

# Serviços de dados centrais: componentes básicos para estabelecer valor de negócios

WHITE PAPER:

**QUALIDADE DE DADOS**



# Serviços de dados centrais: componentes básicos para estabelecer valor de negócios

## INTRODUÇÃO

O USO EFICAZ DE INFORMAÇÕES COMERCIAIS EXCEDEM AS DEMANDAS DE VALIDAÇÃO DA ESTRUTURA DE DADOS OU OS TESTES DE ASSOCIAÇÃO DE DOMÍNIOS DE VALOR. COMO MAIS DECISÕES SÃO DISTRIBUÍDAS ATRAVÉS DA HIERARQUIA ORGANIZACIONAL, E COMO MAIS INFORMAÇÕES SÃO PROCESSADAS E EXPOSTAS PARA O CONSUMIDOR FINAL, EXISTE UMA NECESSIDADE CRESCENTE DE REPRESENTAÇÕES VISUAIS DE VÁRIOS ASPECTOS DAS TRANSAÇÕES EMPRESARIAIS, COMO TIPO, TEMPO, DURAÇÃO E, ESPECIALMENTE, OS LOCAIS DAS TRANSAÇÕES EMPRESARIAIS. O APERFEIÇOAMENTO DO VALOR E DA UTILIDADE DAS INFORMAÇÕES COMERCIAIS, TANTO PARA APLICATIVOS OPERACIONAIS QUANTO ANALÍTICOS, REQUER A COMBINAÇÃO DE PRÁTICA EFICAZ DE GERENCIAMENTO DE DADOS, COMO A GOVERNANÇA DE DADOS, CONTROLE DE QUALIDADE, APRIMORAMENTOS DE DADOS, E CADA VEZ MAIS A CAPACIDADE DE LOCATION INTELLIGENCE COMO A CODIFICAÇÃO GEOGRÁFICA, MAPEAMENTO E ROTEAMENTO.

A PADRONIZAÇÃO DA PRÁTICA EFICAZ DE GERENCIAMENTO DE DADOS NO DESENVOLVIMENTO OU REENGENHARIA DA ARQUITETURA DA EMPRESA DIMINUI O RISCO E AUMENTA A CONSISTÊNCIA; A CONFIANÇA ORGANIZACIONAL DE DADOS SIMPLIFICA O DESENVOLVIMENTO DE APLICATIVOS ENQUANTO AUMENTA A CONFIABILIDADE DOS DADOS. EMBORA MUITO TENHA SE ESPECULADO SOBRE CONCEITOS COMO A INTEGRAÇÃO DE DADOS MESTRE COMO OPORTUNIDADE PARA GERAR UMA MAIOR QUALIDADE DE DADOS, ESTE DOCUMENTO EXPLORA COMO UM CONJUNTO CONSOLIDADO DE DADOS COM VALORES AGREGADOS QUE ENGLOBALAM O ESCOPO DAS PRÁTICAS RECOMENDADAS DE GERENCIAMENTO DE DADOS PODE REDUZIR OS CUSTOS OPERACIONAIS E AUMENTAR A EFICIÊNCIA, ENQUANTO AUMENTA A QUALIDADE E A UTILIDADE DOS DADOS DA EMPRESA. UMA ÚNICA CAMADA DE SERVIÇOS DE DADOS REDUZIRÁ A COMPLEXIDADE DA ESTRUTURA DE GERENCIAMENTO DE DADOS; PADRONIZAR A MANEIRA COMO AS FERRAMENTAS DE DADOS SÃO USADAS, ASSIM COMO AS MANEIRAS COMO AS REGRAS SÃO APLICADAS, É ALGO QUE VAI GARANTIR O VALOR E A CONFIABILIDADE DE INFORMAÇÕES PREVISÍVEIS.

# À MEDIDA QUE MAIS EMPRESAS FAZEM A TRANSIÇÃO PARA A ARQUITETURA ORIENTADA A SERVIÇOS, UMA INTEGRAÇÃO DE DADOS PERFEITA SE TORNA A BASE FUNDAMENTAL PARA OS APLICATIVOS DA EMPRESA.

## Motivadores dos valores de negócios e confiabilidade das informações

Apesar da concordância dos executivos sobre a expectativa geral das informações de alto valor, existem alguns eventos que exigem uma atenção imediata para aumentar o valor e a utilidade de dados. Fusões e aquisições corporativas, mudanças nos requisitos de conformidades regulatórias, aumento de atritos com clientes, aumento notável em atividades de call center, melhorias em migrações de sistemas de processos comerciais de contato direto com clientes, a introdução de CRM ou ERP ou o comprometimento com o MDM são exemplos de eventos que dependem de ajustes organizacionais em práticas de gerenciamentos de dados para serem bem-sucedidos.

Cada um desses eventos demonstra uma necessidade de instituir processos comerciais sólidos e as correspondentes práticas de gerenciamento de informações para aprimorar a experiência do cliente, enquanto mantém simultaneamente economias de escala. Em vez de cegamente jogar tecnologia, talvez várias vezes, em um problema presumido, na expectativa de que alguma coisa funcione, uma abordagem mais planejada busca um retorno sobre o investimento em tecnologia, que é alcançado quando a necessária prática de gerenciamento de informações é suportada por serviços de dados centrais, implementados uma vez, mas implantados várias vezes de maneira consistente.

Esta abordagem combina práticas de gerenciamento de informações fundamentais com aqueles serviços que aumentam o valor e a utilidade das informações comerciais, como o controle de qualidade dos dados, o aprimoramento de dados e a Location Intelligence. Permitir esses serviços de dados centrais requer

- Aprimorar o processo de produção de informação que suporta inúmeros consumidores de dados
- Padronizar as ferramentas e técnicas através da empresa,

- Padronizar a definição e o uso subsequente de regras empresariais e
- Padronizar a análise espacial/serviços de inteligência geográfica associados com as operações e transações.

Prever e padronizar o catálogo de serviços de dados com valor agregado são tarefas que vão eliminar a situação em que os produtos iguais ou competitivos são adquiridos várias vezes, ou quando diferentes grupos implementam a mesma técnica sem terem conhecimento dos planos dos outros. No entanto, esta abordagem finalmente solidificará e reutilizará o investimento em ferramentas de gerenciamento de dados para obter múltiplos retornos sobre o investimento.

## Serviços de dados centrais

Apesar de os critérios que caracterizam a capacidade como serviço “principal” não serem completamente claros e objetivos, eles podem ser resumidos em termos de como suportam as práticas recomendadas de gerenciamento de informações da empresa. Por exemplo, pode-se esperar que qualquer serviço que aprimora a interoperabilidade dos processos comerciais, reduz a complexidade do desenvolvimento de aplicativos, padroniza as técnicas, integra a definição e a aplicação das regras empresariais ou aumenta a utilidade e a qualidade das informações dos clientes possa ser considerado “serviços essenciais”.

Mas, para garantir o cumprimento previsível de uma boa prática de gerenciamento de informações, uma lista razoável deve incluir pelo menos estes serviços de integração de dados, qualidade de dados e Location Intelligence:

- Acesso e integração de dados
- Perfil de dados
- Análise
- Padronização
- Resolução de identidade

## Serviços de dados centrais: componentes básicos para estabelecer valor de negócios

- Ligação e mesclagem de registro
- Limpeza de dados
- Aprimoramento de dados
- Codificação geográfica
- Mapeamento e análise espacial
- Roteamento
- Inspeção e monitoramento de dados

### Acesso e integração de dados

As técnicas empregadas no desenvolvimento de armazém de dados tem se tornado universais, especialmente à medida que cresce a necessidade da interoperabilidade dos sistemas. Como mais pessoas desejam replanejar seus ambientes para permitir o compartilhamento, ou até mesmo a consolidação do conjunto de dados isolados, existe uma necessidade crescente para a integração de serviços de dados básicos, fornecendo a capacidade de conectar com perfeição o acesso de dados de várias fontes diferentes e distribuir esses dados a vários destinos. A capacidade dos servidores remotos de acessar os dados de suas fontes originais suporta a implementação eficiente e a implantação de práticas recomendadas de gerenciamento de dados, que podem ser implementados utilizando-se os serviços de dados centrais descritos neste documento. Os serviços de integração dependem da capacidade de conexão com infraestrutura frequentemente usadas e ambientes de dados aliados com serviços de análise e padronização, para transformar os dados extraídos em formatos que são adequados para a distribuição.

Na empresa moderna, o compartilhamento de dados, a federação e a sincronização estão rapidamente se tornando o padrão. À medida que mais empresas fazem a transição para a arquitetura orientada a serviços, uma integração de dados perfeita se torna a base fundamental para os aplicativos da empresa.

### Perfil de dados

Uma avaliação da qualidade de dados objetiva é um processo de análise e descobrimento, que, por sua vez, se baseia em uma revisão objetiva dos valores de dados que preenchem o conjunto de dados alvo. Uma análise estatística quantitativa ajudará os analistas de dados a identificar possíveis falhas como os valores atípicos, valores ausentes, elementos de dados que não respeitam os padrões definidos ou os domínios de valores, ou que são inconsistentes no que diz respeito a outros elementos de dados.

O perfil de dados é um conjunto de algoritmos para a análise estatística e avaliação de qualidade dos valores de dados em um conjunto de dados, assim como a exploração de relações que existem entre elementos de dados entre os conjuntos de dados. O perfil de dados oferece a capacidade de identificar falha nos dados, assim como uma maneira de comunicar essas falhas a especialistas no assunto que, com os seus conhecimentos no assunto, podem confirmar a existência de problemas nos dados. A ferramenta de perfil de dados verifica todos os valores em uma coluna e fornece uma distribuição de frequência dos valores desta coluna, sugerindo o tipo e o potencial de uso de cada coluna. A análise cruzada de colunas pode expor dependências de valores integrados enquanto a análise entre as tabelas explora o conjunto de valores sobrepostos para identificar as Principais relações internacionais entre as entidades. O perfil de dados pode ajudar na descoberta de regras de negócios integradas nos conjuntos de dados, que podem ser usados para inspeções e monitoramento contínuos.

Parafraçando Lord Kelvin, “Se você não pode medir, não pode melhorar”. O perfil de dados fornece um mecanismo concreto para analisar, avaliar, quantificar e medir a qualidade de dados através de diferentes dimensões, como totalidade, validade e razoabilidade. O perfil de dados é atualmente reconhecido como a primeira etapa na maioria das migração de dados, integração de dados e projetos de gerenciamento de dados mestres, para conhecer quais problemas em potencial existem no conjunto de fonte de dados e auxiliar no planejamento de transformações que limparão os dados quando forem migrados para o destino.

# UM CONJUNTO CONSOLIDADO DE SERVIÇOS DE DADOS COM VALORES AGREGADOS PODE REDUZIR OS CUSTOS OPERACIONAIS E AUMENTAR A EFICIÊNCIA, ENQUANTO APRIMORA A QUALIDADE E A UTILIDADE DOS DADOS EMPRESARIAIS.

## Análise

Em muitos casos, existe a expectativa de que os valores de dados estejam em conformidade com um tamanho e formatos predefinidos. No entanto, pequenas variações ou ambiguidades em representações de valores de dados podem confundir indivíduos, assim como aplicativos automatizados. Por exemplo, considere estes valores diferentes de dados: {Califórnia, CA, Calif., US-CA, Cal, 06}. Alguns desses valores usam caracteres, outros usam dígitos e alguns usam pontuações e caracteres especiais. Na maioria das vezes, uma pessoa lendo isso reconheceria que eles representam o mesmo valor conceitual: o estado da Califórnia. No entanto, a automatização do processo de determinar se esses valores são exatos, ou investigar se existem valores duplicados, estes valores devem ser analisados em seus segmentos de componentes e então transformados em um formato padrão.

Descrever os diferentes componentes e padrões de formatos usados para representar um objeto de dados (nome de pessoa, descrição de produto, etc.) é algo que apoia o serviço principal de análise, que é usado para determinar se um valor está em conformidade com um padrão reconhecido como parte de um processo de avaliação de correspondência e limpeza. A análise baseada em padronização permite o reconhecimento automático e subsequente padronização de componentes de valores significativos. Este tipo de serviço incorpora, por necessidade, a capacidade de definir conjuntos de padrões, listas de valores de dados pertencentes a domínios de valores definidos, ou expressões regulares e gramáticas.

Quando a mesma informação é representada várias vezes de maneira diferente, torna-se difícil determinar que diferentes registros (no mesmo conjunto de dados, ou entre diferentes conjuntos de dados) se referem ao mesma entidade. Mas, mesmo antes que se possam aplicar ligações algorítmicas de registros mais complexas, é necessário analisar que o texto é o campo de dados crítico para identificar e subsequentemente segregar e reordenar os tokens-chave que contêm informações de identificação. Junto com a padronização, a análise ajudará na redução de variações entre múltiplos registros e simplificará correspondências e ligações.

## Padronização

A análise usa padrões definidos, expressões regulares ou gramáticas gerenciadas por um mecanismo de regras junto com uma tabela de consulta para distinguir entre valores válidos e inválidos. Quando os padrões são reconhecidos, outras regras e ações podem ser iniciadas para transformar os dados de entrada em um formulário que possa ser usado com mais eficácia, para padronizar a representação (presumindo-se uma representação válida) ou corrigir os valores (caso sejam identificados erros conhecidos). Continuando com o nosso exemplo, cada um dos valores de Califórnia {Califórnia, CA, Calif., US-CA, Cal, 06} pode ser padronizado com a abreviação de dois caracteres padrão dos correios dos Estados Unidos como CA.

O serviço de padronização utiliza o serviço de análise, que pode ser combinado com uma biblioteca de domínios de dados para dividir valores de dados em vários componentes e reorganizá-los em um formato normalizado. A padronização pode também mudar palavras inteiras para abreviações, ou abreviações para palavras inteiras, transformar apelidos em um formato de nome padrão, traduzir em vários idiomas (por exemplo, espanhol para inglês) e corrigir erros ortográficos comuns.

A padronização fornece dois diferentes tipos de valores. Em primeiro lugar, ao mapear representações variantes de padrões bem definidos, o processo de padronização pode ajudar a normalizar representações diferentes e limpar com eficácia os dados falhos. Em segundo lugar, mapear um valor para uma representação padronizada (mesmo sem a limpeza específica) reduz a variação em estrutura de valor, aperfeiçoando a ligação de registros, com o objetivo de remover duplicações e consolidar os dados principais. Padronizar apelidos (como "Bob", "Rob" ou "Bobby" para "Robert") aumentará a possibilidade de que dois registros que representam a mesma entidade sejam ligados em conjunto, e isso melhora significativamente a resolução unificada, que é necessária para um programa bem-sucedido de gerenciamento de qualidade de dados e dados mestres.

## Serviços de dados centrais: componentes básicos para estabelecer valor de negócios

### Resolução de identidade

Uma vez que os sistemas operacionais têm crescido organicamente para um conjunto de aplicativos empresariais, não é difícil que várias instâncias de dados em sistemas diferentes tenham várias formas diferentes de se referir à mesma entidade no mundo real. Alternativamente, o desejo de consolidar e ligar dados sobre os mesmos conceitos empresariais com o mesmo nível de confiança pode convencer alguém que o registro pode não existir para uma entidade no mundo real, quando de fato ele realmente existe. Esses dois problemas representam o mesmo desafio central: ser capaz de comparar dados de identificação em alguns registros para determinar a semelhança entre este par ou distinguir as entidades representadas nesses registros.

As duas questões são tratadas através de um processo chamado resolução de identidade, em que o nível de semelhança entre qualquer dupla de registros é computada, na maioria das vezes com base no peso da semelhança aproximada entre um conjunto de valores de atributos entre os dois registros. Se a pontuação ficar acima de um nível específico, os dois registros são considerados uma correspondência e são apresentados ao cliente final como provável possibilidade de representar a mesma entidade. A resolução de identidade é usada para reconhecer quando somente pequenas variações sugerem que registros diferentes são conectados e onde os valores podem ser limpos, ou onde diferenças suficientes entre os dados sugerem que os dois registros realmente representam entidades distintas. Como a comparação de cada registro com muitos outros registros é computacionalmente intensa, muitos serviços de dados usam algoritmos avançados para bloquear registros que têm probabilidade de conter correspondências com conjuntos menores, reduzindo o tempo de computação.

A resolução de identidade é um componente crítico da maioria dos aplicativos de gerenciamento de dados essenciais, de qualidade de dados e de inteligência comercial. O desejo de um “panorama de 360 graus do cliente” ou de haver um catálogo de produto abrangente é previsto na capacidade de

localizar todos os registros que contêm informações sobre cada entidade única e sua resolução em uma visualização unificada. A análise e a padronização são etapas preliminares, mas a técnica-chave é a resolução de identificação.

### Ligação, mesclagem e consolidação de registros duplicados

A resolução de identificação fornece a base de um ou mais serviços sofisticados de dados essenciais: análise de registros duplicados e eliminação. A identificação de registros similares no mesmo conjunto de dados provavelmente significa que os registros são duplicados e podem estar sujeitos a limpeza e/ou eliminação. A identificação de registros similares em conjuntos diferentes pode indicar um link através de conjunto de dados, que ajuda a facilitar a mesclagem ou os registros similares com o propósito de se fazer limpeza dos dados e se suportar o gerenciamento de dados essenciais ou a iniciativa de integração de dados do cliente.

A resolução de identidade leva ao processo de mesclagem e consolidação para estabelecer uma única visualização do cliente. Embora alguns aplicativos empresariais consigam continuar consultando um conjunto de registros contendo informações, a automatização do processo de mesclagem selecionará os melhores valores desses registros e permite a criação de uma única “cópia ideal” dos dados que podem suportar tanto o processo operacional como o analítico, aperfeiçoando as maneiras de conduzir o negócio.

### Limpeza de dados

Após determinar que um valor de dados não está em sintonia com as expectativas do usuário final, o analista de dados pode optar por diferentes abordagens para solucionar o problema. Uma abordagem é transformar os dados em um formato que atenda ao nível de aceitação do usuário empresarial. Esse serviço de limpeza de dados utiliza a análise, a padronização e o aprimoramento de serviços, assim como a resolução de identidade e ligação de registros. Ao analisar os valores e iniciar

# O PERFIL DE DADOS É ATUALMENTE RECONHECIDO COMO A PRIMEIRA ETAPA NA MAIORIA DAS MIGRAÇÕES DE DADOS, INTEGRAÇÕES DE DADOS E GERENCIAMENTOS DE PROJETOS DE DADOS ESSENCIAIS

um padrão de erros conhecidos, o serviço de limpeza de dados pode aplicar as regras de transformações para inserir valores de dados, nomes ou endereços corretos, eliminar dados extrínsecos e/ou sem significado e até mesmo mesclar registros duplicados.

Embora a prática recomendada em qualidade de dados seja determinar onde está a fonte de erros e eliminar a raiz do problema da baixa qualidade dos dados, frequentemente a equipe de qualidade de dados não possui o controle administrativo para aplicar mudanças nos aplicativos atuais de produção. Portanto, uma alternativa que pode ser necessária para assegurar a alta qualidade dos dados é a limpeza direta dos dados e, pelo menos, fornecer dados que atendam a um nível de concordância para todos os consumidores de dados, e o serviço de limpeza de dados possibilita isso.

## Aprimoramento de dados

Uma abordagem diferente para o aprimoramento de valores de dados envolve o aperfeiçoamento, no qual as informações adicionais são anexadas a registros existentes. O aperfeiçoamento de dados, que utiliza a análise, a padronização e a ligação de registros, é um processo de aprimoramento que adiciona informações dos conjuntos de dados de terceiros (como padronização de nome, importação de dados demográficos, importação de dados psicográficos e anexos de listas de famílias). A lista de aperfeiçoamento é simplificada significativamente quando o fornecedor tem parceria com os agregadores de dados, cujos dados podem ser usados tanto com um sistema de registro em relação ao qual as instâncias de dados são correspondidas, quanto como um recurso para anexar. Uma atividade típica de aperfeiçoamento de dados, a padronização e a limpeza de endereços, depende da análise, da padronização e da disponibilidade das informações de endereços de terceiros.

Existem alguns setores que não se beneficiam com o uso de fontes de dados adicionais para anexar e aperfeiçoar as informações empresariais. Se isso significa padronização e correção de endereços para reduzir os custos de distribuição e aprimorar taxas de respostas diretas, apêndices geodemográficos que fornecem informações de perfis e segmentação de clientes, dados de créditos usados para

acessar e reduzir risco, informações históricas (como histórico de conserto de automóveis ou eventos médicos) para fornecer ideias melhores sobre o diagnóstico e o cuidado preventivo, entre outros, o aperfeiçoamento dos dados é um processo que otimiza as decisões de negócios em um contexto analítico e operacional.

## Localização geográfica e codificação geográfica

A codificação geográfica é um processo de Location Intelligence para traduzir um local conceitual em coordenadas específicas em um mapa da superfície da terra. O código geográfico é um par de coordenadas para a latitude e longitude, e o serviço de localização geográfica (ou “geolocalização”) fornece um par de coordenadas “latlong” quando for fornecido com um endereço. Quando combinado com o serviço de aperfeiçoamento de dados, é possível determinar valores (codificados geograficamente) de um nome de local, endereço de rua, código postal ou de uma interseção entre duas ruas. Este serviço também funciona de forma inversa se você fornecer os valores da latitude/longitude.

Muitos setores dependem da localização geográfica para suportar a avaliação de risco de processos comerciais, especialmente nos setores financeiros e de seguros. Empresas de varejo de todas as modalidades e formas estão constantemente revisando locais para determinar o lugar ideal para uma nova loja, filial, etc., e a combinação de mobilidade crescente e uma capacidade aprimorada de distribuir informações “a qualquer hora, em qualquer lugar” sugere que os aplicativos empresariais terão uma necessidade crescente de serem “conscientes em termos de localização”.

## Mapeamento e análise espacial

Espera-se que os aplicativos que rastreiam os locais de interações comerciais muitas vezes anotem a posição das informações espaciais (por exemplo, localização, local de interesse, rota ou endereço) no mapa. O serviço de mapeamento permite a edição, a análise espacial e a exibição de informações espaciais em um mapa. Por sua vez, o usuário final pode pensar em localizações adicionais ou pontos de interesses em uma

## Serviços de dados centrais: componentes básicos para estabelecer valor de negócios

proximidade relativa a posições identificadas e um serviço de diretório que corresponde com a análise espacial. Esta análise fornece dados de referência para realizar pesquisas de proximidade em categorias definidas (como um guia de restaurante ou uma lista de parques públicos) e fornecer uma lista das correspondências mais próximas.

A análise espacial, ao fornecer uma revisão quantitativa do fenômeno e das atividades no espaço, permite a derivação de conhecimentos sobre o contexto geográfico das informações existentes. Saber onde um indivíduo mora, ou onde um evento acontece, fornece um vetor de dados; a análise espacial permite que se responda questões mais interessantes, como:

- Quantos clientes semelhantes existem dentro de uma área especificada?
- Quais são as taxas de mortalidade por doenças em diferentes partes de um país?
- Quais locais de varejo melhor atendem a uma região em particular?
- Quantas liquidações ocorreram em um determinado bairro?
- Quantos clientes são atendidos em cada código postal?
- Como o rendimento médio de uma região mudou em um período específico?

Cada evento ou transação comercial acontece em algum lugar; a análise espacial permite que você busque padrões baseados em localizações e utilize os resultados para aprimorar o desempenho da empresa.

### Roteamento

O roteamento é um serviço de Location Intelligence para planejar as rotas entre dois pontos, determinar o caminho ideal entre vários pontos e criar matrizes e polígonos do tempo de duração por carro. O serviço de roteamento pode calcular o caminho com base em diferentes critérios, como distância

e tempo, e pode fornecer resultados baseados em otimização para a menor distância de percurso ou o menor tempo de trajeto por carro. Com o serviço de roteamento, as empresas podem manter níveis altos de serviços e custos eficientes de serviços de despacho, identificar áreas de coletas e aprimorar a tomada de decisões.

De um ponto de vista operacional, o serviço de roteamento pode ajudar a otimizar logísticas e distribuição e aprimorar o serviço ao cliente. Encontrar as melhores rotas reduz os custos de energia e pode permitir um melhor organograma para aumentar a produtividade de distribuição. Ao mesmo tempo, a redução do tempo de distribuição melhora o serviço ao cliente e aumenta a satisfação.

### Inspeção e monitoramento

Uma vez que o perfil de dados expõe as regras de negócios em potencial usadas em diferentes processos comerciais, a análise de dados pode documentar as regras e confirmar sua importância em coordenação com o especialista no setor comercial. Essas regras descrevem essencialmente as expectativas do usuário final e podem ser usadas para medir e monitorar a validade dos dados. A auditoria e o monitoramento das regras de qualidade de dados definidas fornecem uma avaliação proativa da conformidade com as expectativas. Os resultados dessas auditorias podem alimentar métricas de qualidade de dados, preenchendo painéis e cartões de pontos de gerenciamento.

O serviço de auditoria e monitoramento utiliza os serviços de análise e perfil de dados para, de forma proativa, validar os dados em relação a um conjunto de regras de negócios definidas (ou descobertas). Essa inspeção de dados pode notificar os analistas quando as instâncias de dados não estão em conformidade com as expectativas da qualidade dos dados, assim como as medidas da linha de base em relação às quais se podem comparar as inspeções de qualidade de dados em andamento.

## PITNEY BOWES BUSINESS INSIGHT FORNECE UMA SOLUÇÃO EMPRESARIAL QUE TORNA POSSÍVEL CENTRALIZAR O CONTROLE E APRIMORAR O DESEMPENHO

Já observamos que o perfil de dados permite a análise e o descobrimento de problemas de dados em potencial que podem causar impacto no sucesso da empresa. Mas, ao dar suporte à inspeção e ao monitoramento de dados, o perfil de dados preenche outra necessidade crítica dos negócios em relação à garantia contínua da qualidade de dados. O serviço de perfil de dados promove a governança dos dados operacionais e alimenta informações para a proteção sobre as falhas de dados que podem ser revistas e corrigidas antes que qualquer impacto negativo na empresa possa acontecer.

### Apresentando os serviços de informações comerciais

Cada vez mais empresas estão buscando oportunidades para reduzir os custos de planejamento, desenvolvimento e manutenção da tecnologia internamente desenvolvida e externamente adquirida. A combinação dos serviços de dados essenciais com produtos de informações agregadas em serviços padronizados de informações comerciais vai apoiar aplicativos diferentes em toda a empresa. Em geral, esses serviços tratam de necessidades comuns através de uma gama de aplicativos empresariais e fornecem versões padronizadas que ajudarão na redução de custos, enquanto introduzem eficiências operacionais. Eis alguns exemplos comuns:

- **Validação da qualidade de dados** – Quando os dados geralmente usados são adquiridos ou criados por mais de um canal (por exemplo, dados de clientes digitados em um site ou transcritos por um representante de call center), existem oportunidades para variações em ortografia, ordem de nomes ou ausência de títulos ou honorárias. Regras aplicadas inconsistentemente podem levar a inserção duplicada dos dados, ou poderão impedir que os registros sejam localizados nos atuais conjuntos de dados.

Para lidar com essa possível inconsistência, pode-se compor um serviço de validação da qualidade de dados usando a padronização de análise, limpeza e serviços de perfil. Esse serviço de validação de dados pode aplicar conjuntos comuns de padronização de dados e regras de validação a múltiplos canais de aquisição de dados em toda a organização. A padronização das limitações e das transformações da qualidade de dados em todos os pontos de captura e troca de dados reduzirá significativamente os riscos

de duplicação ou os erros de correspondência, e levarão à consistência dos dados em toda a empresa.

- **Avaliação de risco da localização da propriedade** – Qualquer setor que dependa da avaliação dos riscos relacionados à localização e às propriedades se beneficiaria de um serviço de padronização de avaliação de risco. Por exemplo, as seguradoras de acidentes e propriedades avaliam o risco associado em garantir uma apólice de seguros de propriedade, incluindo os riscos de catástrofes naturais como furacões ou terremotos, crimes na região, e disponibilidade e eficácia dos serviços de emergência. As seguradoras de financiamento de imóveis estão interessadas em classificar as propriedades (por exemplo, casa individual em vez de condomínio), ocupação do proprietário, valores de propriedades similares e inadimplência e execução de hipotecas na região. Ainda assim, diferentes aplicativos empresariais dentro da mesma empresa podem aplicar suas regras de avaliações de risco de maneiras diferentes, usando técnicas e ferramentas diferentes.

Um serviço padronizado de avaliação de risco da localização da propriedade poderia incorporar a análise de endereços e aperfeiçoamento de dados para fornecer um endereço padronizado. Esse endereço padronizado pode ser submetido aos serviços de mapeamento e codificação geográfica, aliados aos conjuntos de dados agregados que contêm riscos de propriedade, zoneamento de catástrofe naturais e outros atributos relacionados com a codificação geográfica da localidade.

- **Proteção de dados operacionais** – Em muitas empresas, existe um interesse crescente em operacionalizar as políticas de governança de dados e procedimentos através da introdução da função do protetor de dados, que é responsável por assegurar que as informações atendem às necessidades dos que trabalham com informações. As expectativas com a qualidade dos dados são definidas pelo usuário final da informação, mas, como os conjuntos de dados criados em um contexto são usados em múltiplos processos comerciais, os requisitos de qualidade dos dados coletados precisam ser

## Serviços de dados centrais: componentes básicos para estabelecer valor de negócios

impostos em vários estágios de processamento através de múltiplos processos de ponta a ponta.

O serviço de proteção de dados operacionais pode ser usado para instituir inspeções e monitoramento de qualidade de dados aplicados consistentemente em um processo comercial contínuo. As regras de qualidade de dados podem ser definidas uma vez e implantadas usando os serviços de perfil de dados para validar os registros individuais ou o conjunto de registros, ou os já implantados ou quando são transferidos entre atividades. Um serviço de proteção de dados comum inspecionará automaticamente a validade dos dados e notificará o protetor dos dados quando a expectativa da qualidade não for alcançada. Além disso, estatísticas cumulativas podem ser coletadas e fornecidas através do painel e da pontuação da governança de dados.

- **Perfil demográfico** – Um aspecto do marketing focalizado envolve uma combinação de dados individuais confiáveis e dados demográficos de região. Da mesma forma que “cada macaco fica no seu galho”, a compreensão das características gerais associadas com a localidade de um cliente em potencial informa a abordagem do marketing focalizado.

Fornecendo o nome e o endereço de um cliente em potencial, o serviço de perfil demográfico pode analisar e padronizar o nome e aplicar a resolução de identidade para corresponder a informação com os clientes conhecidos, ou identificar pessoas relacionadas dentro de um conjunto de dados de clientes conhecidos. Em seguida, o endereço pode ser analisado e padronizado. Por sua vez, os dados agregados associados com uma localidade mapeada podem ser usados para adicionar dados demográficos valiosos, relacionados à região em que a localidade está mapeada. A combinação de dados demográficos com o registro do cliente em potencial fornecerá à equipe de vendas

recomendações dirigidas com base no perfil materializado.

- **Gerenciamento da jurisdição fiscal** – Outro aspecto do gerenciamento de conformidade e risco envolve a determinação de taxas fiscais por jurisdição. Embora esse tipo de serviço seja importante para a análise de risco baseada em propriedades, também é necessário para as empresas que tenham responsabilidades fiscais corporativas associadas a várias localidades.

Esse tipo de serviço combina a padronização e os serviços de limpeza, os serviços de codificação geográfica e localização, e dados fornecidos pelo agregador, para fornecer informações sobre o estado, o país, a cidade, o município e o distrito especial de impostos associado a um endereço, monitorando as constantes mudanças na jurisdição fiscal e nas regras municipais.

### Considerações

Esses são apenas alguns exemplos de serviços comerciais que podem ser replanejados com os serviços de dados centrais. O interessante é que eles têm a tendência de exibir essas três características como um subproduto da combinação dos serviços de dados centrais.

- A análise, a padronização e a inspeção de dados e serviços de monitoramento para normalizar e limpar os dados,
- A combinação de limpeza dos dados e de inteligência geográfica para a preparação do aperfeiçoamento de dados e
- O aperfeiçoamento e o aprimoramento dos valores através da adição de informações de conjuntos de dados agregados.

Se existir um consenso de que a prática eficaz de gerenciamento de dados, como a governança de dados, o controle da qualidade de dados e o aperfeiçoamento dos dados pode melhorar a lucratividade, então vale a pena considerar abordagens para a implantação de serviços de informações comerciais

## PITNEY BOWES BUSINESS INSIGHT: SUA FONTE DE PERGUNTAS E RESPOSTAS

padronizadas, utilizando-se da abordagem de serviços de dados. Uma avaliação das necessidades de informações comerciais destaca os serviços comerciais que, se padronizados através da empresa, podem aprimorar a qualidade e a consistência, enquanto diminuem os custos operacionais.

Um pensamento final: podem existir muitas maneiras de planejar e implantar os serviços de dados centrais, passando por desenvolvimento interno, aquisição de suites de produtos, até a total terceirização. Mas, se fica claro que os serviços de informações comerciais essenciais dependem da combinação de qualidade de dados, da Location Intelligence e do conjunto de dados com valores agregados, a qualidade desses aspectos deve ser incorporada em qualquer decisão de desenvolver ou comprar. Em outras palavras, o critério de avaliação da solução deve incorporar até que ponto os produtos suportam os serviços de qualidade dos dados centrais, serviços de Location Intelligence e a integração com o conjunto de dados com valores agregados.

#### ESTADOS UNIDOS

One Global View  
Troy, NY 12180  
telefone: +1.518.285.6000  
1.800.327.8627  
www.pbinsight.com  
la.sales@pb.com

#### AMÉRICA LATINA E CARIBE

ESCRITÓRIO MIAMI  
15100 NW 67th Avenue Suite 403  
Miami Lakes, FL 33014  
U.S.A.  
telefone: +1.305.825.8381  
fax: +1.305.826.1246  
la.sales@pb.com

ESCRITÓRIO BRASIL  
Rua Dom Aguirre, 438 - 3o andar  
04671-903 São Paulo, SP  
Brasil  
telefone: +55.11.5698.5555  
www.pbinsight.com.br  
la.sales@pb.com

ESCRITÓRIO MÉXICO  
San Francisco 1393  
Colonia Del Valle  
Delegación Benito Juárez  
03210 México, DF  
México  
telefone: +52.55.5636.0949  
la.sales@pb.com