

# Informatica Enterprise Data Preparation

## Benefícios

- Encontre os dados certos para seus projetos de análise com base em autoatendimento
- Prepare e compartilhe rapidamente os dados necessários
- Operacionalize facilmente a preparação de dados para reutilizá-los

## Descubra dados brutos e prepare-os para análises avançadas, repetidamente

Os dados são, sem dúvida, a base da transformação digital. As organizações usam novas plataformas de processamento de dados, tais como a Apache Hadoop para extrair ideias, que se não improváveis, antes eram inatingíveis. Graças ao surgimento do Apache Hadoop e da abordagem de data lake, as organizações já podem se dar ao luxo de reunir todos os dados de modo que fiquem acessíveis aos usuários a qualquer momento para qualquer tipo de análise.

O grande volume de dados que é processado nos sistemas Hadoop está sobrecarregando a organização de TI. Os analistas de negócios esperam dados de qualidade do Hadoop, enquanto a equipe de TI fica sobrecarregada com processos manuais e demorados para transformar dados brutos em ativos de dados adequados para a finalidade.

Sem mecanismos inteligentes, repetíveis e dimensionáveis para administrar os dados, todas as oportunidades prometidas pelos data lakes correm o risco de estagnação. A chave para resolver a crise dos chamados data swamps é a tecnologia de inteligência artificial orientada por metadados da Informatica®, conhecida como mecanismo CLAIRE™, que automaticamente descobre, cria perfis e infere relacionamentos dos ativos de dados.

O Informatica Enterprise Data Preparation permite que Big Data bruto seja descoberto sistematicamente, para que os analistas de negócios possam transformar os conjuntos de dados em informações confiáveis em uma base automatizada. Os cientistas de dados e analistas corporativos podem encontrar rapidamente os dados que estiverem procurando usando pesquisa semântica e multifacetada, enquanto entendem automaticamente a linhagem de dados e relacionamentos de dados. Equipes de analistas de dados também podem facilmente colaborar uns com os outros e compartilhar resultados em espaços de trabalho do projeto. À medida que adicionam conjuntos de dados ao seu espaço de trabalho do projeto, algoritmos de aprendizado de máquina trabalham em segundo plano para recomendar conjuntos de dados alternativos que eles possam estar interessados em utilizar.

Você pode abrir conjuntos de dados dentro do espaço de trabalho do projeto a qualquer momento na ferramenta de preparação de dados fácil de usar e semelhante ao Excel, no Enterprise Data Preparation. A abordagem orientada a metadados à preparação de dados é uma forma inteligente de transformar Big Data em ativos de informações confiáveis que fornecem um valor de negócios sustentável.

## Principais recursos

### Pesquisa inteligente e visualizações inteligentes

Encontre dados no data lake e também em outros sistemas empresariais usando a busca semântica inteligente e os resultados baseados em inferência. Você pode filtrar os ativos de dados com base em facetas dinâmicas usando os atributos do sistema e visualizá-las com os gráficos baseados no Apache Zeppelin.

### Exploração abrangente de dados

Obtenha uma visão geral dos ativos de dados, incluindo atributos personalizados, criação de perfil de estatísticas para a qualidade de dados, domínios de dados para conteúdo comercial e informações de utilização. Informações de contribuição colaborativa sobre conjuntos de dados através de marcação e de metadados enriquecidos para adicionar contexto comercial. Entenda seus dados rapidamente, visualizando dados de amostra baseados nas credenciais do usuário. Entenda como o ativo de dados está relacionado com os outros ativos da empresa com base nas associações com tabelas ou exibições, usuários, relatórios e domínios de dados.

Use uma interface intuitiva como Excel para preparar de forma interativa os dados para análise com transformações incorporadas para filtrar, agregar, integrar e combinar dados. Execute a limpeza e transformação de dados de nível de coluna usando string, matemática, data e operações lógicas. Inteligência guiada ajuda a preparar os conjuntos de dados, tais como as recomendações de chaves de junção quando os conjuntos de dados são combinados. Consulte visões gerais de estatísticas descritivas na planilha, incluindo distribuições de valor e distribuições numéricas e de data. Todas as etapas são registradas em receitas que você pode usar para gerar automaticamente fluxos de dados que pode agendar repetidamente para operacionalizar insights de análise.

The screenshot displays a data analysis interface with a table of data, a 'Recipe' panel on the right, and a 'COLUMN OVERVIEW' panel at the bottom. The table has columns: #, vehicletype, education, ethnicity, occupation, incomelev, dwellingty, HighEarn, oldcustid, firstname, lastname, email, and year. The 'Recipe' panel lists transformations like 'Change ethnicity to Text type' and 'Split email into 3 columns'. The 'COLUMN OVERVIEW' panel shows statistics for the 'email' column, including a value frequency chart and suggestions like 'Split by @'.

#	vehicletype	education	ethnicity	occupation	incomelev	dwellngty	HighEarn	oldcustid	firstname	lastname	email	year
1	Truck	Diploma	3	Firefighter	55	Apartment	0	208	Aaron	MORRIS	amorris	
2	2-door	Masters	1	Accountant	150	TwinHm	1	8219	Aaron	THOMAS	athomas	
3	Truck	Diploma	3	Firefighter	55	Apartment	0	208	Aaron	MORRIS	amorris	
4	2-door	Masters	1	Accountant	150	TwinHm	1	8219	Aaron	THOMAS	athomas	
5	Truck	Diploma	3	Firefighter	55	Apartment	0	208	Aaron	MORRIS	amorris	
6	2-door	Masters	1	Accountant	150	TwinHm	1	8219	Aaron	THOMAS	athomas	
7	Truck	Diploma	3	Firefighter	55	Apartment	0	208	Aaron	MORRIS	amorris	
8	2-door	Masters	1	Accountant	150	TwinHm	1	8219	Aaron	THOMAS	athomas	
9	Truck	Diploma	3	Firefighter	55	Apartment	0	208	Aaron	MORRIS	amorris	
10	Truck	Diploma	1	Engineer	80	Condo	0	9294	Aaron	HANSEN	ahansen	
11	Convertible	Masters	1	Engineer	50	Single	0	305	Aaron	BRADLEY	abradley	

Figura 1: Encontrar com rapidez conjuntos de dados com facetas dinâmicas e busca semântica inteligente.

### **Operacionalização da preparação de dados**

Usando milhares de regras de negócios pré-integradas, os analistas de negócios podem melhorar sistematicamente a qualidade de dados dos produtos, enquanto realizam as etapas de preparação de dados para a equipe de TI para uma execução repetida.

### **Colaboração empresarial**

Gerencie publicações de dados de forma automatizada enquanto organiza o trabalho adicionando ativos de dados aos espaços de trabalho do projeto. Colabore com outros analistas, adicionando membros da equipe a projetos com funções diferentes, tais como coproprietário, editor ou visualizador, e com privilégios personalizados.

### **Recomendações de ativos de dados**

Melhore a produtividade e aumente a reutilização de ativos confiáveis com recomendações automatizadas baseadas em algoritmos de aprendizado de máquina aplicados ao comportamento e conhecimento compartilhado de outros usuários. Ativos alternativos e adicionais são recomendados para um projeto baseado em ativos de dados adicionados a esse projeto.

### **Carregamentos de dados baseados em assistente**

Carregue arquivos pessoais delimitados para reservatório de dados usando uma interface baseada em assistente. Tabelas Hive são criadas automaticamente para os carregamentos no formato mais ideal. Você pode criar, acrescentar ou substituir ativos para dados carregados.

## **Benefícios principais**

### **Encontrar e acessar quaisquer dados**

Os analistas de negócios querem uma maneira eficiente de gerenciar o "volume, a variedade e a velocidade" em constante crescimento dos dados, tipicamente associados ao Big Data. Encontre facilmente ativos de dados confiáveis usando a busca semântica inteligente e as facetas dinâmicas para filtrar os resultados. Um processo de descoberta baseado na aprendizagem de máquina automática transforma os ativos de dados relacionados em recomendações inteligentes de novos ativos de dados que podem ser de interesse para o analista. Isso aumenta muito a confiança e reduz conjuntos de dados duplicados que são criados para projetos similares.

### **Colaboração com governança**

Para aumentar a eficiência dos projetos analíticos de Big Data, os analistas de negócios colaboram com conjuntos de dados usando espaços de trabalho do projeto. À medida que os conjuntos de dados são adicionados aos espaços de trabalho de projeto, os analistas podem visualizar estatísticas de perfil, linhagem de ponta a ponta de conjuntos de dados, todos os ativos de dados relacionados, domínios de dados, usuários e muito mais. Isso ajuda na avaliação da qualidade de dados, compartilhamento de dados confiáveis e na descoberta progressiva de outros conjuntos de dados úteis para o projeto. Os analistas de negócios também podem gerenciar publicações de dados de forma automatizada. Segurança baseada na função garante que apenas os analistas adicionados a um projeto tenham acesso aos dados.

## 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.

## Prepare e compartilhe rapidamente os dados necessários

Conforme os ciclos de negócios continuam a encolher, a velocidade é uma das poucas vantagens competitivas com a qual os analistas de dados podem contar na corrida para adicionar valor aos negócios. Prepare e compartilhe rapidamente o ferramental de dados para o fornecimento de análises competitivas. Recursos de preparação de dados de autoatendimento da Informatica fornecem uma interface tipo Excel familiar e fácil de usar para os analistas de negócios, permitindo-lhes rapidamente agrupar, filtrar e combinar dados nas informações necessárias. A marcação e o compartilhamento de ativos de dados de contribuição colaborativa fornecem aos analistas de negócios controle sobre o processo de curadoria de dados, aumentando a eficiência operacional.

## Operacionalizar a preparação de dados em fluxos de trabalho reutilizáveis

Independentemente das ferramentas de automação e de autoatendimento, à sua disposição, os analistas muitas vezes têm de repetir as atividades de preparação de dados em novos conjuntos de dados. Isto desperdiça os ganhos de escala em curso e de reutilização. O Informatica Enterprise Data Preparation registra as etapas de preparação de dados que você pode agendar como pipelines de dados automatizados e repetidos. Isso transforma a preparação de dados de um processo manual em um sistema reutilizável e sustentável.

Para obter mais informações, visite a [página do produto Informatica Enterprise Data Preparation](#).



**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

IN06\_0319\_03237

© Copyright Informatica LLC 2019. Informatica, o logotipo Informatica, e CLAIRE são marcas comerciais ou marcas 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 mudança sem notificação prévia e foram fornecidas "COMO SE ENCONTRAM", sem garantias de qualquer espécie, expressas ou implícitas.