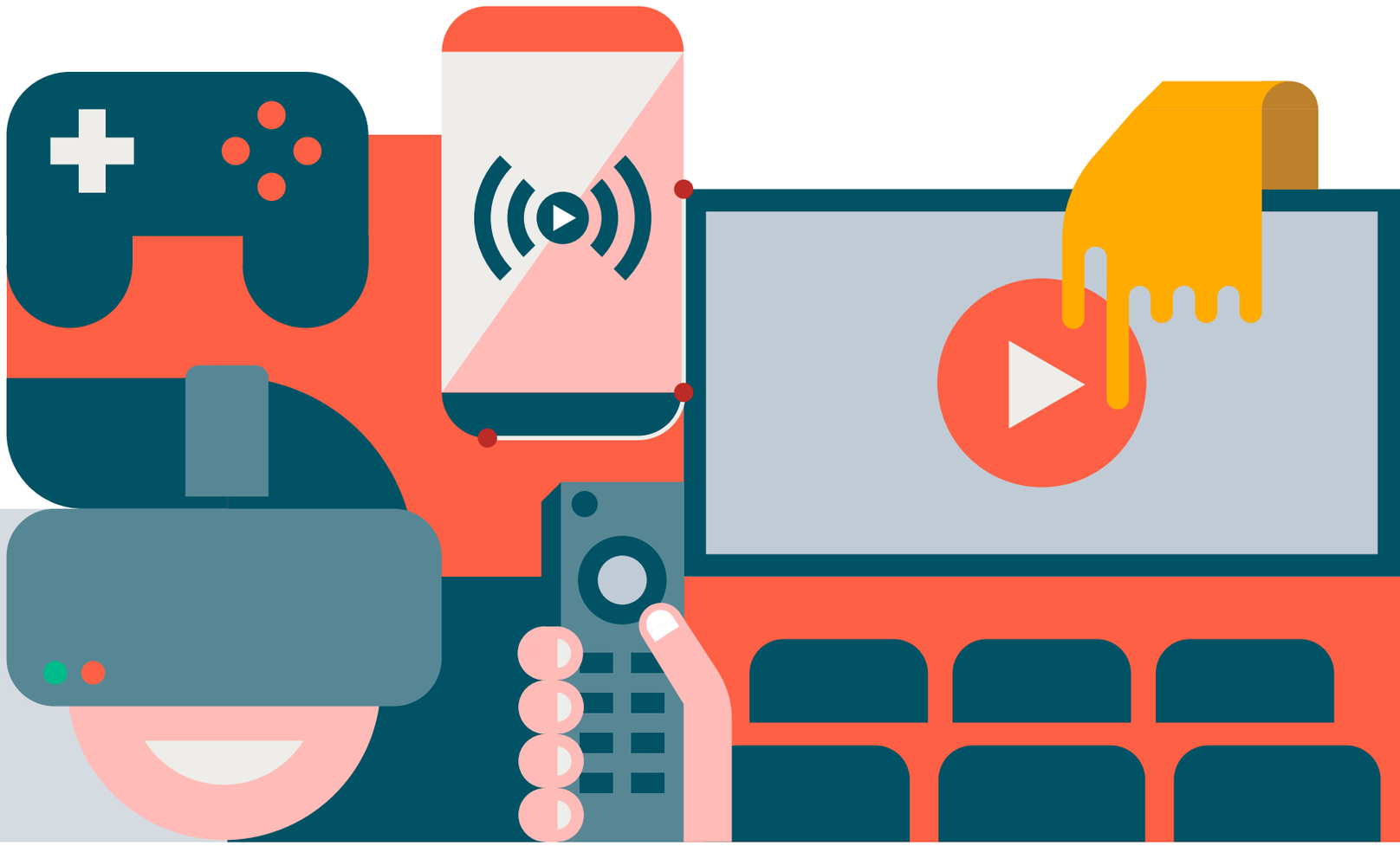
EBOOK

Enable Team Collaboration and Drive Intelligent Outcomes with Data, Analytics and AI

By leveraging Lakehouse for Media & Entertainment



databricks

Contents

databricks

Introduction

There are few industries that have been disrupted more by the digital age than media and entertainment. For decades, media organizations acted as wholesalers for content, which was a vehicle monetized mostly through advertising. With little focus on the consumer experience, the world of broadcasting, outdoor advertising, publishing and entertainment remained largely unchanged until the early 2000s. Then came digital.

As digital took off, the rise of FAANG companies (Facebook, Amazon, Apple, Netflix, Google) heightened consumer expectations around smarter, personalized experiences, making data and AI table stakes for success. Brands have shifted their ad budgets to digital channels such as connected TV, mobile and search advertising to more definitively target their ad spend, while also driving compliance with increasing privacy regulations.

Driving better data, analytics and AI outcomes for consumers, advertisers and employees is now a board-level initiative for most media and entertainment companies. The problem? Traditional data architectures weren't built to support AI/ML use cases, especially across broad teams of data engineers, data scientists and analysts, while supporting the scale and agility media and entertainment companies need to support evolving customer demands.

databricks

The solution? Media and entertainment organizations are making heavy investments in modern data technologies and industry partnerships to support a more thoughtful use of their data to shape the entire consumer, advertising and content lifecycle.

This is achieved by:

- Removing data silos by placing all data — regardless of type or frequency — in a single, open architecture, including unstructured data like video, images and voice content
- Ensuring data is in a ready state for all analytics and AI/ML use cases
- Having a cloud infrastructure environment based on open source and open standards so IT and data teams can move with agility

The Lakehouse for Media & Entertainment is doing just this. It's an ecosystem and business solution approach enabling teams across the entertainment industry to collaborate and innovate around data and AI. It eliminates technical limitations that have impeded collaboration across the value chain and enables data teams to drive greater 1:1 personalization at scale, operate in real time and capitalize on the full value of their content.

Welcome to the Lakehouse for Media & Entertainment.

Transforming Media and Entertainment With Lakehouse

"With Databricks Lakehouse Platform on AWS, Warner Bros. Discovery is powering the future of content discovery and audience experiences. By leveraging data to better predict consumer behavior and provide personalized content recommendations in real time, we are able to customize the viewer experience and improve overall engagement for our customers."

Martin Ma
GVP, Engineering at Warner Bros. Discovery

WARNER BROS. DISCOVERY

databricks

Media and Entertainment Transformation Trends

Media and entertainment companies are not sitting idly by in this environment. Today, successful teams are addressing these challenges by leveraging fast and connected data from all corners of the enterprise. Four trends are driving transformation in media:

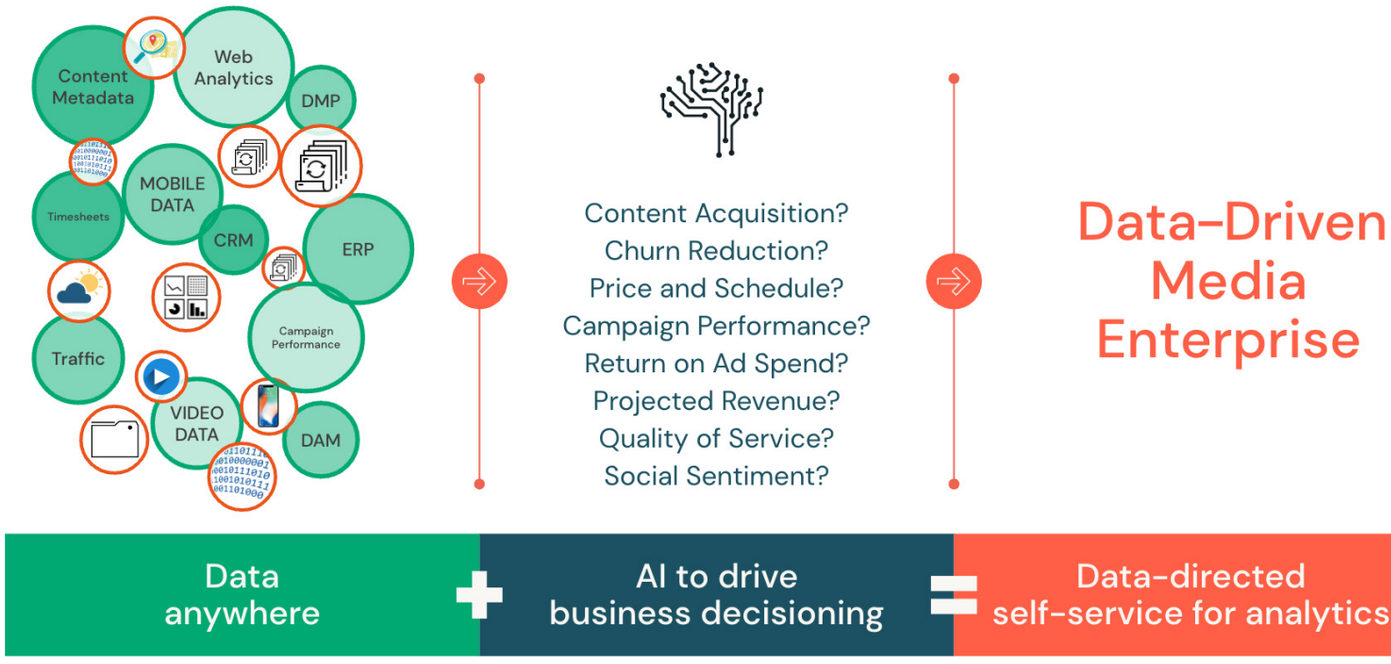
Rise of the subscription economy

The subscription economy was on the rise before the pandemic, but the breadth and depth of subscriptions in nearly every industry is expected to last, even after the pandemic subsides. The financial services firm UBS predicts that this "subscription economy" will grow to $1.5 trillion by 2025, more than double the $650 billion it's estimated to be worth now. Direct-to-consumer is now the most important revenue channel for most media and entertainment companies, with traditional distribution under pressure to evolve.

The dominance of digital triopoly in advertising

Google, Meta (Facebook) and Amazon now own two-thirds of all money made in digital advertising, with their dominance further strengthened by privacy regulations (GDPR, CCPA). In a world where the FAANG companies have unseated the traditional ad incumbents, these incumbents must move quickly to understand who is viewing their ads, drive better outcomes for advertisers through data and ultimately become easier to transact with, since FAANG companies will continue to accelerate their investment in ease of access to data and automation for customers.

Media and entertainment organizations need a platform to connect disparate data sources, apply intelligence, scale governance and use data for competitive advantage



databricks

Privacy and compliance

With Apple removing the Identifier for Advertisers (IDFA) and Google planning to stop supporting third-party cookies on its Chrome browser in 2024, this effectively ends two decades of media- and data-driven-performance targeted advertising. Marketing leaders and their teams must prepare for a future of cookieless advertising and focus on consent-based advertising instead. This will include overhauling first-party data strategy, changing playbooks, resetting measurements and reevaluating spend across Google, Meta and Amazon.

Focus on consumer experience

Consumers want to be connected on their own terms. Data privacy regulations (GDPR, CCPA) and the elimination of third-party cookies have diminished the ability for data teams within media and entertainment organizations to gather customer insights. Meanwhile, the rise of subscriptions is shifting the way consumers engage with content. Reestablishing a direct and primary relationship with those consumers has become a core priority for organizations, and the competitive pressure is only increasing. As a result, many companies are turning to customer data platforms (CDPs) to help overcome the challenge of building rich, behavioral data sets of customer interest and intent.

Creating optimal in-game experiences for SEGA

"With Databricks Lakehouse, we have access to data inputs across our entire player community, gaining insights that we can quickly build into new features they'll love."

Felix Baker
Data Services Manager, SEGA Europe

SEGA®

databricks

Media and Entertainment Data Challenges

Creating a unified audience profile

Audience data has traditionally been captured, stored and managed directly in disparate systems (e.g., DMP, ESP, data lake, data warehouse), depending on size/granularity, intended use case(s) and data types. This siloed approach is incredibly complex, especially when it comes to managing customer data as an asset that can be used to support a variety of use cases (e.g., personalization, content recommendations, next-best action or offer). To address this, media and entertainment teams need a single platform that can provide a 360-degree view of their audience, bringing together product telemetry, business analytics, application health and quality, as well as data outside of their experience such as social media listening.

Delivering a flawless user experience

A byproduct of media consumers having more choice means the customer experience matters more than ever. But this requires being able to identify the quality of service issues in near real-time, a capability that is not directly supported by legacy tech stacks. Data warehouses cannot support data processing at B2C scale, nor are they the right place to handle streaming ML workloads for real-time consumer lifecycle use cases.

Maximizing all your media data

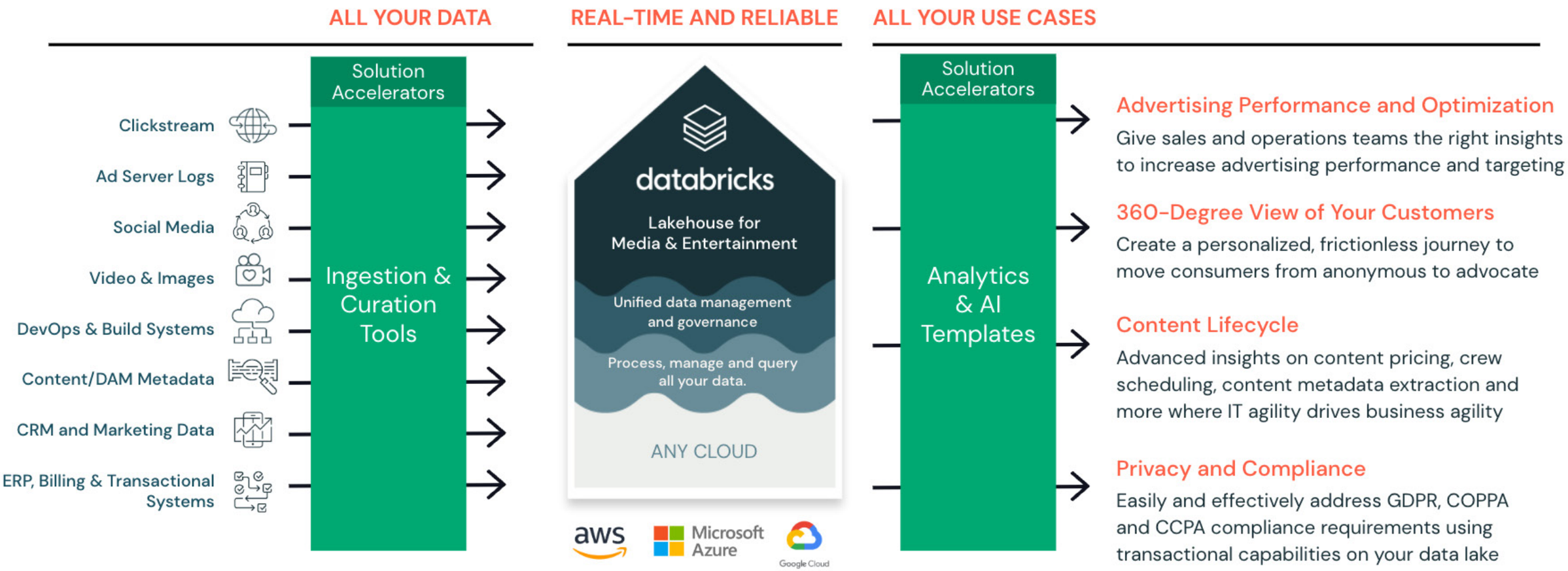
Media companies are built on unstructured data like video, images and audio files. Therefore, a team's ability to analyze unstructured data is essential for effective media asset management. Marketing teams need to be able to leverage archived content for campaigns, production teams need to look for existing content to include in new productions, and sales teams seek IP they can package and sell to other media companies. The challenge with legacy data warehouse solutions is they cannot handle unstructured data workflows with video, image and audio files. This prevents companies from unlocking the potential of their most valuable assets, leaving countless dollars on the table.

Moving beyond aggregation to advanced analytics

Complex AI use cases for the advertising and consumer lifecycle currently require the expensive acquisition and harmonization of massive data sets. Advanced insights help to improve content pricing, crew scheduling, content metadata extraction and more where IT agility drives business agility. This can only be done on an enterprise data platform built from the ground up to support ML/AI.

Lakehouse for Media & Entertainment

Drive more intelligent consumer and advertiser outcomes with data, analytics and AI



The Databricks Lakehouse for Media & Entertainment is the only enterprise data platform that allows you to leverage all your data, from any source, for any workload to always offer greater consumer personalization with more intelligent outcomes for advertisers and partners, at the lowest cost, with the greatest investment protection.

Why Lakehouse for Media & Entertainment?

Accelerate audience and advertiser outcomes on an open, collaborative platform for data, analytics and AI.

Get a real-time 360° view of your audience

Bring together all your structured and unstructured data — clickstream, demographic, social — in a single platform for analytics and AI. With a holistic view of the consumer journey, organizations can understand content preferences that help deliver more personalized experiences and develop more targeted advertising and engagement.

Mitigate churn and increase ARPU

With an agile cloud-based platform, organizations can quickly and reliably process massive amounts of data and feed it to downstream systems to deliver a 1:1 experience on any channel at any time. As consumers' expectations around real-time recommendations keep rising — and media companies fight for consumer attention — ensuring that personalization approaches real-time is becoming a requirement for many organizations.

Drive revenue from your entire content library

Media companies are built on unstructured data like video, images and audio files, so the ability to analyze unstructured data is essential for effective media asset management. The Lakehouse for Media & Entertainment can be used by marketers to leverage archived content for campaigns, production teams looking for existing content to include in new productions and sales teams seeking IP they can sell to other media companies.

Put ML and AI at the core of your business

Unlock the power of machine learning to better understand consumer, employee and advertiser needs. When all your data is centralized and seamlessly connected by a full suite of collaborative analytics and machine learning tools, data teams can work together to build powerful predictive models that drive new innovations in personalization, content monetization and advertiser outcomes.

What makes the Databricks Lakehouse special?

Reliability and performance

Lakehouse offers unparalleled scale, speed and cost savings. The best data warehouse is a lakehouse.

Optimized for real-time

Uncover actionable insights, drive 1:1 personalization and enhance the player experience with machine learning at the core.

All your data on one, agile platform

Extract the full value of your data using the platform built to handle your structured, unstructured and streaming data.

databricks

SOLUTIONS

Ad Performance and Optimization

Advertising businesses are facing more pressure than ever as fewer companies control more of the market based on their ability to deliver easy, self-service buying tools, targeting the right audiences at the right time.

Often, legacy technology is blocking campaign optimization. This can manifest in many ways, including some of the following examples:

- Inability to scale efficiently. Costly, hard-to-scale infrastructure cannot process the diverse audience data (e.g., demographics, clickstream, content metadata, social, call center) required to create a single view across all advertising.
- Teams lack the ability to reliably process streaming data. Optimizing ad placements in real time is an X-factor for media teams. Those who can process streaming data and blend it with historical data for real-time insights create strategically differentiated experiences that increase engagement and monetization.
- Legacy technology creates cumbersome ML workflows. Siloed data teams and complex processes for building, tracking and deploying analytics and ML models mean stale and outdated personalization. And in a world where customer experience is everything, getting this right is key to performance.

Differentiated capabilities

The lakehouse makes your advertising smarter by placing your data on a single platform with machine learning at the core. By providing your ad ecosystem with a data platform that can handle massive data sets, Databricks can give your decision makers the ability to price and manage inventory in near real-time while driving the best outcomes for end clients.

- The Lakehouse for Media & Entertainment supports the largest of data jobs at near real-time intervals

  Example: Customers are tapping billions of data points from ad exchanges and daily ad opportunities.
- The lakehouse event-driven architecture provides a simpler method of ingesting and processing batch and streaming data than legacy approaches, such as lambda architectures. This architecture handles the change data capture and provides ACID compliance to transactions.
- Delta Live Tables simplifies the creation of data pipelines and automatically builds in lineage to assist with ongoing management
- The lakehouse allows for true real-time stream ingestion of data, and even analytics on streaming data. Data warehouses require the extraction, transformation, loading and then additional extraction from the data warehouse to perform any analytics.
- Databricks Photon engine provides record-setting query performance, enabling users to query even the largest of data sets to power real-time decisions in BI tools

Benefits

- **8%–10%** lift in revenue with personalized experience/ cross-sell use cases driven by data and AI
- **15%–25%** sales increase from improved ROI for ads and promotion allocation
- **30%–40%** improved return on ad spend (ROAS) using ML algorithms to optimize campaigns

Common Use Cases

- Multi-touch Attribution
- Sales Forecasting and Ad Attribution
- Customer Segmentation
- Customer Lifetime Value
- Inventory Forecasting
- Real-Time Bidding for Agencies

SOLUTIONS

360-Degree View of Your Customers

Consumer expectation for entertainment everywhere, matched with smarter, more personalized experiences, means data teams need to look at their data as the new content and look at AI as the new market research.

Media and entertainment companies need to build streaming, direct-to-consumer services to stay competitive. The key to a successful streaming service is personalization, including 1:1 marketing at scale. The only way to achieve this is with data. By applying advanced analytics and ML techniques to audience data, media companies can build personalized experiences that better attract and engage subscribers and mitigate churn.

Data warehouses and legacy analytics platforms face two primary limitations when it comes to delivering personalization at scale:

- The lack of stream processing and ability to blend with historical data means teams are severely limited when it comes to optimizing customer experiences in real time
- Orchestrating and managing end-to-end production workflows remains a bottleneck for most organizations, relying on external tools or cloud-specific solutions

Because of these limitations, data teams are forced to make compromises in how they personalize experiences. Whether that's relying on outdated data, creating generic, one-size-fits-all solutions, or missing out on critical customer insights that could lead to deeper engagement and retention.

Personalization is critical to today's media and entertainment organizations. And this starts with having a data platform that can provide a real-time 360-degree view of your audience.

Companies that use Databricks distributed computing to run personalization models are able to more precisely tailor recommendations. In some categories, variance (error) has been reduced from 29% to 3% (71% accuracy to 97% accuracy).

Differentiated capabilities

The Lakehouse for Media & Entertainment empowers teams of all sizes to deliver more effective 1:1 personalized experiences at scale by connecting data silos, giving teams a single platform for all data, analytics and AI workloads — critical to customer patterns and creating recommendations.

- Delta provides optimized storage and querying that reduce the retrieval time of all types of data without the costly extraction that is required of data warehouse systems. This additional cost is reflected in the additional time required and incremental processing charges.
- Databricks provides users with the ability to fully distribute model calculations. The results of these analyses are captured and persisted in Delta for fast retrieval.
- Hyper-parameter optimization enables users to more efficiently find the most optimal hyper-parameter values, leading to better recommendations
- MLflow provides a streamlined repository for tracking the results of experiments and managing the deployment of models

Common Use Cases

- Quality of Service/Experience
- Toxicity Detection
- Recommendation Engine
- Survivorship/Attribution
- Churn Prediction
- Customer Segmentation
- Customer Lifetime Value
- Next-Best Offer/Next-Best Action
- User Cold Start

SOLUTIONS

Content Lifecycle

Media and entertainment embodies the industry of unstructured data. Video, images and audio content seek to engage and inform consumers looking for news and entertainment. While many organizations do an excellent job personalizing the consumer experience, driving greater engagement and monetization of their content, one of the biggest lost opportunities organizations face is not extracting the full value of their unstructured data.

Often, companies have content sitting idle in physical and digital asset management systems. Historically, value is realized at launch, but teams are increasingly finding ways to repackage content, provide recommendations deeper into their catalog, and use machine learning and AI to protect rights and revenues of their unstructured data. This leads to questions such as: "Are we leaving money on the table?" "Are we consistently tracking the rights of my content across channels?" "Are we empowering our teams to be more productive?" With unstructured data, it's not just about repackaging content that exists, but rather how teams can surround that content with more effective marketing and smarter operations.

In summary, companies that create content often sit on mountains of data with limited ability to identify and monetize it.

That's where Lakehouse for Media & Entertainment stands apart. Because the lakehouse supports any type of data (structured, semi-structured, unstructured) on a single platform, teams can more effectively realize value through faster and more accurate content identification, improved customer experiences through richer personalization and automated flagging for potential rights and revenue violations.

Differentiated capabilities

The lakehouse supports the use of all types of structured, unstructured and semi-structured data with Delta, Apache Spark™ and partnerships with companies including Labelbox and John Snow Labs to make your content smarter and work harder for you. As a result, the lakehouse empowers you to extract greater value from more of your unstructured data.

- Unlike EDWs that store unstructured data as blobs, Delta enables companies to store data of all structures and immediately access that unstructured data to transform it into useful information
- Databricks partners with leading companies such as Labelbox, which enable companies to rapidly label images for use in computer vision projects
- Databricks partners with John Snow Labs, which provides enhanced NLP libraries to assist in the classification of text data

Common Use Cases

- Content Repackaging
- Model-Assisted Labeling
- Contextual Advertising
- Recommendation Engine
- Content Localization
- Rights Management and Royalty Reporting
- Sentiment Analysis

Leading Brands Choose Databricks

About Databricks

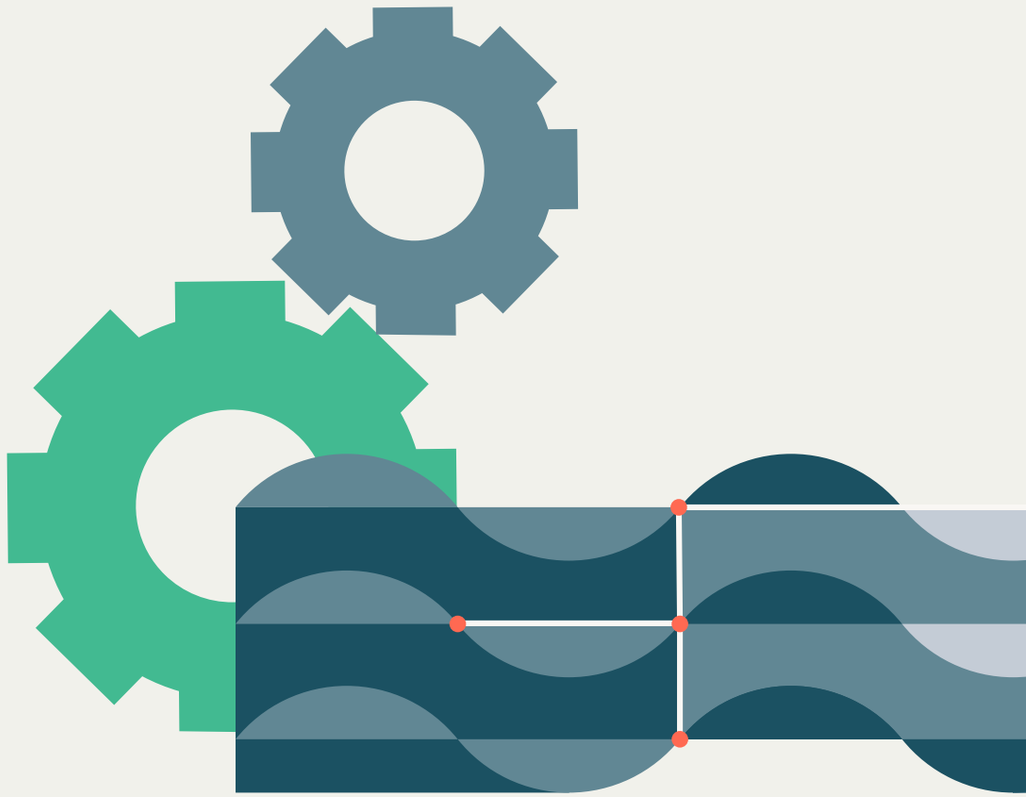
Databricks is the data and AI company. More than 7,000 organizations worldwide — including Comcast, Condé Nast, H&M and over 40% of the Fortune 500 — rely on the Databricks Lakehouse Platform to unify their data, analytics and AI. Databricks is headquartered in San Francisco, with offices around the 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. To learn more, follow Databricks on Twitter, LinkedIn and Facebook.

Get started with a free trial of Databricks and start building data applications today

START YOUR FREE TRIAL

To learn more, visit us at:

Media & Entertainment Solutions



databricks