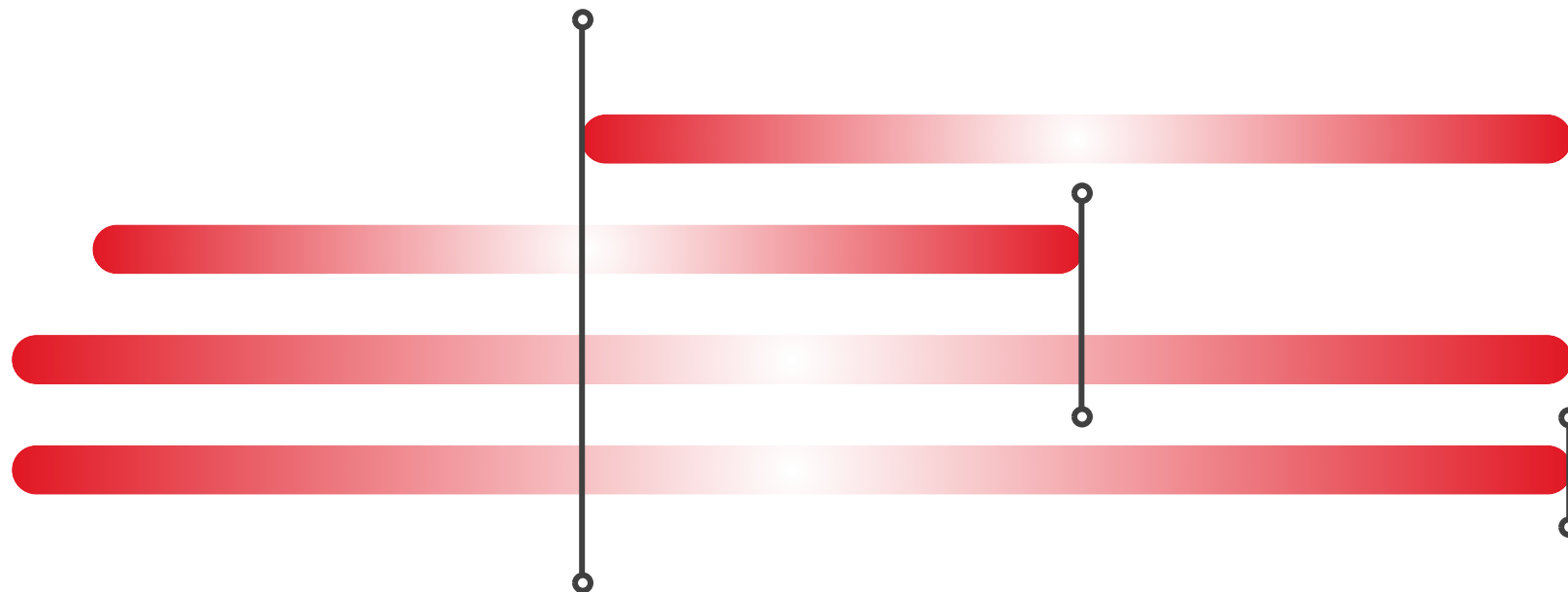




uniesp

Centro Universitário



ENGENHARIA DE DADOS

TÓPICOS ABORDADOS

Introdução à Computação em Nuvem:

- Definição de computação em nuvem.
- Benefícios da computação em nuvem, como escalabilidade, flexibilidade e redução de custos.
- Modelos de serviço em nuvem: IaaS (Infrastructure as a Service), PaaS (Platform as a Service) e SaaS (Software as a Service).

O que é DaaS:

- Definição de DaaS.
- Diferença entre dados tradicionais e DaaS.
- Benefícios do DaaS para organizações, como acesso fácil a conjuntos de dados, eliminação de custos de infraestrutura e capacidade de escalabilidade.

Arquitetura do DaaS:

- Componentes principais do DaaS, incluindo armazenamento de dados, ferramentas de processamento de dados e serviços de entrega de dados.
- Descrição do processo de provisionamento e gerenciamento de serviços de dados.

Implementação do DaaS:

- Opções de implementação do DaaS, incluindo provedores de serviços em nuvem e soluções de software específicas para DaaS.
- Considerações ao migrar para o DaaS, como requisitos de segurança, conformidade regulatória e custos.

Cenários de Uso do DaaS:

- Exemplos de casos de uso do DaaS em diferentes setores, como análise de dados, ciência de dados, aprendizado de máquina, IoT (Internet das Coisas) e análise preditiva.
- Demonstração de como os usuários podem interagir com os dados fornecidos pelo DaaS e incorporá-los em suas próprias aplicações e processos de negócios.



CLOUD COMPUTING

CLOUD COMPUTING

INTRODUÇÃO



É a entrega de recursos computacionais por demanda, ou seja conforme são necessários.

Essa escala é entregue através de plataformas, como por exemplo:

- AWS
- Azure
- IBM Cloud
- Oracle OCI
- Google Cloud Plataforma

Cloud abrangem tudo desde aplicativo até datacenter. E é pago conforme o uso.

Esse é o conceito de PaaS (Platform as a Service), onde ao invés do desenvolvedor demandar tempo para preparo de infraestrutura, isso já é fornecido em um modelo padrão e configurável.

CLOUD COMPUTING

INTRODUÇÃO

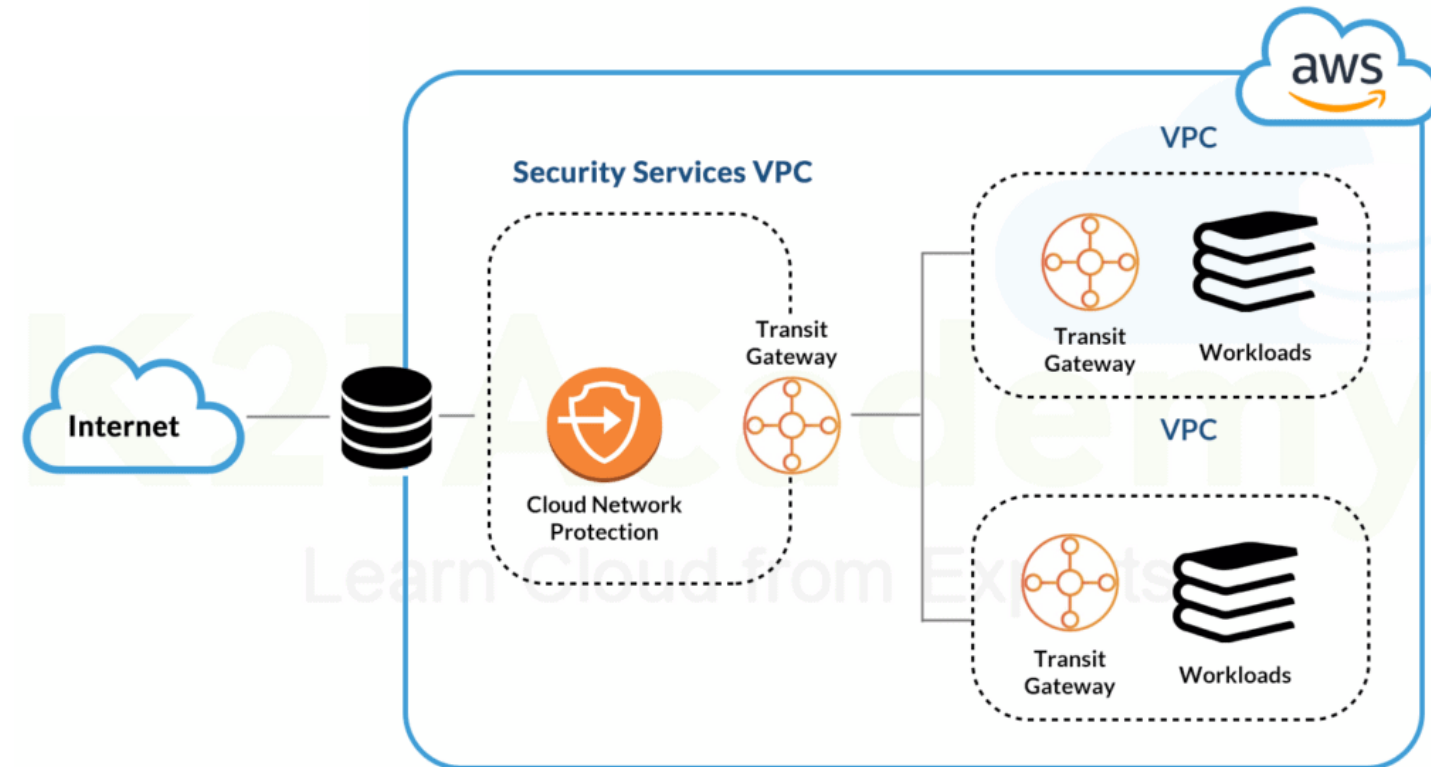


Exemplos básicos de recursos computacionais que operam em Cloud Computing e podem ser rapidamente provisionados são:

- Redes (Networks)
- Servidores (Servers)
- Armazenamentos (Storages)
- Aplicações (Applications)
- Serviços (Services)

CLOUD COMPUTING

NETWORKS



Networks: redes se referem às infraestruturas de comunicação que conectam todos os recursos e usuários dentro de um ambiente de computação em nuvem.

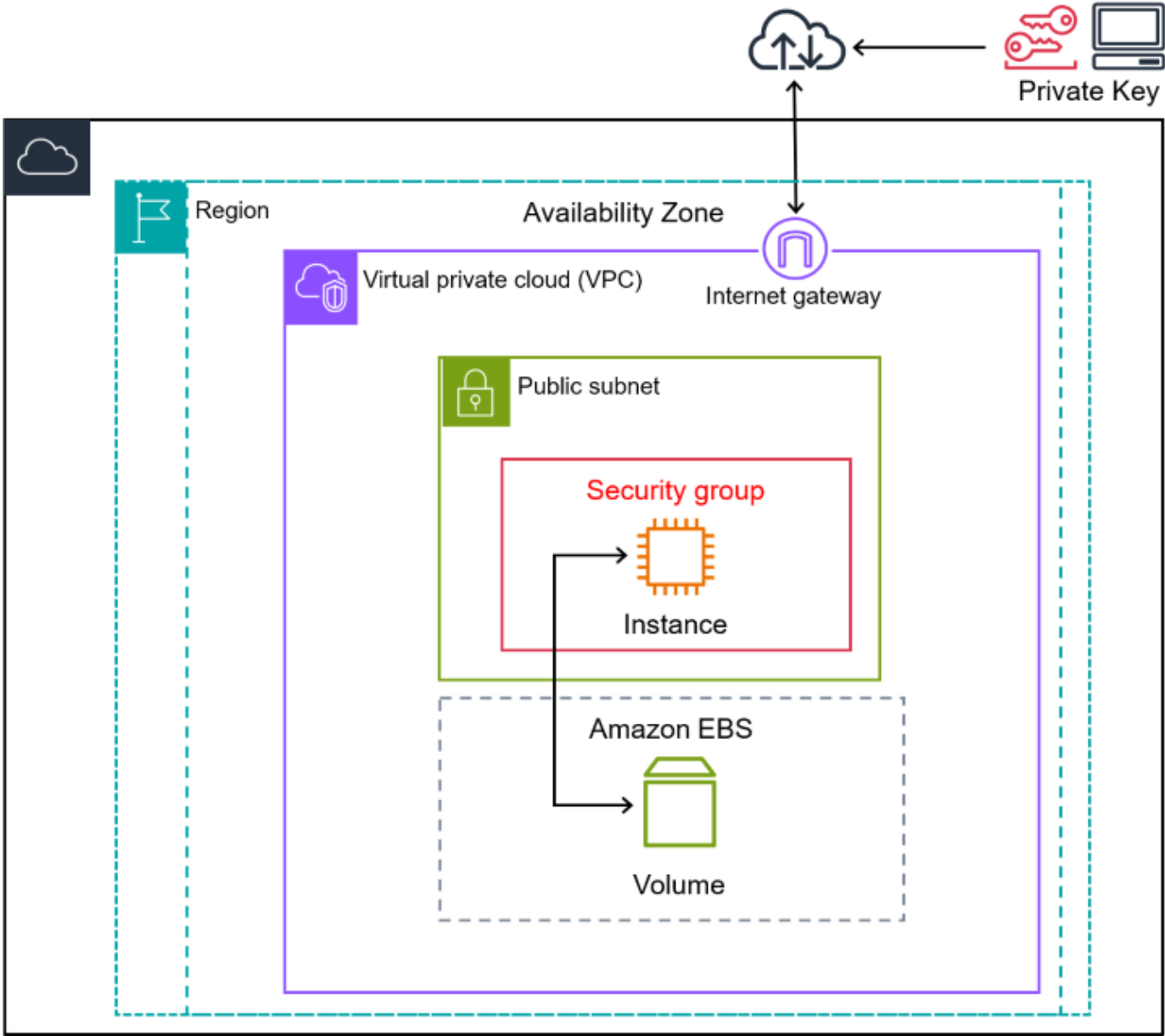
Exemplo:

Amazon Virtual Private Cloud (VPC) é um serviço da AWS que permite criar uma rede virtual na nuvem, isolada logicamente, onde você pode lançar recursos da AWS, como instâncias EC2 (servidores virtuais) e RDS (banco de dados relacional).

Dentro de uma VPC, os usuários podem criar sub-redes, que são segmentos lógicos da rede que podem ser configurados com diferentes regras de acesso e políticas de segurança. Além disso, é possível configurar rotas de rede personalizadas para direcionar o tráfego entre sub-redes e para a internet.

CLOUD COMPUTING

SERVICES



Servers: Servidores na nuvem são instâncias de computação virtual que executam aplicativos ou serviços para os usuários.

Exemplo: Amazon EC2 (Elastic Compute Cloud) é um serviço de computação em nuvem da AWS que oferece capacidade de computação escalável na nuvem. Os usuários podem lançar instâncias de servidores virtuais sob demanda, sistema operacional, quantidade de CPU, memória, armazenamento e outras características de configuração.

O Amazon EC2 oferece recursos de escalabilidade e resiliência incorporados. Os usuários podem dimensionar horizontalmente suas instâncias EC2 conforme necessário para lidar com picos de carga de trabalho e podem aproveitar as opções de armazenamento e backup para garantir a resiliência dos dados e aplicativos.

CLOUD COMPUTING

STORAGES



Amazon S3

Durable object storage for all types of data



Amazon Glacier

Archival storage for infrequently accessed data



Amazon EBS

Block storage for use with Amazon EC2



Amazon EFS

File storage for use with Amazon EC2

Economics

Pay as you go
No upfront investment
No commitment

Easy to Use

Self service administration
SDKs for simple integration

Reduce risk

Durable and Secure
Avoid risks of physical media handling

Agility, Scale

Reduce time to market
Focus on your business, not your infrastructure

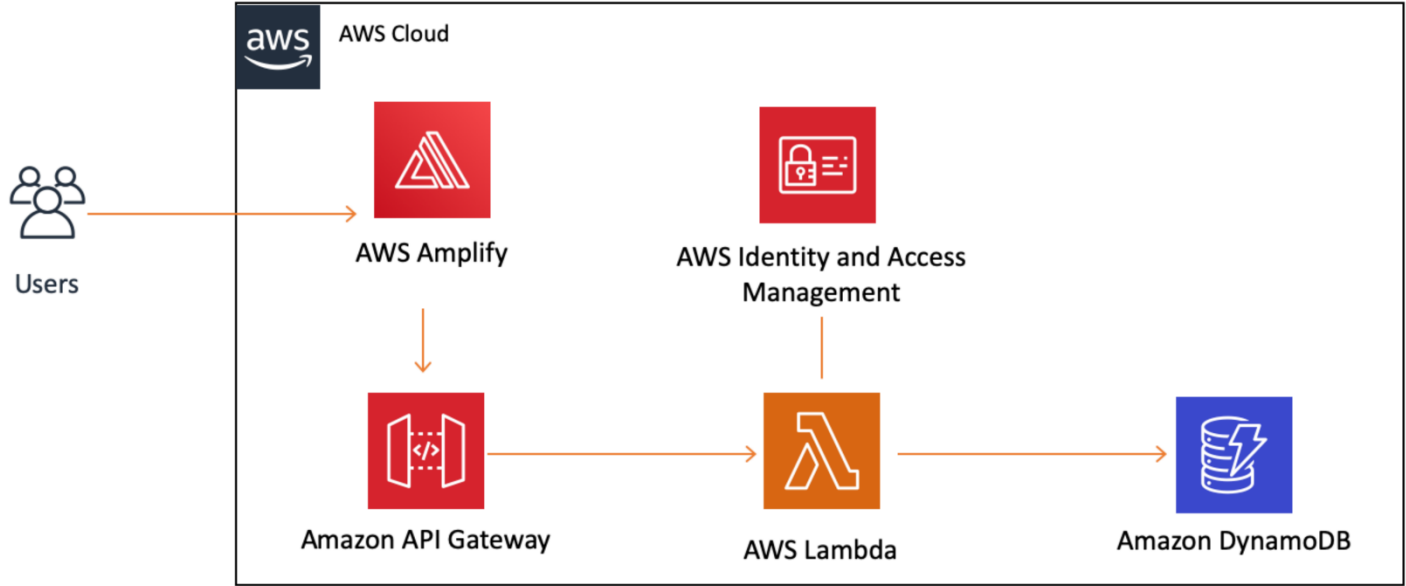
Storage: Armazenamento em cloud computing refere-se aos recursos de armazenamento de dados disponíveis na nuvem para os usuários, permitindo o armazenamento e a recuperação de dados de forma segura e escalável.

Exemplo: Amazon S3 (Simple Storage Service) é um serviço de armazenamento de objetos da AWS, projetado para armazenar e recuperar grandes quantidades de dados de forma durável e altamente disponível.

O Amazon S3 se integra perfeitamente com outros serviços da AWS, permitindo aos usuários construir soluções completas e integradas na nuvem. Por exemplo, o S3 pode ser usado para armazenar arquivos de mídia para serem acessados por aplicativos hospedados em instâncias EC2, ou pode ser usado como origem de dados para serviços de análise de dados, como Amazon Athena e Amazon Redshift.

CLOUD COMPUTING

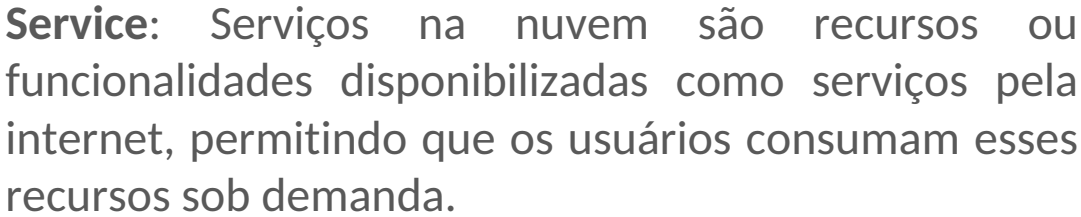
APPLICATIONS



Applications: Aplicativos na nuvem são programas ou software que são executados em infraestrutura de computação em nuvem e são acessíveis aos usuários pela internet.

Exemplo: O Amazon WorkMail é um serviço de e-mail e calendário gerenciado na nuvem que oferece uma alternativa ao Gmail. Ele permite que você crie e gerencie caixas de correio de forma escalável e segura, com integração com outros serviços da AWS, como o Amazon Simple Storage Service (S3) para armazenamento de anexos.

Ao combinar esses serviços da AWS, é possível criar uma suíte de aplicativos de produtividade baseada em nuvem semelhante ao Google Workspace. Cada serviço é executado na infraestrutura de nuvem da AWS e pode ser acessado pelos usuários através de navegadores web, proporcionando uma experiência de colaboração eficiente e escalável.



Exemplo: AWS Lambda é um serviço de computação serverless da AWS que permite executar código sem provisionar ou gerenciar servidores. é uma ferramenta poderosa e flexível que permite aos desenvolvedores executar código de forma eficiente e econômica na nuvem, sem se preocupar com a infraestrutura subjacente. Ele oferece escala automática, pagamento por uso e integração perfeita com outros serviços da AWS, tornando-o uma escolha popular para uma ampla variedade de cargas de trabalho e aplicativos na nuvem.

CLOUD COMPUTING SERVICES



IaaS

Infrastructure-as-a-Service

host



PaaS

Platform-as-a-Service

build



SaaS

Software-as-a-Service

consume

Os serviços mais comuns disponíveis na nuvem são:

- **IaaS** (Infrastructure as a Service);
- **SaaS** (Software as a Service);
- **PaaS** (Platform as a Service);
- **CaaS** (Container as a Service);

Como essas ferramentas da plataforma são acessadas livremente pela Internet, e não através de um sistema operacional ou pacote instalado em uma máquina local, os desenvolvedores não precisam se preocupar com a logística de montar algo que será instalado em um computador do cliente, pois tudo é mantido na própria nuvem. Qualquer pessoa com um navegador da web pode acessar o aplicativo.

CLOUD COMPUTING

IAAS (INFRASTRUCTURE AS A SERVICE)



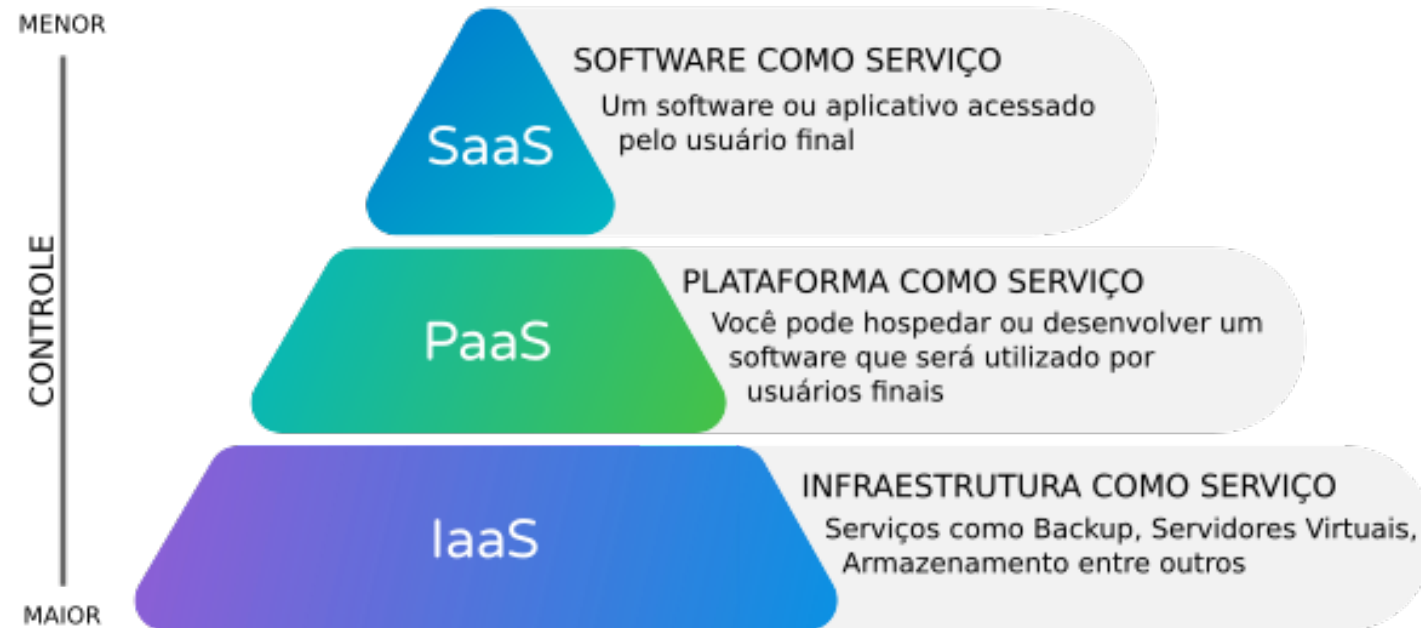
A **Infraestrutura como Serviço**, como o próprio nome sugere, fornece infraestrutura de computação em nuvem, incluindo servidores, rede, sistemas operacionais e armazenamento por meio da tecnologia de virtualização.

Esses servidores em nuvem geralmente são fornecidos à organização por meio de um painel ou uma API, proporcionando aos clientes controle total sobre toda a infraestrutura. O IaaS fornece as mesmas tecnologias e recursos que um data center tradicional, sem a necessidade de manter ou gerenciar fisicamente tudo.

Os clientes de IaaS ainda podem acessar seus servidores e armazenamento diretamente, mas tudo é terceirizado por meio de um “datacenter virtual” na nuvem.

CLOUD COMPUTING

IAAS (INFRASTRUCTURE AS A SERVICE)



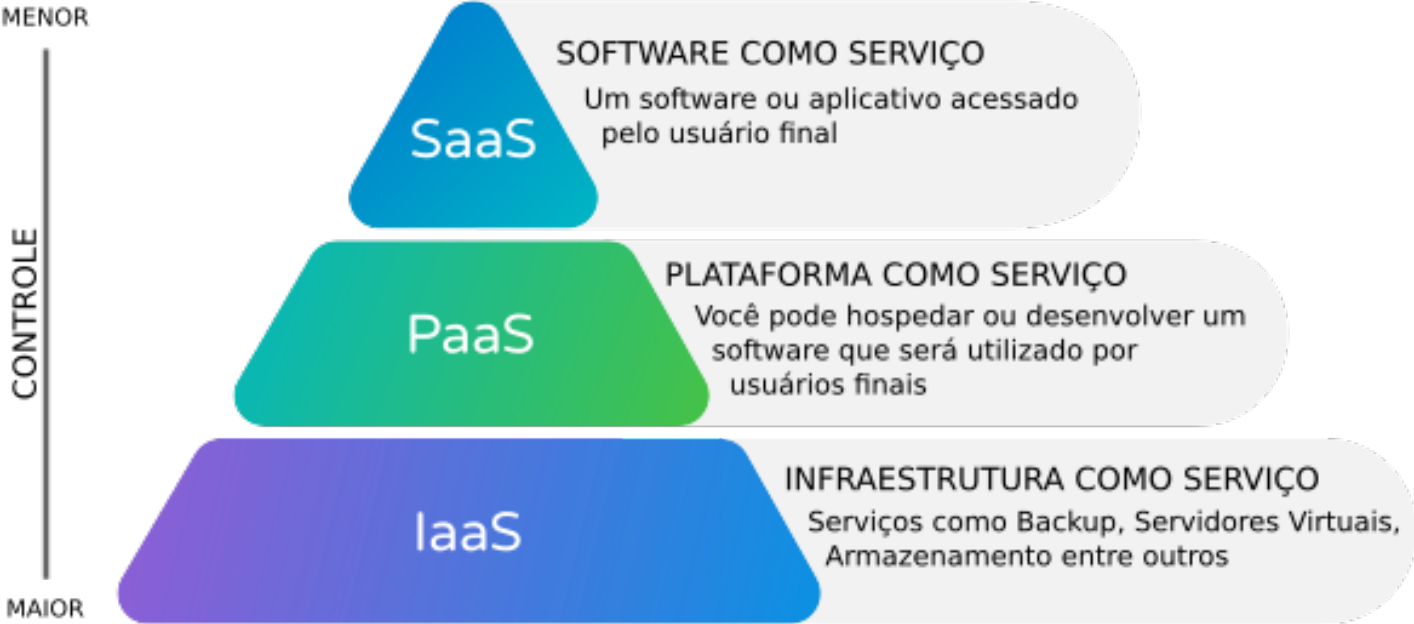
A interligação entre IaaS e Engenharia de Dados pode ser feita de várias maneiras para suportar efetivamente o fluxo de trabalho de engenharia de dados. Aqui estão algumas abordagens:

- I. *Provisionamento de Infraestrutura;*
- II. *Implantação de Ferramentas de Engenharia de Dados;*
- III. *Escala Automática;*
- IV. *Armazenamento de Dados;*
- V. *Segurança e Conformidade;*

Podemos construir uma infraestrutura flexível, escalável e segura para suportar todas as etapas do ciclo de vida dos dados, desde a ingestão até a análise e a tomada de decisões.

CLOUD COMPUTING

IAAS (INFRASTRUCTURE AS A SERVICE)

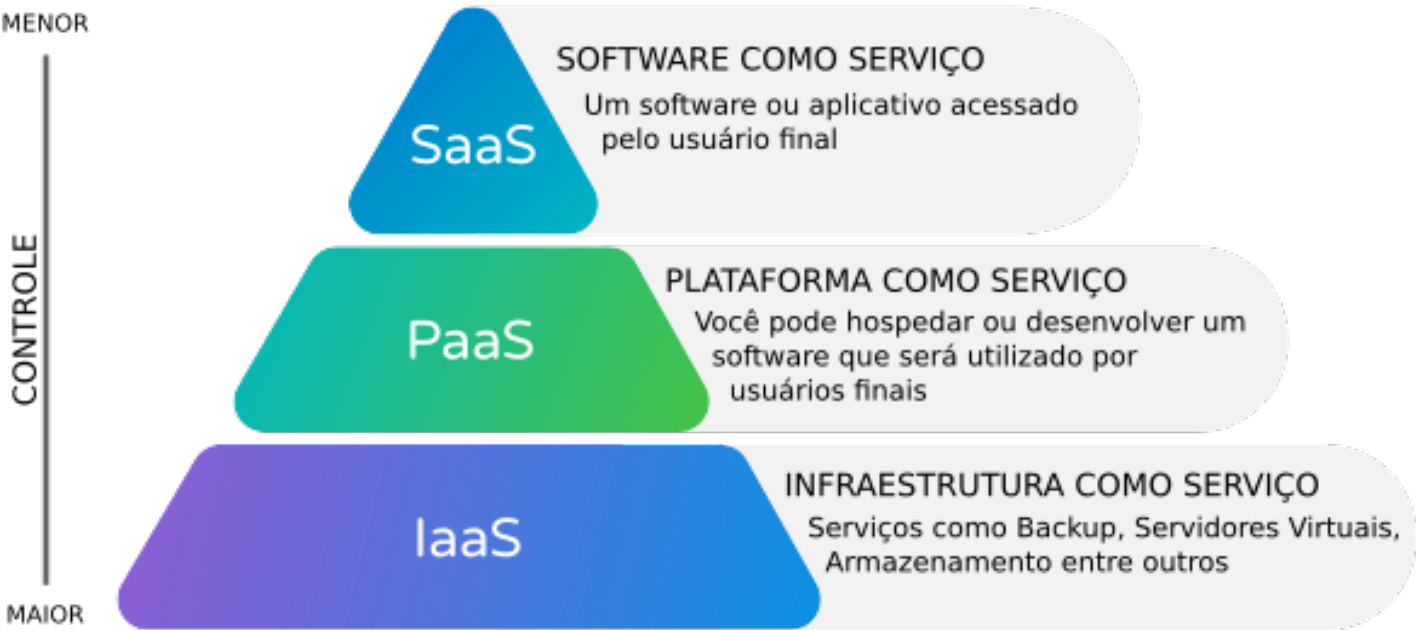


Cenário: Uma empresa de comércio eletrônico precisa lidar com grandes volumes de dados de transações de clientes e deseja construir um sistema de análise de dados para entender o comportamento do cliente e otimizar as recomendações de produtos.

Aplicação de IaaS: A empresa pode usar serviços de IaaS, como Amazon Web Services (AWS) ou Microsoft Azure, para provisionar recursos de computação e armazenamento conforme necessário. Eles podem implantar máquinas virtuais para executar clusters de processamento de big data, como Apache Hadoop ou Apache Spark. Além disso, eles podem usar serviços de armazenamento escalável, como Amazon S3 ou Azure Blob Storage, para armazenar grandes conjuntos de dados brutos e processados.

CLOUD COMPUTING

SAAS (SOFTWARE AS A SERVICE)



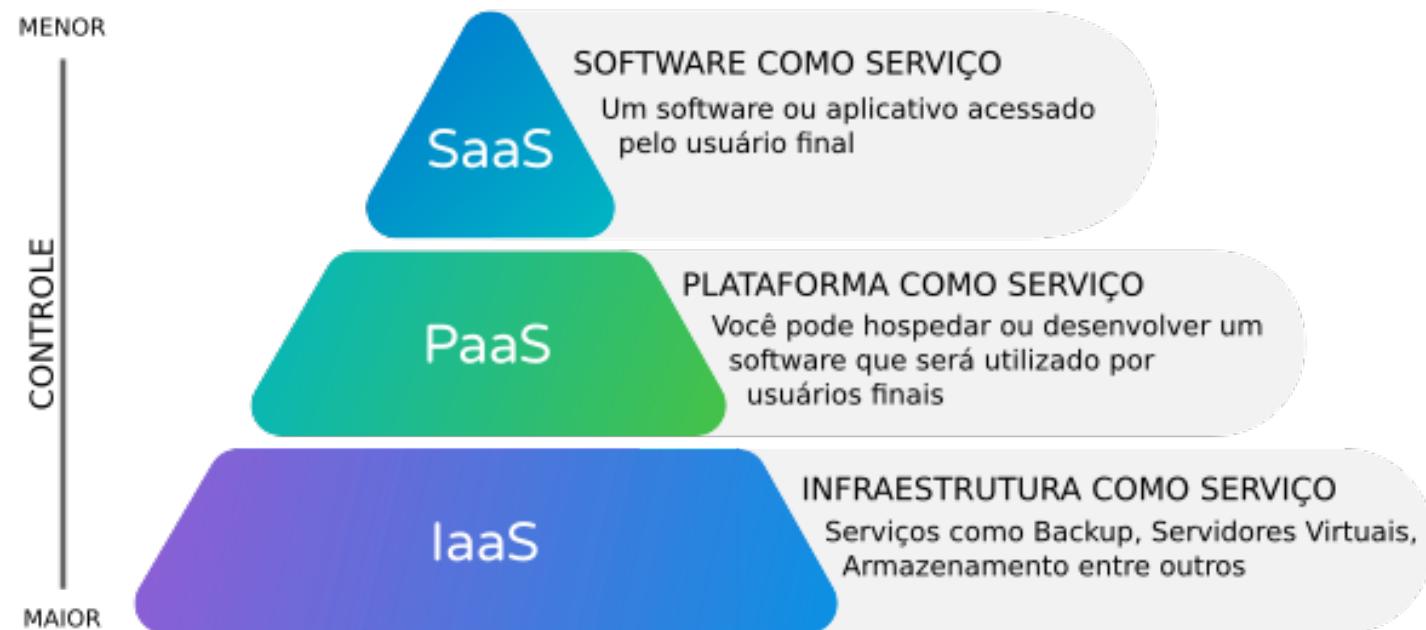
O **Software como Serviço** é a opção mais utilizada para empresas no mercado de nuvem.

O SaaS utiliza a Internet para entregar aplicativos aos clientes, sendo gerenciados por um fornecedor terceirizado.

A maioria dos aplicativos SaaS é executado diretamente no navegador de web, o que significa que eles não exigem nenhum download ou instalação por parte do cliente.

CLOUD COMPUTING

SAAS (SOFTWARE AS A SERVICE)



SaaS e Engenharia de dados se interligam de várias maneiras, onde são aproveitados recursos para facilitar processos.

Ao interligar SaaS com engenharia de dados, você pode aproveitar as vantagens dos aplicativos SaaS para coletar, gerenciar e analisar dados, melhorando assim a eficiência operacional e impulsionando a tomada de decisões baseada em dados.

- I. *Integração de Dados;*
- II. *Análise de Dados;*
- III. *Automação de Processos;*
- IV. *Gerenciamento de Dados;*
- V. *Colaboração e Compartilhamento de Dados;*

CLOUD COMPUTING

SAAS (SOFTWARE AS A SERVICE)

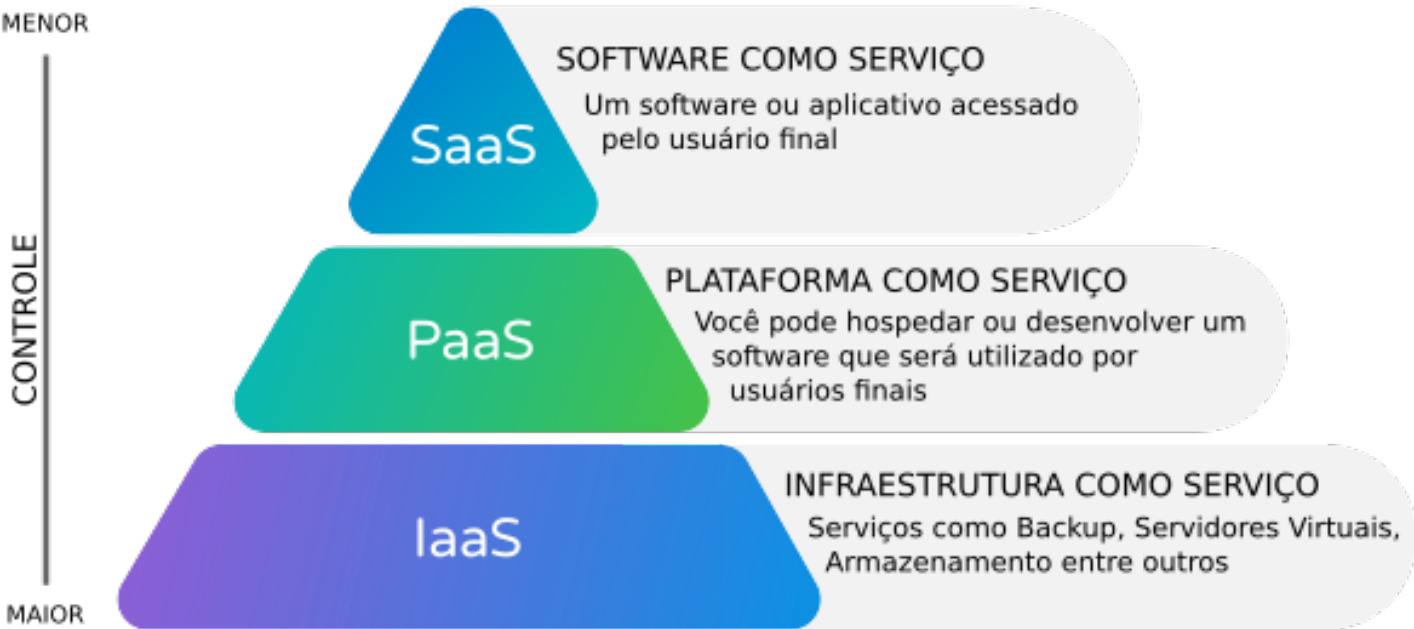


Cenário: Uma empresa de marketing precisa analisar o desempenho de suas campanhas de marketing em várias plataformas de mídia social e anúncios online.

Aplicação de SaaS: A empresa pode utilizar plataformas de análise de marketing como Google Analytics, HubSpot ou Adobe Analytics, que são exemplos de SaaS. Essas plataformas oferecem recursos poderosos de análise de dados, permitindo que a empresa acompanhe métricas-chave, como tráfego do site, conversões, ROI de campanhas e engajamento nas redes sociais. Os dados podem ser facilmente acessados através da interface web do SaaS, sem a necessidade de configuração de infraestrutura adicional.

CLOUD COMPUTING

PAAS (PLATAFORMA AS A SERVICE)



A **Plataforma como Serviço** fornece componentes em nuvem para determinado software enquanto são usados principalmente para aplicativos.

O PaaS fornece uma estrutura para desenvolvedores desenvolver e usar para criar aplicativos personalizados. Todos os servidores, armazenamento e rede podem ser gerenciados pela empresa ou por um provedor de terceiros, enquanto os desenvolvedores podem manter o gerenciamento dos aplicativos. O modelo do PaaS é semelhante ao SaaS, exceto que, ao invés de entregar o software pela Internet, o PaaS fornece uma plataforma para criação de software. Essa plataforma é fornecida via web, oferecendo aos desenvolvedores a liberdade de se concentrar na criação do software sem se preocupar com sistemas operacionais, atualizações de software, armazenamento ou infraestrutura.

CLOUD COMPUTING

PAAS (PLATAFORMA AS A SERVICE)



PaaS fornece uma plataforma completa para desenvolver, implantar e gerenciar aplicativos, o que pode ser aproveitado para facilitar o trabalho de Engenharia de Dados:

- I. Ferramentas de Processamento de Dados;
- II. Bancos de Dados Gerenciados;
- III. APIs para Integração de Dados;
- IV. Automação de Processos;
- V. Desenvolvimento de Aplicações de Análise;
- VI. Colaboração e Compartilhamento de Dados;

Ao interligar PaaS com Engenharia de Dados, podemos criar um ambiente integrado e eficiente para desenvolver, implantar e gerenciar pipelines de dados, análises e aplicativos de dados personalizados. Isso pode ajudar a melhorar a eficiência, acelerar o ciclo de desenvolvimento e fornecer insights valiosos para o seu negócio.

CLOUD COMPUTING

PAAS (PLATAFORMA AS A SERVICE)



Cenário: Uma startup de tecnologia está desenvolvendo um aplicativo de análise de dados para processar e visualizar grandes conjuntos de dados de sensores IoT (Internet das Coisas).

Aplicação de PaaS: A startup pode usar uma plataforma PaaS, como Microsoft Azure IoT Hub ou AWS IoT Core, para gerenciar a ingestão e processamento de dados de sensores IoT. Essas plataformas fornecem serviços gerenciados para coleta, armazenamento e análise de dados de dispositivos IoT. Além disso, a startup pode utilizar serviços de análise de dados em tempo real, como Azure Stream Analytics ou AWS Kinesis, para processar e analisar os dados dos sensores. Eles também podem usar serviços de banco de dados gerenciados, como Azure Cosmos DB ou AWS DynamoDB, para armazenar os dados processados e permitir consultas rápidas e escaláveis.

CLOUD COMPUTING

VANTAGENS

SERVIÇO	VANTAGEM	APLICAÇÃO
IaaS	O IaaS é o modelo de Computação em Nuvem mais flexível, pois os recursos de hardware podem ser adquiridos conforme a necessidade. Além disso, o modelo é fácil de automatizar e é altamente escalável. Os clientes, portanto, mantêm o controle completo de sua infraestrutura.	AWS EC2 Google Compute Engine (GCE) Huawei Kunpeng Cloud Services.
SaaS	Devido ao seu modelo de entrega na web, o SaaS elimina a necessidade de a equipe de TI fazer o download e instalar aplicativos em cada computador. Com o SaaS, os fornecedores gerenciam todos os possíveis problemas técnicos como dados, middleware, servidores e armazenamento, resultando em uma manutenção e suporte otimizados para o seu negócio.	G Suite, Salesforce, Mailchimp, Google Analytics, Dropbox.
PaaS	Os benefícios do PaaS incluem desenvolvimento e implementação simples e econômica, que é altamente disponível. Além disso, os desenvolvedores podem focar na personalização dos aplicativos.	AWS Beanstalk, AWS RDS, GCP App Engine.

ENGENHARIA DE DADOS

TÓPICOS ABORDADOS

Introdução à Computação em Nuvem:

- ~~Definição de computação em nuvem.~~
- ~~Benefícios da computação em nuvem, como escalabilidade, flexibilidade e redução de custos.~~
- ~~Modelos de serviço em nuvem: IaaS (Infrastructure as a Service), PaaS (Platform as a Service) e SaaS (Software as a Service).~~

O que é DaaS:

- Definição de DaaS.
- Diferença entre dados tradicionais e DaaS.
- Benefícios do DaaS para organizações, como acesso fácil a conjuntos de dados, eliminação de custos de infraestrutura e capacidade de escalabilidade.

Arquitetura do DaaS:

- Componentes principais do DaaS, incluindo armazenamento de dados, ferramentas de processamento de dados e serviços de entrega de dados.
- Descrição do processo de provisionamento e gerenciamento de serviços de dados.

Implementação do DaaS:

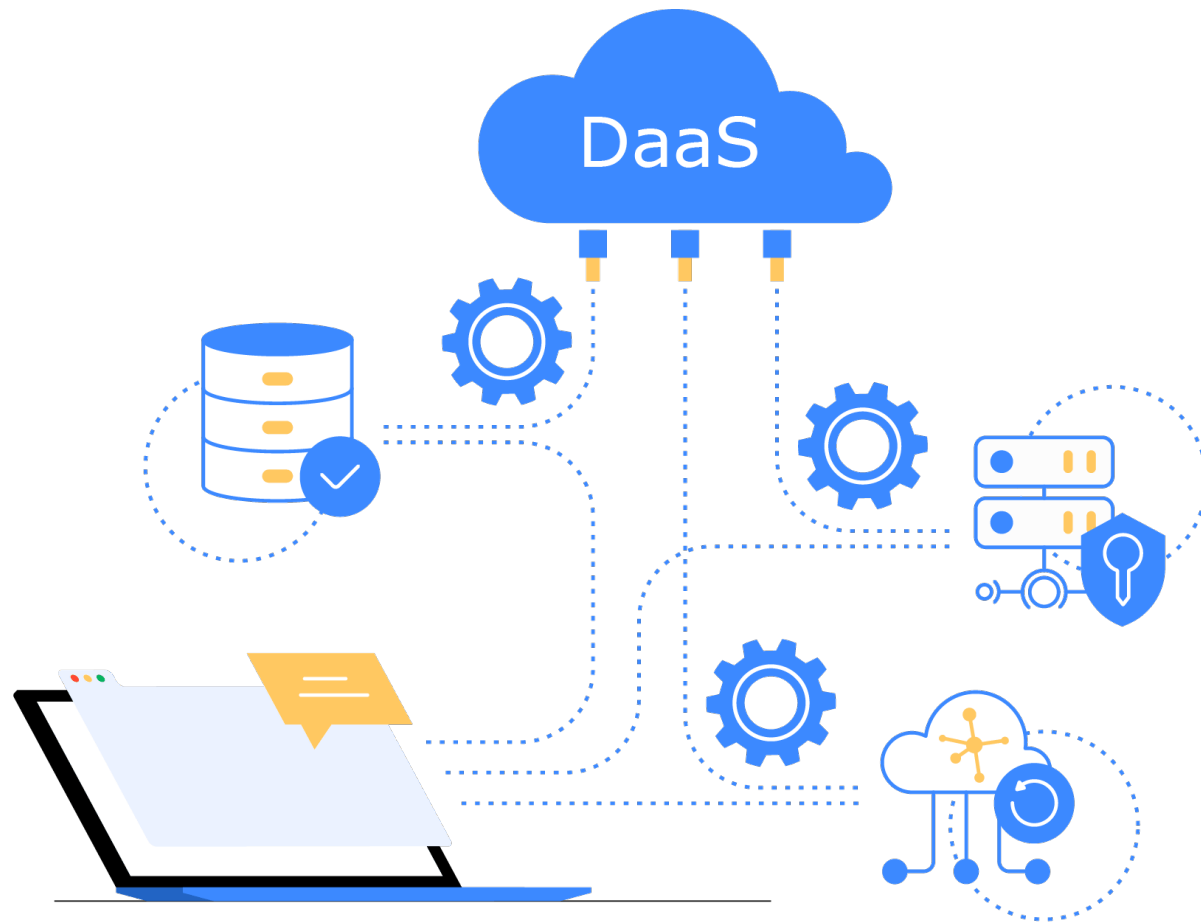
- Opções de implementação do DaaS, incluindo provedores de serviços em nuvem e soluções de software específicas para DaaS.
- Considerações ao migrar para o DaaS, como requisitos de segurança, conformidade regulatória e custos.

Cenários de Uso do DaaS:

- Exemplos de casos de uso do DaaS em diferentes setores, como análise de dados, ciência de dados, aprendizado de máquina, IoT (Internet das Coisas) e análise preditiva.
- Demonstração de como os usuários podem interagir com os dados fornecidos pelo DaaS e incorporá-los em suas próprias aplicações e processos de negócios.

DATA AS A SERVICE

DEFINIÇÃO



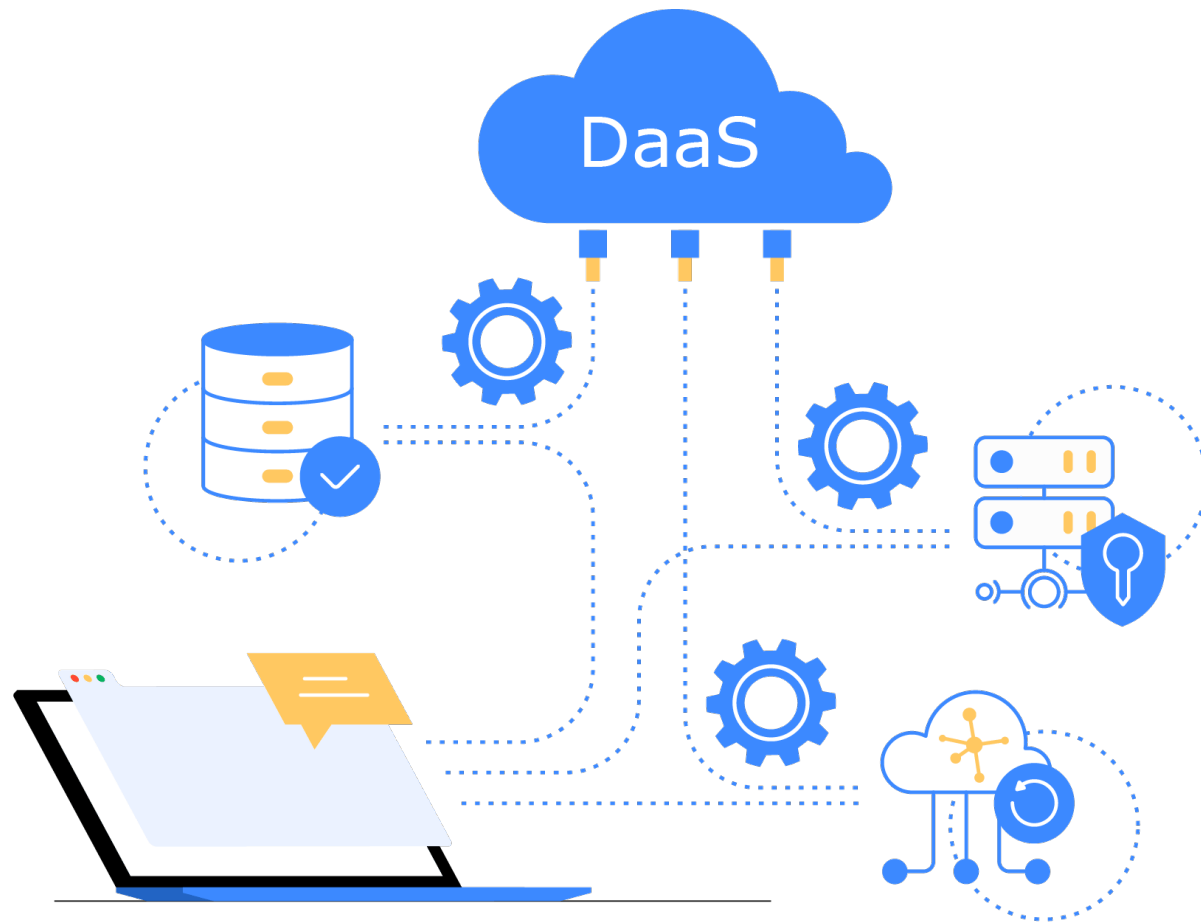
Dados como serviço é uma variedade de soluções baseadas em nuvem usadas para trabalhar com dados. Essas soluções são oferecidas por fornecedores de DaaS, que disponibilizam dados de negócios prontamente sob demanda pela Internet.

DaaS é frequentemente descrito como uma estratégia de gerenciamento de dados que emprega a nuvem para fornecer armazenamento de dados, realizar monitoramento, processar dados e assim por diante.

Semelhante ao SaaS (Software as a Service), o DaaS elimina a necessidade de baixar e gerenciar software localmente. Essa característica torna o serviço em nuvem popular entre empresas que desejam migrar seu armazenamento de dados de servidores locais para a nuvem.

DATA AS A SERVICE

DEFINIÇÃO



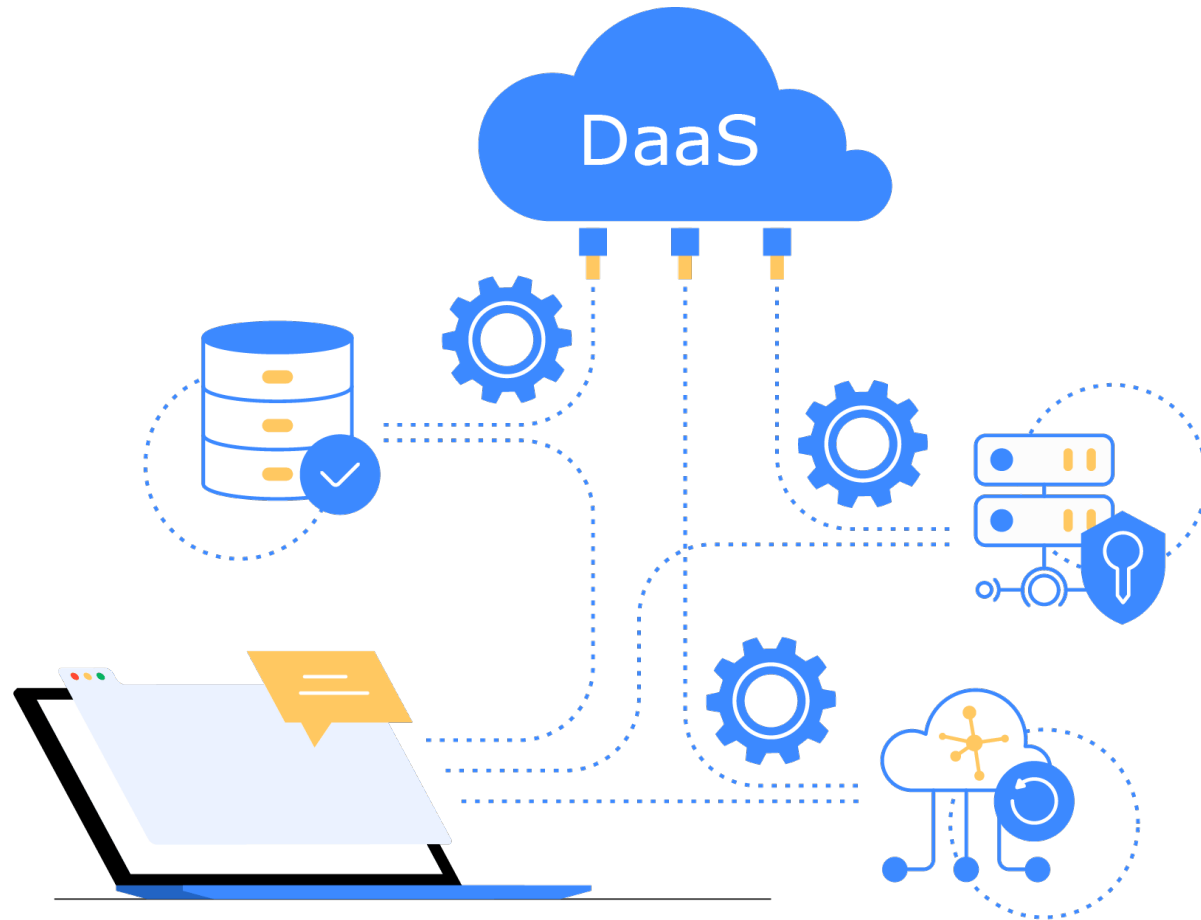
Numerosos benefícios do DaaS que podem influenciar significativamente o desempenho dos negócios atraem grandes e pequenas empresas em todo o mundo.

Maior acessibilidade – As empresas que adotam DaaS não estão vinculadas a sistemas e servidores de armazenamento locais, uma vez que armazenam dados remotamente. Portanto, podem colaborar com parceiros globais de forma eficiente. Simultaneamente, os funcionários podem acessar dados sem precisar ir ao escritório, economizando tempo e energia.

Produtos melhores – o DaaS ajuda as empresas a perceber os pontos fracos e fortes de seus produtos. Com base nisso, eles podem minimizar os pontos fracos e aproveitar os pontos fortes, levando a melhorias no produto.

DATA AS A SERVICE

DEFINIÇÃO



Custo-benefício – Quem entende perfeitamente o que é DaaS enfatiza que sua vantagem mais considerável é o custo-benefício. Com o DaaS, não há necessidade de investir recursos financeiros na manutenção e gerenciamento de dados porque tudo é feito pelos fornecedores de DaaS.

Pouco tempo de configuração – As empresas podem começar a armazenar e processar dados imediatamente ao empregar DaaS. Por outro lado, as empresas que mantêm os dados localmente gastam muito mais tempo na configuração.

Experiência do usuário aprimorada – As empresas que estão cientes do que é DaaS tomam decisões baseadas em dados que melhoram a experiência do cliente. Portanto, eles observam melhorias na retenção de clientes e no valor da vida útil do cliente.

DATA AS A SERVICE

TRADICIONAL X MODERNO

ABORDAGEM	TRADICIONAL	DAAS
Acesso aos Dados:	Nos métodos tradicionais, os dados são armazenados localmente em servidores ou infraestrutura de TI da própria organização. O acesso aos dados geralmente requer uma infraestrutura de hardware e software dedicada dentro da organização.	Os dados são disponibilizados como um serviço através da nuvem. Isso significa que os usuários podem acessar os dados remotamente pela internet, sem a necessidade de manter uma infraestrutura de TI local. Os dados são hospedados e gerenciados por um provedor de serviços em nuvem, que fornece acesso sob demanda aos dados conforme necessário.
Armazenamento de Dados:	Nos métodos tradicionais, os dados são armazenados localmente em bancos de dados, servidores de arquivos ou sistemas de armazenamento próprios da organização. A capacidade de armazenamento é limitada pela infraestrutura de TI disponível.	Os dados são armazenados na nuvem pelo provedor de serviços. Isso permite uma escalabilidade quase ilimitada do armazenamento, conforme a demanda, sem a necessidade de investimentos em infraestrutura adicional por parte da organização.
Gerenciamento e Manutenção:	A organização é responsável pelo gerenciamento e manutenção dos seus próprios dados. Isso inclui tarefas como backup, atualizações de software, segurança e monitoramento de desempenho.	O provedor de serviços em nuvem é responsável pelo gerenciamento e manutenção dos dados. Isso inclui garantir a disponibilidade, segurança, conformidade e desempenho dos dados hospedados na nuvem.
Custos e Flexibilidade:	Os custos de infraestrutura de TI própria podem ser significativos, incluindo a compra de hardware, licenças de software e custos operacionais contínuos. A escalabilidade pode ser limitada pela capacidade da infraestrutura existente.	Custos são geralmente baseados no uso, o que significa que as organizações podem pagar apenas pelo que realmente utilizam. Além disso, a escalabilidade é facilitada, permitindo que as organizações aumentem ou diminuam facilmente a capacidade de armazenamento e processamento de dados conforme necessário.

DATA AS A SERVICE

BENEFÍCIOS PARA EMPRESAS



Os benefícios do DaaS para organizações são significativos e podem transformar a maneira como lidam com dados.

Acesso Fácil a Conjuntos de Dados: Com o DaaS, as organizações podem acessar uma ampla variedade de conjuntos de dados de forma fácil e rápida, muitas vezes por meio de APIs simples ou interfaces web amigáveis.

Exemplo de Uso: Uma empresa de entrega de utiliza as API do Google Maps Platform para integrar recursos de mapeamento em seu aplicativo. Eles usam a API de Geocodificação para converter endereços em coordenadas geográficas, a API de Roteamento para calcular rotas otimizadas e a API de Visualização de Mapa para exibir mapas interativos em seu app. Com isso a empresa fornece aos entregadores e clientes informações precisas de localização em tempo real, melhorando a eficiência e a experiência do usuário.



Eliminação de Custos de Infraestrutura: O DaaS elimina a necessidade de investimento em infraestrutura de TI dedicada para armazenamento, processamento e análise de dados.

Exemplo de Uso: Uma empresa de comércio eletrônico usa o Amazon Redshift para armazenar e analisar grandes volumes de dados de transações de clientes, histórico de compras e comportamento de navegação na web. Eles podem executar consultas complexas e gerar insights valiosos sobre padrões de compra, preferências do cliente e tendências de mercado sem precisar investir em infraestrutura de data warehousing dedicada. Isso permite que a empresa tome decisões de negócios mais informadas e eficazes para melhorar a experiência do cliente e impulsionar as vendas.



O Amazon Redshift é um serviço de data warehousing totalmente gerenciado que permite às organizações armazenar e analisar grandes volumes de dados de maneira eficiente e escalável na nuvem da AWS.

Ele oferece capacidade de armazenamento expansível, alta disponibilidade e desempenho rápido para cargas de trabalho de análise de dados.

DATA AS A SERVICE

BENEFÍCIOS PARA EMPRESAS



Agilidade e Flexibilidade: As organizações podem facilmente adicionar novos conjuntos de dados, alterar os requisitos de processamento ou análise e integrar-se a novas ferramentas ou plataformas conforme necessário.

Exemplo de Uso: Uma empresa de varejo online usa o Snowflake Data Cloud para armazenar e analisar dados de vendas, estoque, comportamento do cliente e marketing. Eles podem facilmente adicionar novos conjuntos de dados, como dados de mídia social, análises de sentimentos ou dados de localização, conforme necessário para melhorar suas análises e insights. Além disso, eles podem integrar-se a novas ferramentas de análise de dados e visualização de dados para fornecer uma visão abrangente de suas operações e desempenho de negócios.



O Snowflake Data Cloud é uma plataforma de dados em nuvem que oferece armazenamento, processamento e compartilhamento de dados de maneira escalável e flexível.

Ele permite que as organizações armazenem e analisem uma ampla variedade de dados, incluindo dados estruturados, semiestruturados e não estruturados, em uma arquitetura de dados unificada.

Introdução à Computação em Nuvem:

- ~~Definição de computação em nuvem.~~
- ~~Benefícios da computação em nuvem, como escalabilidade, flexibilidade e redução de custos.~~
- ~~Modelos de serviço em nuvem: IaaS (Infrastructure as a Service), PaaS (Platform as a Service) e SaaS (Software as a Service).~~

O que é DaaS:

- ~~Definição de DaaS.~~
- ~~Diferença entre dados tradicionais e DaaS.~~
- ~~Benefícios do DaaS para organizações, como acesso fácil a conjuntos de dados, eliminação de custos de infraestrutura e capacidade de escalabilidade.~~

Arquitetura do DaaS:

- Componentes principais do DaaS, incluindo armazenamento de dados, ferramentas de processamento de dados e serviços de entrega de dados.
- Descrição do processo de provisionamento e gerenciamento de serviços de dados.

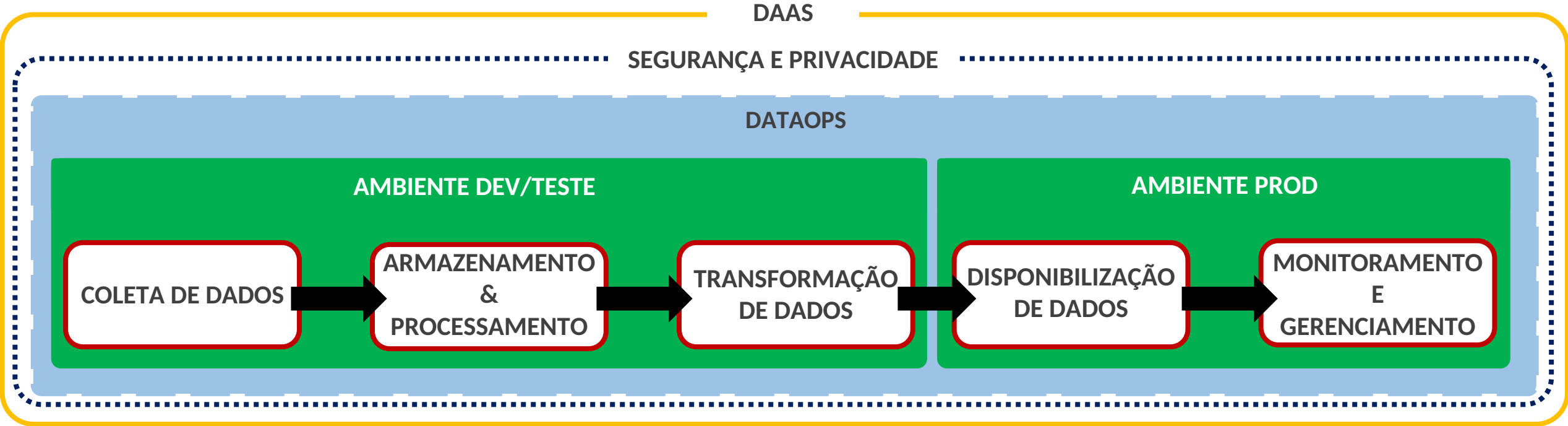
Implementação do DaaS:

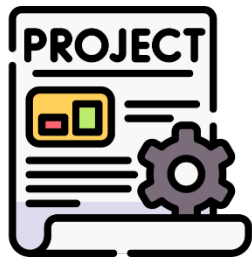
- Opções de implementação do DaaS, incluindo provedores de serviços em nuvem e soluções de software específicas para DaaS.
- Considerações ao migrar para o DaaS, como requisitos de segurança, conformidade regulatória e custos.

Cenários de Uso do DaaS:

- Exemplos de casos de uso do DaaS em diferentes setores, como análise de dados, ciência de dados, aprendizado de máquina, IoT (Internet das Coisas) e análise preditiva.
- Demonstração de como os usuários podem interagir com os dados fornecidos pelo DaaS e incorporá-los em suas próprias aplicações e processos de negócios.

DATA AS A SERVICE
ARQUITETURA





AMBIENTE DEV/TESTE

COLETA DE DADOS

ARMAZENAMENTO
&
PROCESSAMENTO

TRANSFORMAÇÃO
DE DADOS

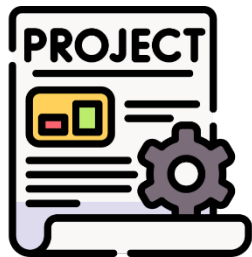
DESCRIÇÃO DO PROCESSO DE PROVISIONAMENTO E GERENCIAMENTO DE SERVIÇOS DE DADOS

Definição de Requisitos:

- O processo começa com a definição clara dos requisitos de dados da organização. Isso pode incluir os tipos de dados necessários, os volumes esperados, os requisitos de segurança e privacidade, entre outros.

Planejamento de Infraestrutura:

- Com base nos requisitos definidos, é feito um planejamento da infraestrutura necessária para armazenar, processar e disponibilizar os dados. Isso pode incluir a seleção de tecnologias de armazenamento, processamento e segurança adequadas.



AMBIENTE DEV/TESTE

COLETA DE DADOS

ARMAZENAMENTO
&
PROCESSAMENTO

TRANSFORMAÇÃO
DE DADOS

DESCRIÇÃO DO PROCESSO DE PROVISIONAMENTO E
GERENCIAMENTO DE SERVIÇOS DE DADOS

Provisionamento de Recursos:

- Nesta etapa, os recursos de hardware e software são provisionados de acordo com o planejamento feito anteriormente. Isso pode envolver a compra ou alocação de servidores, sistemas de armazenamento, software de banco de dados, ferramentas de processamento de dados, entre outros.

Configuração e Instalação:

- Uma vez provisionados, os recursos são configurados e instalados conforme as especificações do projeto. Isso pode incluir a configuração de servidores, a instalação e configuração de software, a configuração de redes e firewalls, entre outros.



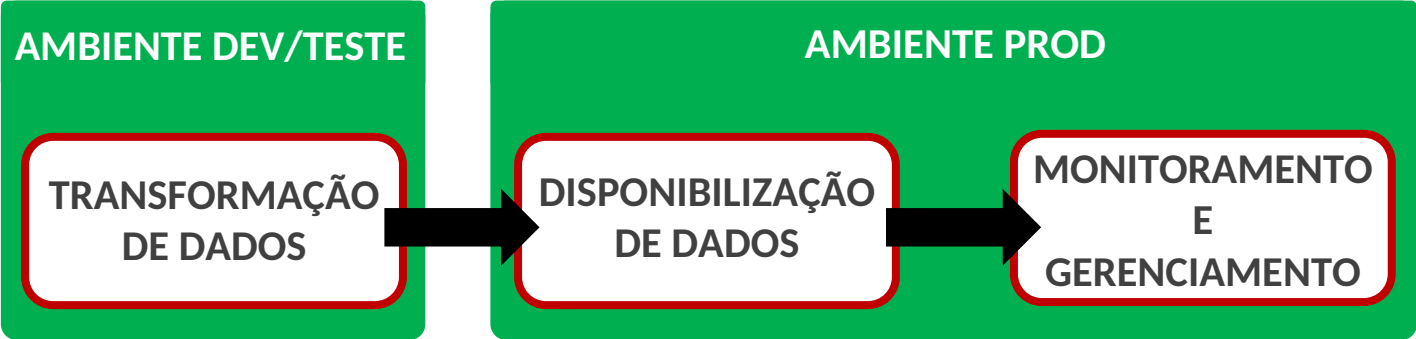
DESCRIÇÃO DO PROCESSO DE PROVISIONAMENTO E GERENCIAMENTO DE SERVIÇOS DE DADOS

Integração de Dados:

- Os dados são integrados a partir de diversas fontes, como bancos de dados internos, sistemas legados, feeds de dados externos, etc. Isso pode envolver o desenvolvimento de pipelines de dados para extrair, transformar e carregar (ETL) os dados para a infraestrutura de armazenamento.

Implementação de Segurança:

- As medidas de segurança são implementadas para proteger os dados contra acesso não autorizado, perda de dados e outros riscos de segurança. Isso pode incluir a configuração de políticas de acesso, criptografia de dados, monitoramento de atividades suspeitas, entre outros.



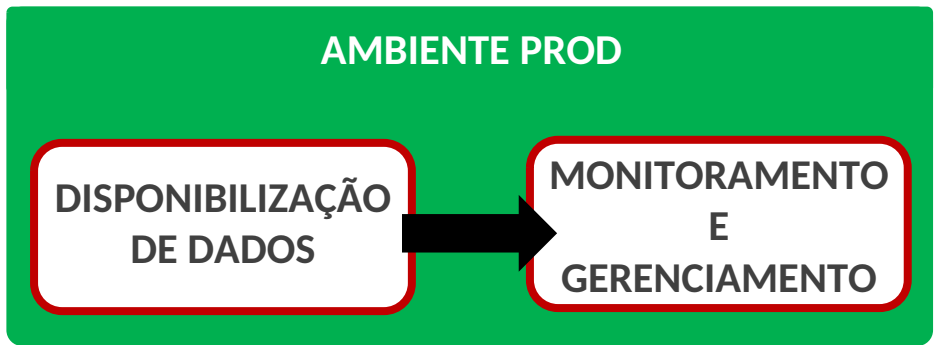
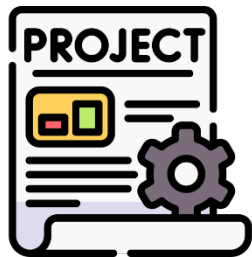
DESCRIÇÃO DO PROCESSO DE PROVISIONAMENTO E GERENCIAMENTO DE SERVIÇOS DE DADOS

Testes e Validação:

- Antes de disponibilizar os serviços de dados para uso, são realizados testes e validações para garantir que tudo esteja funcionando conforme o esperado. Isso pode incluir testes de integridade de dados, testes de desempenho, testes de segurança, entre outros.

Implantação e Disponibilização:

- Uma vez que os testes sejam bem-sucedidos, os serviços de dados são implantados e disponibilizados para uso pelos usuários finais. Isso pode envolver a configuração de interfaces de consulta, APIs, ferramentas de visualização de dados, entre outros.



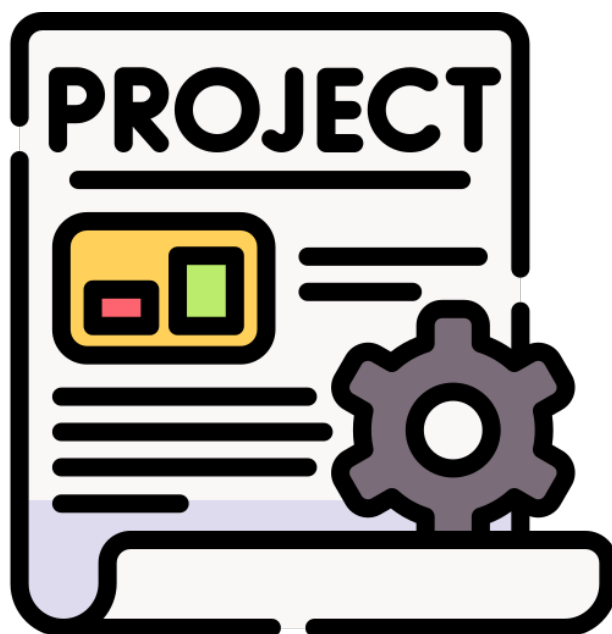
DESCRIÇÃO DO PROCESSO DE PROVISIONAMENTO E GERENCIAMENTO DE SERVIÇOS DE DADOS

Monitoramento e Manutenção:

- Após a implantação, os serviços de dados são monitorados continuamente para garantir seu desempenho, disponibilidade e segurança. São realizadas atividades de manutenção preventiva, correção de problemas e atualizações conforme necessário.

Otimização Contínua:

- O processo de provisionamento e gerenciamento de serviços de dados é um ciclo contínuo. À medida que os requisitos e as tecnologias evoluem, são feitas otimizações contínuas na infraestrutura, nos processos e nos serviços de dados para garantir que atendam às necessidades em constante mudança da organização.



DESCRIÇÃO DO PROCESSO DE PROVISIONAMENTO E GERENCIAMENTO DE SERVIÇOS DE DADOS

1. Definição de Requisitos;
2. Planejamento de Infraestrutura;
3. Provisionamento de Recursos;
4. Configuração e Instalação;
5. Integração de Dados;
6. Implementação de Segurança;
7. Testes e Validação;
8. Implantação e Disponibilização;
9. Monitoramento e Manutenção;
10. Otimização Contínua;

ENGENHARIA DE DADOS

TÓPICOS ABORDADOS

Introdução à Computação em Nuvem:

- Definição de computação em nuvem.
- Benefícios da computação em nuvem, como escalabilidade, flexibilidade e redução de custos.
- Modelos de serviço em nuvem: IaaS (Infrastructure as a Service), PaaS (Platform as a Service) e SaaS (Software as a Service).

O que é DaaS:

- Definição de DaaS.
- Diferença entre dados tradicionais e DaaS.
- Benefícios do DaaS para organizações, como acesso fácil a conjuntos de dados, eliminação de custos de infraestrutura e capacidade de escalabilidade.

Arquitetura do DaaS:

- Componentes principais do DaaS, incluindo armazenamento de dados, ferramentas de processamento de dados e serviços de entrega de dados.
- Descrição do processo de provisionamento e gerenciamento de serviços de dados.

Implementação do DaaS:

- Opções de implementação do DaaS, incluindo provedores de serviços em nuvem e soluções de software específicas para DaaS.
- Considerações ao migrar para o DaaS, como requisitos de segurança, conformidade regulatória e custos.

Cenários de Uso do DaaS:

- Exemplos de casos de uso do DaaS em diferentes setores, como análise de dados, ciência de dados, aprendizado de máquina, IoT (Internet das Coisas) e análise preditiva.
- Demonstração de como os usuários podem interagir com os dados fornecidos pelo DaaS e incorporá-los em suas próprias aplicações e processos de negócios.

ENGENHARIA DE DADOS

VERDADEIRO OU FALSO

- 1 - O modelo DaaS permite aos usuários acessar dados sob demanda através da internet.
- 2 - DaaS é exclusivamente implementado no modelo de nuvem pública.
- 3 - No modelo DaaS público, os dados são armazenados e processados em infraestrutura própria da organização.
- 4 - DaaS Privado oferece maior controle sobre a segurança e privacidade dos dados.
- 5 - DaaS Híbrido combina elementos de infraestrutura local e serviços de nuvem pública.
- 6 - DaaS é exclusivamente utilizado por grandes empresas de tecnologia.
- 7 - O modelo DaaS público é mais adequado para organizações que precisam de controle total sobre seus dados.
- 8 - DaaS pode incluir serviços de integração de dados, limpeza de dados e enriquecimento de dados.
- 9 - No modelo DaaS Privado, os recursos de dados são compartilhados entre várias organizações.
- 10 - O modelo DaaS Híbrido oferece flexibilidade ao combinar infraestrutura local e serviços de nuvem pública.



ENGENHARIA DE DADOS

TÓPICOS ABORDADOS

Introdução à Computação em Nuvem:

- Definição de computação em nuvem.
- Benefícios da computação em nuvem, como escalabilidade, flexibilidade e redução de custos.
- Modelos de serviço em nuvem: IaaS (Infrastructure as a Service), PaaS (Platform as a Service) e SaaS (Software as a Service).

O que é DaaS:

- Definição de DaaS.
- Diferença entre dados tradicionais e DaaS.
- Benefícios do DaaS para organizações, como acesso fácil a conjuntos de dados, eliminação de custos de infraestrutura e capacidade de escalabilidade.

Arquitetura do DaaS:

- Componentes principais do DaaS, incluindo armazenamento de dados, ferramentas de processamento de dados e serviços de entrega de dados.
- Descrição do processo de provisionamento e gerenciamento de serviços de dados.

Implementação do DaaS:

- Opções de implementação do DaaS, incluindo provedores de serviços em nuvem e soluções de software específicas para DaaS.
- Considerações ao migrar para o DaaS, como requisitos de segurança, conformidade regulatória e custos.

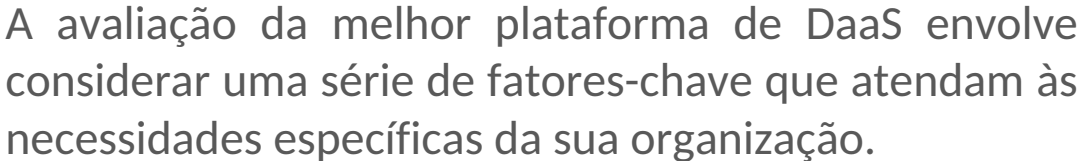
Cenários de Uso do DaaS:

- Exemplos de casos de uso do DaaS em diferentes setores, como análise de dados, ciência de dados, aprendizado de máquina, IoT (Internet das Coisas) e análise preditiva.
- Demonstração de como os usuários podem interagir com os dados fornecidos pelo DaaS e incorporá-los em suas próprias aplicações e processos de negócios.



é uma abordagem que permite às organizações armazenar dados de forma conveniente e eficiente, sem a necessidade de possuir ou gerenciar a estrutura subjacente.

Em vez disso, os dados são disponibilizados como um serviço, geralmente por meio da nuvem, permitindo que as empresas os consumam conforme necessário, pagando apenas pelo que utilizam.



Como:

- Identificação das Partes Interessadas (Stakeholders);
- Elicitação de Requisitos (Coleta de informações);
- Análise de Requisitos (Analisar informações);
- Documentação de Requisitos;
- Validação de Requisitos;
- Gestão de Mudanças;

Analise SWOT



Analise os Objetivos de Negócios: Identifique os objetivos estratégicos da organização e como os dados podem ajudar a alcançá-los. Por exemplo, se a meta é aumentar as vendas, os requisitos de dados podem incluir informações sobre clientes, produtos e desempenho de vendas.

Como:

- Identificação dos Objetivos de Negócios (Estratégias);
- Entendimento do Contexto Organizacional (Missão);
- Análise de Stakeholders (Interessados)
- Mapeamento de Objetivos para Requisitos de Dados;
- Priorização de Objetivos e Requisitos de Dados;
- Análise de Lacunas;
- Definição de Métricas de Sucesso (KPI - indicadores-chave de desempenho);

DATA AS A SERVICE

IMPLEMENTAÇÃO



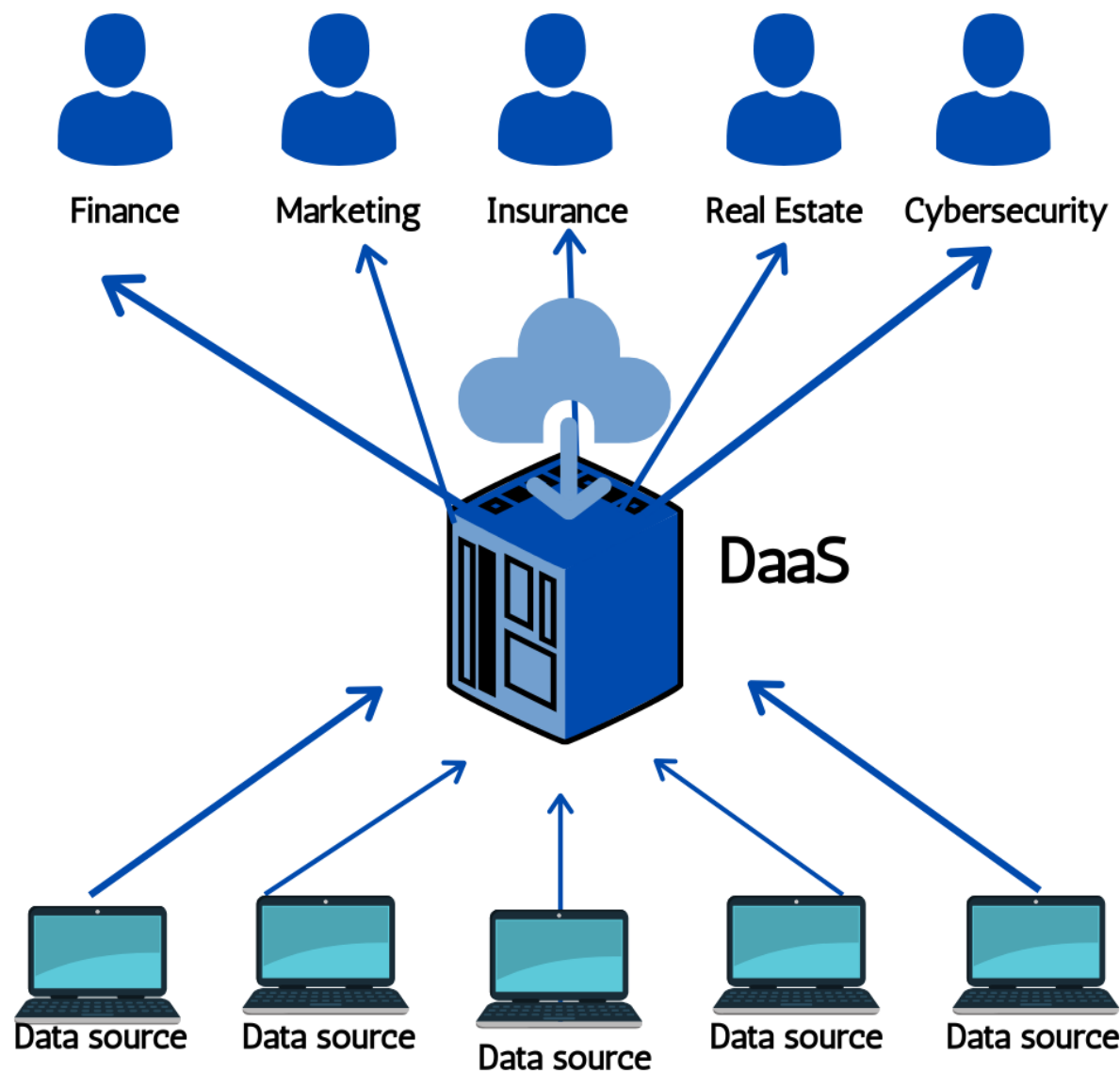
Avalie o Volume e os Tipos de Dados: Analise o volume de dados que a organização lida atualmente e prevê lidar no futuro. Além disso, identifique os diferentes tipos de dados, como dados estruturados (por exemplo, bancos de dados SQL), dados não estruturados (por exemplo, documentos, mídia social) e dados semi-estruturados (por exemplo, XML, JSON).

Como:

- Identificação das Fontes de Dados;
- Análise do Volume de Dados (Armazenados);
- Previsão de Crescimento de Dados (Custos futuros);
- Classificação dos Tipos de Dados (Estruturados, não-estruturados ou semi-estruturados);
- Avaliação da Capacidade de Gerenciamento de Dados Atual (identificação de pontos de atenção ou risco);

DATA AS A SERVICE

IMPLEMENTAÇÃO



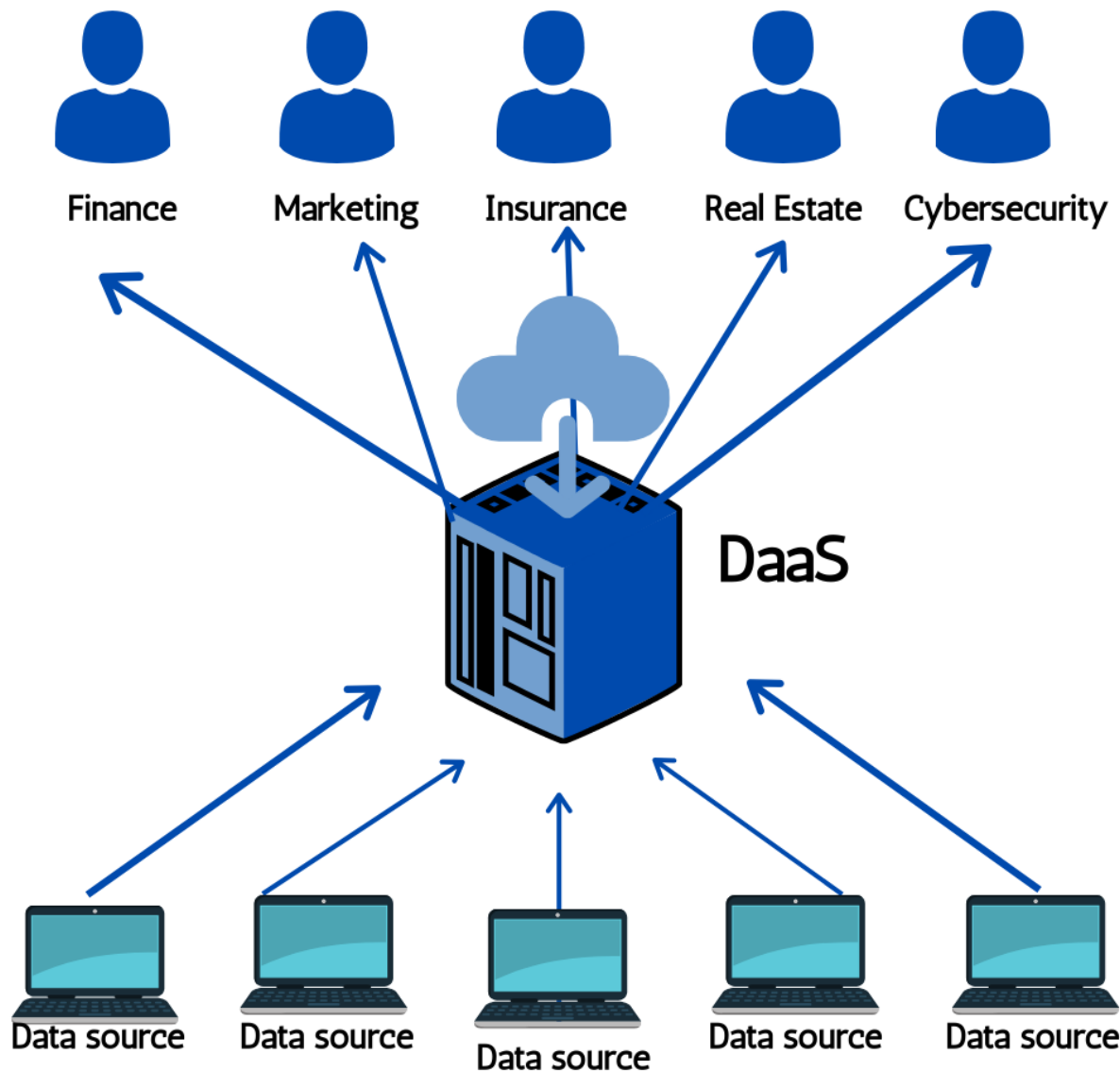
Avalie a Integração com Sistemas Existentes: Entenda quais sistemas e aplicativos existentes na organização precisam integrar-se com a plataforma de DaaS. Isso pode incluir sistemas de CRM, ERP, sistemas de gestão de conteúdo, entre outros.

Como:

- Mapeamento dos Sistemas e Aplicativos Existentes;
- Entendimento dos Fluxos de Dados;
- Análise das Integrações Atuais;
- Priorização das Integrações;
- Avaliação da Viabilidade Técnica;
- Definição de Requisitos de Integração;
- Comunicação com as Partes Interessadas;

DATA AS A SERVICE

IMPLEMENTAÇÃO



Identifique Necessidades de Segurança e Conformidade: Avalie os requisitos de segurança e conformidade da sua organização, incluindo regulamentos específicos, políticas internas de segurança de dados e necessidades de privacidade dos clientes.

Como:

Levantamento dos Regulamentos e Políticas Relevantes;

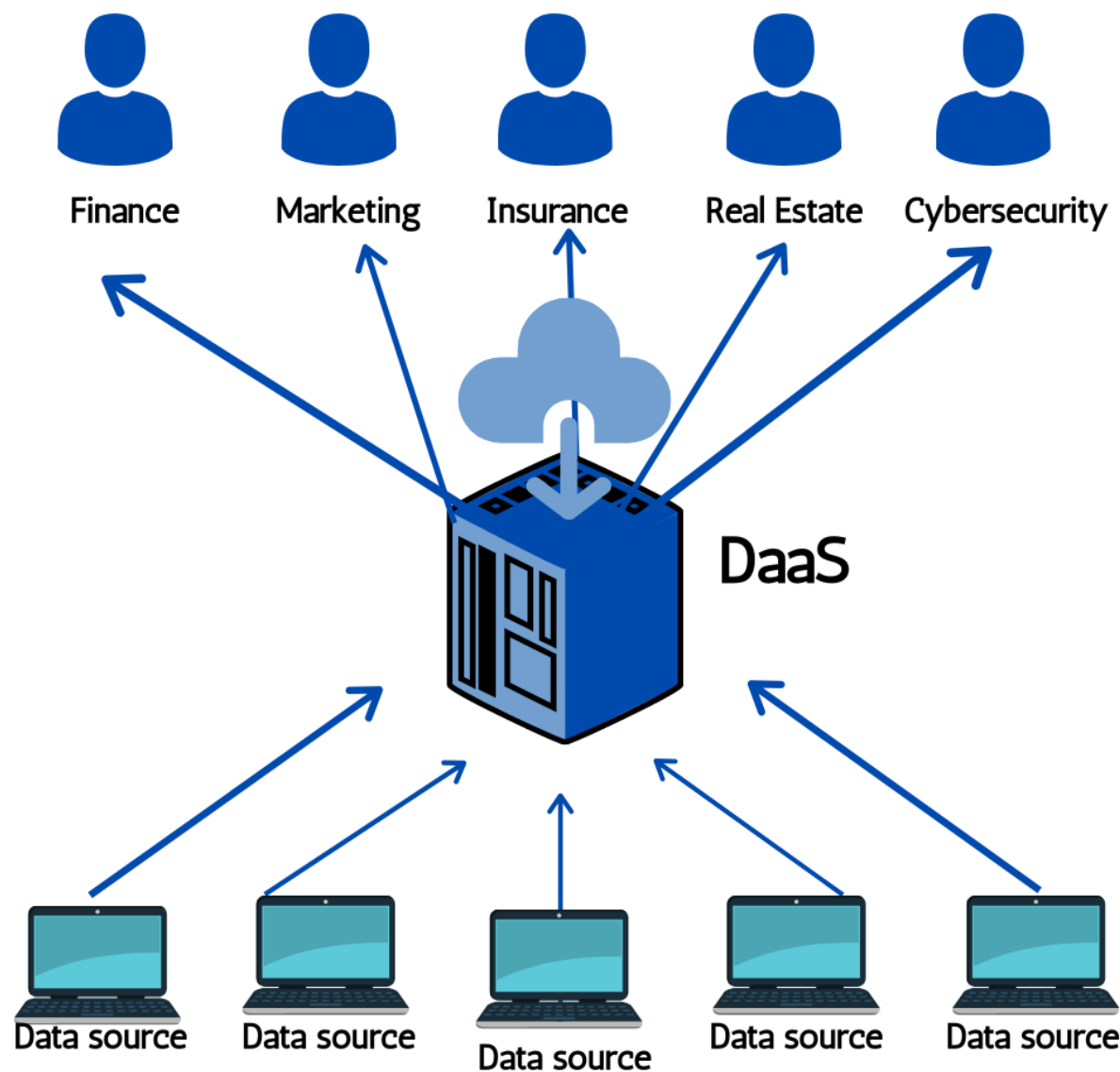
Análise das Políticas Internas de Segurança de Dados;
Identificação dos Dados Sensíveis e PII (Informações Pessoais Identificáveis);

Avaliação de Riscos de Segurança de Dados;
Definição de Requisitos de Segurança e Conformidade;
Implementação de Medidas de Segurança e Conformidade;

Revisão e Atualização Regular;

DATA AS A SERVICE

IMPLEMENTAÇÃO



Considerações ao Escolher um Provedor de DaaS:

Escalabilidade: A capacidade do serviço de lidar com volumes crescentes de dados conforme a demanda.

Custo: Modelos de precificação variáveis, dependendo do volume de dados, frequência de acesso e complexidade das consultas.

Integração: Facilidade de integração com sistemas existentes e outras ferramentas de análise e processamento de dados.

Segurança: Medidas de segurança implementadas para proteger os dados em trânsito e em repouso.

Facilidade de Uso: Interface intuitiva e facilidade de configuração e uso.

A escolha ideal depende das necessidades específicas da sua organização, incluindo orçamento, volume de dados, requisitos de segurança e preferências de integração.

Introdução à Computação em Nuvem:

- Definição de computação em nuvem.
- Benefícios da computação em nuvem, como escalabilidade, flexibilidade e redução de custos.
- Modelos de serviço em nuvem: IaaS (Infrastructure as a Service), PaaS (Platform as a Service) e SaaS (Software as a Service).

O que é DaaS:

- Definição de DaaS.
- Diferença entre dados tradicionais e DaaS.
- Benefícios do DaaS para organizações, como acesso fácil a conjuntos de dados, eliminação de custos de infraestrutura e capacidade de escalabilidade.

Arquitetura do DaaS:

- Componentes principais do DaaS, incluindo armazenamento de dados, ferramentas de processamento de dados e serviços de entrega de dados.
- Descrição do processo de provisionamento e gerenciamento de serviços de dados.

Implementação do DaaS:

- Opções de implementação do DaaS, incluindo provedores de serviços em nuvem e soluções de software específicas para DaaS.
- Considerações ao migrar para o DaaS, como requisitos de segurança, conformidade regulatória e custos.

Cenários de Uso do DaaS:

- Exemplos de casos de uso do DaaS em diferentes setores, como análise de dados, ciência de dados, aprendizado de máquina, IoT (Internet das Coisas) e análise preditiva.
- Demonstração de como os usuários podem interagir com os dados fornecidos pelo DaaS e incorporá-los em suas próprias aplicações e processos de negócios.



Caso de Uso: Personalização de Campanhas Publicitárias em uma Empresa de Marketing:

Contexto

Uma empresa de marketing deseja personalizar campanhas publicitárias para diferentes segmentos de clientes com base em dados comportamentais, demográficos e de compra.

Para isso, a empresa precisa acessar, integrar e analisar grandes volumes de dados provenientes de várias fontes, incluindo:

- Redes sociais;
- Histórico de compras;
- Interações em sites;
- Dados demográficos.

DATA AS A SERVICE

CENÁRIOS DE USO



Caso de Uso: Personalização de Campanhas Publicitárias em uma Empresa de Marketing:

Solução DaaS

A empresa decide utilizar um serviço DaaS para centralizar e fornecer os dados necessários.

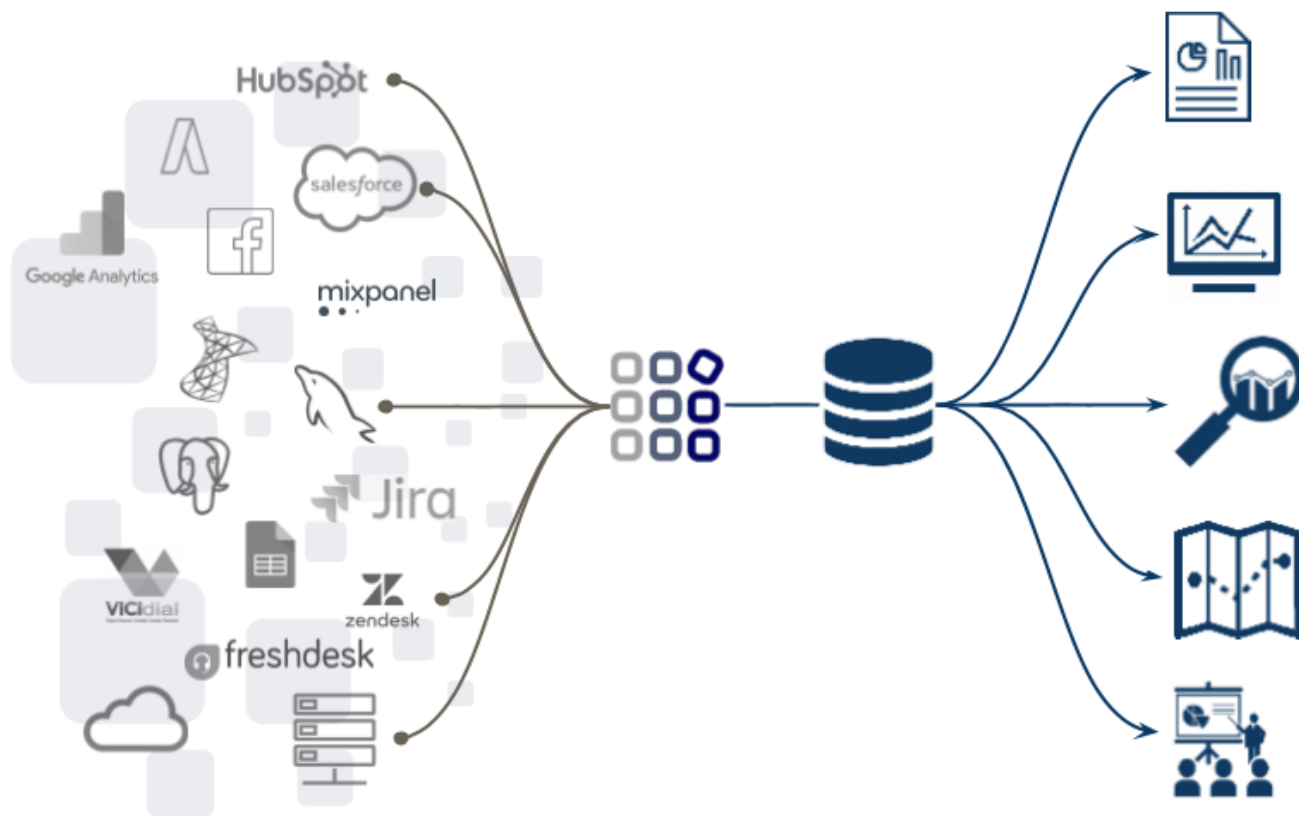
O DaaS facilita o acesso a dados atualizados em tempo real, integra dados de múltiplas fontes e oferece ferramentas de análise robustas.

Detalhes Técnicos:

- I. Coleta e Integração de Dados
- II. Armazenamento de Dados
- III. Processamento e Análise de Dados
- IV. Segurança e Governança de Dados
- V. Entrega de Dados e Integração com Ferramentas de Marketing

DATA AS A SERVICE

CENÁRIOS DE USO



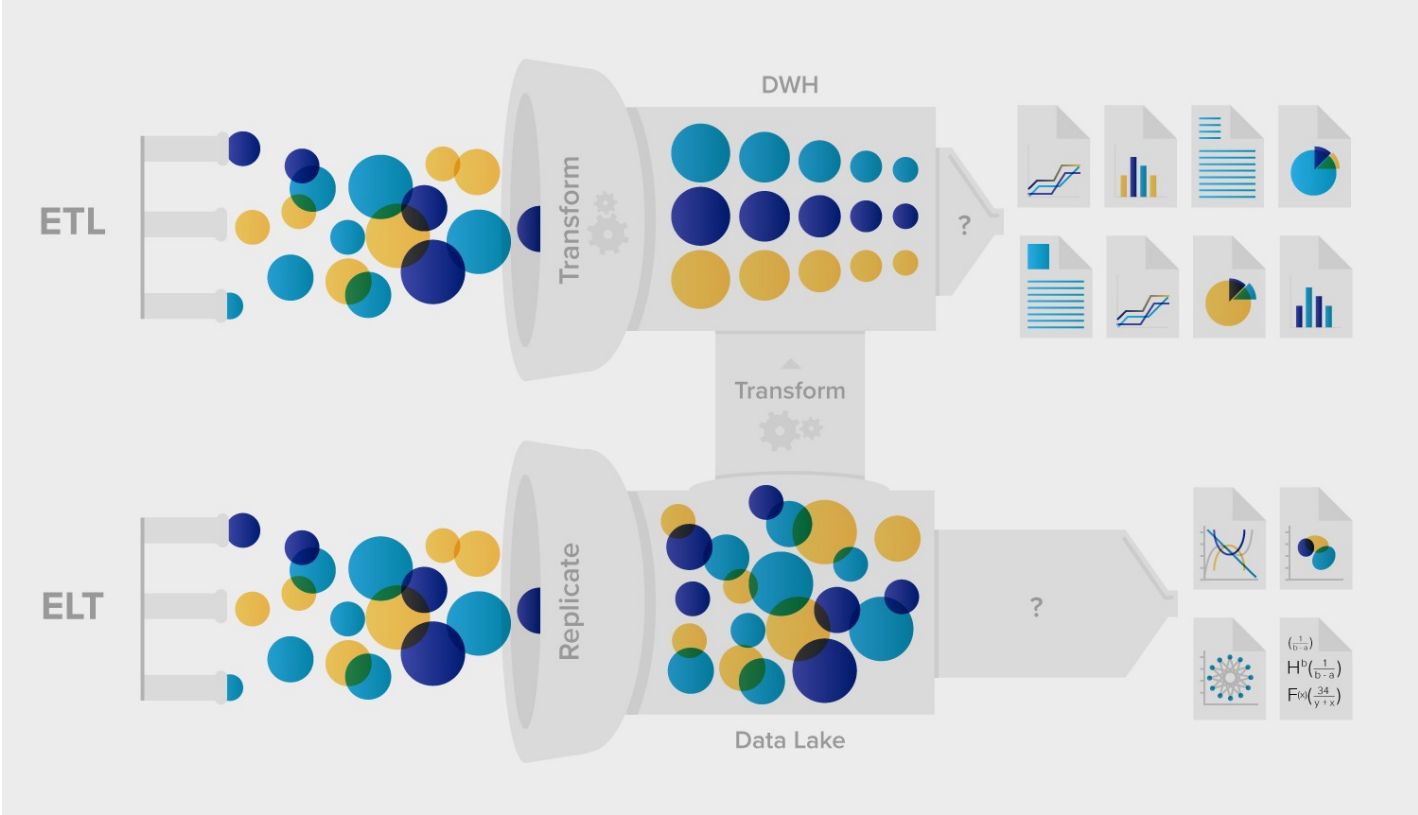
Caso de Uso: Personalização de Campanhas Publicitárias em uma Empresa de Marketing:

Coleta e Integração de Dados

- **Fontes de Dados:** Redes sociais (Facebook, Twitter), CRM, sistemas de e-commerce, Google Analytics, bancos de dados demográficos.
- **APIs e ETL (Extract, Transform, Load):** O DaaS utiliza APIs para coletar dados em tempo real de redes sociais e outras plataformas. Processos ETL são implementados para transformar e carregar dados de sistemas legados e bancos de dados locais para a plataforma DaaS.

DATA AS A SERVICE

ETL & ELT



Nas últimas décadas, o ETL (extrair, transformar, carregar) tem sido a abordagem tradicional para armazenamento e análise de dados. A abordagem ELT (extrair, carregar, transformar) muda o antigo paradigma.

Mas, o que realmente acontece quando o “T” e o “L” são trocados?

ETL e ELT resolvem a mesma necessidade:

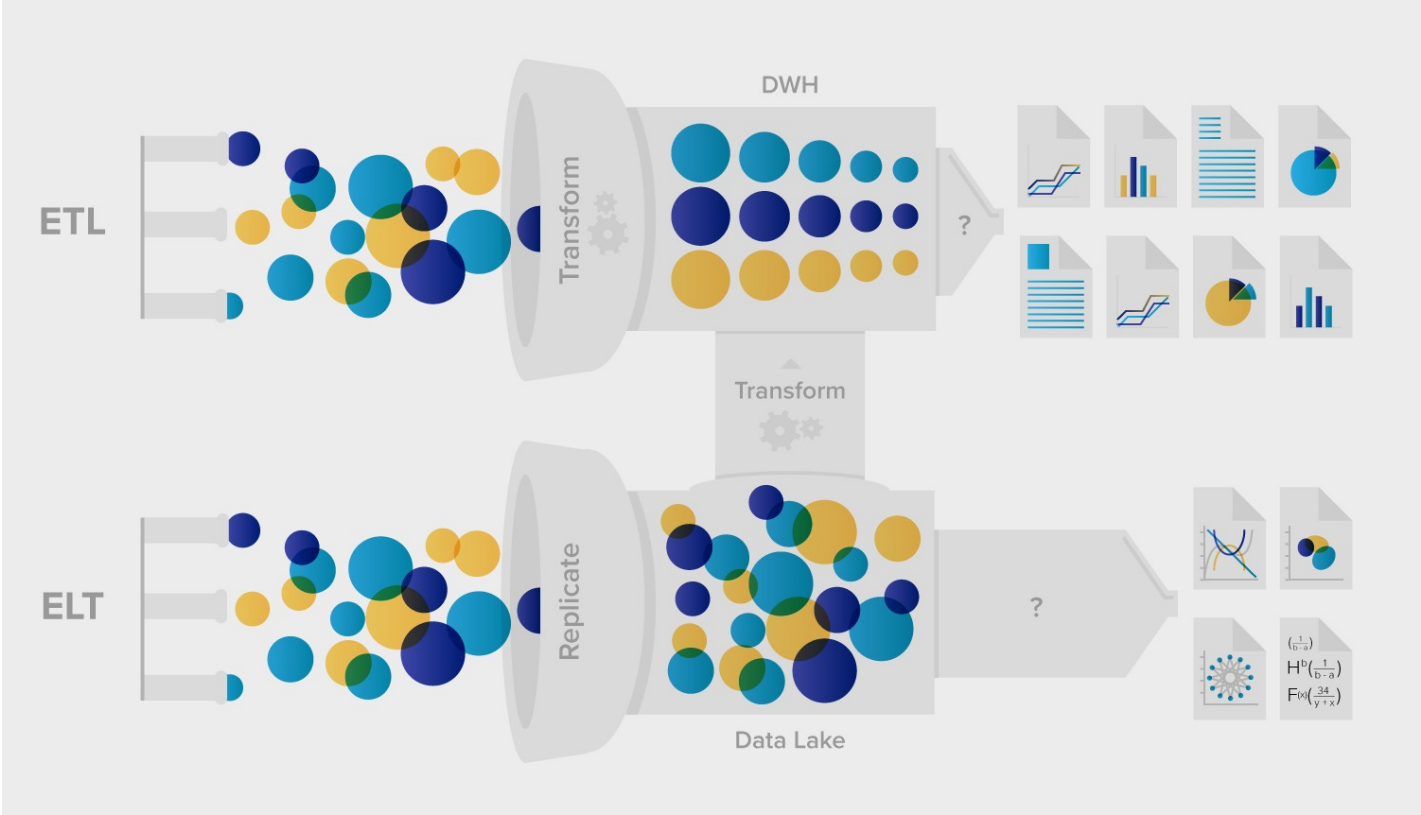
Bilhões de dados e eventos precisam ser coletados, processados e analisados pelas empresas.

Os dados precisam estar limpos, gerenciáveis e prontos para análise.

Precisam ser enriquecidos, moldados e transformados para torná-los significativos.

DATA AS A SERVICE

ETL & ELT



As diferenças tecnológicas: vamos primeiro alinhar os 3 estágios – E, T, L:

Extração : recuperar dados brutos de um pool de dados não estruturado e migrá-los para um repositório de dados temporário.

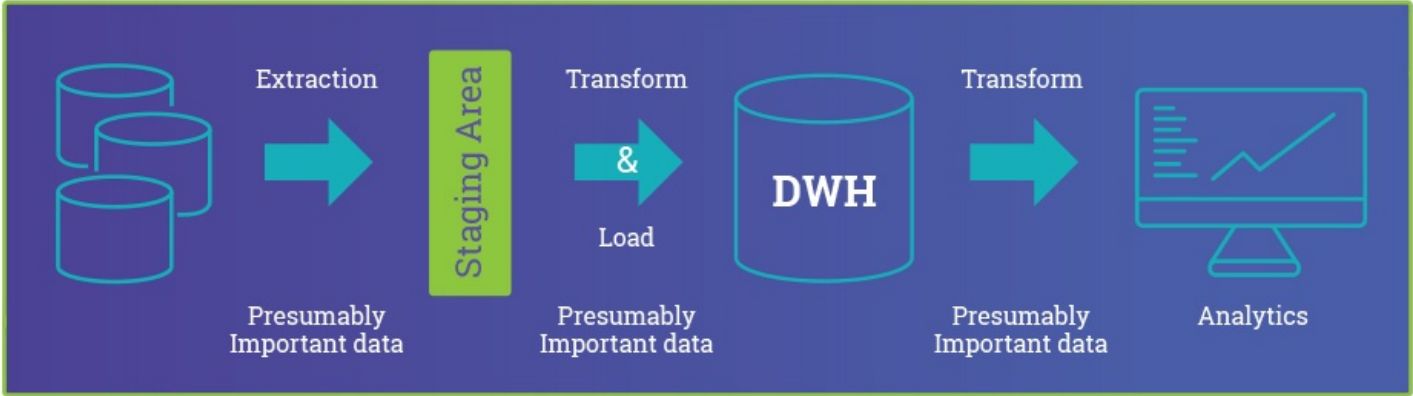
Transformação : Estruturar, enriquecer e converter os dados brutos para corresponder à fonte de destino.

Carregando : Carregando os dados estruturados em um data warehouse para serem analisados e usados por ferramentas de business intelligence (BI).

DATA AS A SERVICE

ETL & ELT

ETL



O ETL requer gerenciamento dos dados brutos, incluindo a extração das informações necessárias e a execução das transformações corretas para atender às necessidades de negócios.

Cada estágio – extração, transformação e carregamento – requer interação de engenheiros e desenvolvedores de dados e lidar com as limitações de capacidade de data warehouses tradicionais.

Usando ETL, analistas e outros **usuários de BI se acostumaram a esperar**, já que o simples acesso às informações não está disponível até que todo o processo de ETL seja concluído.

DATA AS A SERVICE

ETL & ELT

ELT



Na abordagem ELT, depois de extrair seus dados, você inicia imediatamente a fase de carregamento – movendo todas as fontes de dados em um único repositório de dados centralizado.

Com as tecnologias de infraestrutura atuais que usam a nuvem, os sistemas agora podem oferecer suporte a grande armazenamento e computação escalonável.

Portanto, um grande conjunto de dados em expansão e processamento rápido é virtualmente infinito para manter todos os dados brutos extraídos.

Dessa forma, a abordagem ELT oferece uma alternativa moderna ao ETL .

DATA AS A SERVICE

ETL & ELT

ELT

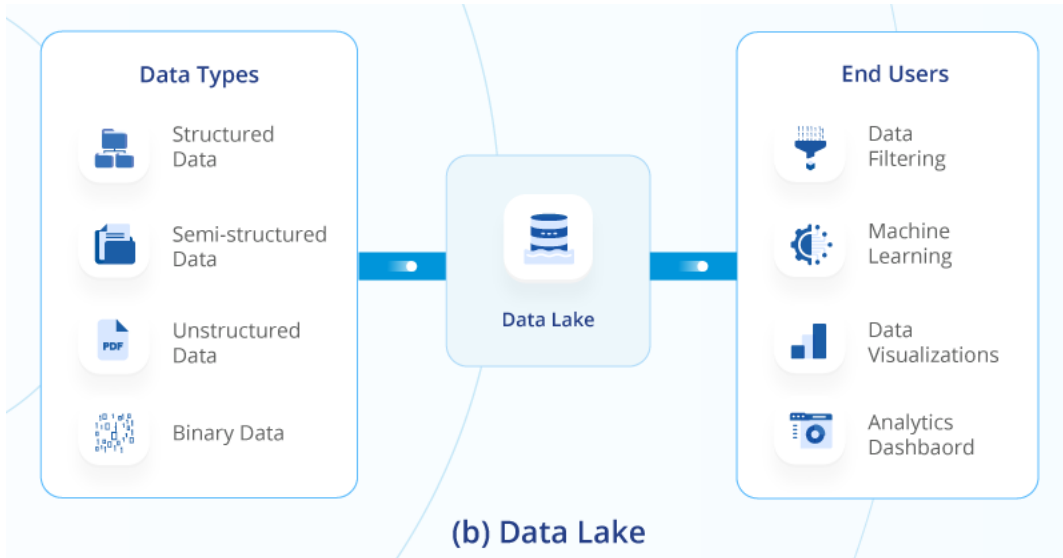
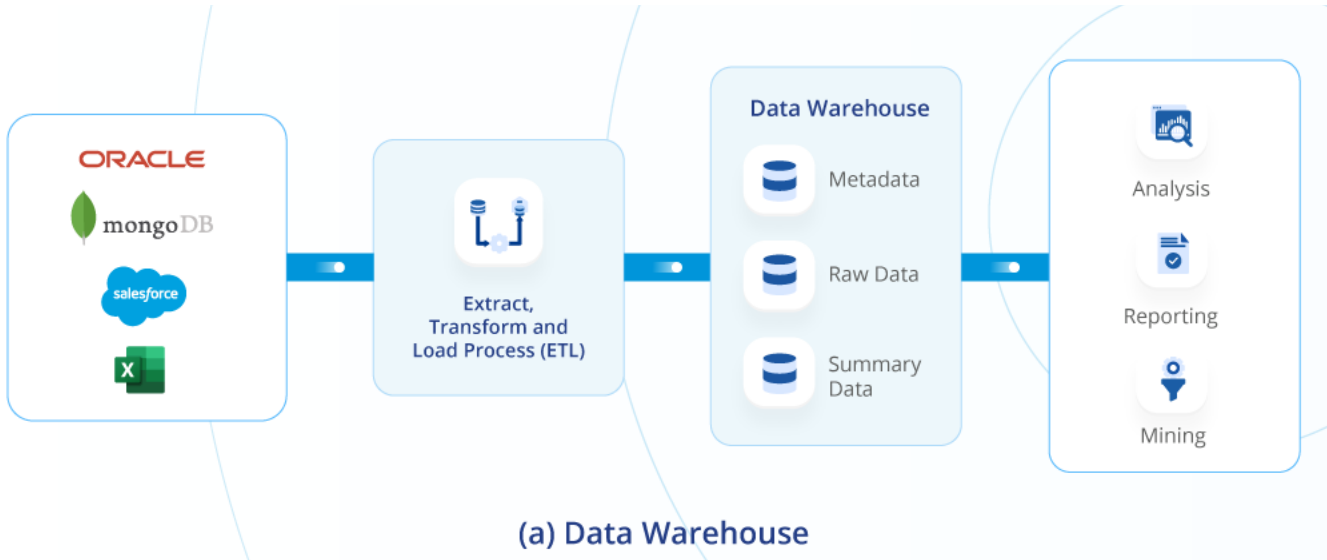


No entanto, ainda está evoluindo. Portanto, as estruturas e ferramentas de suporte ao processo ELT nem sempre são totalmente desenvolvidas para facilitar o carregamento e o processamento de grande quantidade de dados.

O lado positivo é muito promissor: Permitindo acesso ilimitado a todos os seus dados a qualquer momento e economizando esforços e tempo dos desenvolvedores para analistas e usuários de BI .

DATA AS A SERVICE

CENÁRIOS DE USO



Caso de Uso: Personalização de Campanhas Publicitárias em uma Empresa de Marketing:

Armazenamento de Dados

- **Data Lake e Data Warehouse:** O DaaS armazena dados brutos em um Data Lake para permitir análise flexível e em larga escala. Dados estruturados e preparados para análise são armazenados em um Data Warehouse.
- **Tecnologias:** Utilização de serviços como Amazon S3 (para Data Lake) e Amazon Redshift ou Google BigQuery (para Data Warehouse).

DATA AS A SERVICE

DATA LAKE X DATA WAREHOUSE

DATA LAKE

Dados



Desestruturado

Usuários



Cientista de dados
Engenharia de dados

Utilidade



Processamento de streaming
Aprendizado de Máquina
Análise em tempo real

Dados Brutos

Os Data Lakes contêm dados não estruturados, semi estruturados e estruturados com processamento mínimo. Ele pode ser usado para conter dados não convencionais, como logs e dados de sensores

Grandes

Os Data Lakes contêm vastas quantidades de dados na ordem de petabytes. Como os dados podem estar em qualquer forma ou tamanho, grandes quantidades de dados não estruturados podem ser armazenadas indefinidamente e podem ser transformadas quando em uso somente

Indefinidos

Os dados nos Data Lakes podem ser usados para uma ampla variedade de aplicações, como aprendizado de máquina, análise de fluxo e AI

DATA WAREHOUSE

Dados



Estruturado

Usuários



Analista de dados
Analista de Negócios
Business Intelligence

Utilidade



Processamento por lotes
BI e relatórios

Dados Refinados

Os Data Warehouses contêm dados altamente estruturados que são limpos, pré-processados e refinados. Este dados é armazenado para usos muito específicos, como BI.

Menores

Os Data Warehouses contêm menos dados na ordem de terabytes. Para manter a limpeza dos dados e a saúde do armazém, os dados devem ser processados antes da ingestão e é necessário uma purga periódica de dados

Relacional

Os Data Warehouses contêm dados históricos e relacionais, como sistemas de transações, operações etc.





Caso de Uso: Personalização de Campanhas Publicitárias em uma Empresa de Marketing:

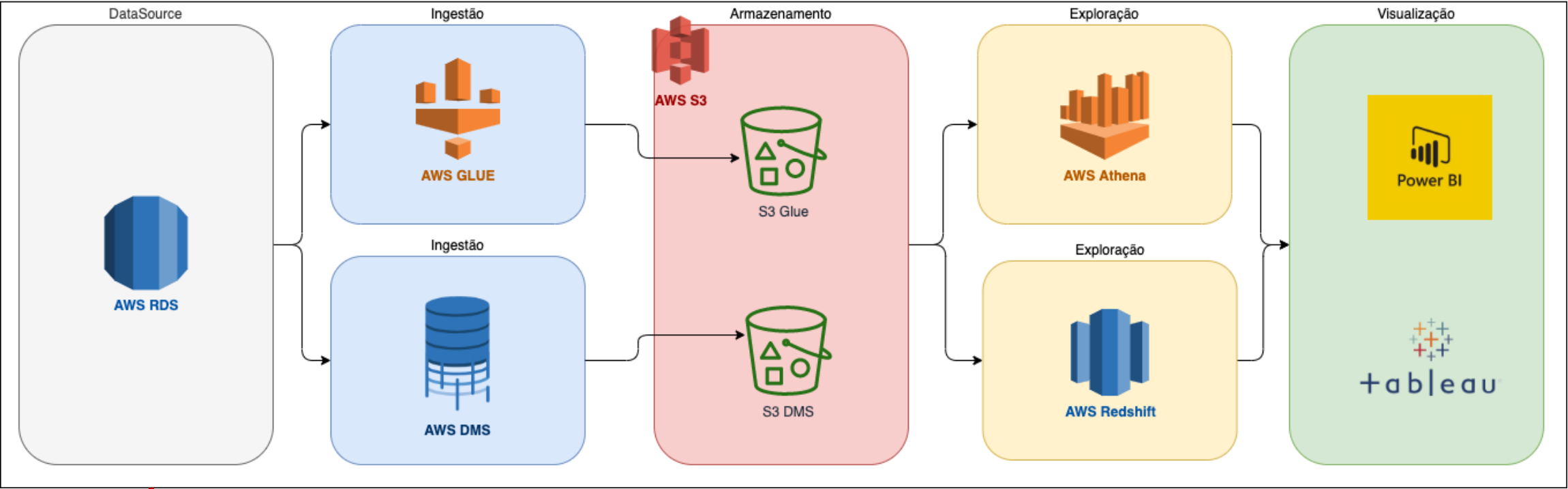
Processamento e Análise de Dados

- **Ferramentas de Análise:** Ferramentas de Business Intelligence (BI) como Tableau, Power BI, e ferramentas de machine learning como TensorFlow ou Scikit-learn são integradas ao DaaS.
- **Pipeline de Dados:** Uso de Apache Spark para processamento distribuído de grandes volumes de dados. Pipelines de dados são configurados para processar dados em tempo real e batch.

DATA AS A SERVICE

PIPELINE DE DADOS

ARQUITETURA DE DADOS AWS

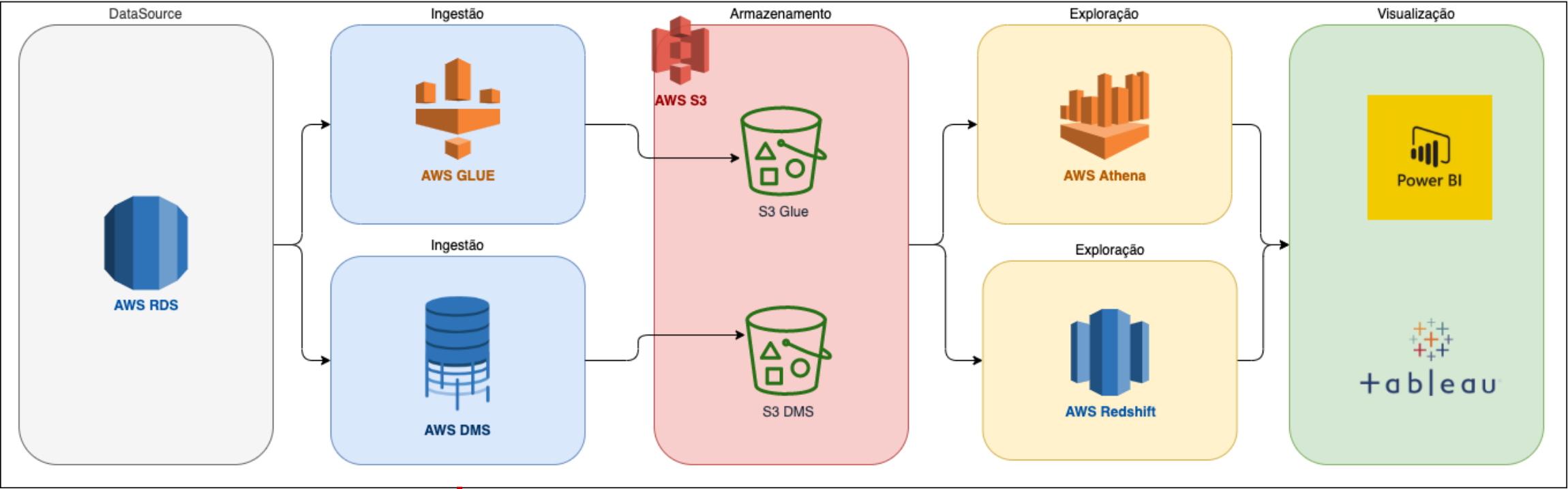


Amazon RDS (Relational Database Service) é usado como a fonte de dados onde os dados transacionais são armazenados. Pode ser um banco de dados como MySQL, PostgreSQL, SQL Server, etc.

DATA AS A SERVICE

PIPELINE DE DADOS

ARQUITETURA DE DADOS AWS



AWS Glue:

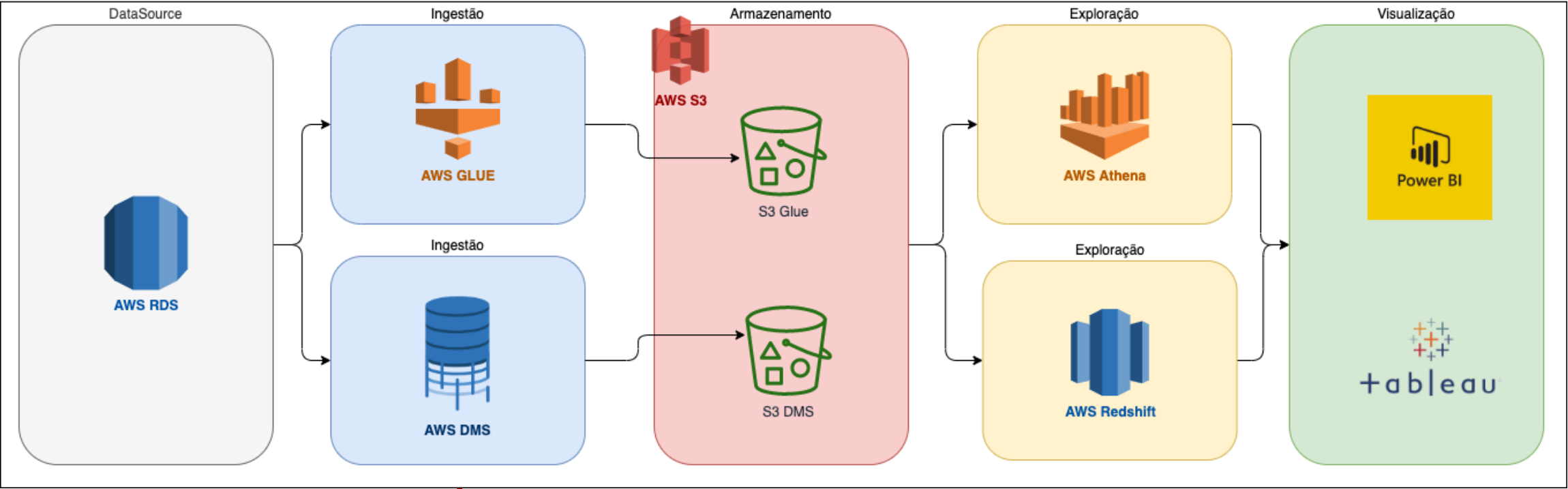
Crawlers: AWS Glue crawlers são configurados para catalogar os dados em Amazon RDS. Crawlers escaneiam os dados e criam metadados no AWS Glue Data Catalog, facilitando a descoberta de dados.

ETL Jobs: Jobs ETL (Extract, Transform, Load) em AWS Glue são criados para extrair dados do RDS, transformar conforme necessário (limpeza, normalização, agregação), e carregar os dados transformados para Amazon S3.

DATA AS A SERVICE

PIPELINE DE DADOS

ARQUITETURA DE DADOS AWS



AWS DMS (Database Migration Service):

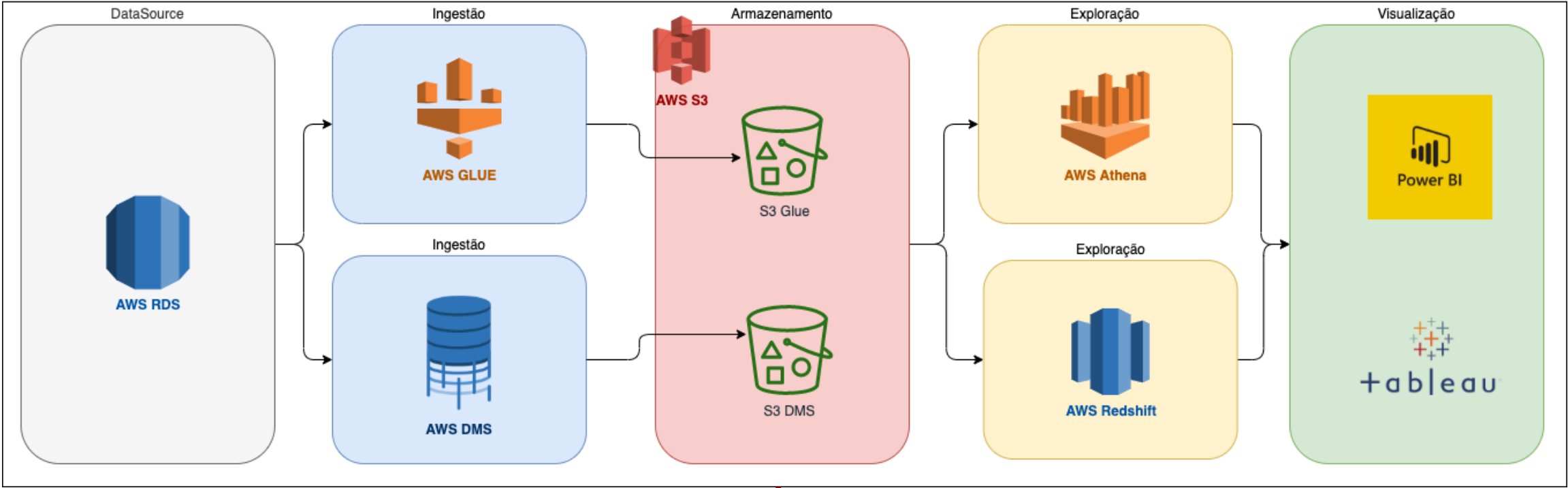
Replicação de Dados: AWS DMS é configurado para replicar dados de Amazon RDS para Amazon S3. Isso pode ser feito através de uma carga completa inicial, seguida de replicação contínua (CDC - Change Data Capture) para capturar mudanças em tempo real.

Transformação de Dados: Embora o DMS seja focado na replicação de dados, ele suporta transformações básicas durante a migração, como alteração de tipos de dados e filtragem de colunas.

DATA AS A SERVICE

PIPELINE DE DADOS

ARQUITETURA DE DADOS AWS

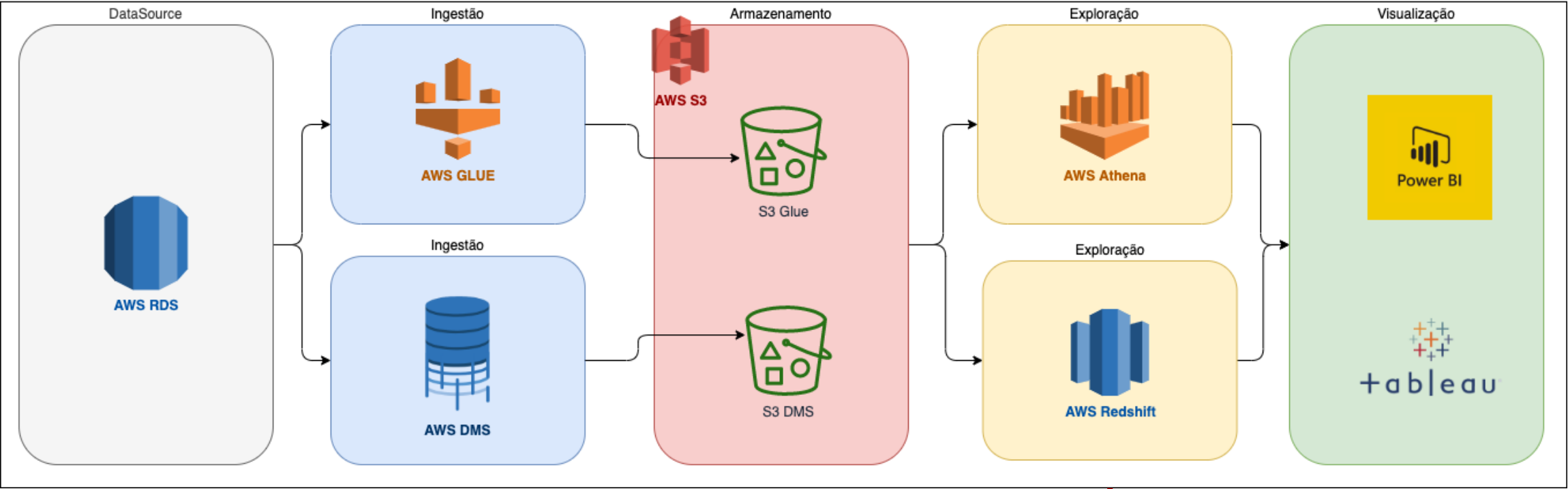


Amazon S3 (Simple Storage Service) é usado para armazenar os dados ingeridos. Os dados são salvos em formatos otimizados para análise, como Parquet ou ORC, organizados em buckets e pastas específicas.

DATA AS A SERVICE

PIPELINE DE DADOS

ARQUITETURA DE DADOS AWS

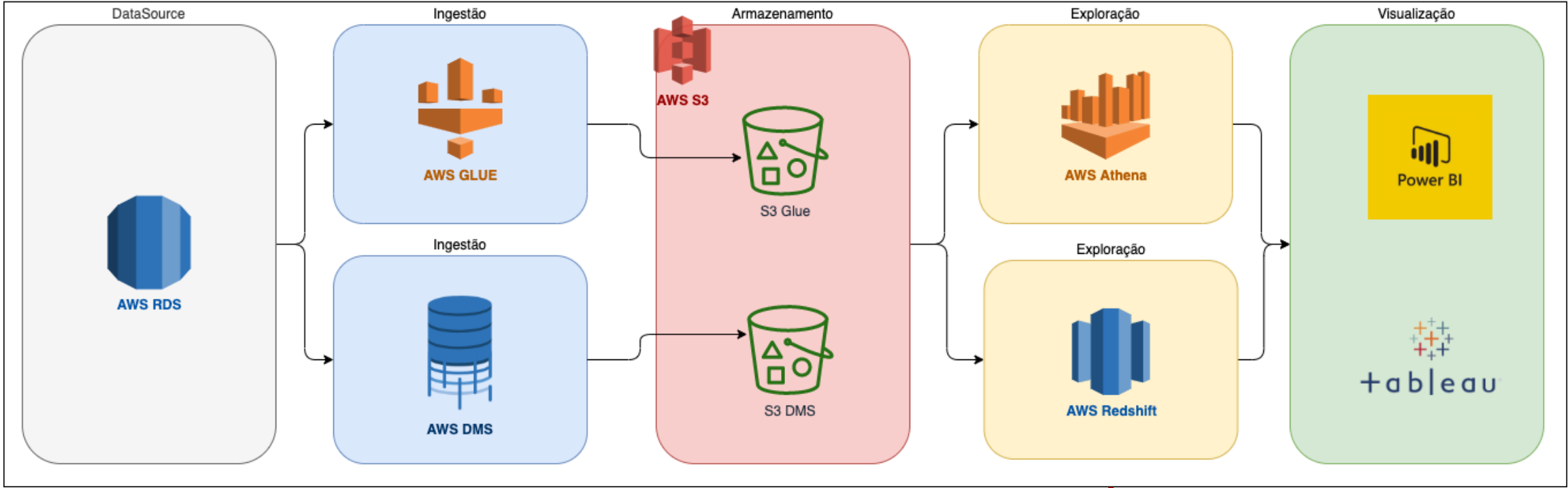


AWS Athena:
Consulta de Dados: AWS Athena permite consultar os dados armazenados no S3 diretamente usando SQL. Como Athena é um serviço de consulta sem servidor, você paga apenas pelas consultas que executa.
Integração com Glue Catalog: Athena usa o Glue Data Catalog para armazenar os metadados dos dados no S3, facilitando a descoberta e consulta dos dados.

DATA AS A SERVICE

PIPELINE DE DADOS

ARQUITETURA DE DADOS AWS



Amazon Redshift:

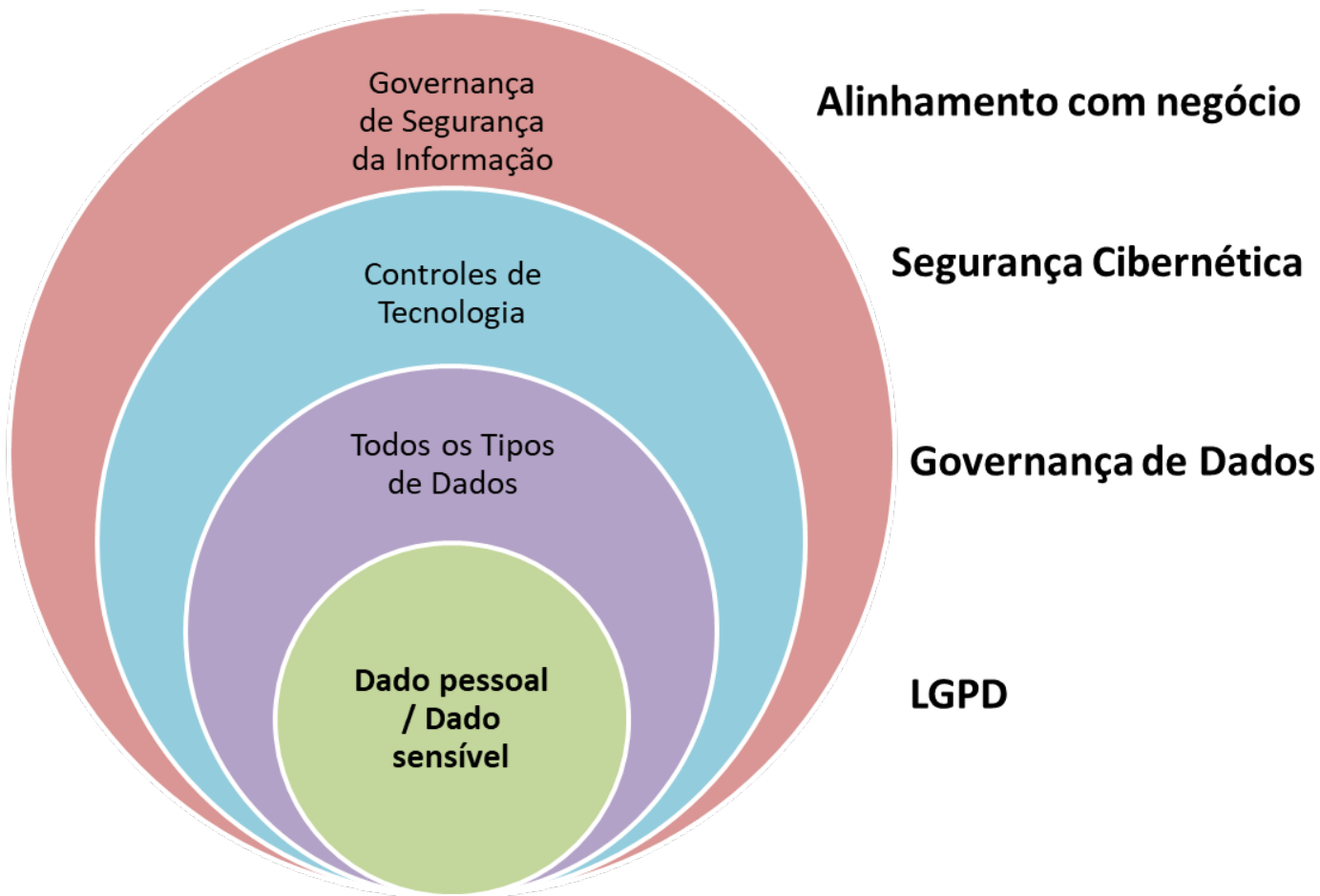
Armazenamento e Análise: Amazon Redshift é um data warehouse totalmente gerenciado para armazenar e analisar grandes volumes de dados.

Redshift Spectrum: Redshift Spectrum permite a consulta direta dos dados no S3 sem mover os dados para o cluster Redshift, usando SQL.

Carga de Dados: Dados podem ser carregados do S3 para Redshift usando comandos COPY ou usando AWS Glue para ETL.

DATA AS A SERVICE

CENÁRIOS DE USO



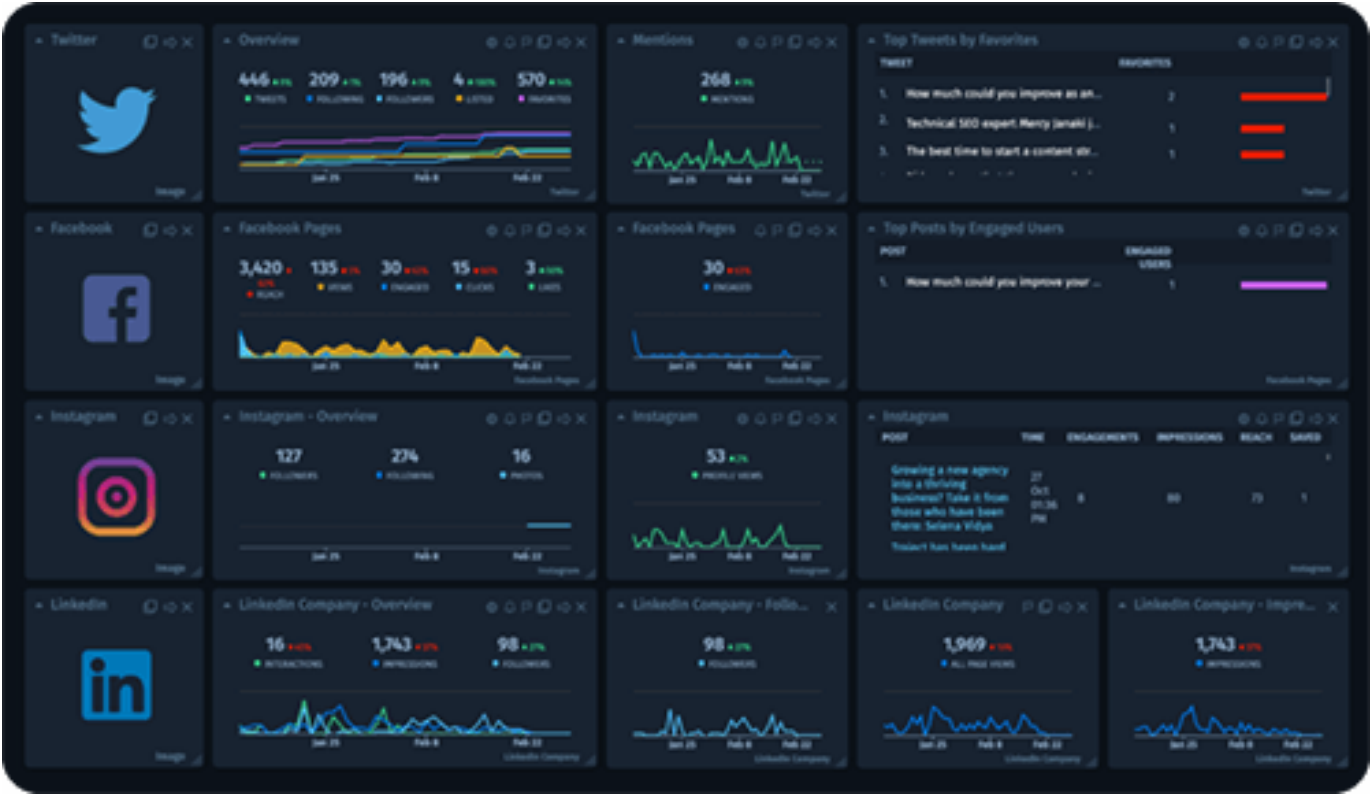
Caso de Uso: Personalização de Campanhas Publicitárias em uma Empresa de Marketing:

Segurança e Governança de Dados

- **Controle de Acesso:** Implementação de políticas de controle de acesso baseadas em roles (RBAC) para garantir que apenas usuários autorizados possam acessar dados sensíveis.
- **Compliance:** Conformidade com regulamentações como GDPR e LGPD para proteger dados pessoais dos clientes.

DATA AS A SERVICE

CENÁRIOS DE USO



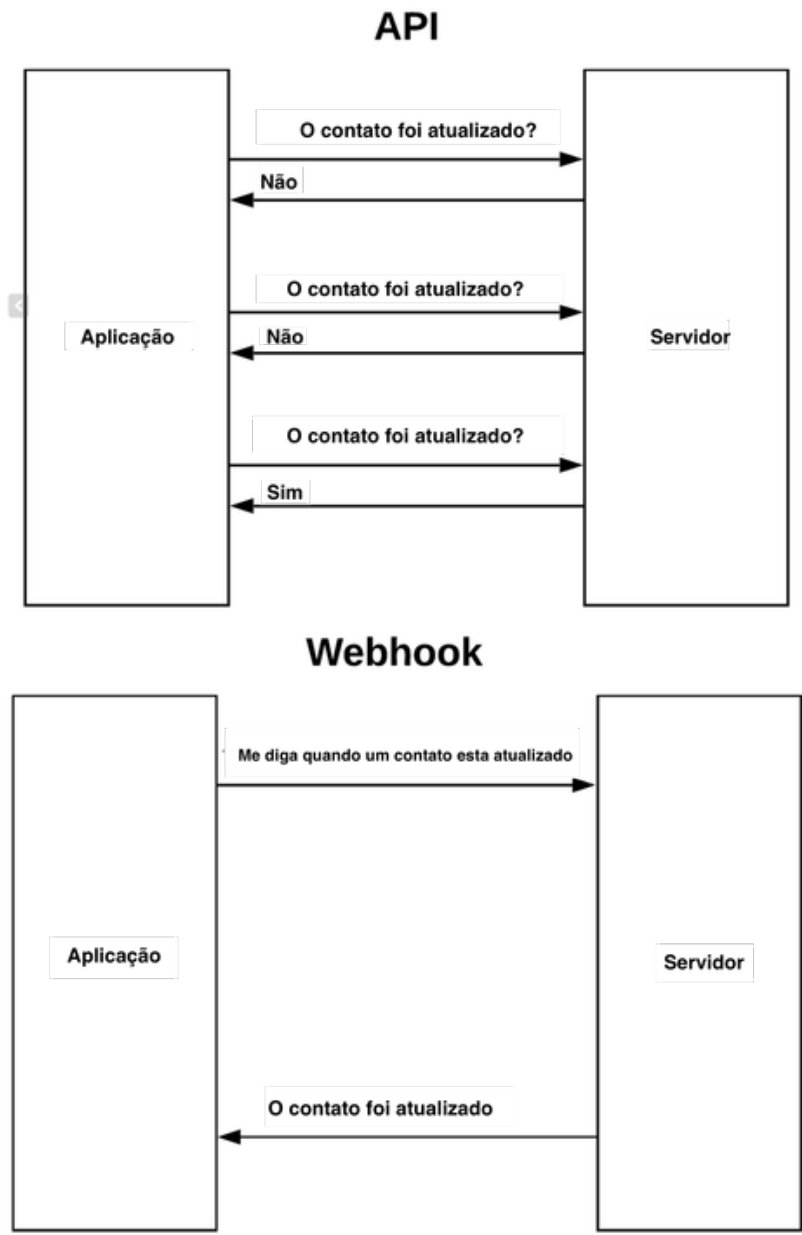
Caso de Uso: Personalização de Campanhas Publicitárias em uma Empresa de Marketing:

Entrega de Dados e Integração com Ferramentas de Marketing

- **APIs e Webhooks:** APIs RESTful são utilizadas para entregar dados em tempo real para plataformas de marketing e automação de campanhas (como HubSpot, Marketo).
- **Painéis de Controle:** Dashboards interativos são criados utilizando ferramentas BI para permitir que os gerentes de marketing visualizem insights e métricas de desempenho das campanhas em tempo real.

DATA AS A SERVICE

WEBHOOKS



O termo webhook foi criado em 2007, pelo desenvolvedor Jeff Lindsay, para definir o tratamento de callback, ou o retorno de uma requisição HTTP iniciada por um evento.

Na prática, quando um evento acontece em um sistema, ele envia uma notificação para outro sistema, que será o receptor desse evento. Portanto, esse é um recurso utilizado para permitir a troca de dados entre duas aplicações.

Em engenharia de dados, um webhook pode ser utilizado para diversas finalidades, especialmente para automatizar processos e integrar sistemas.

Webhooks podem ser usados para interagir com APIs externas e serviços de terceiros. Por exemplo:

- Enviar dados processados para sistemas externos, como plataformas de análise ou serviços de nuvem.
- Atualizar registros em sistemas externos com base em mudanças nos dados internos.

ENGENHARIA DE DADOS

TÓPICOS ABORDADOS

Introdução à Computação em Nuvem:

- Definição de computação em nuvem.
- Benefícios da computação em nuvem, como escalabilidade, flexibilidade e redução de custos.
- Modelos de serviço em nuvem: IaaS (Infrastructure as a Service), PaaS (Platform as a Service) e SaaS (Software as a Service).

O que é DaaS:

- Definição de DaaS.
- Diferença entre dados tradicionais e DaaS.
- Benefícios do DaaS para organizações, como acesso fácil a conjuntos de dados, eliminação de custos de infraestrutura e capacidade de escalabilidade.

Arquitetura do DaaS:

- Componentes principais do DaaS, incluindo armazenamento de dados, ferramentas de processamento de dados e serviços de entrega de dados.
- Descrição do processo de provisionamento e gerenciamento de serviços de dados.

Implementação do DaaS:

- Opções de implementação do DaaS, incluindo provedores de serviços em nuvem e soluções de software específicas para DaaS.
- Considerações ao migrar para o DaaS, como requisitos de segurança, conformidade regulatória e custos.

Cenários de Uso do DaaS:

- Exemplos de casos de uso do DaaS em diferentes setores, como análise de dados, ciência de dados, aprendizado de máquina, IoT (Internet das Coisas) e análise preditiva.
- Demonstração de como os usuários podem interagir com os dados fornecidos pelo DaaS e incorporá-los em suas próprias aplicações e processos de negócios.

DATA AS A SERVICE
CULTURA DATA-DRIVEN



A era da informação evoluiu para um paradigma onde dados não são apenas recursos valiosos, mas também o epicentro das decisões estratégicas em empresas.

O cenário empresarial global está sendo profundamente transformado pela ascensão do paradigma data-driven.

Nesta nova era, dados não são apenas um recurso, mas o alicerce fundamental sobre o qual as estratégias de negócios são construídas e executadas.

Para empresas de todos os setores, tornar-se data-driven significa evoluir para um modelo onde decisões são embasadas em análises de dados profundas e insights que levam a tomada de decisões, uma mudança que não é apenas benéfica, mas essencial para a sobrevivência e o sucesso no mercado competitivo atual.

DATA AS A SERVICE

CULTURA DATA-DRIVEN



Entre os principais benefícios de se tornar uma empresa data-driven está a capacidade de tomar decisões baseadas em evidências.

Esta abordagem minimiza riscos e aumenta a precisão, permitindo que as empresas respondam de forma mais rápida e eficaz às mudanças do mercado.

Além disso, a análise de dados possibilita uma personalização sem precedentes de produtos e serviços, satisfazendo as necessidades dos clientes de forma mais direta e eficiente.

A otimização de processos é outra grande vantagem, onde a identificação de ineficiências operacionais através de análises de dados pode levar a economias significativas e aumento da produtividade. Integrar e harmonizar dados de diversas fontes em um formato coeso e acessível é uma tarefa complexa, que requer habilidades técnicas avançadas.



BENEFÍCIOS

Tomada de Decisão Baseada em Evidências: Com uma abordagem data-driven, as decisões são embasadas em análises de dados concretos, reduzindo a incerteza e aumentando a precisão.

Personalização de Serviços e Produtos: A análise de dados permite entender melhor os clientes, adaptando produtos e serviços para atender às suas necessidades específicas.

Otimização de Processos: A análise de dados ajuda a identificar gargalos operacionais, aprimorando a eficiência e reduzindo custos.

DESAFIOS

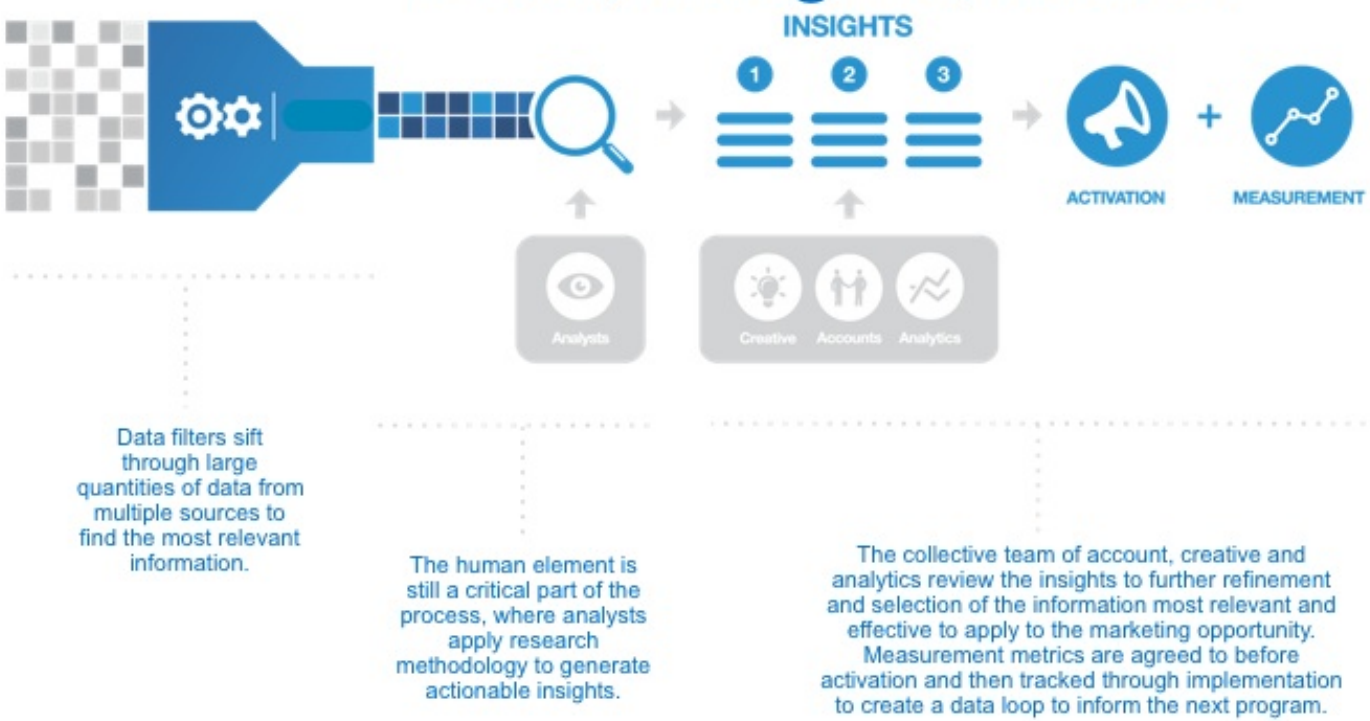
Integração de Dados: A consolidação de dados de diversas fontes em um formato coeso e útil é um desafio técnico significativo.

Privacidade e Segurança de Dados: Manter a segurança dos dados coletados e respeitar a privacidade dos indivíduos é essencial e desafiador.

Resistência à Mudança: A transformação cultural para uma abordagem data-driven pode enfrentar resistência interna, especialmente em organizações tradicionais.

DATA AS A SERVICE
CULTURA DATA-DRIVEN

Data, Insights, Action



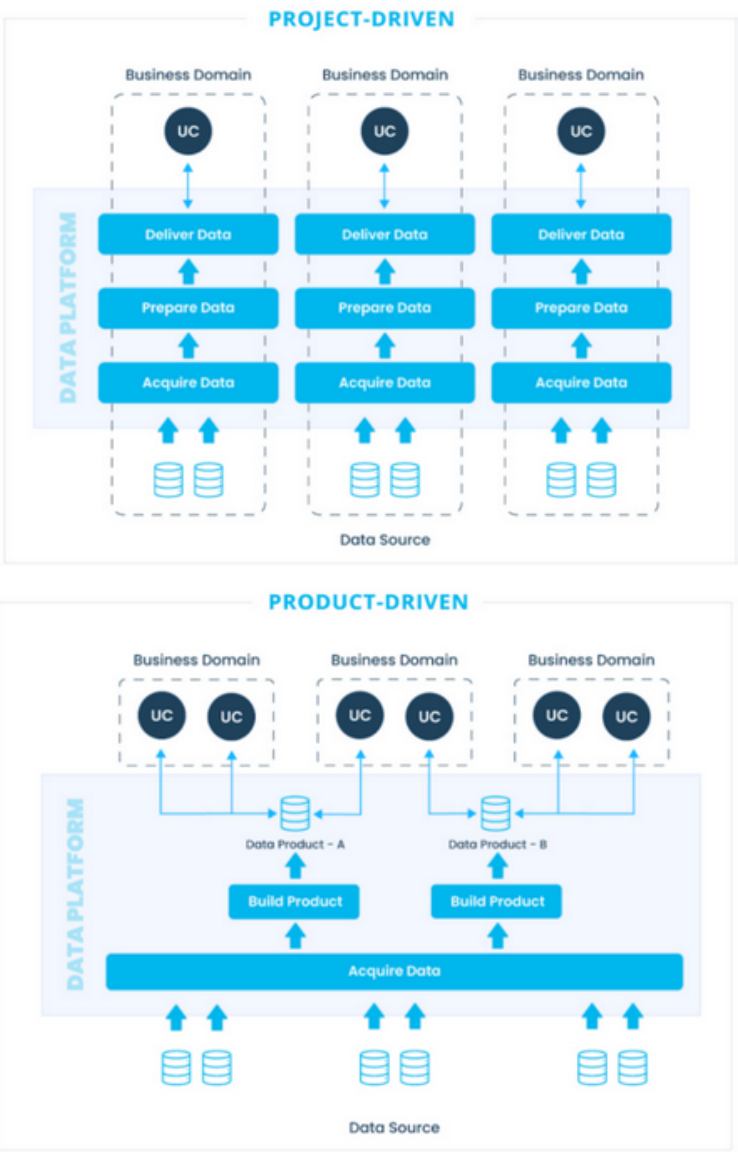
Empresas data-driven têm uma vantagem significativa no mercado.

Elas são capazes de antecipar tendências, responder rapidamente a mudanças no mercado e inovar com base em insights extraídos de dados. Essa agilidade e capacidade de inovação são essenciais para manter a competitividade em um ambiente de negócios cada vez mais volátil.

Esta flexibilidade e capacidade de inovação contínua são fundamentais para manter uma posição de liderança em um ambiente de negócios em constante evolução.

DATA AS A SERVICE

CULTURA DATA-DRIVEN



Comparativo entre uma gestão por projeto e uma gestão por produto

DaaS vem sendo considerado como um caminho para remodelar o mundo dos negócios por meio da inteligência competitiva. A importância crescente dos dados e análises está impulsionando a importância dos dados como um serviço.

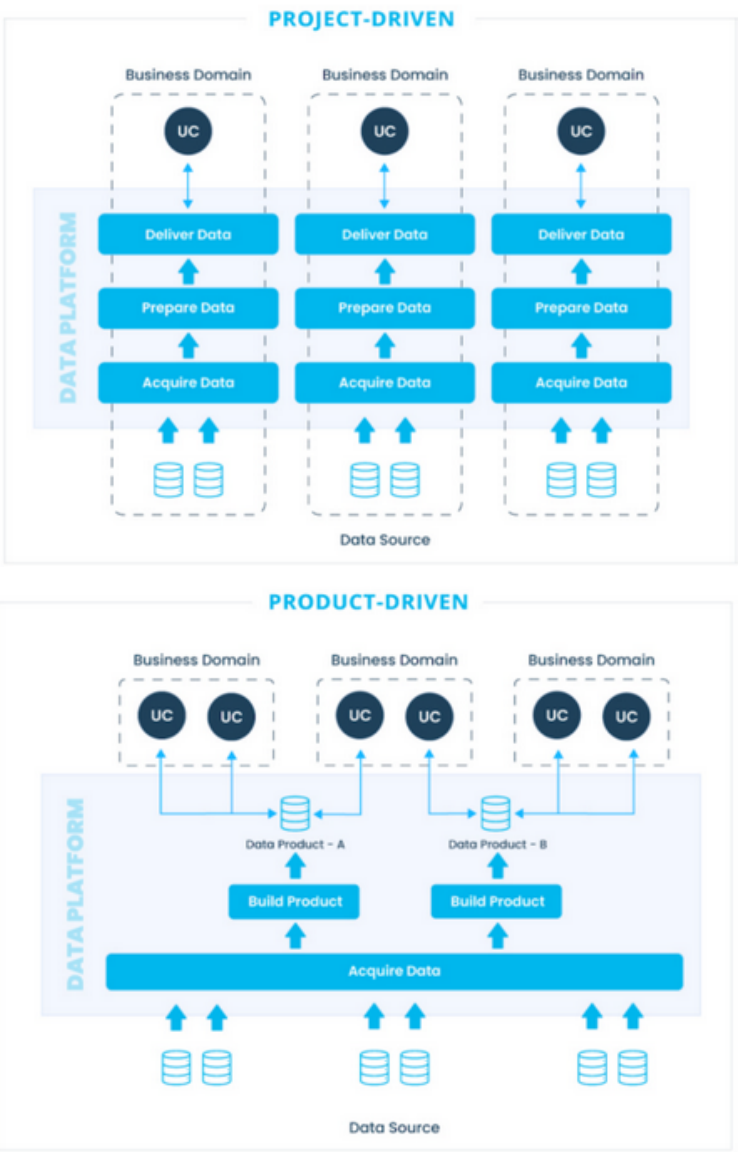
Os serviços DaaS permitem que as empresas acessem facilmente dados externos, bem como tornam mais fácil para as empresas democratizar a análise e capacitar seus usuários de negócios.

Além disso é um grande facilitador na implantação de cultura Data-Driven.

Porém quanto mais você democratizar seus dados e análises, mais precisará ajustar sua abordagem às necessidades dos especialistas de domínio.

DATA AS A SERVICE

CULTURA DATA-DRIVEN



Comparativo entre uma gestão por projeto e uma gestão por produto

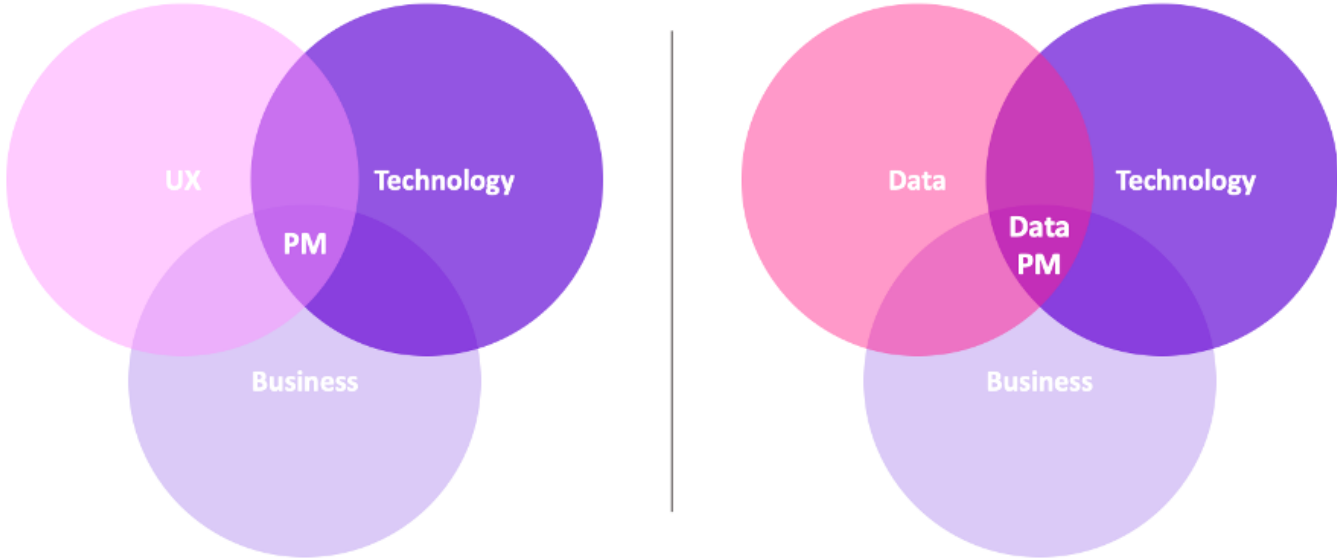
Os produtos de dados entrarão em jogo: de conjuntos de dados confiáveis a aplicações, modelos preditivos e painéis analíticos, os produtos de dados são configurados para se multiplicar e atender às necessidades do maior número possível de pessoas, sem exigir conhecimento avançado ou experiência em dados e IA.

Ao contrário dos projetos com início e fim definidos, os produtos de dados são cíclicos.

Como tal, “produto de dados” abrange vários elementos, como propriedade, lançamento, manutenção e princípios de melhoria contínua. Trabalhar data products na mesma plataforma dos próprios dados já significa manter os mesmos princípios e regras de transparência, reforçando assim a confiança.

DATA AS A SERVICE

CULTURA DATA-DRIVEN



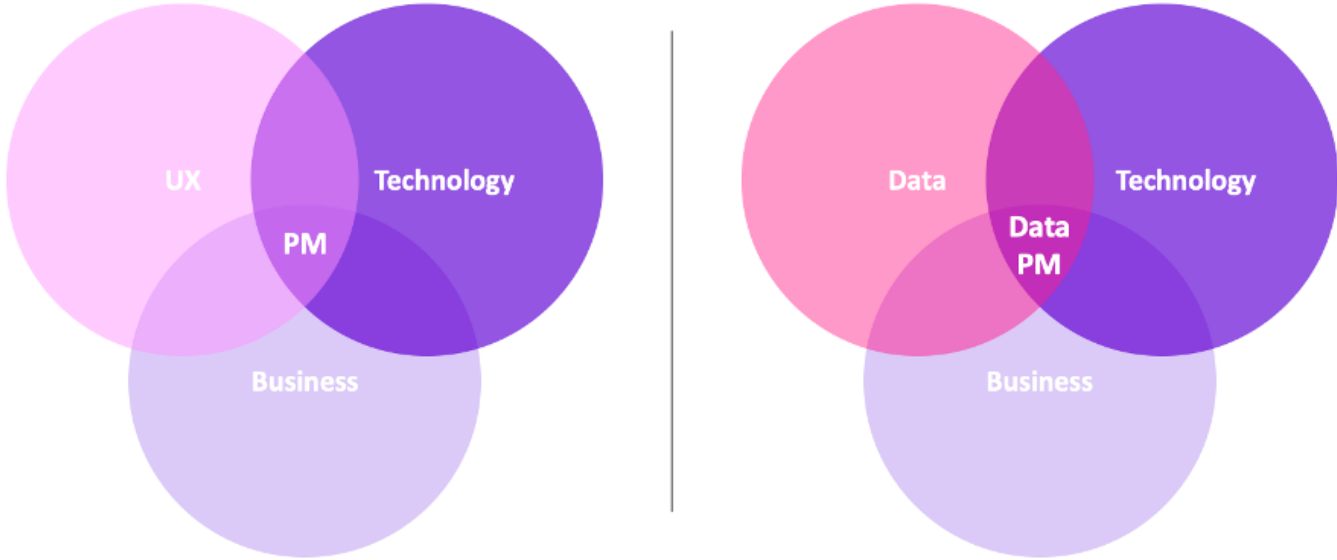
Para quem se interessa pelo universo de produtos digitais, novas possibilidades de atuação têm ganhado espaço no mercado.

Entre elas, a função de **Data Product Manager (DPM)** certamente chama atenção. Por mais que PMs sejam comuns em times de produto, a adição do termo "data" traz confusão para algumas pessoas.

Uma pessoa Data Product Manager pode ter backgrounds distintos, especializações únicas, mas sua importância em times de trabalho é construída por algumas habilidades.

Essas soft skills essenciais vão garantir que quem ocupe essa posição seja realmente capaz de entregar os resultados necessários e conectar pessoas.

DATA AS A SERVICE
CULTURA DATA-DRIVEN



Quando falamos em executar a estratégia de produtos de dados, podemos pensar em inúmeras maneiras de fazer isso, mas é necessário que faça sentido para a equipe.

Sendo assim, você deve ter em mente que o objetivo final destas ações é:

Um maior senso de controle dos dados, responsabilidade pelos dados e democratização dos dados em toda a organização.

E para que isso aconteça é importante implementar alguns tópicos, como:

DATA AS A SERVICE

CULTURA DATA-DRIVEN



Business data owners



Data engineering team



Data consumers

Orientado à demanda: os produtos de dados devem ter uma função clara, isso permitirá uma maior contribuição e colaboração entre os produtores e consumidores;

Reutilizável e escalável: os produtos de dados devem ser projetados para ter fácil utilização, fazendo com que eles possam ser reutilizados em quaisquer outras situações. Isso faz com que a democratização de dados seja difundida;

Descobrável e acessível: os produtos de dados devem ser simples, pois a ideia é que as informações sejam encontradas facilmente, garantindo assim maior fluidez na tomada de decisão;

Sentimento de dono: como o próprio tópico já diz, o dono destes dados deve estar comprometido em gerenciar estes produtos adequadamente durante todo o ciclo de vida daquela informação.