

Informatica Enterprise Data Catalog

Benefícios

- Catalogar e classificar automaticamente todos os tipos de dados por toda a empresa usando um catálogo orientado por IA
- Fornecer um sistema de metadados de registro para a empresa com um catálogo de catálogos
- Extrair automaticamente os metadados mais granulares de uma matriz ampla de fontes de dados, incluindo sistemas corporativos complexos
- Localizar ativos de dados por meio de pesquisas semânticas eficientes similares às do Google
- Descobrir e entender seus ativos de dados com uma visão holística que inclui linhagem, visualização de relacionamentos, estatísticas de criação de perfil de dados e pontuações de qualidade
- Identificar domínios e entidades com uma curadoria inteligente
- Enriquecer ativos de dados com anotações, classificações e revisões governadas e colaborativas
- Associar automaticamente termos de glossário a ativos de dados técnicos
- APIs abertas para integrar em seu ambiente e exibir os metadados inteligentes em qualquer lugar

Libere o poder dos dados com um catálogo de dados inteligente

Dados são a parte vital de nossa economia, e as empresas orientadas por dados convertem seus ativos de dados em receita e lucros. O primeiro passo em qualquer iniciativa de transformação digital orientada por dados é gerenciar seus dados como um ativo da empresa: faça seu inventário, avalie seu valor e maximize seu uso, assim como faz com outros investimentos de capital e operacionais importantes.

Os dados são diversificados e distribuídos por diversos departamentos, aplicativos, armazéns de dados e data lakes diferentes (alguns locais, outros na nuvem), o que torna um desafio saber exatamente quais dados você tem e onde estão. À medida que as fontes de dados proliferam, o cenário dos dados se torna ainda mais complexo.

O Informatica® Enterprise Data Catalog é um catálogo de dados orientado por IA que oferece um mecanismo de descoberta baseado em aprendizagem automática que digitaliza e cataloga ativos de dados por toda a empresa, na nuvem e no local, e Big Data em qualquer lugar. O Enterprise Data Catalog é alimentado pelo mecanismo CLAIRE®, que oferece inteligência ao aproveitar metadados para oferecer recomendações, sugestões e automação de tarefas de gerenciamento de dados. Isso permite que os usuários de TI sejam mais produtivos e os usuários de negócios sejam parceiros completos no gerenciamento e uso dos dados.

O Informatica Enterprise Data Catalog fornece aos analistas de dados e usuários de TI uma pesquisa semântica eficiente e facetas dinâmicas para filtrar resultados de pesquisa, linhagem de dados, estatísticas de perfil, pontuações de qualidade de dados, visões holísticas de relacionamento, recomendação de similaridade de dados e um glossário comercial integrado.

Os recursos de colaboração aproveitam a experiência no assunto e a curadoria social combinadas com o poder da IA para orientar a experiência do usuário e automatizar a curadoria de dados. Agora você pode gerenciar ativos de dados corporativos com facilidade e eficiência para maximizar o valor deles dentro da empresa. Os usuários podem rapidamente localizar dados e gerenciar facilmente o ciclo de vida de termos de negócios, definições, dados de referência, e muito mais.

O Informatica Enterprise Data Catalog é um catálogo de dados orientado por IA que oferece um mecanismo de descoberta baseado em aprendizagem automática para digitalizar e catalogar ativos de dados por toda a empresa, na nuvem e no local, e Big Data em qualquer lugar.

Principais recursos

APIs de metadados para integrar ao seu ambiente

O Informatica Enterprise Data Catalog inclui APIs baseadas em REST que permitem integrá-lo ao seu ambiente e utilizar o conteúdo do catálogo em qualquer lugar. As organizações podem compartilhar quaisquer metadados inteligentes — aplicativos, relatórios de BI e painéis — com os usuários de negócios. Os usuários podem exportar e compartilhar conteúdo de catálogo selecionado e metadados para enriquecimento selecionados.

Extensão Tableau do Enterprise Data Catalog para análises governadas de autoatendimento

A extensão Tableau para o Informatica Enterprise Data Catalog oferece análises ágeis e de autoatendimento com dados governados. Ela permite que os usuários do Tableau acessem os recursos completos do Informatica Enterprise Data Catalog a partir da interface nativa de usuário do Tableau. Os usuários de negócios que são consumidores dos relatórios do Tableau podem aproveitar uma barra de pesquisa inteligente no Tableau para encontrar ativos de dados certificados por organizadores de dados, além de acessar o contexto técnico e empresarial, a qualidade dos dados, as estatísticas de perfil e as recomendações dos pares. Os usuários também podem colaborar com seus pares nos ativos de dados por meio de revisões, classificações e perguntas e respostas, tudo sem sair da interface do Tableau.

Pesquisa semântica com facetas inteligentes

Localize e descubra os conjuntos de dados mais relevantes para suas análises usando a pesquisa semântica eficiente com facetas inteligentes. Pesquisas de palavras-chave avançadas com correspondência de token localizam os ativos de dados mais relevantes no catálogo. Pesquisas semânticas são aplicadas até mesmo nos domínios de dados implícitos, assim, nenhum ativo de dados fica sem ser descoberto. Facetas inteligentes baseadas em resultados de pesquisa permitem que os usuários alterem a pesquisa para os conjuntos de dados de interesse.

Provisionamento de dados de autoatendimento

Depois de encontrar os conjuntos de dados relevantes para sua análise, mova-os facilmente para o destino de sua escolha com um simples provisionamento por clique no Informatica Enterprise Data Catalog. Você pode escolher entre um extenso leque de fontes e destinos, incluindo Amazon Redshift, Azure Synapse Analytics (anteriormente Azure SQL Data Warehouse), Google BigQuery, Snowflake e ferramentas de BI, como o Tableau. Esse recurso aproveita a integração do Informatica Enterprise Data Catalog com o Informatica Cloud Data Integration.

Linhagem de dados e análise de impacto

Rastreie interativamente a origem dos dados através de visualizações de linhagem em qualquer nível — desde visualizações em nível de sistema favoráveis aos negócios que destacam os endpoints a visualizações granulares, que incluem todos os detalhes complexos envolvidos. Uma visualização de linhagem detalhada expande qualquer caminho de linhagem para mostrar linhagens granulares em nível métrico e de colunas. Os usuários podem realizar uma análise de impacto detalhada em ativos de dados nas fases upstream e downstream.

Descoberta holística de relacionamento

Obtenha uma visão holística dos dados com um gráfico de conhecimento que permite que você pesquise, descubra e compreenda os dados corporativos e os relacionamentos de dados significativos. Descubra automaticamente conjuntos de dados relacionados, relacionamentos técnicos, corporativos, semânticos e baseados em utilização. A visão holística de dados mostra conjuntos de dados, tabelas, visualizações, domínios de dados, relatórios e usuários relacionados. Isso auxilia na descoberta progressiva de outros conjuntos de dados de interesse.

Classificações automáticas com domínio inteligente e reconhecimento de entidade

Classifique e identifique domínios e entidades automaticamente, tais como cliente, produto e pedido, entre outros, em todos os ativos de dados, estruturados e não estruturados, no nível do campo, coluna ou tabela. Este é um passo essencial na capacidade das empresas de catalogar, governar e extrair valor de seus ativos de dados. Esses dados classificados permitem uma melhor pesquisa, filtragem de resultados de pesquisa e recomendações do glossário comercial. A Informatica fornece mais de 60 domínios de dados empacotados, como e-mail, número de cartão de crédito, número de previdência social, país, cidade, URL e razão social. Os usuários também podem adicionar seus próprios domínios personalizados. Ativos de dados podem ser classificados usando regras de dados, ou seja, colunas com dados que correspondam a lógicas específicas definidas na regra ou regras de nomes de coluna, ou seja, localizam colunas que correspondam a lógicas de nomes de coluna definidas na regra.

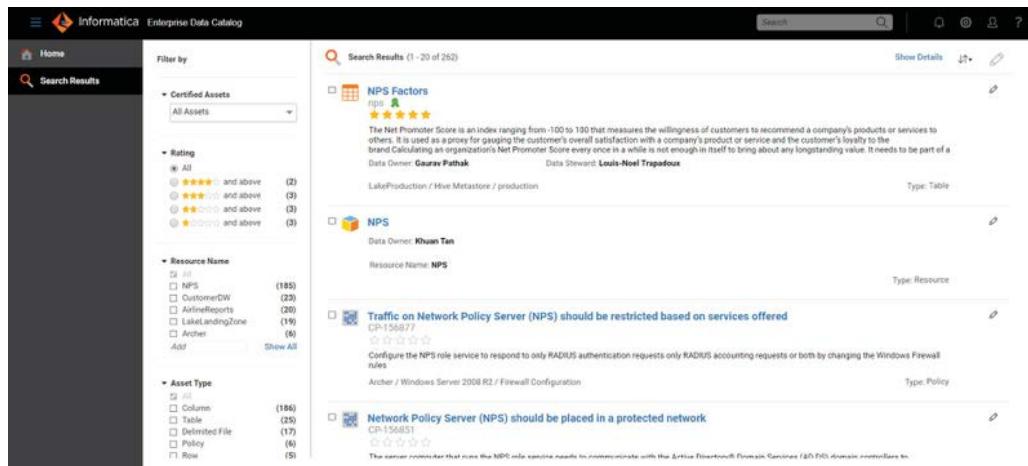


Figura 1: Encontrar com rapidez conjuntos de dados com pesquisa semântica inteligente e facetas dinâmicas. Visualizar classificações e conjuntos de dados certificados.

Colaboração e curadoria social

O Informatica Enterprise Data Catalog capacita os analistas e cientistas de dados a encontrar facilmente os dados mais relevantes e confiáveis para análise utilizando a combinação eficiente da IA e a experiência e colaboração humanas. Proprietários de dados e especialistas no assunto podem certificar conjuntos de dados. Os consumidores de dados podem fornecer classificações e revisões para conjuntos de dados que possibilitam a curadoria social de dados.

Os usuários podem acompanhar conjuntos de dados de interesse e serem notificados de mudanças; uma plataforma de perguntas e respostas permite que os especialistas no assunto respondam a perguntas comuns dos usuários. Além disso, os usuários podem adicionar atributos personalizados e anotações a folhas de dados, aprimorando ainda mais a colaboração da TI de negócios e resultados de pesquisa.

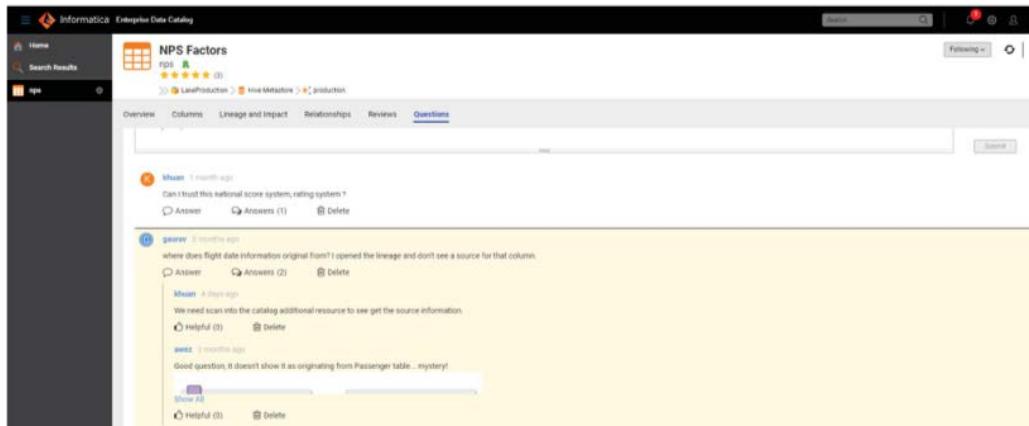


Figura 2: Permitir colaboração com recursos de perguntas e respostas.

Qualidade de dados integrada

Visualize estatísticas de perfil de dados, regras de qualidade de dados, pontuações e grupos métricos junto com metadados técnicos para entender a qualidade dos ativos de dados antes de usar os dados para análise. Estatísticas de perfil incluem distribuições de valor, padrões, tipos de dados e deduções de domínios de dados.

Associação automática de termos de glossário comercial

O Informatica Enterprise Data Catalog facilita a importação de ativos de glossário comercial, como termos, políticas e classificações, do Informatica Axon™. Adicione contexto empresarial valioso aos dados ao associar os termos comerciais com os metadados técnicos certos.

O Informatica Enterprise Data Catalog permite a associação automática de termos de negócios com os metadados técnicos, o que ajuda a eliminar um processo manual maçante para governança de dados. Ele permite que administradores de negócios e de TI gerenciem, de modo colaborativo, os metadados de negócios que incluem uma automação eficiente do fluxo de trabalho humano. O Informatica Enterprise Data Catalog também é compatível com importação de ativos do glossário comercial do Informatica Business Glossary e ferramentas de terceiros.

Similaridade de dados inteligentes

Algoritmos estatísticos e de aprendizado de máquina avançados identificam dados semelhantes e subconjuntos de dados. Este poderoso recurso ajuda os usuários a localizar os dados mais relevantes e confiáveis que eles precisam. Um analista de telecom interessado em análises de atritos de clientes, por exemplo, podem consultar dados que contêm atividades de clientes pré-pagos do trimestre atual. O Informatica Enterprise Data Catalog pode recomendar uma versão mais limpa dos dados (dados substitutos), dados que contêm atividades de clientes durante o trimestre anterior (dados unificáveis) e uma tabela de detalhes de clientes para enriquecer o conjunto de dados (dados combináveis).

Conectividade de metadados universal com scanners avançados

Extraia metadados de qualquer tipo de fonte de dados em toda a empresa, como bancos de dados, armazéns de dados, data lakes baseados em nuvem, ferramentas de BI, clusters do Hadoop, NoSQL e sistemas corporativos complexos, incluindo sistemas legados e de mainframe, ferramentas ETL de vários fornecedores, dialetos de SQL e vários aplicativos corporativos — em ambientes híbridos e multinuvem.

Com os scanners avançados do Enterprise Data Catalog, você pode inspecionar visualmente todos os scripts, procedimentos ou processos para entender totalmente sua lógica e o fluxo de dados internos. Você pode obter a linhagem de dados completa em nível de coluna, além de um inventário com todas as possíveis fontes de linhagem com riqueza de detalhes. Os scanners possibilitam a leitura de códigos estáticos e dinâmicos, além de análises de linguagem para obter a linhagem de dados automatizada.

A seguir estão alguns exemplos de fontes de dados compatíveis com a extração de metadados:

- **Bancos de dados/Armazéns de dados:** Oracle, Microsoft SQL Server, SQL Scripts, Sybase ASE, IBM Netezza, Teradata, JDBC, SAP HANA, SAP BW, SAP BW/4HANA, Procedimentos armazenados
- **Big Data:** Cloudera Navigator, Hive (Cloudera/Hortonworks/MapR/IBM BigInsights/EMR), HDFS, Hortonworks Atlas, Cassandra, MongoDB, Kafka
- **Mainframes:** DB2 z/OS, DB2 i5/OS, COBOL, JCL
- **BI e análise:** SAP BusinessObjects, Tableau, Microsoft Power BI, Cognos, MicroStrategy, OBIEE, QlikView, Qlik Sense, Microsoft SSRS e SSAS, SAS
- **ETL:** Informatica PowerCenter, Informatica Data Engineering Integration, Informatica Intelligent Cloud Services, Informatica Data Integration Hub, Microsoft SSIS, Oracle Data Integrator, IBM InfoSphere DataStage, AWS Glue
- **Glossário comercial:** Informatica Axon Data Governance, Informatica Business Glossary
- **Modelagem de dados:** Erwin Data Modeler, SAP PowerDesigner
- **Aplicativos corporativos:** Salesforce, Oracle, Workday, Informatica MDM, SAP ECC
- **Sistemas de arquivos:** Microsoft SharePoint, Microsoft OneDrive, Windows/Linux Filesystems
- **Formatos de arquivo:** MS Excel, MS Word, MS PowerPoint, Adobe PDF, Arquivos Simples, CSV Delimitado por vírgulas, XML, JSON, Avro, Parquet
- **Plataformas de nuvem:** AWS S3, AWS Redshift, Azure SQL DB, Azure SQL DW, Azure ADLS, Azure ADLS Gen 2, Azure Blob, Google Cloud Storage, Google BigQuery, Snowflake

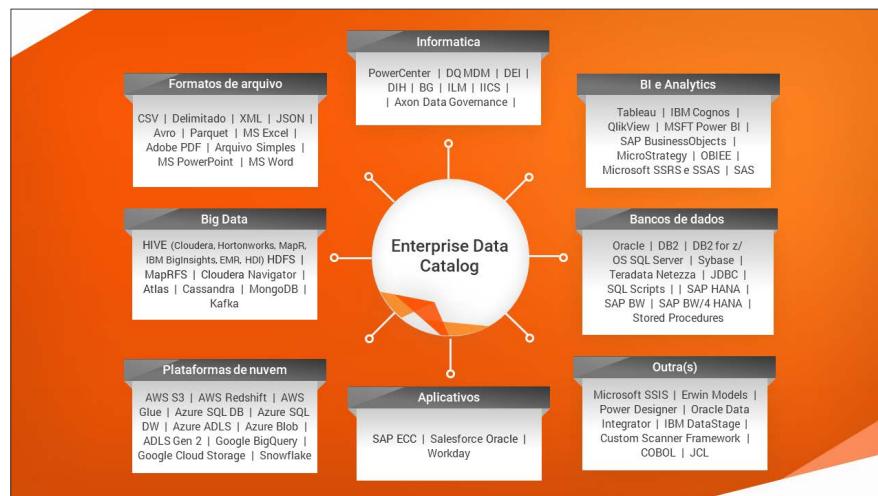


Figura 3: O Informatica Enterprise Data Catalog é compatível com conectividade de metadados universal

Segurança de nível de recursos

Conceda permissões de leitura/gravação para usuários e grupos no nível de recursos a fim de permitir que eles visualizem ou editem atributos personalizados, realizem a curadoria dos domínios e associem os termos de glossários comercial.

Implantações em escala corporativa

O Informatica Enterprise Data Catalog foi desenvolvido para implantações reais em escala de Big Data com a capacidade de digitalizar dezenas de milhões de conjuntos de dados em centenas de fontes de dados. Ele oferece suporte à ingestão de metadados paralelos e indexação distribuída de alta velocidade para atualizar rapidamente o conteúdo e oferecer desempenho de pesquisa incomparável e alta disponibilidade tolerante a falhas para implementações 24 horas por dia, 7 dias por semana. Com a criação de perfis de dados baseada no Spark, é possível criar o perfil de quantidades imensas de dados em escala para obter um melhor entendimento dos dados corporativos.

Administração unificada

Gerencie e monitore os recursos do catálogo, programações de extrações de metadados, execuções de perfil e muito mais a partir de um console de administração unificado. Um painel de controle de trabalho fornece widgets para monitorar tarefas e exibir recursos. Alertas de e-mail auxiliam administradores a responder proativamente aos problemas do catálogo.



Figura 4: Entenda seus dados com visões holísticas de relacionamento de dados.

Benefícios

Catalogar, de maneira inteligente, todos os tipos de dados na empresa

O Informatica Enterprise Data Catalog descobre, de maneira inteligente, muitos tipos de dados e seus relacionamentos na empresa. Scanners pré-desenvolvidos coletam metadados de bancos de dados, armazéns de dados, data lakes, armazenamentos de dados em nuvem, aplicativos, ferramentas de BI, ferramentas de ETL, catálogos de metadados de terceiros, NoSQL, e muito mais.

Sobre a Informatica

A transformação digital muda expectativas: melhor serviço, entrega mais rápida e com menos custos. As empresas precisam mudar para continuar relevantes, 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.

Todos os metadados são indexados e catalogados em um banco de dados de gráficos altamente dimensionável para atualizações rápidas, pesquisas inteligentes e consultas rápidas. À medida que mais e mais dados são criados e propagados dentro da empresa, conjuntos de dados semelhantes e duplicados inevitavelmente surgem. O Informatica Enterprise Data Catalog utiliza algoritmos estatísticos e avançados de aprendizagem automática para descobrir dados e subconjuntos de dados semelhantes, o que ajuda os usuários a localizar os dados necessários, relevantes e confiáveis.

Localizar ativos de dados por meio de pesquisas semânticas eficientes

Algumas vezes, tentar localizar os dados necessários entre centenas de sistemas corporativos pode parecer em vão. Apenas com pesquisas semânticas eficientes criadas a partir de inteligência orientada por metadados abrangentes e uma infraestrutura dimensionável é possível ter alguma esperança de localizar dados relevantes. O Informatica Enterprise Data Catalog oferece uma pesquisa semântica com facetas inteligentes para refinar ainda mais os resultados de pesquisa. Como a Informatica associa de forma única metadados comerciais, técnicos e operacionais, os usuários de negócios podem pesquisar termos comerciais para localizar seus dados e depois procurar por relacionamentos holísticos para localizar ativos de dados relacionados.

Descubra e entenda seus ativos de dados com visões holísticas de relacionamento e linhagem

O ditado popular, "Você não pode gerenciar o que você não pode medir" é verdadeiro quando se trata de gerenciar ativos de dados. Para obter o valor máximo dos dados, é preciso entender o que você tem, de onde eles vieram, como foram transformados e qual é o nível de sua confiança nesses dados. O Informatica Enterprise Data Catalog responde a essas e outras perguntas com um resumo de ponta a ponta completo e linhagem detalhada, estatísticas de criação de perfil, pontuações de qualidade de dados e visões holísticas de relacionamento, fornecendo uma imagem clara dos seus dados.

Enriqueça ativos de dados com contexto de negócios por meio de anotações governadas e colaborativas (crowdsourced)

O Informatica Enterprise Data Catalog maximiza a reutilização e o valor dos dados classificando automaticamente os ativos de dados corporativos até o nível de campo/coluna. Para aumentar ainda mais o valor dos dados, o Informatica Enterprise Data Catalog capta o contexto de quem está usando os dados e para qual objetivo, junto com etiquetas, anotações, classificações e revisões colaborativas. Essa "sabedoria das multidões" ajuda a enriquecer e organizar os dados, tornando-os ainda mais valiosos dentro da empresa. O Informatica Enterprise Data Catalog se integra com o Informatica Axon para facilitar a importação dos ativos do glossário comercial, tais como termos de negócios, definições e políticas do Axon. Esses metadados de negócios são associados automaticamente a metadados técnicos e operacionais para que analistas de negócios, administradores de dados e outros usuários possam rapidamente localizar, entender e colaborar no escopo de ativos de dados.

Saiba mais

Para saber mais sobre o Informatica Enterprise Data Catalog, visite

<https://www.informatica.com/products/data-catalog/enterprise-data-catalog.html>.



Sede América Latina 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

IN06_0620_03238

© Copyright Informatica LLC 2020. Informatica, o logotipo da Informatica, CLAIRE, Axon, PowerCenter, Big Data Management e Informatica Cloud são marcas comerciais ou marcas comerciais registradas da Informatica LLC nos Estados Unidos e em 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.