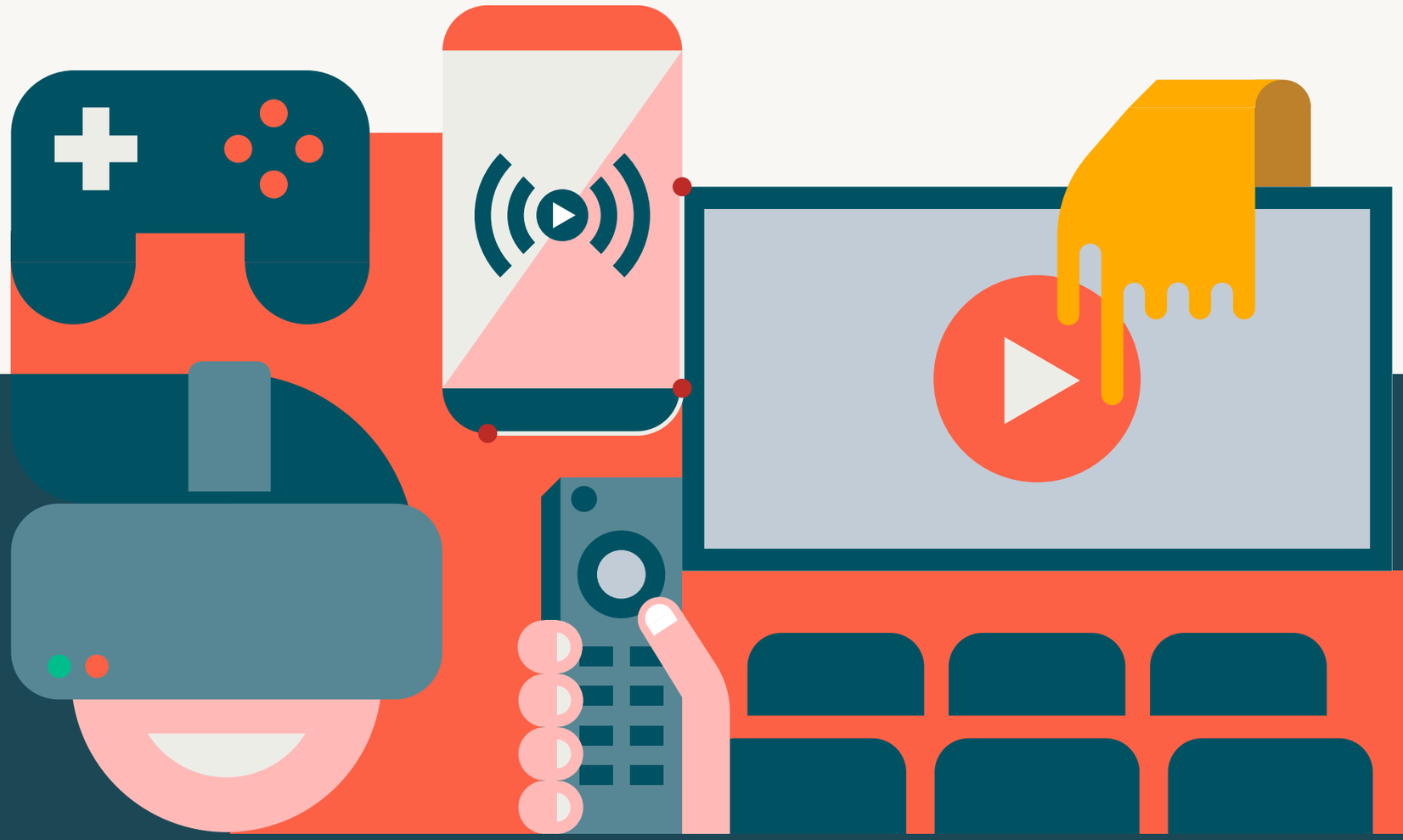
# Enable Team Collaboration and Drive Intelligent Outcomes With Data, Analytics and AI

By leveraging the Data Intelligence Platform for Media & Entertainment



databricks

# Contents


Introduction ..... 3

Media & Entertainment Transformation Trends ..... 5
- The rise of generative AI: Capturing everyone's mindshare ..... 5
- The increasing importance of data collaboration ..... 6
- Reducing the cost of the data estate ..... 6

Focus on Consumer Experience ..... 8

Media & Entertainment Data Challenges ..... 9
- Creating a unified audience profile ..... 9
- Delivering a flawless user experience ..... 9
- Maximizing all your media data ..... 10
- Moving beyond aggregation to advanced analytics ..... 10

Unlocking Innovation With the Data Intelligence Platform for Media & Entertainment ..... 11

Why the Data Intelligence Platform for Media & Entertainment? ..... 12
- Create more meaningful customer engagements ..... 12
- Extract the full value of your data and AI ..... 12
- Unlock productivity with AI and real-time streaming data ..... 12
- Optimize your content and advertising ..... 12

Solutions ..... 14
- Ad performance and optimization ..... 14
- 360-degree view of your customers ..... 16
- Content lifecycle ..... 18

Customers That Choose Databricks ..... 20

# Introduction

The Media & Entertainment industry is experiencing a period of unparalleled innovation, driven by rapid advancements in digital technologies and the growing influence of generative AI (GenAI). Trends leading the charge include the broad adoption of streaming platforms, the proliferation of content across multiple channels and the increasing importance of personalized user experiences. As a result, organizations have to play both offense and defense when it comes to their data and AI strategy.

In this new era, data is the backbone of the industry, enabling Media & Entertainment companies to better understand their audiences, optimize their content and advertising strategies and drive competitive innovation. However, the traditional data architectures that served the industry well in the past are struggling to scale to meet the needs of both internal and external customers.

One such force is GenAI, which is further amplifying the need for teams to re-evaluate the agility of their enterprise data platform. There is a GenAI use case in virtually every aspect of business, with content creation, personalization and enhancing customer support popping to the top of areas of investment. Yet, the integration of generative AI into existing data ecosystems presents its own set of challenges, ranging from control to production quality and cost.

This eBook explores the evolving data landscape in the Media & Entertainment industry, with a particular focus on the opportunities and challenges presented by generative AI. It provides a guide to building robust, scalable and secure data platforms that harness the power of these transformative technologies, while also addressing the critical considerations around cost-effective and responsible AI deployment.

Whether you're a data engineer, a data scientist or a leader in the organization, this eBook will equip you with the insights and strategies you need to navigate the future of data-driven innovation in your industry.

## Generative AI is taking the world by storm

**91%**

of organizations are experimenting with or investing in GenAI[1]

**75%**

of CEOs say companies with advanced GenAI will have a competitive advantage[2]

**40%**

increase in performance of employees who used GenAI[3]

1. Laying the foundation for data and AI-led growth, MIT Technology Review
2. CEO decision-making in the age of AI, IBM Institute for Business Value
3. How generative AI can boost highly skilled workers' productivity, MIT Management Sloan School

Data is the strategic differentiator when it comes to mitigating costs and out-innovating the competition. It's no wonder Media & Entertainment organizations are putting data at the center of their consumer, advertising and content lifecycle strategies.

How are teams doing this?

- Removing data silos by placing all data — regardless of type or frequency — in a unified, 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 that empowers IT and data teams to move with agility

The Databricks Data Intelligence Platform for Media & Entertainment is doing just this. The Databricks Platform empowers your entire organization to maximize the value of data and AI. It's built on lakehouse architecture to provide an open, unified foundation for all data and governance, and it's powered by a Data Intelligence Engine that understands the uniqueness of your data. As a result, the Data Intelligence Platform 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 Data Intelligence Platform for Media & Entertainment

## Transforming media and entertainment with the Data Intelligence Platform

"We selected Databricks Data Intelligence Platform to efficiently manage large-scale data processing in one place. It would have been impossible to develop ML models using data from siloed data lake and data warehouse environments, so with Databricks we have been able to simplify and unify our data architecture and accelerate time to market for ML solutions."

— SeungYong Oh, Head of Data Platform, Devsisters

DEVSISTERS

# Media & Entertainment Transformation Trends

Media & 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:

## The rise of generative AI: Capturing everyone's mindshare

Rapid advancements in generative AI, particularly large language models (LLMs) like ChatGPT, have fundamentally disrupted the fields of content creation, customer support, personalization and many more. Databricks has supported the development and deployment of AI agents supporting customer service teams in recommending actions and offers in real time to improve win-back and upsell. We've also seen massive investment around hyper-personalized customer experiences that adapt and learn audience behavior over time, as well as the deployment of fraud detection models that stay ahead of commercial cheaters. As the hype around generative AI continues to grow, media organizations are racing to integrate these technologies into their workflows. The focus is shifting from simply building competency around LLMs to getting them into production efficiently and, more importantly, into the hands of customers.

However, the integration of generative AI is not without its challenges. Media organizations must navigate business and ethical considerations, such as the potential for AI-generated content to be misused or to displace human creative talent. Careful governance and responsible deployment of these technologies are crucial to ensuring that the benefits of generative AI are realized while mitigating the risks.

## The increasing importance of data collaboration

Cookie degradation is here. This effectively ends two decades of media- and data-driven performance-targeted advertising. In response, organizations are exploring new creative ways to leverage third-party data to enhance customer identity through data collaboration. This includes overhauling first-party data strategy, changing playbooks, resetting measurements and reevaluating spend across Google, Meta and Amazon.

Data clean rooms have emerged as a key strategy, allowing media companies to maintain data silos while creating a secure environment for data exchange and collaboration. This helps to augment customer profiles, improve audience targeting and deliver more personalized experiences. The sharing and collaboration of data assets, both internally and externally, helps to break down silos and foster a culture of data-driven decision-making.

The shift toward data collaboration is a critical step in the Media & Entertainment industry's journey beyond the walled gardens of the past. Organizations can now unlock the full potential of their data assets and drive more value for their business.

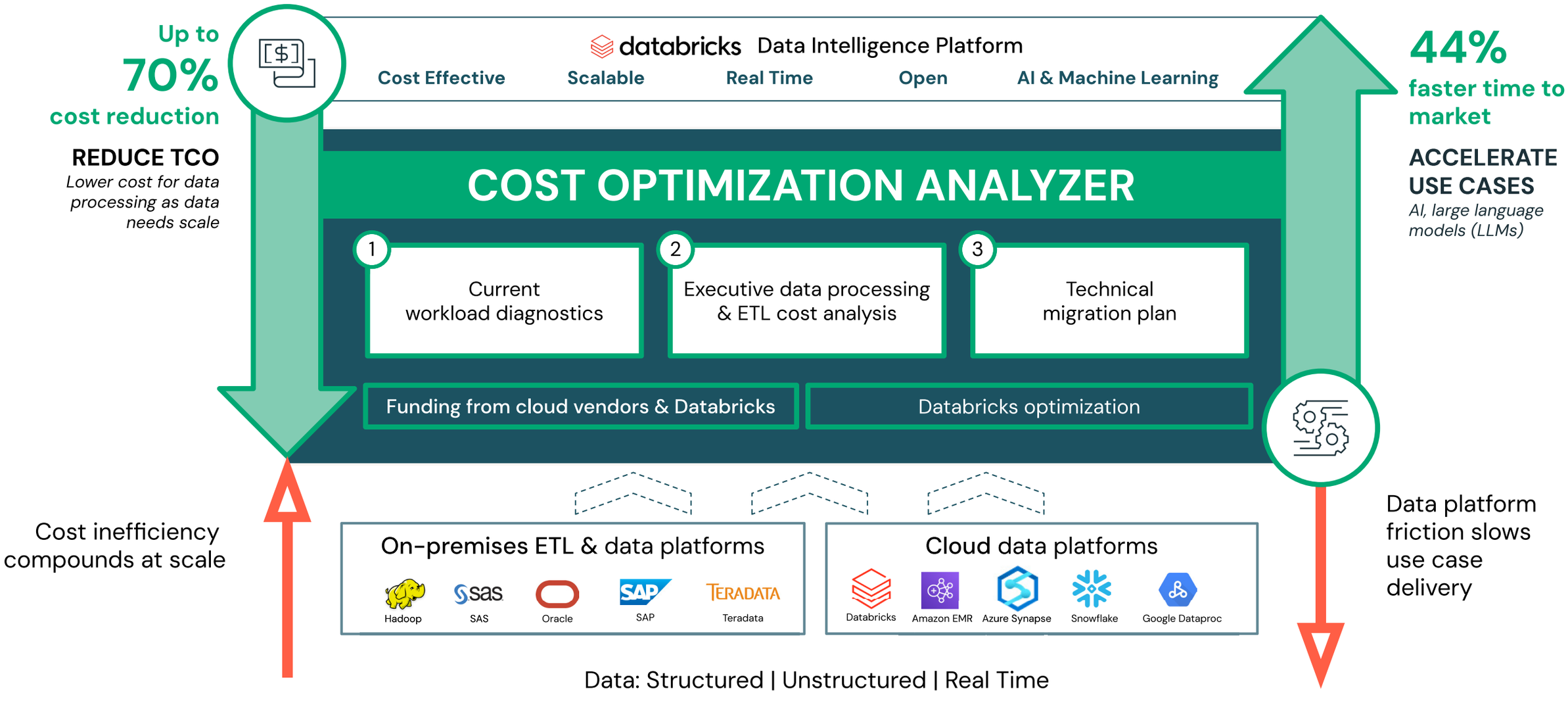
## Reducing the cost of the data estate

As organizations navigate the evolving data landscape, a key priority is to drive more value from their data platforms and reduce the overall cost of their data estate. One of the ways organizations are achieving this is by focusing on data platform interoperability. By ensuring seamless integration and data flow between different systems and applications, media companies can optimize their data infrastructure and unlock new insights that drive business value.

Additionally, media organizations are exploring ways to leverage cloud-based data solutions and serverless architectures to reduce the cost and complexity of their data management. A unified data platform better empowers teams to scale their data capabilities more efficiently, reduce maintenance overhead and free up resources for more strategic initiatives.

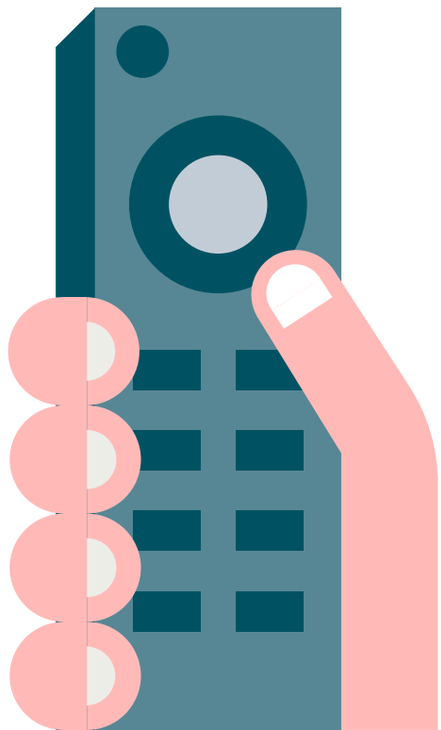
The drive to take cost out of the data estate isn't just about cutting expenses. Teams must consider how their data platform helps to improve the effectiveness of other platform investments such as a composable customer data platform, as well as aligning the data platform to better support the evolving needs of the business.

# Databricks Data Intelligence Platform delivers **lower costs and faster innovation**



# Focus on Consumer Experience

Consumers want to be connected on their own terms. Data privacy regulations such as the General Data Protection Regulation (GDPR) and the California Consumer Privacy Act (CCPA), and the elimination of third-party cookies have diminished the ability of data teams within Media & Entertainment organizations to gather robust customer insights. Meanwhile, the rise of subscriptions is shifting the way consumers engage with content, and by association, brands. 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 datasets of customer interest and intent.



## Discover how innovative companies across every industry are leveraging the Databricks Data Intelligence Platform for success

"Adobe is using the Databricks Data Intelligence Platform to help bring creativity to life, with end-to-end data management that unifies all data and AI at scale for over 92 teams and with 20% faster performance."

— Bin Mu, VP of Enterprise Data and Analytics, Adobe

Adobe

# Media & Entertainment Data Challenges

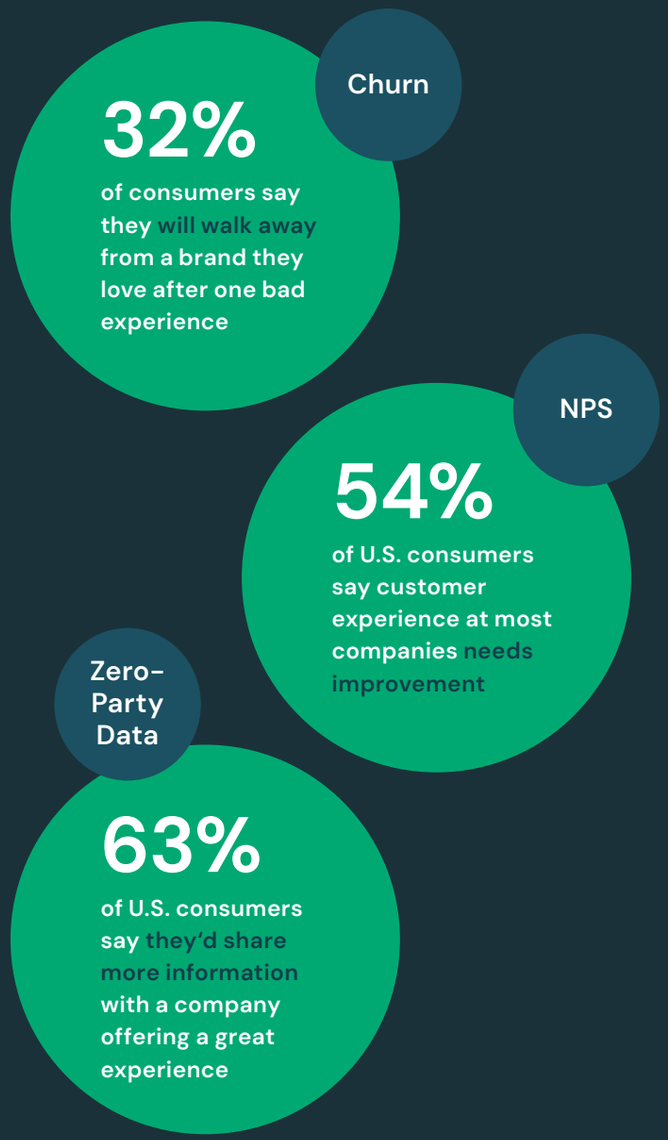
## Creating a unified audience profile

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 cases and data types. This siloed approach is incredibly complex, especially when it comes to managing customer data used to support a variety of use cases (e.g., personalization, recommendations, next best action) and teams across organizations. To address this, Media & Entertainment teams need a single, unified data 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. With a unified view of the customer, teams can better share insights and collaborate on actions that result in direct business value.

## Delivering a flawless user experience

A byproduct of more choices for media consumers is that the customer experience (CX) matters more than ever. CX and content are some of the best forms of strategic differentiation, which means personalization and quality of service (QoS) use cases require data insights delivered in near real-time, a capability not directly supported by legacy tech stacks. A data warehouse can't support data processing at B2C scale, nor is it the right place to handle streaming ML workloads for real-time consumer lifecycle use cases.

## Customer experience (CX) is now table stakes for success




Sources: PWC - Future of CX 2022 and Gartner 2023 CIO and Technology Executive Survey

## Maximizing all your media data

Media & Entertainment is the industry of unstructured data, like video, images and audio files. Therefore, a team's ability to analyze unstructured data quickly and efficiently 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 their inability to handle unstructured data workflows with video, image and audio files. This often 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 require the acquisition and harmonization of massive datasets. All of which depends on strong governance and security, especially with the advent of generative AI. When teams build on this strong foundation, they unlock the ability to pursue AI initiatives — from using APIs like OpenAI to custom-built models — without compromising data privacy and IP control, turning data into profitability. This can only be done on an enterprise data platform built from the ground up to support ML/AI.

# Unlocking Innovation With the Data Intelligence Platform for Media & Entertainment

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

The Databricks Data Intelligence Platform for Media & Entertainment was designed to meet the dynamic needs of the Media & Entertainment ecosystem at scale — while delivering enterprise-grade security and intelligently reducing costs to operate. The Data Intelligence Platform is built with AI at the core and is the only platform to combine all your data seamlessly — from BI to generative AI. With it, you can leverage cutting-edge AI applications across your entire organization at an operational speed and cost efficiency never before possible.

databricks

# Why the Data Intelligence Platform for Media & Entertainment?

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

## Create more meaningful customer engagements

The Data Intelligence Platform gives you a single, unified view of your audience so you can deliver 1:1 personalized experiences in real time, at scale. Cultivate relationships with your customers on their terms by understanding sentiment across channels and personalizing recommendations. Bring together all your structured and unstructured data — clickstream, demographic, social — in a single platform for analytics and AI.

Learn how Minecraft is driving gamer personalization for millions of global gamers in real time

## Extract the full value of your data and AI

Most media organizations struggle to harmonize data engineering, data science and business analytics so that they can work from a common set of business tools. This means you're sitting on mountains of data with limited ability to identify and monetize it. The Data Intelligence Platform helps you move at the speed of business, combining your consumer, content, advertiser and operational data on a single platform with common tools and a full suite of capabilities.

Learn how Showtime is creating TV hits using AI

## Unlock productivity with AI and real-time streaming data

Whether your employees are creating data assets or consuming them, ensuring they have the right data at the right time is critical to success. With the Data Intelligence Platform, teams across the company — not just data teams — can collaborate more effectively. The result? Employees are freed up to self-serve and focus on high-impact customer scenarios — from personalization to churn mitigation.

Learn how theTradeDesk enables natural language translation at scale

## Optimize your content and advertising

When it comes to buying or selling ads, driving engagement or mitigating churn, identifying the right person, at the right moment, with the right message is critical. The Data Intelligence Platform makes your advertising and content smarter. By placing all your data on a single platform, with machine learning at the core, every individual touching data can work simultaneously and collaboratively to deliver better outcomes for advertisers and derive greater revenue from your entire content catalog.

Learn how Condé Nast serves up multimedia content on a global scale

# Media & Entertainment Solution Areas

## Audience Experience

Create a personalized, frictionless journey to move consumers from anonymous to advocate

## Advertising Performance and Optimization

Give sales and operations teams the right insights to increase advertising performance and targeting

## Content Lifecycle

Advanced insights on content pricing, scheduling, metadata extraction and more where IT agility drives business agility

### Use Cases

- Audience Targeting and Personalization With Composable CDPs
- Personalization Engines
- GDPR/CCPA Compliance
- Churn Management
- Forum Moderation
- Streaming Quality of Service
- Data Clean Rooms
- Ad Attribution
- Inventory & Yield Management
- Segmentation & Targeting
- Campaign Performance
- Ad Verification
- Product Development Lifecycle
- Content Pricing and Analytics
- Crew/Resource Optimization
- Image/Profanity Recognition
- Slate Managernent & Monetization

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 for targeting the right audiences at the right time.

Often, legacy technologies block 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 AI and 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.

SOLUTIONS

# Ad Performance and Optimization

## Differentiated capabilities

A data intelligence platform makes your advertising smarter by placing your data on a single, unified platform with AI at the core. Here, Databricks can empower your decision-makers with the ability to price and manage inventory in near real-time while driving the best outcomes for end clients.

- The Data Intelligence Platform for Media & Entertainment supports the largest of data jobs at near real-time intervals. Learn how Acxiom is tapping billions of data points from ad exchanges and daily ad opportunities.
- The Data Intelligence Platform is based on lakehouse event-driven architecture, which provides a simpler method of ingesting and processing batch and streaming data than legacy approaches, such as lambda architectures. Lakehouse 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 Data Intelligence Platform 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, a process that adds time, cost and complexity.
- Databricks Photon engine provides record-setting query performance, enabling users to query even the largest of datasets to power real-time decisions in BI tools

## Customer benefits

- **8–10%** lift in revenue with personalized experience and 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 AI/ML 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 expectations for entertainment everywhere, matched with smarter, more personalized experiences, mean data teams need to look at their data as the new content and AI as the new market research.

Organizations 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, and increasingly with AI. By applying advanced analytics and AI use cases to audience data, media companies can build hyper-personalized experiences that better attract and engage audiences and subscribers while mitigating churn.

Data warehouses and legacy analytic 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.

Because personalization use cases are important to Media & Entertainment organizations, teams need a data platform capable of providing a real-time 360-degree view of audiences. This adds demonstrable value in the form of targeting and personalization to composable customer data platforms (CDPs).

SOLUTIONS

# 360-Degree View of Your Customers

## Differentiated capabilities

The Data Intelligence Platform for Media & Entertainment empowers teams of all sizes to deliver more effective 1:1 personalized experiences at scale by connecting data silos. This gives teams a single platform for all data, analytic and AI workloads, which is critical to customer patterns and creating recommendations.

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

## Customer benefits

- Companies that use Databricks 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).

## 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 & 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 have 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 ML and AI to protect the rights and revenues of their unstructured data. They're asking 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 the Databricks Data Intelligence Platform for Media & Entertainment stands apart. Because the Databricks Platform 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.

SOLUTIONS

# Content Lifecycle

## Differentiated capabilities

The Data Intelligence Platform supports the use of all types of structured, unstructured and semi-structured data types 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 Databricks Platform empowers you to extract greater value from more of your unstructured data.

- Unlike EDWs that store unstructured data as blobs, Delta Lake 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 enables 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

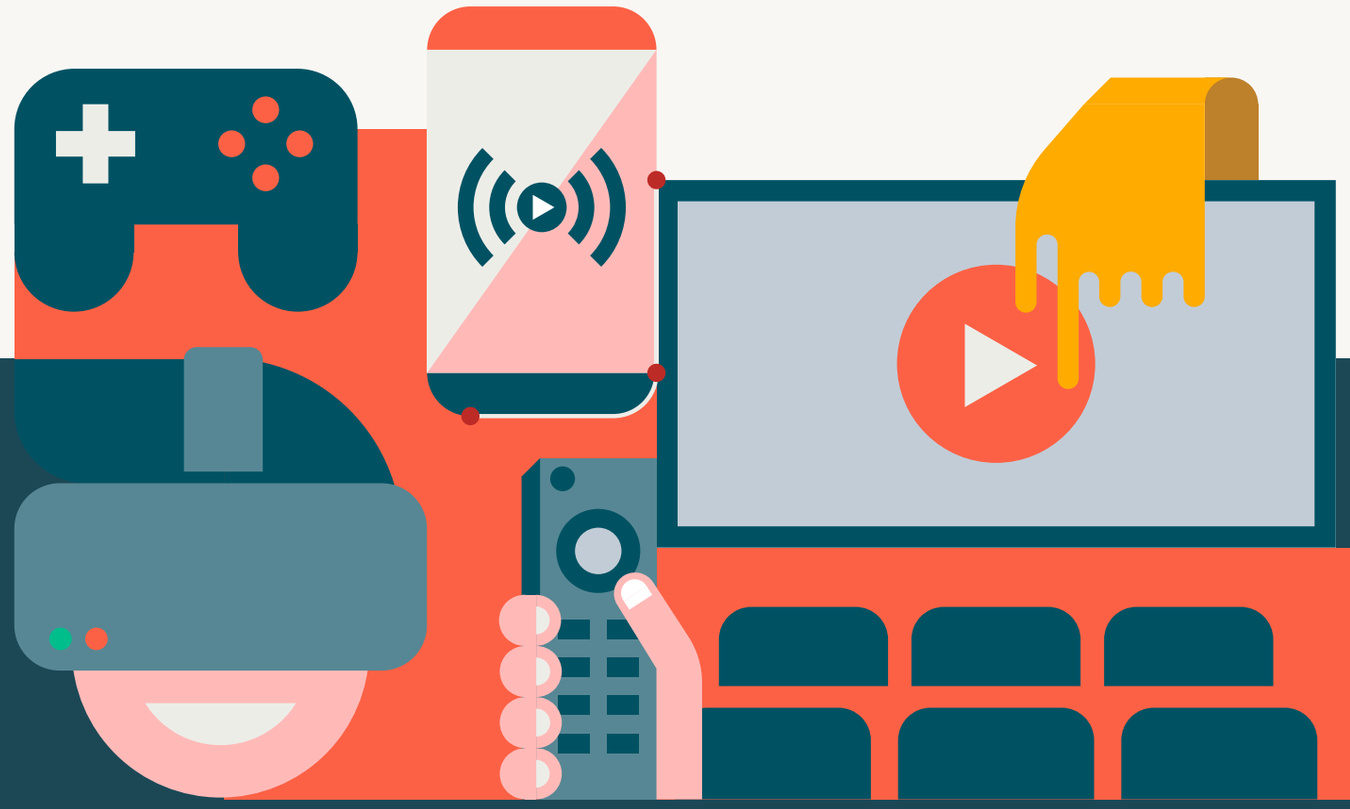
## Common use cases

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

# Customers That Choose Databricks

## About Databricks

Databricks is the data and AI company. More than 10,000 organizations worldwide — including Comcast, Condé Nast, Grammarly and over 60% of the Fortune 500 — rely on the Databricks Data Intelligence Platform to unify and democratize data, analytics and AI. Databricks is headquartered in San Francisco, with offices around the globe, and was founded by the original creators of Lakehouse, Apache Spark™, Delta Lake and MLflow. To learn more, follow Databricks on LinkedIn, X 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

© Databricks 2024. All rights reserved. Apache, Apache Spark, Spark and the Spark logo are trademarks of the Apache Software Foundation. Privacy Policy | Terms of Use