

Catálogo de dados inteligentes para armazém de dados, data lakes e lakehouses na nuvem

Principais benefícios

- Maximize o valor do armazém de dados na nuvem, data lake, e lakehouse sem interferências
- Acelere até mesmo as mais complexas iniciativas de migração e modernização das nuvens
- Ganhe visibilidade em todos os ativos de dados da sua empresa com gerenciamento de metadados líder da indústria
- Permita que os usuários facilmente encontrem dados relevantes na nuvem

Acelere a descoberta de dados a partir da IA e da integração de dados na nuvem

Atualmente, as empresas estão transferindo rapidamente sua infraestrutura de dados e analítica para a nuvem. Embora este movimento para capitalizar a agilidade, escalabilidade e eficiência de custos da nuvem afete todos os aspectos da análise de dados, é particularmente urgente para armazenamento de dados e data lakes. Durante anos, os armazéns de dados das empresas foram instrumentos estáveis, potencializando os sistemas analíticos e de relatórios das empresas. Agora, está em andamento uma grande mudança para modernizá-los na nuvem, a fim de alcançar melhorias significativas no desempenho e na vantagem competitiva. Ao mesmo tempo, as organizações estão adotando data lakes para acompanhar seus armazéns de dados na nuvem. E, mais recentemente, as organizações estão construindo lakehouses na nuvem, que mesclam o melhor dos armazéns de dados e data lakes para fornecerem possibilidades analíticas que impulsionem, desde dashboards de BI a projetos avançados de IA e aprendizagem de máquinas.

No entanto, a adoção de armazéns de dados, data lakes e lakehouses na nuvem pode apresentar novos desafios. Apesar de inicialmente ser simples suportar um novo armazém de dados ou data lake, maximizar o valor do seu investimento requer estratégia e planejamento.

Quer esteja construindo um novo armazém de dados, data lake ou lakehouse, ou modernizando dados e cargas de trabalho na nuvem ao longo do tempo, é essencial entender e avaliar o panorama atual dos seus dados e certificar-se de que você tenha as ferramentas e as melhores práticas para gerenciar seus dados na nuvem.

O Informatica® Intelligent Cloud Lakehouse Data Management Solution for Cloud Analytics foi desenvolvido sobre uma base de integração de dados, qualidade e governança de dados, e gerenciamento de metadados líderes do setor. O Informatica Enterprise Data Catalog fornece a base de gerenciamento de metadados para que você possa entender o panorama dos seus dados com ativos de descoberta de dados baseados na aprendizagem de máquinas e total visibilidade. O Informatica Intelligent Cloud ServicesSM (IICS) fornece serviços empresariais iPaaS líderes de mercado para integração de dados na nuvem, ingestão em massa na nuvem e qualidade de dados na nuvem.

Com as possibilidades de aprendizagem por máquinas construídas sobre um gerenciamento abrangente de metadados, o Informatica Enterprise Data Catalog fornece uma base comum de metadados empresariais para uma visibilidade total e uma linhagem inteligente, automatizada em todo o seu ambiente. A solução permite estabelecer uma relação entre os armazéns de dados on-premise e o seu moderno ecossistema de nuvens para alcançar o valor máximo sem prejudicar seus consumidores de dados.

Principais recursos

Ampla conectividade de metadados

A solução de catalogação de dados inteligentes da Informatica integra-se com todos os principais ecossistemas em nuvem e modernos armazéns de dados, data lakes e lakehouses, incluindo o Amazon Redshift, Amazon S3 Synapse Analytics, Google BigQuery, Snowflake e Databricks. Também inclui conectividade out-of-the-box de metadados para centenas de aplicações na nuvem, sistemas on-premise, aplicações empresariais e de middleware, armazenamento de dados e ferramentas analíticas de BI.

Descoberta de dados com base na aprendizagem automática

O mecanismo de inteligência artificial baseado em metadados da Informatica, o CLAIRE[®], aplica algoritmos de IA e de aprendizagem mecânica aos dados técnicos, comerciais, operacionais e de uso, compreendendo e analisando automaticamente metadados e arquivos complexos, permitindo descobrir dados independentemente do local onde residem em um ambiente híbrido e multinuvel.

Linhagem automatizada de dados completos

A linhagem fornece exibições resumidas e detalhadas que lhe permitem entender de onde seus dados se originaram, como se movimentaram pela organização e como foram transformados ao longo do caminho. A linhagem de dados de ponta a ponta através de ambientes híbridos e multinuvel permite rastrear dados, desde a origem até o alvo, para que você possa saber como os dados estão sendo utilizados e o que depende deles. O Informatica Enterprise Data Catalog permite que se realize uma análise de impacto detalhada em ativos de dados ascendentes e descendentes. Dessa forma, você pode entender facilmente o impacto da migração através dos ativos de dados, recursos e usuários. Além disso, a análise de impacto permitida pela linhagem de dados detalhada pode ajudar a demonstrar os benefícios de custo da transferência de certos ativos de dados e cargas de trabalho para a nuvem.

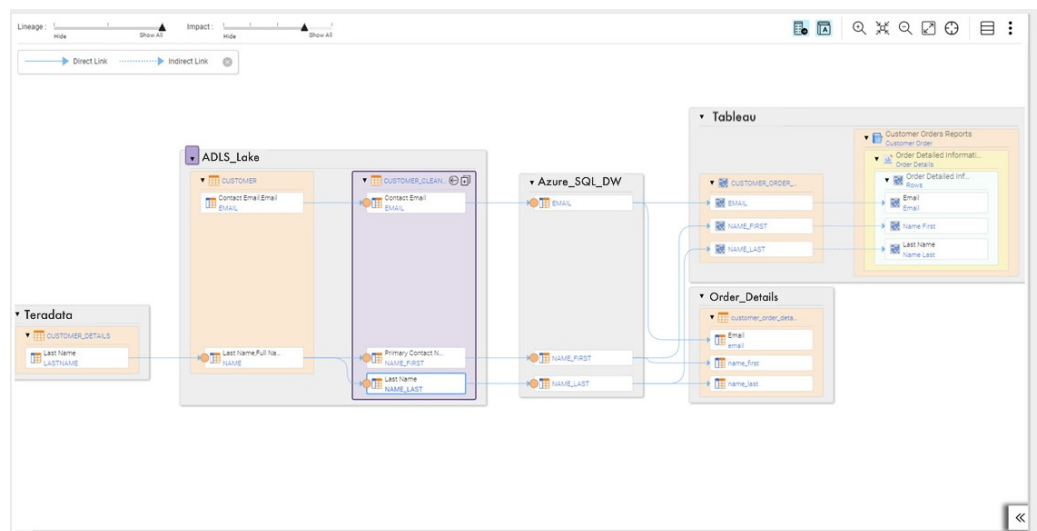


Figura 1. Visualize facilmente a linhagem de dados de ponta a ponta pelos on-premises, armazéns de dados na nuvem, data lakes e lakehouses.

Colaboração e curadoria social

A migração de dados para um armazém de dados, data lake ou lakehouse na nuvem geralmente envolve diversos interessados e usuários. O Informatica Enterprise Data Catalog fornece possibilidades de colaboração e curadoria social que permitem aos proprietários dos dados e aos especialistas da área autenticar os ativos de dados e ajudar a identificar os dados que devem ser migrados. Além disso, os consumidores de dados podem classificar e rever os ativos de dados, acrescentando atributos e anotações personalizadas que facilitam a procura de dados confiáveis durante e após a migração.

Integração com serviços de gerenciamento de dados na nuvem

O Informatica Enterprise Data Catalog, juntamente com o IICS, integram serviços de gerenciamento de dados que fornecem possibilidades essenciais para armazéns de dados, data lakes e lakehouses na nuvem. O IICS Cloud Data Integration fornece ETL, ELT, ingestão de dados, sincronização e replicação de alta performance para ambientes multinuvm. O IICS Cloud Mass Ingestion suporta streaming, arquivo e ingestão de base de dados para os armazéns de dados, data lakes e filas de mensagens. O IICS Cloud Data Quality permite identificar, limpar e padronizar dados para que você possa corrigir e monitorar problemas de qualidade de dados com rapidez. A partir do IICS, você pode utilizar o Informatica Enterprise Data Catalog para descobrir e selecionar dados, ver fontes de dados e mapeamentos, ver regras de qualidade de dados, quadros de resultados, e ganhar visibilidade abrangente na linhagem de dados de ponta a ponta.

The screenshot displays the Informatica Enterprise Data Catalog interface. The search bar at the top contains the text 'web orders'. The left sidebar shows navigation options like 'Home', 'Explore', 'Data Catalog', 'Bundles', 'My Jobs', and 'My Import/Export Logs'. The main area is divided into 'Filters' and 'Search Results (1-25 of 117)'. The 'Filters' section includes categories for Asset Type, Resource Type, Resource Name, and Rows. The 'Search Results' section lists several data assets with their types, sources, and last updated dates. The assets listed are:

Asset Name	Type	Source	Last Updated
web_order_lineitem.json	Delimited File	Amazon S3	2020-03-08T10:56:57.765Z
web_order.dat	Delimited File	Azure Data Lake Store	2019-11-04T06:50:09.470Z
Customer Web Report	Report	Tableau	2019-07-25T09:38:24.447Z
Web Orders Staging	Table	Oracle	2019-07-25T09:38:38.071Z
rpt_party	Table	Hive	2019-07-25T09:38:38.071Z
WEB_ORDERS	Table	Oracle	2019-07-01T11:58:40.170Z
FactCallCenter	AWS	AWS Glue	2019-07-01T11:59:57.599Z

Figura 2. Descubra e selecione os dados utilizando o Informatica Enterprise Data Catalog com o IICS.

Mapeamentos pré-configurados

Modelos de mapeamento pré-configurados, assistentes de tarefas fáceis de usar para replicação e sincronização de dados e ingestão de dados em massa para altas cargas de volume de trabalho permitem que se mova mais dados para a nuvem em menos tempo. A integração perfeita entre o Informatica Enterprise Data Catalog e o IICS permite encontrar os dados que você pretende migrar e movê-los facilmente para o seu armazém de dados, data lake ou lakehouses na nuvem utilizando uma única interface sem código. "Vá às compras" no seu catálogo de dados com o designer visual do IICS Cloud Data Integration, e depois construa cargas de dados simples a complexas em apenas alguns cliques para conectar diversos pontos finais através da nuvem e on-premises, transformar dados e programar cargas de trabalho analíticas.

Principais benefícios

Maximize o valor do armazém de dados, data lake e lakehouse na nuvem sem interferências

Avalie exaustivamente o panorama do seus dados para entender o que deve migrar ou consolidar, e onde irá gerar mais valor a partir de um armazém de dados, data lake ou lakehouse na nuvem. Com a linhagem de dados de ponta a ponta, ganhe visibilidade nas relações e dependências de dados. Capacite os consumidores de dados com acesso de autoatendimento para que encontrem facilmente dados na nuvem para análise. O Informatica Enterprise Data Catalog garante que as estruturas e esquemas do seu armazém de dados on-premises sejam transferidas para a nuvem, evitando assim potenciais interferências durante e após a migração. Após a migração, as exibições de linhagem de ponta a ponta ajudam a validar se os dados foram corretamente mapeados. Além disso, a curadoria e a descoberta de dados automatizados orientados por metadados suportam uma governança de dados escalável e permitem a entrega de dados confiáveis na nuvem.

Acelere até mesmo as mais complexas iniciativas de migração e modernização das nuvens

Diminua significativamente o tempo necessário para a migração e modernização da nuvem. Descubra os dados que pretende migrar ou consolidar, em seguida, construa cargas de integração de dados simples a complexos utilizando um designer de mapeamento visual que conecta diversos pontos finais através da nuvem e on-premises com transformações pré-construídas para acelerar o desenvolvimento.

Ganhe visibilidade em todos os ativos de dados da sua empresa com gerenciamento de metadados líder da indústria

É preciso descobrir e entender seus dados antes de poder obter valor a partir deles. A solução da Informatica impulsiona as capacidades de gerenciamento de metadados líderes da indústria, para que você possa potencializar centenas de conectores para sistemas baseados na nuvem e on-premises. Com a inteligência orientada por metadados, a solução ajuda a identificar todos os locais em todos os sistemas da sua organização onde os dados são armazenados e utilizados, depois é possível mover todos esses dados facilmente para o seu armazém de dados, data lake ou lakehouse na nuvem corporativa. Além disso, ela captura e armazena metadados importantes sobre contingências e relações entre conjuntos e subconjuntos de dados para manter as conexões entre eles durante e após a migração e modernização dos dados na nuvem.

Sobre a Informatica

A transformação digital muda expectativas: melhor serviço, entrega mais rápida e com menos custos.

As empresas devem mudar para continuar competitivas, e a solução está nos dados.

Como líder mundial em gerenciamento de dados corporativos na nuvem, estamos preparados para ajudar você a ocupar de maneira inteligente uma posição de liderança — em qualquer setor, categoria ou nicho. A Informatica oferece a perspectiva para que você se torne mais ágil, aproveite novas oportunidades de crescimento ou invente coisas novas. Estamos 100% focados em todos os tipos de dados para oferecer a versatilidade de que você precisa para prosperar.

Convidamos você a explorar tudo o que a Informatica tem a oferecer — e estimular o poder dos dados para impulsionar sua próxima revolução com inteligência.

Permita que os usuários encontrem facilmente dados relevantes e confiáveis na nuvem

Quando os dados são transferidos durante uma migração, os consumidores de dados podem ter dificuldade em localizar os dados de que dependem para fazer o seu trabalho. Os links favoritos e os relatórios recorrentes podem ser rompidos. E os consumidores de dados podem continuar utilizando os dados de uma origem conhecida on-premise, embora a nuvem forneça uma alternativa melhor. Os analistas e cientistas de dados também precisam estar cientes de onde se encontram os dados, e entender quais os ativos de dados mais adequados às suas necessidades. Ao consultar o Informatica Enterprise Data Catalog, os usuários podem facilmente encontrar os dados de que necessitam sem ter de pedir ajuda à TI, evitando assim congestionamentos e ineficiências.

Próximas etapas

Saiba mais lendo a [Folha de dados do Informatica Enterprise Data Catalog](#) e o resumo da solução [Modernize seu armazém de dados e data lake na nuvem](#).



Sede América Latina Av. das Nações Unidas, 12.901 - 3º andar - Torre Norte - Brooklin Novo - 04578-000 - São Paulo, SP Tel.: 0800-878-3060
Ligação gratuita nos EUA: 1.800.653.3871

IN17_0620_03625

© Copyright Informatica LLC 2020. Informatica, o logotipo da Informatica, Informatica Intelligent Cloud Services e CLAIRE são marcas comerciais ou marcas registradas da Informatica LLC nos Estados Unidos e outros países. Uma lista atualizada de marcas comerciais da Informatica está disponível na Web em <https://www.informatica.com/br/trademarks.html>. Outros nomes de empresas e produtos podem ser nomes comerciais ou marcas comerciais de seus respectivos proprietários. As informações contidas nesta documentação estão sujeitas a alterações sem notificação prévia e foram fornecidas "COMO SE ENCONTRAM", sem garantias de qualquer espécie, expressas ou implícitas.