

Softwares especializados em data cleansing, disponíveis comercialmente, podem, de forma automática, pesquisar arquivos de dados, corrigir erros e integrá-los em formato consistente para toda a empresa.

A Seção Interativa sobre Pessoas ilustra a experiência da American Water em gerenciar os dados como um recurso. À medida que lê esse caso, tente identificar as políticas, procedimentos e tecnologias que foram necessárias para melhorar a gestão de dados nessa empresa.

## SEÇÃO INTERATIVA: PESSOAS

## A AMERICAN WATER MANTÉM OS DADOS FLUINDO

A American Water, fundada em 1886, é a maior concessionária de serviços de água nos Estados Unidos. Sediada em Voorhees, Nova Jersey, a empresa emprega mais de 7 mil profissionais dedicados que fornecem serviços de água potável, águas residuais e outros serviços relacionados a aproximadamente 16 milhões de pessoas em 35 estados, bem como Ontário e Manitoba, no Canadá. A maior parte dos serviços da American Water suporta concessionárias subsidiárias administradas localmente, que são reguladas pelo estado norte-americano em que cada uma opera, bem como pelo governo federal. A American Water também possui subsidiárias que administram, sob contrato, os sistemas municipais de água potável e de águas residuais e outras que fornecem produtos e serviços referentes à gestão de água a empresas e comunidades residenciais.

Até recentemente, os sistemas e os processos de negócios da American Water eram muito concentrados e muitos desses processos eram manuais. Ao longo do tempo, o ambiente dessas informações tornou-se cada vez mais difícil de gerenciar. Muitos sistemas não eram integrados, de modo que a elaboração de qualquer tipo de relatório que tivesse que fornecer informações sobre mais de uma região era um processo totalmente manual. Os dados tinham que ser extraídos dos sistemas de suporte de cada região e, em seguida, combinados manualmente para criar a saída desejada. Quando a empresa estava se preparando para realizar uma oferta pública inicial de suas ações em 2006, os seus sistemas de software não podiam lidar com os controles regulatórios necessários, de forma que cerca de 80% desse trabalho tinha que ser realizado manualmente. Era quase um pesadelo.

A administração queria que a empresa mudasse de um grupo descentralizado de empresas regionais independentes para uma organização mais centralizada com processos de negócio padronizados e elaboração de relatórios contemplando toda a empresa. O primeiro passo para alcançar esse objetivo era implementar um sistema de planejamento de recursos empresariais (*Enterprise Resource Planning* — ERP) desenvolvido para substituir os diversos sistemas por uma única plataforma de software integrado. A empresa selecionou a SAP como o seu fornecedor de sistema ERP.

Um passo importante desse projeto foi a migração dos dados dos sistemas antigos da American Water para a nova plataforma. Os dados da empresa estavam armazenados em vários sistemas diferentes em diversos formatos. Cada empresa regional mantinha alguns dos seus próprios dados em seus próprios sistemas, e uma parte deles era redundante e inconsistente. Havia, por exemplo, partes de materiais importantes duplicados porque um material poderia receber um nome na operação da empresa em Missouri e outro nome na empresa de Nova Jersey. Esses nomes tiveram que ser padronizados de modo que o mesmo nome dado para um material fosse utilizado em todas as unidades da empresa. Os usuários corporativos da American Water precisaram adotar essa nova visão de dados contemplando toda a empresa.

A migração de dados implica muito mais do que apenas transferi-los entre os sistemas novos e antigos. Os usuários corporativos precisam saber que os dados não são apenas uma responsabilidade do departamento de sistemas de informação: a empresa é a "dona" dessas informações. São as necessidades da empresa que determinam as regras e normas para a gestão dos dados. Portanto, cabe aos usuários corporativos catalogar e rever todas as porções de dados em seus sistemas para determinar precisamente quais porções do antigo sistema serão utilizadas no novo e quais dados não precisam ser trazidos. Os dados também precisam ser revistos para garantir que estejam corretos e consistentes e que os redundantes sejam eliminados.

O mais provável é que algum tipo de data cleansing (limpeza e padronização de dados) será necessário. A American Water, por exemplo, possuía informações sobre mais de 70 mil fornecedores em seu arquivo principal de dados de fornecedor. Andrew Clarkson, líder de inteligência empresarial da American Water, pediu para que os usuários corporativos definissem um fornecedor ativo e usassem essa definição para identificar quais dados deveriam ser migrados. Ele também trabalhou com vários grupos funcionais para padronizar como apresentar os dados desejados.

Um dos objetivos do trabalho de gestão de dados da American Water foi apoiar um programa de inteligência empresarial para toda a empresa, com base

em uma única visão do negócio. Um sistema analítico e um data warehouse seriam capazes de combinar os dados a partir do sistema ERP SAP com os de outras fontes, incluindo informações sobre novos clientes e de sistemas de gestão de ativos da empresa. Isso significava que os usuários corporativos da American Water tiveram que refletir sobre quais tipos de relatórios eles queriam. A empresa havia planejado, originalmente, que o sistema fornecesse 200 relatórios, mas depois reduziu esse número pela metade. Os usuários corporativos foram treinados para gerar esses relatórios e depois customizá-los. A maioria dos usuários financeiros inicialmente tentou criar seus relatórios usando o software de planilha Microsoft Excel. Ao longo do tempo, entretanto, eles aprenderam a fazer a mesma coisa usando as ferramentas do Business Objects Web Intelligence da SAP que vinha com o sistema. Esse produto é um conjunto de ferramentas que permite aos usuários corporativos visualizar, classificar e analisar

dados de inteligência empresarial. Isso inclui ferramentas para a geração de consultas, relatórios e painéis interativos.

No momento, a American Water está se concentrando em promover a ideia de que os dados devem ser "limpos" para serem eficazes e investiu um esforço incrível no trabalho de data cleansing — identificar as porções incompletas, incorretas, imprecisas e irrelevantes e, em seguida, substituir, modificar ou eliminar os dados "sujos".

De acordo com Clarkson, assim como as estações de tratamento de água têm medidas e medidores para verificar a qualidade da água conforme ela é tratada, a gestão de dados precisa assegurar a qualidade dos dados em todas as etapas para certificar-se de que o produto final será verdadeiramente útil para a empresa.

Fontes: David Hannon, "Clean, Smooth, Flowing Data at American Water", SAP Insider Profiles, jan.-mar. 2013 e Disponível em: <[www.amwater.com](http://www.amwater.com)>, acesso em: 2 mar. 2013.

## PERGUNTAS SOBRE O ESTUDO DE CASO

1. Discuta o papel da política de informação, da administração dos dados e dos esforços para assegurar a qualidade deles na melhoria da gestão de dados da American Water.
2. Descreva os papéis desempenhados pelos especialistas em sistemas de informação e pelos usuários finais no projeto de transformação dos sistemas da American Water.
3. Por que a participação dos usuários corporativos foi tão importante? Se eles não tivessem desempenhado esse papel, o que teria acontecido?
4. Como a implementação de um data warehouse ajudou a American Water a se tornar uma organização mais centralizada?
5. Dê alguns exemplos de problemas que teriam ocorrido na American Water se os seus dados não fossem "limpos".
6. Como o data warehouse da American Water melhoraria as operações e as tomadas de decisões gerenciais?

## Resumo

**1** Como um banco de dados relacional organiza os dados? O banco de dados relacional é, atualmente, o método mais usado para organizar e manter dados em sistemas de informação. Ele organiza-os em tabelas bidimensionais, com linhas e colunas chamadas relações. Cada tabela contém dados sobre uma entidade e seus atributos. Cada linha representa um registro e cada coluna, um atributo ou campo. Cada tabela também contém um campo-chave, que confere identificação exclusiva a cada registro para posterior recuperação ou manipulação. O diagrama entidade/relacionamento descreve graficamente a relação entre as entidades (tabelas) em um banco de dados relacional. O processo de fragmentar agrupamentos complexos de dados e simplificá-los, a fim de minimizar a redundância e as relações muitos para muitos inadequadas, é denominado normalização. Bancos de dados não relacionais estão se tornando populares para gerenciar os tipos de dados que não podem ser facilmente manipulados pelo modelo de dados relacional.