

Modernize seu armazém de dados e data lake na nuvem

Benefícios principais

- Aumente a flexibilidade com o gerenciamento de dados em nuvem baseado em metadados, que fornece integração e qualidade dos dados com governança
- Reduza custos com recursos de IA que eliminam a necessidade de codificação manual
- Melhore a agilidade com integração sem código e conectividade pronta para uso para centenas de aplicativos
- Impulsione a produtividade em todo o ciclo de desenvolvimento de software com automação e inteligência
- Habilite o acesso a dados de autoatendimento e a análise usando recursos de pesquisa em idioma natural

Concretize as possibilidades de suas iniciativas de análise com o gerenciamento inteligente e automatizado de dados de cloud lakehouse

Hoje, as organizações se tornaram mais orientadas a dados à medida que embarcam em transformações digitais para acelerar a obtenção de valor, reduzir custos, melhorar a eficiência e fornecer insights confiáveis para a tomada de decisões de negócios. Para ajudar a atingir esses objetivos, as empresas estão modernizando suas iniciativas de análise criando armazéns de dados e data lakes na nuvem ou consolidando armazéns de dados on-premise na nuvem. As empresas também estão adotando modernas tecnologias de nuvem que mesclam armazéns de dados e data lakes em uma única plataforma de dados, às vezes chamada de "lakehouse".

O gerenciamento de dados é essencial para maximizar o valor dos projetos de análise na nuvem. Para tirar o máximo proveito da agilidade, flexibilidade e escalabilidade da nuvem e evitar repetir as dificuldades que afetaram seus ambientes on-premise, as empresas precisam adotar uma abordagem sistemática à qualidade e ao gerenciamento de dados em seus armazéns de dados, data lakes e lakehouses.

Ao implantar uma solução inteligente e automatizada de gerenciamento de dados de cloud lakehouse, para criar e gerenciar cargas de trabalho na nuvem, as organizações podem melhorar a transparência dos dados, conectar-se a diversas fontes de dados e gerenciar ambientes híbridos e com várias nuvens cada vez mais complexos. Essa abordagem permite que as pessoas da sua empresa, desde os analistas de negócios aos cientistas e engenheiros de dados, acessem com rapidez e facilidade dados de alta qualidade para suas iniciativas de análise, impulsionando a inovação e proporcionando às organizações uma vantagem competitiva.

Como adotar uma abordagem em fases para a modernização

Para começar, as empresas geralmente montam um armazém de dados em nuvem para um projeto departamental para testar a solução e avaliar suas opções em nuvem. Depois que isso for entregue com sucesso e alguns benefícios comerciais forem alcançados, a próxima fase é migrar os projetos mais tradicionais de armazém de dados local para a nuvem ou adicionar um data lakes na nuvem para obter dados brutos para a exploração da ciência de dados e a curadoria de novas fontes de dados.

Algumas organizações, ao iniciarem sua transição para a nuvem, optam por processos manuais, como codificação manual, mas essa abordagem rapidamente sobrecarrega a equipe de TI, tornando-se cara de manter e difícil de escalar. Outras organizações podem optar por soluções limitadas para atender às necessidades específicas de gerenciamento de dados, mas depois acabam com um conjunto de abordagens fragmentadas que aumentam a complexidade e introduzem riscos e custos desnecessários.

A Informatica® ajuda as organizações a evitar esses erros e a acelerar suas iniciativas de armazém de dados e data lake na nuvem com a solução de gerenciamento de dados de cloud lakehouse moderna, independente e líder do mercado — [Informatica Intelligent Cloud ServicesSM](#) (IICS). O IICS é uma solução de gerenciamento de dados nativa da nuvem moderna, baseada em microserviços, orientada por metadados, alimentada por IA e criada especificamente para armazéns de dados e data lakes na nuvem, independentemente da plataforma ou provedor de tecnologia de nuvem (incluindo Amazon, Microsoft, Snowflake, Databricks e Google).

Principais funções

Gerenciamento de dados de cloud lakehouse

Para ter sucesso com armazéns de dados e data lakes na nuvem, a Informatica fornece a única solução completa do setor para catalogar, ingerir, integrar, limpar e governar dados. Outras soluções requerem vários componentes diferentes, geralmente de fornecedores separados. O IICS fornece recursos abrangentes, nativos da nuvem, baseados em IA e integrados para Cloud Data Integration, Cloud Data Quality e Metadata Management, todos construídos em uma plataforma inteligente. O IICS também oferece automação baseada em IA com instalação e configuração mínimas, atualizações automáticas, uma pilha de tecnologia totalmente integrada para alta disponibilidade e recursos críticos. [Certificações de confiança](#) (por exemplo, SOC, SOC2, HIPAA, ISO / IEC 27001, Cloud Security Alliance, Privacy Shield Framework e mais).

Integração de dados

Informe e integre rapidamente todos os tipos de dados usando um ambiente de desenvolvimento visual intuitivo para criar facilmente pipelines de dados que alimentam seu armazém de dados e data lake na nuvem. O IICS Cloud Data Integration fornece conectividade nativa da nuvem pré-criada para praticamente qualquer tipo de dados corporativos, seja em multinuvem ou on-premise. [O Cloud Data Integration](#) foi projetado para escalabilidade e desempenho e oferece recursos críticos de otimização, incluindo otimização push-down para que você possa processar os dados com eficiência. O Cloud Data Integration Elastic fornece processamento Spark sem servidor para maior escalabilidade e capacidade sob demanda.

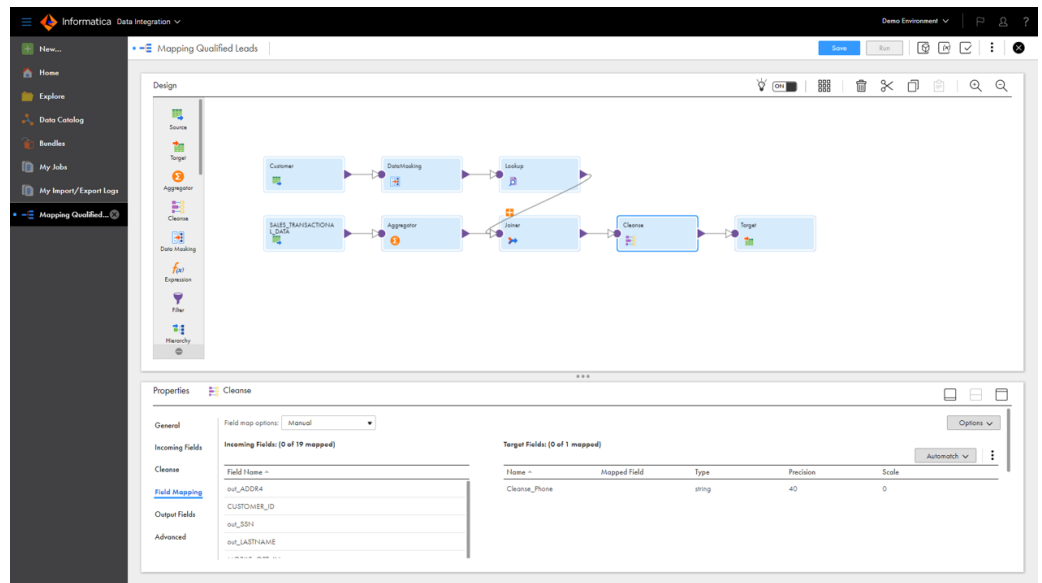


Figura 1: O Cloud Data Integration ajuda a criar pipelines de dados com clicar e arrastar.

Você pode criar rapidamente pipelines de dados com assistentes alimentados por integrações predefinidas e prontas para uso, descoberta de dados inteligente, análise automatizada de arquivos complexos e recomendações de transformação com IA. Isso facilita a descoberta de dados para entrada no armazém de dados e no data lake da nuvem e fornece a capacidade de reutilizar os pipelines de dados para outros projetos. O IICS Cloud Mass Ingestion permite que você ingira dados de uma variedade de fontes, incluindo arquivos, bancos de dados, captura de dados alterados e streaming de dados em tempo real.

Data Quality

Verifique se o armazém de dados na nuvem possui dados confiáveis e se o data lake não se torna um data swamp. [Informatica Cloud Data Quality](#) fornece recursos nativos da nuvem para que você possa adotar uma abordagem holística para criar um perfil rápido de seus dados, permitindo identificar, corrigir e monitorar problemas de qualidade de dados antes de mover os dados para seu armazém de dados ou data lake na nuvem. O Cloud Data Quality ajuda a transformar os dados para que você possa limpar, padronizar e enriquecer todos os dados. Você pode usar um conjunto extenso de regras de qualidade de dados pré-criadas sem codificação adicional para garantir a entrega e análise confiáveis de dados.

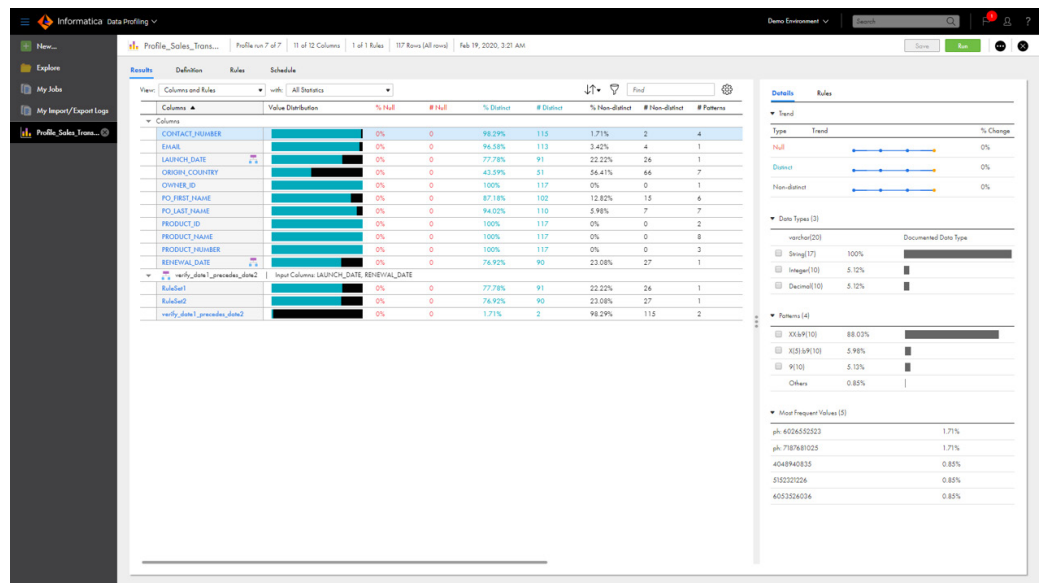


Figura 2: O Cloud Data Quality identifica automaticamente regras para seus perfis.

O Cloud Data Quality possui regras de negócios, padronização, correspondência, limpeza mundial de endereços e gerenciamento versátil de qualidade de dados para todos os tipos de projetos. O mecanismo CLAIRE[®] fornece inteligência artificial baseada em metadados, o que permite recomendações inteligentes de regras de qualidade de dados baseadas em como foram gerenciados os dados parecidos.

Gerenciamento de metadados

Apenas a Informatica possui uma fundação de gerenciamento de metadados unificada para facilitar a colaboração e permitir fluxos de trabalho de ponta a ponta para integração de dados com IA, qualidade de dados, catalogação de dados e governança de dados. O IICS fornece uma abordagem orientada a metadados para a construção de pipelines de dados usando um ambiente de desenvolvimento visual, com a capacidade de executar em todas as principais plataformas de nuvem (Amazon Web Services, Microsoft Azure e Google Cloud Platform) e nos mais recentes mecanismos de processamento (Spark), sem exigir que você codifique novamente seus pipelines existentes. Com essa abordagem, você pode descobrir, marcar, relacionar e provisionar dados automaticamente no seu armazém de dados e data lake na nuvem.

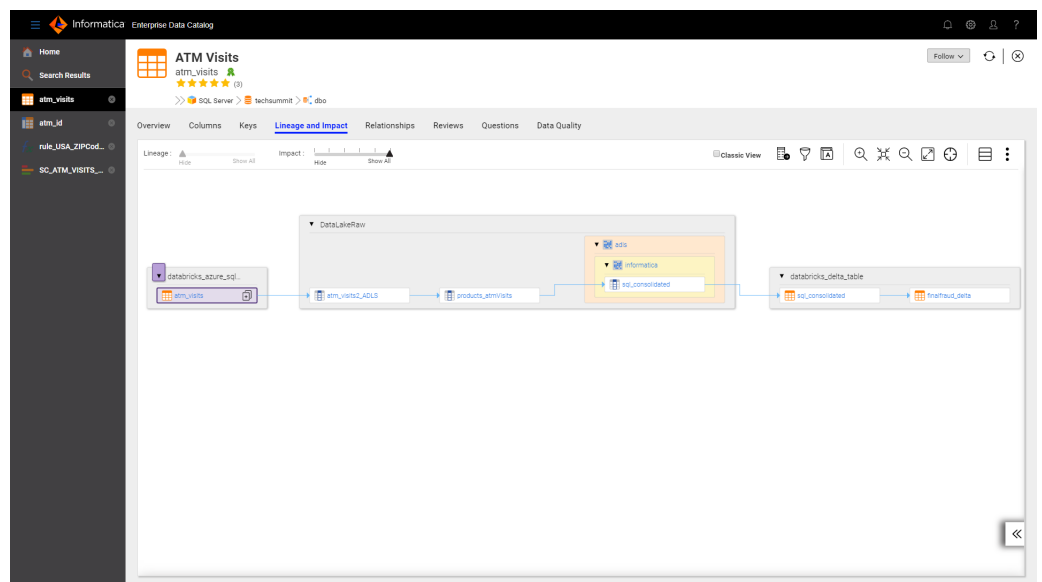


Figura 3: O Enterprise Data Catalog fornece linhagem de dados de ponta a ponta.

A Informatica fornece um catálogo de dados corporativos de seus ativos de dados e seus relacionamentos, varrendo e coletando os metadados de todos os sistemas corporativos, selecionando e aumentando os metadados com o contexto comercial e inferindo relacionamentos e linhagem. A seguir, são apresentadas as categorias de metadados coletadas e analisadas:

- **Técnico:** Esquemas de banco de dados, mapeamentos e códigos, transformações e verificações de qualidade
- **Negócios:** Termos de glossário, processos de governança e contexto de aplicativos e negócios
- **Operacional e infraestrutura:** Estatísticas em tempo de execução, registros de data e hora, métricas de volume, informações de log e informações de sistema e localização
- **Uso:** Classificações de usuários, comentários e padrões de acesso

Os metadados nessas quatro categorias se tornam a fundação para uma base de metadados comum. A Informatica ativa esses metadados aplicando IA e aprendizado de máquina e integrando-os à pilha de tecnologias.

Benefícios principais

Impulsione a produtividade e reduza os custos

A solução abrangente e integrada de gerenciamento de dados do Informatica Cloud Lakehouse pode aumentar a produtividade e diminuir o risco de projetos, ao mesmo tempo em que fornece inteligência, automação e eficiência. Você pode reduzir as despesas operacionais e maximizar a utilização de recursos com o Spark sem servidor no processamento em escala. Ao todo, você pode fornecer uma imagem completa para a tomada de decisões críticas aos negócios; a solução ajuda a descobrir, ingerir, integrar, limpar, dominar, governar, proteger e fornecer dados de qualidade com governança para todos os seus usuários.

- Analistas de dados, cientistas de dados e organizadores de dados podem obter uma compreensão mais rápida dos dados por meio de linhagem de dados detalhada de ponta a ponta, perfis de qualidade de dados e marcação de multidão e de IA, permitindo que eles gastem mais tempo analisando os dados e construindo os modelos de aprendizado de máquina.
- Os engenheiros de dados podem encontrar muitas tarefas de implementação, parcial ou totalmente automatizadas, com as próximas melhores recomendações de transformação, geração automatizada de pipeline de dados, hidratação automatizada de armazém de dados diretamente do catálogo de dados corporativos e análise automatizada de arquivos complexos, como IoT e dados de log da web.
- Os profissionais de segurança e privacidade de dados acham mais simples detectar o uso indevido de dados, proteger dados confidenciais e demonstrar que os controles apropriados são mantidos por meio da detecção e proliferação automatizada de dados confidenciais entre sistemas e limites geográficos.

Aumente a flexibilidade e a escalabilidade

A Informatica é o único fornecedor independente de gerenciamento de dados em nuvem da empresa que fornece os melhores recursos líderes do setor em integração de dados, qualidade de dados, governança de dados e gerenciamento de metadados. Você não precisa se preocupar em integrar várias ferramentas de solução pontual que possam limitar significativamente seus recursos de gerenciamento de dados e aumentar riscos e custos.

O gerenciamento de dados do Informatica Cloud Lakehouse prova o futuro dos seus ambientes multinuvem que mudam rapidamente e das pilhas analíticas em evolução, para que você possa implantar com sucesso um armazém de dados ou data lake na nuvem com flexibilidade. Você pode conectar-se a diferentes aplicativos e fontes de dados on-premise e na nuvem, integrando perfeitamente grandes volumes de dados e gerenciando facilmente cargas de trabalho analíticas com alto desempenho.

A Informatica se conecta de maneira nativa a todos os tipos de fontes de dados e ajuda a ingerir ETL e enviar o processamento para os armazéns de dados na nuvem; portanto, se você decidir alterar sua plataforma em nuvem, será mais fácil migrar sem recodificar, ajudando a proteger seu investimento. Com a integração de dados na nuvem sem servidor Spark, você configura apenas seus requisitos; o Cloud Data Integration Elastic fornece a capacidade necessária com gerenciamento e escalabilidade otimizados de recursos.

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.

Melhore a agilidade

Com o IICS, você obtém a rápida implantação de tarefas, instalação e configuração mínimas, atualizações automáticas, integração rápida de dados e uma pilha de tecnologia completa e integrada para desenvolvimento ágil e tempo de inatividade minimizado. Comece rapidamente sem treinamento avançado, aproveitando a conectividade pronta para uso com visibilidade para centenas de aplicativos, integração sem código e ingestão em massa em escala para cargas de trabalho de alto volume.

Adquira agilidade na nuvem para integrar novos recursos, reduzindo o tempo de colocação no mercado e aumentando o ROI de seus investimentos em tecnologia com maior rapidez na primeira avaliação de valor. Obtenha os benefícios do projeto mais rapidamente com a conclusão mais rápida do projeto, forneça autoatendimento sob controle, reduza os custos do projeto e reduza os riscos com a continuidade dos negócios incorporada.

Analistas de negócios e cientistas de dados podem acessar facilmente uma grande quantidade de dados confiáveis usando a pesquisa semântica semelhante ao Google. O CLAIRE auxilia no fornecimento de recomendações relevantes de novos ativos de dados que podem aumentar a análise. Isso aumenta a agilidade e permite que os consumidores de dados realizem análises ad-hoc e de ciência de dados sem precisar contar com o suporte da organização de TI.

Agora você pode garantir o sucesso do seu armazém de dados ou projeto de data lake na nuvem, acelerando primeiramente obtenção de valor e fornecendo retorno sobre o investimento (ROI) com o gerenciamento de dados do cloud lakehouse, que é alimentado por inteligência e automação.

Próximas etapas

[Entre em contato conosco](#) para saber mais sobre a Informatica Cloud Lakehouse Data Management.

Experimente uma avaliação gratuita de 30 dias do Informatica Cloud Data Warehouse — parte do serviço Informatica Cloud Data Integration — [inscreva-se hoje mesmo](#).



Informatica

Sede mundial Av. Das Nações Unidas, 12901 - 3º, Andar - Torre Norte - Brooklin Novo, 04578-000 - São Paulo, SP Tel.: 0800-878-3060
Ligação gratuita nos EUA: 1.800.653.3871

IN17_0420_03569

© 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.