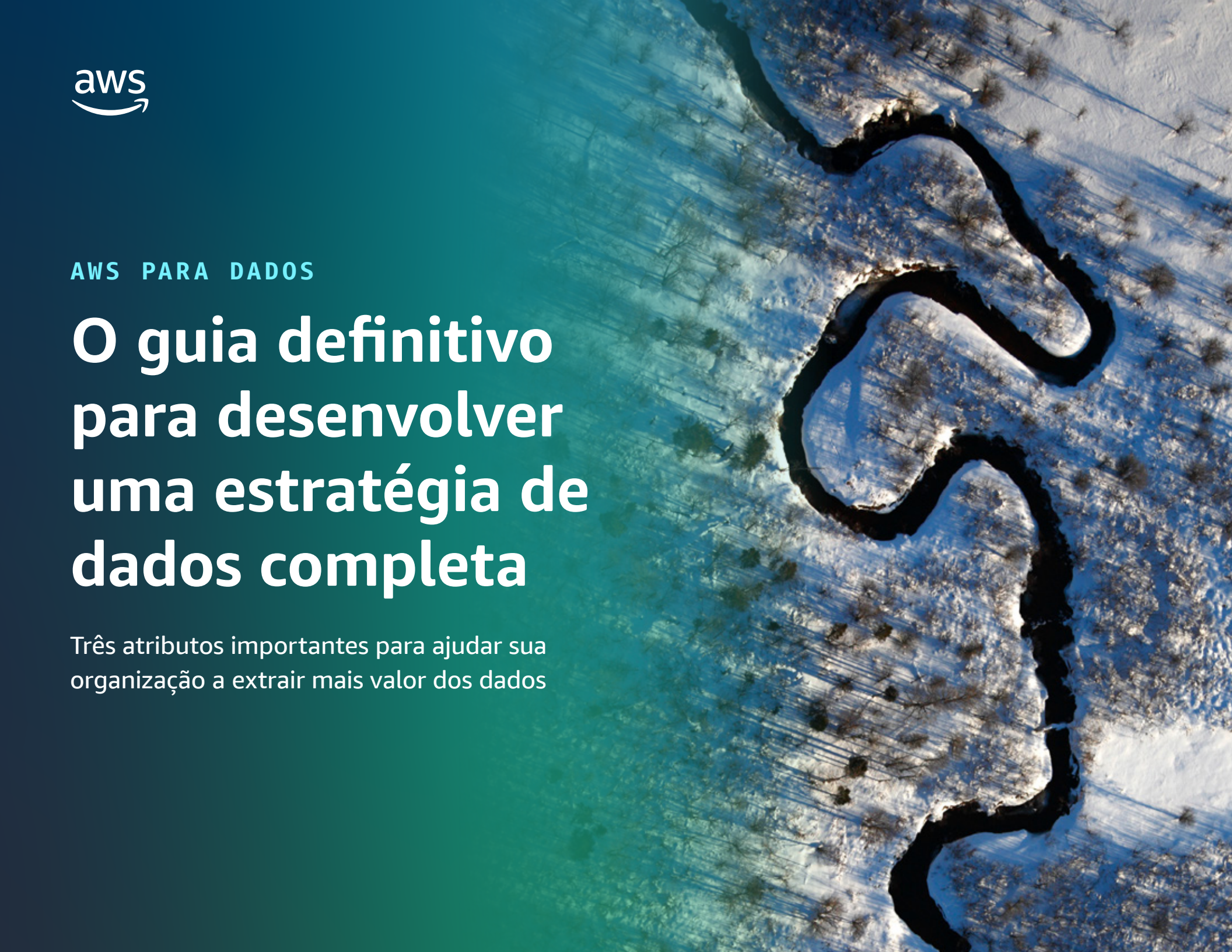




AWS PARA DADOS

# O guia definitivo para desenvolver uma estratégia de dados completa

Três atributos importantes para ajudar sua  
organização a extrair mais valor dos dados



# Sumário

Dados são a centelha de inovações relevantes.....	3
Principais desafios e considerações .....	4
Três atributos importantes que podem ajudar sua organização a extrair mais valor dos dados .....	5
Abrangência.....	6
Integração .....	11
Governança .....	15
Como tornar a segurança mais estratégica .....	19



## INTRODUÇÃO

# Dados são a centelha de inovações relevantes

Hoje, mais do que nunca, os dados são essenciais para toda aplicação, processo e decisão de negócios. Eles são a **gênese da invenção moderna**. E, no cenário atual de complexidade e rápida evolução, a maneira como você utiliza os dados de sua organização pode ser uma oportunidade de ouro para acelerar a inovação e alcançar suas metas organizacionais. Há muita coisa em jogo. De acordo com a Forrester Research, as organizações que têm um sistema para promover insights com dados são 140% mais propensas a criar vantagens competitivas sustentáveis e 78% mais propensas a fomentar um ambiente favorável ao crescimento da receita.<sup>1</sup>

Com uma necessidade urgente de capacitar a organização inteira a usar dados para tomar decisões melhores e mais rápidas, que propiciem novas ideias e aumentem a agilidade dos negócios, os líderes estão admitindo uma verdade fundamental: a jornada para a inovação começa com dados e, para tornar-se uma **organização orientada por dados** bem-sucedida, parte-se da implementação de uma estratégia de dados completa.

## Existe uma estratégia de dados comprovada? A resposta é “sim”

A boa notícia é que você não precisa reinventar a roda. As organizações já estão desfrutando os benefícios da estratégia de dados completa criada na Amazon Web Services (AWS). Por exemplo, a **AstraZeneca** está integrando e escalando seus recursos de dados e inteligência artificial por toda a empresa para inovar com maior rapidez e melhorar os resultados para os pacientes. Com os serviços de dados da AWS, a empresa já consegue executar mais de 51 bilhões de testes estatísticos em menos de 30 horas, facilitando a disponibilização de insights genômicos para projetos de descoberta de medicamentos. O **BMW Group** e outras montadoras usam dados para otimizar a cadeia de suprimentos e melhorar a capacidade de produção. E a **ENGIE**, bem como outras empresas do setor de energia, usa os dados para descobrir novas formas de reduzir os custos para os clientes e, ao mesmo tempo, concretizar seus ousados e ambiciosos objetivos de sustentabilidade.

Embora as conquistas sejam ilimitadas, há um desafio central: muitas organizações detêm um rico manancial de dados, mas não sabem como tirar proveito deles. Neste e-book, você aprenderá os conceitos básicos de como desenvolver uma estratégia de dados completa para acompanhar suas necessidades associadas a dados, no presente e no futuro, possibilitando uma vantagem sustentável proveniente da revelação do valor dos dados.

<sup>1</sup> “Creating a data-driven culture” (Como criar uma cultura orientada por dados), CIO.com, 31 de março de 2022

# Principais desafios e considerações

## Nunca foram gerados e armazenados tantos dados quanto atualmente

As ferramentas on-premises e os datastores legados não conseguem atender às demandas atuais, assim as organizações precisam de novos datastores capazes de escalar e crescer conforme as necessidades dos negócios mudam, comportando os gigabytes e terabytes que são processados atualmente ou os petabytes e exabytes que serão gerenciados no futuro.

## Dados distribuídos em várias fontes geram perdas de produtividade e altos custos

As organizações modernas precisam de facilidade para acessar e analisar diversos tipos de dados, como arquivos de log, fluxo de cliques, voz e vídeo. No entanto, esses vários tipos de dados costumam ser armazenados em silos em diversos datastores. Para extrair informações, as organizações precisam desfazer essa compartimentação e, assim, conseguir unificar todos os tipos de dados. Essa importante otimização de custos e de operações está transformando a infraestrutura, que passa de uma fonte de complexidade e gastos a um mecanismo de geração de valor.

## O estado atual da tomada de decisões não é sustentável

O Gartner aponta que 65% das decisões tomadas atualmente são mais complexas (envolvem mais partes interessadas ou escolhas) do que eram há cinco anos.<sup>2</sup> Para agilizar e melhorar as decisões, as organizações precisam ser capazes de realizar operações de análise e machine learning (ML) com rapidez e bom custo-benefício, usando ferramentas ideais e performance para escalar de acordo com cada caso. As organizações não podem mais desperdiçar tempo precioso em constantes reimplementações e reconfigurações da infraestrutura para aumentar a performance e a capacidade.

## A adoção de análises e machine learning ainda é dificultada por falta de habilidades e inércia

Muitas empresas têm dificuldades para avançar no uso de ferramentas de análise e ML. O Gartner revelou que as organizações que investem em IA colocaram apenas 54% de seus pilotos de prova de conceito de IA em produção.<sup>3</sup> A constante insuficiência de habilidades em dados e ML e a quantidade ou qualidade dos dados para realizar treinamentos são apenas alguns dos problemas que retardam o progresso nessa área tão importante. Mesmo assim, a necessidade de ajudar os usuários empresariais a tomar decisões orientadas por dados é cada vez maior.

## Manter a governança de dados é um trabalho que exige dedicação total

Arquiteturas de dados tradicionais exigem procedimentos de gerenciamento arriscados e complicados, pois os dados são acessados de muitos locais. Conceder, monitorar, auditar e remover o acesso dos funcionários, sem deixar de cumprir com o número cada vez maior de regulamentações, é um trabalho que exige dedicação total. A automação dessas tarefas obrigatórias de governança de dados libera as equipes modernas para que voltem a se concentrar na inovação.

## Está cada vez mais difícil proteger os dados

Houve uma época em que as equipes de TI escolhiam entre desenvolver arquiteturas rápidas ou seguras. Hoje, precisam de ambos os aspectos. Os ataques à segurança aumentaram em 31% de 2020 para 2021, conforme aponta o relatório *State of Cybersecurity Resilience* (Estado da resiliência cibernética) da Accenture, enquanto a média de ataques sofridos por organização aumentou de 206 para 270 de um ano para o outro.<sup>4</sup> Mas como as organizações podem maximizar a privacidade e a segurança?

<sup>2</sup> “How to make better business decisions” (Como tomar melhores decisões de negócios), Gartner, 20 de outubro de 2021

<sup>3</sup> “Half of AI models never make it to production” (Metade dos modelos de IA nunca chega à produção), EnterpriseAI, 23 de agosto de 2022

<sup>4</sup> “How aligning security and the business creates cyber resilience” (Como o alinhamento entre a segurança e o negócio cria resiliência cibernética), Accenture, 2021

## COMO SE TORNAR UMA ORGANIZAÇÃO ORIENTADA POR DADOS

# Três atributos importantes que podem ajudar sua organização a extrair mais valor dos dados

De acordo com uma pesquisa da PwC com mais de mil executivos seniores, organizações altamente orientadas por dados são três vezes mais propensas a obter melhorias significativas na tomada de decisões em comparação com aquelas que utilizam menos os dados.<sup>5</sup>

A AWS pode ajudar sua organização a implementar uma estratégia completa que facilita o gerenciamento dos dados em todas as etapas da jornada: da ingestão, armazenamento e consulta dos dados à análise, visualização e execução de modelos de ML. Independentemente dos seus desafios de negócios, sua estratégia de dados deve ter:

1. **Abrangência:** equipada com as ferramentas certas, com a relação ideal de preço/performance para qualquer usuário, tipo de dado e caso de uso.
2. **Integração:** capacidade de integrar os dados que são armazenados e analisados em ferramentas e sistemas diferentes para obter uma melhor compreensão de seu negócio e prever o que vai acontecer.
3. **Governança:** políticas de controle de todos os seus dados para disponibilizá-los de forma segura quando e onde os usuários precisarem deles para acelerar a inovação.

Uma mentalidade orientada por dados também pode exigir uma mudança cultural mais ampla, na qual tanto as metas quanto as decisões são respaldadas pela estratégia de dados.

Acesse o link abaixo para conferir por que os dados desempenham um papel fundamental para viabilizar essa mudança cultural. Descubra também por que um número cada vez maior de empresas está utilizando recursos orientados por dados para automatizar diversos casos de uso essenciais para os negócios.

**Leia o e-book *Cinco elementos-chave de uma cultura orientada por dados* »**



<sup>5</sup> “Data and Intuition: Good Decisions Need Both” (Dados e intuição: as boas decisões necessitam de ambos), Harvard Business Publishing, 6 de janeiro de 2023



**Abrangência**



**Integração**



**Governança**



# 1

## Abrangência

**Equipada com as ferramentas certas, com a relação ideal de preço/performance para qualquer usuário, caso de uso e tipo de dado**

As organizações precisam criar estratégias de dados preparadas para o futuro que possam satisfazer as necessidades empresariais no presente e no futuro. É preciso mais do que apenas um único data lake, ou data warehouse, ou ferramentas de business intelligence (BI) para aproveitar os dados de maneira eficaz. É necessária uma estratégia de dados completa com um conjunto abrangente de ferramentas que leve em consideração a escala e a variedade dos dados e as diversas finalidades para as quais você deseja usá-los. Na verdade, 94% dos mil principais clientes da AWS usam mais de dez serviços de bancos de dados e análises da AWS.

Desenvolver com um provedor de nuvem que inova para oferecer continuamente todas as ferramentas de dados necessárias e muito mais, com a relação certa de preço/performance para o seu caso de uso, garante que você tenha uma estratégia de dados que cresce no seu ritmo. A AWS tem o conjunto mais amplo e aprofundado de recursos de dados para atender a qualquer workload de dados ou caso de uso. De bancos de dados para aplicações até armazenamento para data lakes, incluindo ferramentas de análise, ML e usuário final, a AWS fornece o recurso certo em cada área, para que você não precise comprometer a performance, o custo ou os resultados. A AWS está sempre acelerando o ritmo das inovações para satisfazer todas as suas necessidades de dados.



## ABRANGÊNCIA

### Escalabilidade de aplicações orientadas por dados

Desenvolva aplicações preparadas para o futuro em uma infraestrutura de dados moderna pelo melhor preço e performance para o seu caso de uso em larga escala. Os bancos de dados da AWS incluem o Amazon Aurora, que fornece a performance e a disponibilidade dos bancos de dados comerciais por um décimo do custo. Mais de 100 mil clientes da AWS usam o Amazon Aurora em prol de seus bancos de dados relacionais. A AWS oferece oito mecanismos de bancos de dados específicos para casos de uso que incluem grafos, streaming e documentos, cada um projetado especificamente para proporcionar performance ideal para suas aplicações, transformando o modelo econômico da propriedade de bancos de dados.

### Alimentação de dados e data warehouses

As organizações usam data lakes e data warehouses para unificar conjuntos de dados e aplicam análises e ML para ajudar na tomada de decisões e criar experiências orientadas por dados para os clientes.

A AWS tem ajudado centenas de milhares de clientes a construir uma base sólida para data lakes com serviços como Amazon Simple Storage Service (Amazon S3), AWS Glue e AWS Lake Formation durante anos. Os clientes também utilizam o Amazon Redshift, um serviço de data warehouse rápido e que opera na escala de petabytes, o qual oferece uma relação preço/performance até cinco vezes melhor do que outros serviços de data warehouse na nuvem. O Amazon Redshift ajuda você a ingerir, combinar e executar com segurança análises históricas, em tempo real e preditivas em todos os dados com apenas alguns cliques, proporcionando resultados de consulta super-rápidos.

### Fornecimento de análises para todos os casos de uso

A agilidade real ajuda organizações a se adaptarem rapidamente às mudanças nas necessidades de negócios. Para impulsionar essas ações rápidas, os serviços de análises da AWS permitem que as equipes de sua organização consumam, combinem e executem análises históricas, preditivas e em tempo real em todos os seus dados. Isso inclui serviços para consultas SQL, análises de logs, streaming e Apache Spark. Para consultas de big data, o Amazon EMR é compatível com mais frameworks de big data que qualquer outro provedor e oferece insights em até metade do tempo. Os clientes da AWS obtêm até o triplo de performance com o Apache Spark quando iniciam ambientes de execução compatíveis e otimizados para a AWS ao usar Amazon EMR, AWS Glue e Amazon SageMaker. Para tomar decisões em tempo real, você vai precisar de serviços de streaming de dados como o Amazon Kinesis Data Streams (Amazon KDS), que lhe permite desenvolver aplicações para dados de eventos de alta frequência, como dados de fluxo de cliques, e obtenha insights em questão de segundos.



### Dica para a estratégia de dados

Alcance uma estratégia de dados com bom custo-benefício sem comprometer a performance. Possibilite que sua organização maximize seus recursos atuais por meio da otimização dos custos:

#### Samsung

A Samsung economizou 44% em custos operacionais mensais e 22% em taxas de manutenção ao migrar para o Amazon Aurora PostgreSQL.

#### Carrier

A Carrier conectou sua rede logística de cadeia de frio para ajudar os clientes a otimizar as operações da cadeia de frio, reduzir o consumo de energia e melhorar os resultados com redução de custos, atrasos, perdas de carga e deterioração em trânsito.

#### United Airlines

A United Airlines criou um aeroporto inteligente com mais de 20 mil sensores que produzem dados para gerar insights em tempo real, otimizando os recursos dos equipamentos em terra, o que economizou US\$ 120 milhões em equipamentos que deixaram de ser necessários.

## ABRANGÊNCIA

O Amazon Kinesis Data Firehose carrega de maneira simples e confiável fluxos de dados em serviços de data lake, data warehouse e análises, sem a necessidade de extração, transformação e carregamento (ETL) ou preparação complicada de dados.

## Implantação de ciência de dados e machine learning

O ML adiciona inteligência aos processos existentes, automatiza tarefas manuais demoradas e acelera a inovação por meio da criação de produtos e serviços. Com a AWS, você tem acesso ao conjunto mais abrangente de serviços de IA e ML. Com o Amazon SageMaker, serviço usado por dezenas de milhares de clientes, você pode criar, treinar e implantar modelos de ML para qualquer caso de uso com infraestrutura, ferramentas e fluxos de trabalho totalmente gerenciados.

## Insights de dados em toda a organização

Os indivíduos experientes no uso de dados não são mais os únicos que conseguem extrair rapidamente insights valiosos e relevantes dos dados para ajudar a embasar a tomada de decisões. Soluções de BI alimentadas por ML, como o Amazon QuickSight, facilitam a conexão com fontes de dados. Analistas de negócios podem utilizar esses dados para demonstrar novas tendências e insights preditivos em visualizações e painéis interativos de BI.

O QuickSight Q utiliza ML, permitindo que os usuários consultem seus dados em linguagem simples, sem precisar escrever qualquer linha de código. Os usuários corporativos podem até questionar o porquê, para entender melhor os fatores que estão impactando as tendências de dados. Também podem prever métricas com instruções como “prever as vendas para os próximos 12 meses” e, assim, receber uma resposta imediata com base nos insights de dados passados e sazonalidade. A interface visual do tipo “apontar e clicar” permite que analistas de negócios gerem previsões precisas com ML sem experiência prévia. Com poucos cliques, os analistas podem importar dados de várias fontes, preparar dados automaticamente e desenvolver e analisar modelos de ML.

## Aumento da proficiência em dados

Ter funcionários capazes de usar os dados de maneira eficiente vai ajudar sua organização a atingir seus objetivos com os dados. Invista na formação e qualificação de sua equipe em dados, análises e ML com o treinamento da AWS.

# 94%

Com o mais amplo e mais aprofundado conjunto de serviços de dados do setor, 94% dos mil principais clientes da AWS usam, cada um, mais de dez serviços de bancos de dados e análises da AWS.

## Escale a tomada de decisões com dados em toda a organização

- **Amazon QuickSight:** satisfaça as diversas necessidades de análises usando a mesma fonte de referência por meio de painéis modernos e interativos, relatórios paginados, análises integradas e consultas em linguagem natural.
- **Amazon SageMaker Canvas:** interface da AWS sem uso de código que permite que analistas de negócios gerem previsões precisas com ML sem experiência prévia.
- **Amazon DataZone:** simplifica o acesso governado aos dados para os usuários corporativos.
- **Treinamento e Certificação da AWS:** mais de 150 cursos de desenvolvimento profissional relacionados a dados, análises e ML.





## HISTÓRIA DO CLIENTE

# ADP toma 312 trilhões de decisões por mês com processos analíticos

A ADP ajuda mais de 900 mil empresas na gestão de 70 milhões de pessoas em seus processos de quadro de funcionários e folha de pagamento. Essa gestão gera quantidades enormes de dados. A ADP chega a processar mais de 2,5 petabytes de dados com mais de 25 bilhões de pontos de dados individuais representados. A ADP usa o Amazon Redshift e o Amazon Neptune para executar alguns aspectos de seu processamento geral de dados. Esses serviços de dados da AWS ajudam as empresas na medição, comparação, previsão e aplicação de insights relacionados às equipes. Com os serviços da AWS, a ADP também permite que as organizações criem painéis de igualdade salarial, ajudando mais de dois terços das empresas a obterem melhoria nessa área.

**“Este é o momento de usarmos os dados para ajudar as pessoas a entender quais medidas podem tomar para criar um ambiente de trabalho mais diverso, igualitário e inclusivo e moldar o futuro que todos almejamos.”**

Jack Berkowitz, CDO da ADP

[Leia mais sobre a experiência da ADP ›](#)



## HISTÓRIA DO CLIENTE

# BMW Group democratiza o uso de dados em grande escala

O BMW Group migrou para um data lake centralizado e baseado na AWS por causa da agilidade, flexibilidade e capacidade de processar terabytes de dados de telemetria provenientes de milhões de veículos diariamente. Os dados anonimizados de sensores de veículos e outras fontes de toda a empresa agora podem ser facilmente acessados pelas equipes internas que criam aplicações internas e voltadas aos clientes. A criação de um catálogo de dados legível e a exibição clara dos recursos de dados se mostraram essenciais, aumentando a produtividade de analistas de dados, cientistas de dados e engenheiros.

**“Estamos apenas começando nossa jornada com a AWS e estamos ansiosos para ajudar nossa empresa a implementar sua estratégia de impulsionar a inovação para o futuro.”**

Kai Demtröder, vice-presidente de transformação com dados, inteligência artificial, dados e plataformas de DevOps do BMW Group

**Descubra como organizações líderes estão usufruindo os benefícios de uma estratégia de dados completa criada na AWS ›**

# 2

## Integração

**Elimine a compartimentação para colocar os dados em ação de forma eficaz**

Existem oportunidades para transformar os negócios com o uso de dados em toda a cadeia de valor. Mas essa transformação exige que você veja o quadro completo do cliente e da empresa. Como os dados estão espalhados por vários departamentos, serviços, bancos de dados on-premises e aplicações de terceiros, você precisa ser capaz de integrá-los facilmente para obter os melhores insights. As empresas têm várias abordagens para unificar dados, como data mesh, lake house, data fabric e assim por diante. No entanto, isso geralmente envolve um data lake como elemento fundamental. Os data lakes permitem que você colete, armazene, organize e processe dados valiosos armazenados em silos e os disponibilize, de maneira controlada, para ferramentas de análise, visualização e machine learning.





## INTEGRAÇÃO

### Sem ETL

Muitas organizações têm vários data lakes além dos data warehouses, ferramentas de análise, ferramentas de ML e aplicações de SaaS. A integração dos dados entre silos exige pipelines complexos de ETL, que podem levar horas ou até mesmo dias. Isso está longe de ser rápido o suficiente para a tomada de decisões moderna. As organizações precisam adotar tecnologias que automatizam ou eliminam o processo de ETL sempre que possível.

A AWS está investindo em um futuro sem ETL, para que as organizações integrem todos os seus dados. Isso inclui incorporar ML à fonte de dados com integração do SageMaker ao Amazon Redshift, Amazon Aurora, Amazon Athena e Amazon Neptune, bem como integrar o Amazon Aurora ao Amazon Redshift para fazer análises em tempo real e fornecer uma integração direta entre o Amazon S3 e o Amazon Redshift para lidar com fluxos de dados em tempo real. Além disso, você pode executar consultas em dados armazenados em bancos de dados operacionais, data warehouses e data lakes para fornecer insights entre várias fontes de dados sem movimentar nenhum dado. Tudo isso usando o Amazon Athena e o Amazon Redshift.

### Análise de dados próprios e de terceiros

Para eliminar os silos de dados, não é viável estabelecer conexões somente com suas fontes de dados. Você precisa ser capaz de se conectar com facilidade a todas elas, não importa se estão na AWS, em aplicações de terceiros, em ambientes on-premises ou até mesmo em outro ambiente de nuvem. A AWS possibilita que você integre automaticamente centenas de fontes de dados da AWS e de terceiros.

Cada vez mais, as organizações estão utilizando dados de terceiros para aprofundar os insights, unindo esses dados externos aos seus próprios dados. O AWS Data Exchange permite que os clientes da AWS acessem dados de terceiros em arquivos, tabelas e APIs de mais de 300 provedores de dados e mais de 3,5 mil produtos de dados, tudo de um só lugar. Dados externos de parceiros e clientes também estão sendo usados, aumentando a necessidade de políticas de governança abrangentes para proteção dos dados. As salas limpas de dados, ambientes protegidos onde várias partes podem analisar dados combinados sem expor os conjuntos de dados brutos, surgiram como uma solução. O AWS Clean Rooms ajuda empresas e seus parceiros de negócios a analisarem e colaborarem com segurança usando seus conjuntos de dados, sem compartilhar ou revelar os dados subjacentes.



### Conecte-se a centenas de fontes de dados

- **Amazon AppFlow:** integre data lakes e data warehouses a mais de 50 fontes de dados.
- **AWS Data Exchange:** acesse mais de 350 provedores de terceiros e mais de 3,5 mil produtos de dados públicos.
- **Amazon SageMaker Data Wrangler:** crie modelos de ML com mais de 40 fontes de dados com um único clique.



### Dica para a estratégia de dados

#### Busque eliminar as operações de ETL

Evite desenvolver e manter pipelines de dados complexos para realizar operações de ETL. O Amazon Aurora já se integra ao Amazon Redshift sem uso de ETL, possibilitando análises e ML quase em tempo real. Dados transacionais gravados no Amazon Aurora ficam disponíveis no Amazon Redshift quase instantaneamente. Analise dados de vários clusters de banco de dados do Aurora na mesma instância do Amazon Redshift, nova ou existente, e obtenha insights holísticos em várias aplicações.



**Goldman  
Sachs**

## HISTÓRIA DO CLIENTE

# Como a entrega de dados ajuda o Goldman Sachs a operar de forma mais inteligente

O Goldman Sachs percebeu um aumento exponencial na demanda por dados do mercado financeiro, dobrando ano após ano. O AWS Data Exchange é um serviço gerenciado que permite que os processos de dados sejam centralizados, otimizados e automatizados sempre que possível. A migração de feeds de dados do mercado para o AWS Data Exchange pode aumentar a eficiência geral do consumo de dados de terceiros, liberando mais tempo para as análises de dados que agregam valor e reduzindo o tempo gasto na preparação dos dados para usá-los. A entrega de dados sem atrito pelo AWS Data Exchange é essencial para preparar uma organização para o futuro.

**“O AWS Data Exchange é um componente fundamental da estratégia de nuvem financeira do Goldman Sachs porque reduz o atrito na obtenção de dados financeiros de provedores externos, além de possibilitar que nosso foco esteja na entrega dos nossos principais serviços e de análises de dados diferenciadas para atender melhor os clientes.”**

Marco Argenti, codiretor de informações do Goldman Sachs



## HISTÓRIA DO CLIENTE

# ENGIE acelera sua transição ao carbono zero

Empresa global de serviços de energia em processo de transição ao carbono zero, a **ENGIE desenvolveu seu data lake Common Data Hub na AWS**. A ENGIE recebeu suporte da AWS Professional Services no design e implementação da solução e na formação de uma equipe interna para supervisionar a plataforma. Com mais de mil projetos do mundo todo atualmente no Common Data Hub, esta solução coesa desfaz a compartimentação, dando a todos os departamentos acesso igualitário a um framework de dados comum.

**“Fomos convencidos de que a AWS seria uma boa solução por diversos motivos, incluindo o modelo de custo, mas, principalmente, em termos de armazenamento de dados.”**

Gregory Wolowiec, líder da equipe de tecnologia nos programas de dados da ENGIE

**Conheça como os clientes da AWS integraram seus dados eliminando os dados compartimentados e criando soluções transformadoras para suas organizações ›**



# 3

## Governança

**Libere suas equipes para se moverem mais rapidamente com acesso controlado a dados em todo o ciclo de vida**

Além de adotar uma estratégia abrangente e integrada, é igualmente importante garantir que seus usuários possam acessar os dados onde e quando forem necessários com o nível correto de controle. Com a estratégia certa de governança de dados, você pode agir mais rapidamente para capacitar os usuários com o acesso aos dados de que precisam, quando precisam.

À medida que mais dados migram para a nuvem, graças à potência e à escala quase infinitas da nuvem, é fundamental que os modelos de governança de dados corporativos evoluam em sincronia. Os líderes de TI e de negócios precisam de políticas atualizadas para proteger os dados que transitam entre diferentes repositórios e para acomodar as mudanças nas regulamentações de privacidade e segurança de dados sobre onde os dados podem ser armazenados.<sup>5</sup>

**“O segredo para uma boa governança é descobrir como definir o acesso e deixar livre o caminho aberto. Ou seja, criar processos de exceção, em vez de adotar uma abordagem em que, toda vez que necessitamos de dados, precisamos solicitá-los a alguém em uma organização central.”<sup>6</sup>**

Rahul Pathak, vice-presidente de análises da AWS



<sup>6</sup> “[A unified approach to data governance](#)” (Uma abordagem unificada à governança de dados), CIO Magazine, agosto de 2021



## GOVERNANÇA

### Simplificação das permissões de acesso aos dados

A implementação de uma estratégia de governança bem-sucedida continua apresentando um conjunto único de desafios. Proporcionar o nível correto de acesso a conjuntos de dados específicos para clientes internos e externos é um processo demorado e desafiador para as organizações. Elas costumam realizar tarefas pesadas, como scripts manuais ou investigação de clusters de dados individuais, para descobrir quais clientes têm acesso a quais dados.

O trabalho manual também pode gerar problemas caros de qualidade de dados entre equipes e departamentos diferentes. Sem ferramentas de governança centralizadas, os dados ficam presos em silos, ou seja, não será possível acessar e analisar todos os dados necessários para resolver problemas ou identificar grandes oportunidades.

### Desenvolvimento de uma estratégia de governança de dados

Um novo estudo da AWS em parceria com o MIT que contou com a participação de mais de 350 profissionais de dados aponta que a governança de dados é a principal prioridade dos diretores de dados (CDOs): mais de 50% dos CDOs indicam “estabelecer uma governança de dados clara e eficaz” como principal responsabilidade. Governança também é a área em que os CDOs gastam mais tempo, já que mais de 66% dos entrevistados afirmaram que as iniciativas de governança de dados são uma das principais prioridades.<sup>7</sup>

Sem uma abordagem de governança que apoie a inovação, as organizações terão dificuldade para se tornarem orientadas por dados e, em última análise, para permanecerem competitivas. Afinal, quanto mais tempo os funcionários passam preparando os dados, menos tempo passam inovando com eles.

A AWS está investindo na jornada de dados com o intuito de reduzir o esforço da governança de dados de ponta a ponta. O AWS Lake Formation facilita o processo de governar e auditar as ações tomadas com dados em seu data lake no AWS S3, e o AWS Lake Formation também pode ser usado para governar o compartilhamento de dados no Amazon Redshift. O Amazon DataZone é um novo serviço de gerenciamento de dados para catalogar, descobrir, compartilhar e governar os dados, de modo que todos os membros da organização consigam realizar ações baseadas nos dados. Para os modelos de ML, o Amazon SageMaker conta com recursos para ajudar você a governar e auditar todo o ciclo de desenvolvimento de ML.



<sup>7</sup> “The Best Offense Is a Great Defense” (O melhor ataque é uma grande defesa), TechCrunch Brand Studio, pela AWS, 2022

<sup>8</sup> “The Economic impact of data innovation” (Impacto econômico da inovação com dados), Splunk.com, 2022

## Governança holística com a AWS

- **AWS Lake Formation:** facilita o processo de governar e auditar as ações tomadas com dados em seu data lake no Amazon S3.
- **Amazon DataZone:** serviço de gerenciamento de dados para catalogar, descobrir, compartilhar e governar os dados.



### Dica para a estratégia de dados

Um relatório da Splunk sobre os benefícios econômicos de práticas consolidadas para dados reforça a importância do acesso. Ao se concentrarem em aspectos como classificação e compartilhamento de dados entre as áreas de negócios, os “líderes em inovação de dados” alcançam níveis mais altos de fidelidade à marca e satisfação dos clientes, além de conseguir operacionalizar melhor seus dados. Os líderes de dados são “4,6 vezes mais propensos a obter mais de 20% de sua receita com produtos e serviços inovadores” que os iniciantes em dados.<sup>8</sup>



**1. ONEFOOTBALL**

## HISTÓRIA DO CLIENTE

# Ao simplificar a governança, a OneFootball alcançou um aumento de 40% na utilização de sua plataforma de análise

A **OneFootball** cresceu rapidamente para se tornar uma das plataformas de mídia digital mais populares do mundo entre entusiastas do futebol. Com o intuito de usar melhor os dados em benefício da empresa e de 70 milhões de fãs desse esporte que é uma verdadeira paixão, a OneFootball desenvolveu uma solução mais ágil na AWS em poucos dias. Desde a integração dos dados de seus ineficientes bancos de dados de back-end em um data lake baseado na nuvem, a OneFootball simplificou radicalmente a ingestão de dados e conseguiu eliminar workloads de ETL herdadas. É para se apaixonar mesmo.

Com o AWS Lake Formation, a OneFootball conseguiu simplificar a governança e o gerenciamento da segurança em grande escala para que pudesse ampliar o acesso aos dados para toda a empresa. Isso ajudou a equipe a obter um aumento significativo na quantidade de usuários ativos semanais e em outros indicadores de performance essenciais, aumentando o uso da plataforma de análise em 40%.

**“Agora podemos passar mais tempo entendendo os problemas de negócios em vez de manter diferentes tipos de extrações de banco de dados. Ver cada vez mais pessoas da organização usando análises diariamente é uma grande conquista. Ter todas as fontes de dados relevantes integradas de maneira confiável era um pré-requisito.”**

Stephan Durry, diretor de dados e insights da OneFootball





## HISTÓRIA DO CLIENTE

# Pinterest usa governança para colocar os clientes em primeiro lugar

Um sistema escalável de controle de acesso refinado (FGAC) desenvolvido usando o Amazon S3 garantiu que o volume de dados cada vez maior do Pinterest não ultrapassaria os controles existentes da empresa. O FGAC controla o acesso aos dados e se baseia em opções de ofertas de vários critérios, como controle de acesso baseado em função com segurança para conjuntos de dados em escala de petabytes. Também permitiu que criadores e empresas que usam a plataforma se autoidentificassem como membros de minorias ao mesmo tempo que garantiu que dados confidenciais não seriam usados para nenhuma outra finalidade, como publicidade.

**“Os impactos dos esforços de governança do Pinterest para os clientes incluem o uso de dados de autoidentificação de ‘modo muito controlado’ com o intuito de apoiar empresas de pessoas negras no Juneteenth [comemoração da abolição da escravidão nos EUA]. Os criadores também podem adicionar emblemas aos seus perfis, permitindo que o conteúdo apareça em espaços temáticos no Pinterest, para mostrar que a empresa pertence a uma pessoa que se identifica como membro de uma minoria.”**

David Chaiken, arquiteto-chefe do Pinterest

**Descubra como os clientes da AWS empoderaram seus funcionários com estratégias corretas de governança ›**

# Como tornar a segurança mais estratégica

A AWS prioriza a segurança desde o início, com infraestrutura eficiente, resiliente, de alta performance e continuamente protegida para suas workloads e aplicações. Os especialistas em segurança de nível global que monitoram a infraestrutura da AWS também desenvolvem e mantêm uma ampla seleção de serviços de segurança inovadores, que podem ajudar a simplificar as complexidades de seus próprios requisitos regulatórios e de segurança.

Os serviços e as soluções de segurança da AWS podem combinar vantagens importantes:

- **Obter insights com mais rapidez:** forneça o nível certo de acesso aos seus recursos o tempo todo com a certeza de que seus dados estão protegidos. A segurança da AWS é desenvolvida com a performance em mente, para que você receba proteção máxima e governança de dados que não atrase você.
- **Reduzir o tempo de inatividade:** uma segurança de nuvem mais robusta e moderna ajuda sua empresa a prosseguir, de modo que você não precise deixar de analisar os dados para realizar um processo de segurança separado. Ela pode ser integrada em todas as etapas da jornada.
- **Permanecer dentro do orçamento:** a AWS mantém o custo-benefício da segurança e acompanha a evolução das necessidades de seus riscos e requisitos de segurança, protegendo os investimentos de sua organização e seu compromisso com as iniciativas de dados.
- **Manter o foco:** de infraestrutura a serviços, a AWS é segura porque considera a segurança em cada etapa da jornada, para que você possa passar mais tempo transformando os dados em decisões melhores que geram resultados de negócios e menos tempo se preocupando com segurança e governança.

## Uma história de confiabilidade e segurança inigualáveis



### Amazon S3

Armazene e recupere qualquer volume de dados com a melhor segurança



### AWS Lake Formation

Crie um data lake seguro em questão de dias com controle de acesso refinado



### Regiões multi-AZ

Garanta failovers ininterruptos se uma zona de disponibilidade for afetada



## CONCLUSÃO

# A próxima onda de inovação será orientada por dados

Os líderes e responsáveis por decisões que querem entrar na próxima onda de reinvenção precisam de obstinação para chegar aonde os fatos residem. Eles também precisam de ferramentas essenciais para se manterem ágeis o suficiente e mudarem de rumo quando necessário, a fim de aproveitar novas oportunidades. Em outras palavras, a organização precisa se tornar orientada por dados.

As organizações orientadas por dados procuram conhecer os fatos tratando os dados não como propriedade exclusiva de departamentos compartimentados, e sim como ativo organizacional disponível para todos. Colocar uma estratégia de dados moderna em prática na sua organização é possível, independentemente do porte, da localização ou das necessidades de negócios. A AWS oferece o conjunto de serviços mais abrangente para a jornada de dados completa de todas as workloads, tipos de dados e resultados desejados.

Descubra por que a AWS é o melhor lugar para extrair valor de seus dados e transformar em tempo real os insights em inovações relevantes. Saiba como podemos ajudar suas equipes com infraestrutura, ferramentas e suporte a implementação com a principal rede de parceiros e serviços profissionais do mundo. Quando o assunto é dados, os clientes da AWS saem na frente.

## Descubra os principais casos de uso de dados para maximizar o valor empresarial

No e-book [Como usar os dados para maximizar o valor empresarial: seis casos de uso orientados por dados e essenciais aos líderes](#), conheça exemplos de clientes da AWS e descubra como tirar proveito dos dados para melhorar as experiências dos clientes, otimizar e reinventar as cadeias de suprimentos, melhorar a tomada de decisões, preparar as aplicações para o futuro e muito mais

**Saiba mais sobre como reinventar sua organização para que seja orientada por dados ›**