, de ser posible. El enfoque es distinto al de otros entornos de pruebas de privacidad, puesto que permite a los participantes diseñar el nivel de protección y transparencia adecuado para crear públicos, activar campañas o realizar mediciones.

NBCU Audience Insights Hub incluye las siguientes funciones:

- Exploración digital del público. En el nuevo entorno de clean room centralizado de NBCUniversal, los partners podrán explorar cómo se superponen las audiencias y los clientes, y obtener información agregada valiosa sin que los datos subyacentes de ninguna de las partes queden expuestos. Esto permitirá a los partners de marca y a las agencias encontrar clientes y conocerlos mejor para impulsar los resultados, al tiempo que protegen la información de los espectadores.
- Planificación multiplataforma. NBCUniversal combinará el nuevo entorno de clean room con sus API de TV lineal propias, lo que proporcionará a los partners, por primera vez en la historia, acceso en régimen de autoservicio a los datos lineales y digitales agregados de NBCUniversal (todos los ingredientes necesarios para la planificación de medios y multiplataforma) en una sola integración.
- Medición del alcance y la frecuencia. El nuevo centro incorporará modelos de medición del alcance certificados, que permitirán a los partners utilizar los datos de exposición a anuncios y realizar sus propios análisis. Esto permitirá desduplicar el alcance y la frecuencia de las campañas para mejorar la eficiencia de la planificación y la medición de los medios.
- Atribución multiplataforma. NBCUniversal continuará ampliando sus funciones de medición interoperables, lo que permitirá a los partners identificar su propia atribución multiplataforma en autoservicio. A largo plazo, esto ofrecerá capacidades de atribución en circuito cerrado y siempre activas para cada campaña.

CÓMO FUNCIONAN LAS DATA CLEAN ROOM DE SNOWFLAKE EN SIETE PASOS



1. ALMACENAMIENTO

Un responsable de marketing y un editor que ya son clientes de Snowflake deciden iniciar una colaboración de datos. Mantienen los datos de sus clientes en sus propias cuentas de Snowflake, sin necesidad de procesos de ETL adicionales ni de enviar copias de esos datos a otro lugar. Esto hace imposible que la otra parte pueda ver o copiar los datos sin procesar sin permiso expreso.



2. COMBINACIÓN

Ambos determinan qué datos se combinarán y cómo, teniendo en cuenta las posibles restricciones de datos (personales) y cada uno de sus objetivos. La manera más simple es usar algo que ambos ya tengan, como direcciones IP o de correo electrónico en minúsculas y recortadas. También pueden utilizar combinaciones avanzadas de identidades, gráficos, combinaciones en cascada y con expresiones booleanas, o pueden recurrir a un proveedor de identidades externo de Snowflake Data Marketplace, como Acxiom, LiveRamp o Neustar.



3. CALIBRACIÓN

Averiguan cómo consultar los datos combinados según el caso de uso previsto. Por ejemplo, si un minorista desea consultar los datos de interacciones anonimizados de un editor por categoría de contenido y encontrar superposiciones con sus propios datos de compras, debe definir una consulta con las cláusulas SELECT, GROUP BY, JOIN y WHERE adecuadas para comprender mejor qué tipos de contenido consumen sus clientes.



4. APROBACIÓN

A continuación, la otra parte debe aprobar la consulta solicitada, lo que automáticamente garantiza que la consulta cumple las normas que ha establecido respecto al uso conjunto de los datos. Por ejemplo, una de estas reglas podría ser no permitir ningún análisis o suprimir las filas de salida que agreguen menos de 75 personas o dispositivos diferentes para evitar la reidentificación accidental del usuario.



5. EJECUCIÓN

Ahora, la parte solicitante puede ejecutar la consulta aprobada tanto en sus datos como en los de la otra parte, siempre que esa consulta respete las reglas definidas por la otra parte en el paso anterior. Snowflake gestiona el cálculo de datos entre dos (o más) partes.



6. ACTIVACIÓN

Los resultados a nivel agregado contribuirán a definir cómo se orienta la publicidad a los públicos y cómo se segmentan estos en las próximas campañas. Por ejemplo, el minorista del ejemplo anterior podría descubrir que el 7% de los usuarios que consumían contenido de cocina en el sitio del editor compraron zapatillas de deporte a través del editor en el pasado. Por tanto, pueden orientarse anuncios de los últimos modelos de zapatillas de deporte a ese segmento en el futuro.



7. MEDICIÓN

Después de la activación y la finalización de la campaña (o mientras está en curso), una parte puede realizar otro análisis JOIN para medir varias métricas de rendimiento de la campaña.

EN QUÉ SE DIFERENCIAN LAS DATA CLEAN ROOMS DE SNOWFLAKE DE OTROS ENTORNOS DE PRUEBAS DE PRIVACIDAD

Tradicionalmente, proteger la privacidad de los usuarios y disponer de potentes funciones de focalización y medición se han considerado imperativos en conflicto. En realidad, el seguimiento del comportamiento dio lugar a estrategias y tácticas basadas en pocos recorridos de consumidores o en muestras diminutas de datos. Con las data clean rooms de Snowflake, los responsables de marketing y los editores pueden mejorar el rigor estadístico de su planificación y compra de anuncios mediante la combinación y el análisis de enormes conjuntos de datos (todos los datos) de manera anónima y protegiendo la privacidad.

Cinco diferenciadores clave:

1. Secure Data Sharing

Los responsables de marketing y los editores pueden compartir conjuntos de datos actualizados de forma segura en diversas nubes y regiones sin que la otra parte pueda copiarlos. Dichos datos permanecerán actualizados, ya que solo existen en un lugar, en vez de estar distribuidos en silos. Los datos (cifrados por Snowflake) se almacenan solo en la cuenta de Snowflake del propietario de los datos.

Esto difiere notablemente del funcionamiento de los entornos de clean room de tipo “jardín vallado” que ofrecen otras grandes plataformas, en los que los responsables de marketing tienen que entregar sus datos propios a fin de combinarlos con datos de nivel agregado de esas plataformas para, a continuación, modelar el conjunto de datos resultante con fines de focalización y medición.

2. Controles de privacidad

Snowflake cuenta con potentes funciones de procesamiento multipartito, lo que posibilita análisis seguros de conjuntos de datos combinados sin revelar los datos de una empresa a las otras partes.

El problema de los millonarios de Yao proporciona un marco útil para entender cómo funciona este método. Este conocido problema de criptografía se centra en la forma en que dos millonarios, Alice y Bob, pueden determinar quién posee el mayor patrimonio neto sin revelar al otro cuánto dinero tiene. Tradicionalmente, las posibles soluciones al problema han requerido cálculos complejos y costosos, pero Snowflake tiene una forma sencilla de resolverlo. Además, es una solución aplicable a múltiples partes y billones de filas de datos, así como a problemas que van más allá de calcular si A es mayor que B.

A continuación, presentamos una forma sencilla de comparar de forma segura el patrimonio neto de Alice y Bob utilizando data clean rooms de Snowflake:

1. Alice crea una tabla llamada “Mi patrimonio” e indica el valor del mismo: 1,2 millones de dólares. Después, crea una vista segura de esa tabla en SQL y aplica protecciones de clean room a esa vista. Concede a Bob acceso a esa vista, pero solo puede utilizarla para hacer determinados tipos de análisis permitidos, es decir, no puede ejecutar una sentencia `SELECT *` en esa vista.
2. Bob hace lo mismo indicando en su tabla su patrimonio: 1,1 millones de dólares.
3. Bob ejecuta una consulta en Snowflake para determinar qué cifra es mayor. Las reglas de la clean room de Alice permiten esta consulta porque es un tipo aprobado de análisis conjunto desde la perspectiva de Alice (sigue un patrón de consulta preaprobado o las reglas que Alice ha establecido sobre lo que Bob puede hacer con los datos de ambos combinados).
4. Bob descubre que Alice es más rica que él, pero ninguno conoce las cifras exactas que subyacen a esa respuesta.
5. Alice sabe que Bob ha ejecutado esta consulta permitida en sus datos combinados, pero ella no puede conocer el resultado. Para conocerlo, tendría que ejecutar la misma consulta de antes, suponiendo que Bob también haya configurado su clean room para permitir ese tipo de consulta.

Este concepto está directamente relacionado con la forma en que los responsables de marketing y los editores utilizan las data clean rooms de Snowflake para encontrar segmentos superpuestos en sus públicos sin comparar directamente los identificadores compartidos, como direcciones de correo electrónico y direcciones IP. Simplemente aplican el mismo enfoque, pero a millones, miles de millones o, incluso, billones de filas de datos, en lugar de solo a dos cifras. Además, pueden realizar análisis más complejos que meras comparaciones entre dos entidades y hasta analizar los datos de más de dos partes simultáneamente. La escalabilidad de estos cálculos es la misma que si todos los datos estuvieran en manos de una de las partes. La diferencia es que las partes no pueden ver los datos de la otra y solo pueden ejecutar ciertos tipos de análisis permitidos.

La tecnología subyacente también se puede aplicar a problemas mucho más complejos, como la segmentación y activación de públicos basadas en una variedad de parámetros y el entrenamiento de un modelo de aprendizaje automático. Utiliza una fracción de los recursos de procesamiento que necesitaban las soluciones anteriores a problemas de computación segura multipartita, lo que ayuda a las empresas a controlar sus costes y el consumo de energía, y hace posibles iniciativas que antes habrían sido demasiado lentas, inseguras o caras.

Es importante destacar que los derechos de acceso de la otra parte son totalmente revocables. Desde el punto de vista de la seguridad, esto significa que se pueden retirar en cualquier momento, con lo que la otra parte perdería el acceso de inmediato. En cuanto a la actualización de los datos, se trata de una mejora enorme con respecto al enfoque tradicional de compartir copias de datos, ya que una copia no se puede mantener actualizada sin incurrir en una complejidad y un coste continuos.

3. Precios por segundo con elasticidad rápida y escalabilidad lineal

Con Snowflake, las empresas solo pagan por los recursos que utilizan, y pueden ampliar y reducir la prestación según sus necesidades, en lugar de pagar una cantidad fija por recursos de procesamiento que puede que no usen. Además, los costes de la empresa no aumentarán si esta permite que un partner combine una vista de los datos de la empresa con los suyos propios o los consulte. La parte que ejecuta la consulta paga el coste derivado del procesamiento que Snowflake realiza para ejecutarla, aunque puede estar relacionada total o parcialmente con los datos de la otra parte.

4. Enriquecimiento de datos seguro

Siempre habrá casos en los que dos partes carezcan de un identificador común. Esto puede hacer que la combinación de datos y la búsqueda de solapamientos sobre la base de una sola clave (como la dirección de correo electrónico) no ofrezca todas las correspondencias posibles entre los datos de las dos partes. En casos extremos, puede que el responsable de marketing solo tenga direcciones de correo electrónico y que el editor solo tenga direcciones IP, pero que, aun así, deseen combinar sus datos para realizar análisis permitidos. Cuando los clientes de Snowflake se encuentran en esta situación, pueden recurrir a proveedores de Snowflake Data Marketplace de enriquecimiento de datos y unión de identidades, como Acxiom, Epsilon y Neustar. Estos proveedores cuentan con una gran cantidad de datos de clientes de terceros que los responsables de marketing y los editores pueden utilizar para afinar y activar sus públicos, así como para aumentar las tasas de combinación y coincidencia. El paso de enriquecimiento de datos solo tiene lugar en la data clean room, y los datos de terceros no se incorporan a las cuentas de Snowflake de ninguna de las partes.

5. Efecto del Data Cloud

Dado que miles de empresas de todo el mundo ya confían en Snowflake para almacenar de manera segura sus datos confidenciales, Snowflake ofrece una forma sencilla de compartir datos externamente y establecer una clean room con otra parte.

Snowflake se ejecuta en los tres principales proveedores de nube (AWS, Microsoft Azure y Google Cloud Platform) y en todas las regiones, por lo que las partes nunca tendrán que trasladar los datos fuera sus cuentas de Snowflake para combinarlos. Por otra parte, una [estructura de gestión global unificada](#) y una [estructura de gobernanza de acceso](#) permiten combinar y analizar con fluidez los datos entre varias zonas geográficas y regiones específicas del proveedor de nube.

LO QUE DISTINGUE A SNOWFLAKE: MEJORA DEL RENDIMIENTO DE ANUNCIOS Y ROI, Y PRIORIZACIÓN DE LA PRIVACIDAD DE USUARIO

Los responsables de marketing y los editores deberían considerar la inminente desaparición de las cookies de terceros como una oportunidad, en lugar de como una catástrofe. La adopción de data clean rooms para detectar solapamientos de públicos y la aplicación de técnicas avanzadas de aprendizaje automático dentro de ellas para mejorar el modelado de datos pueden ayudarles a ser más rigurosos a la hora de segmentar y activar audiencias. Se trata de una alternativa a la omnipresente práctica actual de sobreindexar muestras diminutas de datos capturados mediante el seguimiento del comportamiento para extrapolar los recorridos de los clientes en públicos de gran tamaño.

En resumen, las data clean rooms de Snowflake son eficaces en cinco ámbitos clave:

- **data sharing seguro** de conjuntos de datos actualizados listos para consultar sin ETL;
- **controles de privacidad** para posibilitar análisis protegidos de billones de filas de conjuntos de datos combinados por dos o más partes sin revelar los datos de una empresa a las otras partes;
- **valor** derivado de la capacidad de ampliar y reducir según las necesidades, en lugar de pagar una cantidad fija de recursos de procesamiento;
- **enriquecimiento seguro de datos** mediante proveedores de identidades que ayuden a establecer correspondencias cuando sea necesario;
- **efecto del Data Cloud** debido a la escala de Snowflake, que cuenta con miles de clientes, y a su capacidad de ejecución en diversas nubes y zonas geográficas

Con las data clean rooms de Snowflake, todos los participantes en la cadena de valor se benefician. Los responsables de marketing pueden obtener un mayor ROAS gracias a una orientación más precisa de la publicidad y a un potente modelado de públicos similares. Los editores pueden reducir los CPM mejorando el rendimiento de su inventario y demostrando el impacto mediante comparaciones directas (y anónimas) entre las estadísticas de exposición a anuncios y los datos de compra de los responsables de marketing. Por último, se respeta la privacidad de los usuarios en Internet, y estos reciben ofertas más relevantes y personalizadas.

ACERCA DE SNOWFLAKE

Snowflake permite a cualquier organización movilizar sus datos con Snowflake Data Cloud. Los clientes utilizan Data Cloud para unificar, descubrir y compartir datos de forma segura, y ejecutar diversos workloads analíticos. Independientemente de la ubicación de los datos o de los usuarios, Snowflake ofrece una experiencia de datos única que abarca varias nubes y regiones geográficas. Miles de clientes de numerosos sectores, incluidas 543 de las empresas que figuran en Forbes Global 2000 (G2K) (2022), a fecha del 31 de octubre de 2022, utilizan Snowflake Data Cloud para impulsar sus negocios. Más información en snowflake.com.



© 2021 Snowflake Inc. Todos los derechos reservados. Snowflake, el logotipo de Snowflake y el resto de nombres de productos, funciones y servicios de Snowflake mencionados en este documento son marcas registradas o marcas comerciales de Snowflake Inc. en Estados Unidos y otros países. El resto de logotipos o nombres de marcas mencionados o utilizados en este documento se usan únicamente con fines identificativos, y pueden ser las marcas comerciales de sus respectivos titulares. Snowflake puede no estar asociado con, patrocinado o apoyado por cualquiera de dichos titulares.

NOTAS FINALES

¹ bit.ly/3x2icxl

² bit.ly/3x4DGZy

³ bit.ly/361OFrL

⁴ bit.ly/2Trdj26

⁵ cnet.co/3A5pUsl

⁶ bit.ly/367KBpO