

Fornecendo informações confiáveis

Os benefícios de dados de qualidade e do IBM System z



Resumo executivo

Você decidiria se submeter a uma operação ou procedimento médico para salvar sua vida com base em dados imparciais ou incompletos? Imagine que sim – e, mais tarde, você descobrisse que o diagnóstico estava incorreto e o tratamento não era necessário. Você teria corrido risco de morte por causa de dados de má qualidade.

No mundo dos negócios, problemas de qualidade de dados podem não ser uma questão de vida ou morte, mas, ainda assim, representam uma grave ameaça. Um número cada vez maior de empresas e organizações governamentais está descobrindo que dados de má qualidade afetam sua capacidade de dar suporte aos processos de negócios, cumprir com os regulamentos e tomar decisões precisas.

Embora a maioria dos líderes de negócios entenda a necessidade de dados de alta qualidade, muitas vezes não sabe com certeza como obtê-los. No entanto, tornou-se claro que é necessário investir em infraestrutura para assegurar uma qualidade de dados mensuravelmente aceitável.

É aí que o IBM® InfoSphere® Information Server para IBM System z® pode fazer uma grande diferença. O InfoSphere Information Server é a base de muitas iniciativas de qualidade de dados bem-sucedidas, pois ajuda as organizações a obterem mais valor a partir das informações complexas e heterogêneas espalhadas pelos seus sistemas. Enquanto isso, o System z oferece uma plataforma de alto desempenho resiliente e confiável para dados críticos; estima-se que 95% das empresas listadas na Fortune 1000 armazenem dados de negócios no System z1.

Data quality can determine business

Entender com clareza os clientes, parceiros e fornecedores pode ser a diferença entre fazer um negócio crescer e não conseguir competir. Para que iniciativas críticas de governança da informação, conformidade e integração de dados principais funcionem, é necessário que a qualidade dos dados contidos nos sistemas seja entendida com clareza e gerenciada de modo ativo.

Em outras palavras, dados de má qualidade são como um vírus. Os vírus são pequenos; contudo, se não forem detectados ou se forem diagnosticados incorretamente, podem se espalhar e se tornar mais agressivos, acabando por danificar o host. Imagine que um dos seus principais sistemas de informação contém alguns dados não verificados ou imprecisos. Os dados podem começar em parte do seu sistema estratégico de gerenciamento de dados principais (MDM) utilizado por sistemas de departamentos, data warehouses, sistemas de inteligência de negócios (BI), subsidiárias, parceiros comerciais ou equipes de relatórios regulamentares para tomar decisões de negócios essenciais. Os dados de má qualidade também podem ser usados por profissionais em toda a cadeia de fornecimento de informações – e estes irão consumir, processar e, em seguida, disseminar as informações "infectadas" para outras pessoas. Ao longo do caminho, irão distorcer métricas, reduzir a precisão dos relatórios e, por fim, afetar as decisões de negócios.

Os efeitos da má qualidade dos dados incluem falhas nos processos de negócios, menor produtividade e desperdício de materiais. Informações perdidas, imprecisas ou incompletas também geram custos mais elevados e trabalho extra, como localização de informações ou reconciliação adicional.

O IBM System z é uma excelente plataforma para o compartilhamento de localização de aplicativos, processos, transações, sistemas de BI e data warehouses críticos para os negócios; além disso, pode gerar muitas vantagens operacionais, de desempenho e de segurança. Entretanto, se não forem resolvidos, dados de má qualidade influirão nos benefícios da plataforma do System z.

Governança da informação

A maioria das organizações ainda não desenvolveu seus processos, políticas e infraestrutura para que possa ajudar a assegurar a alta qualidade dos dados. Como resultado, as organizações estão começando a adotar a governança da informação, isto é, uma disciplina de controle de qualidade que busca aumentar o rigor do processo de definição da terminologia comum e de gerenciar, utilizar, melhorar e proteger informações.

A governança eficiente da informação pode melhorar a qualidade, disponibilidade e integridade dos dados de uma empresa ao promover a colaboração entre as organizações e a

elaboração de políticas estruturadas. Ela equilibra silos faccionários com interesses organizacionais, o que afeta diretamente os quatro fatores que mais interessam às organizações:

- Aumento da renda
- Diminuição dos custos
- Redução dos riscos
- Aumento da confiança dos dados

A governança da informação permite que uma organização monitore sua cadeia de fornecimento de informações como um sistema de ponta a ponta, ajudando a assegurar que as informações sejam definidas e bem-entendidas de modo consistente; de alta qualidade; gerenciadas ao longo do seu ciclo de vida; e protegidas e asseguradas onde quer que estejam. Com a governança da informação, as organizações atingem muitas metas, desde uma melhor tomada de decisões até simplificar e fortalecer a conformidade regulamentar.

Um fórum para a governança da informação

Mais do que nunca, o desafio de proteger e gerenciar dados se tornou uma preocupação universal para as organizações. Para ajudar a entender melhor o setor emergente, a IBM criou, em novembro de 2004, um fórum de liderança para diretores executivos de dados e executivos de segurança, risco, conformidade e privacidade que estão preocupados com problemas de governança da informação.

Desde então, o IBM Information Governance Council tem crescido de forma gradual e passou a abranger quase 55 empresas, universidades e Parceiros de Negócios IBM líderes, incluindo instituições financeiras de grande porte, organizações de telecomunicação, varejistas e até mesmo órgãos do governo. O conselho desenvolveu uma estrutura de governança da informação para ajudar os negócios a entenderem as disciplinas e atores – tanto os principais quanto os de suporte – da governança da informação. Também produziu um modelo de maturidade para ajudar a avaliar a governança da informação dentro de uma organização.

Para mais informações sobre o IBM Information Governance Council, visite: ibm.com/software/data/system-z/data-governance/workshops.html

IBM InfoSphere Information Server para System z

O sucesso de um programa de governança da informação e de iniciativas de qualidade de dados de suporte depende de uma infraestrutura tecnológica robusta para a integração de dados. Por isso que existe o InfoSphere Information Server para System z, uma plataforma de software da IBM que proporciona produtividade e desempenho revolucionários para entender, purificar, transformar e mover informações por toda a empresa, de maneira consistente e segura.

O InfoSphere Information Server para System z ajuda a aproveitar informações em todas as origens. O software oferece as funções necessárias para integrar, validar, enriquecer e entregar informações confiáveis para suas principais iniciativas de negócios. Ele permite:

- *Definir* uma linguagem de negócios comum para suas informações
- *Entender* todas as origens de informações dentro do negócio, analisando seu uso, qualidade e relações
- *Purificar e padronizar* informações a fim de assegurar sua qualidade e consistência
- *Transformar* informações com o objetivo de fornecer informações ricas e sob medida
- *Federar e fornecer* informações de forma a disponibilizá-las com transparência para pessoas, processos e aplicativos

Essas funções se baseiam em uma infraestrutura de processamento paralelo que proporciona aproveitamento e automação em toda a plataforma (vide Figura 1). O InfoSphere Information Server para System z também oferece conectividade para quase todas as origens de dados ou conteúdo, bem como a capacidade de fornecer informações por meio de vários mecanismos.

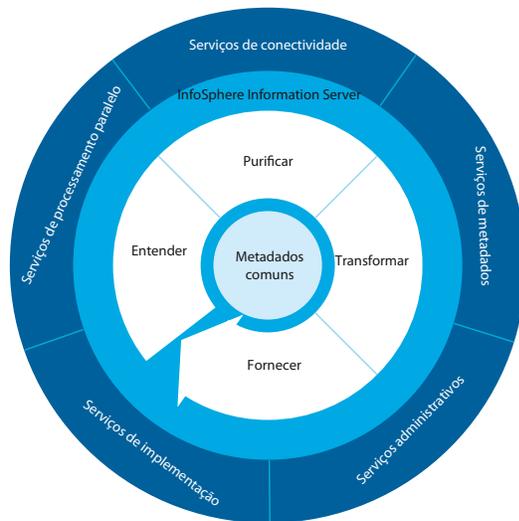


Figura 1: O InfoSphere Information Server, desenvolvido sobre uma base de processamento paralelo e outros serviços, ativa funções críticas para a criação de dados de confiança

Uma abordagem de ciclo de vida de projeto

O InfoSphere Information Server para System z utiliza uma base unificada de gerenciamento de metadados que permite o compartilhamento contínuo de conhecimento ao longo do ciclo de vida do projeto – além da compreensão detalhada do quê as informações significam, de onde elas vêm e como se relacionam com informações em outros sistemas.

A infraestrutura de metadados comum facilita a compreensão compartilhada em domínios de negócios e técnicos, ajudando a reduzir o tempo entre a especificação e o desenvolvimento. Também proporciona um registro persistente da compreensão, capaz de reduzir significativamente o tempo de entrega de um projeto de recebimento de dados e auxiliar a melhorar os insights e a confiança gerais das informações.

Todas as funções do InfoSphere Information Server para System z compartilham um modelo de metadados que abrange metadados de design e operacionais, facilitando para que diferentes papéis e funções colaborem de forma contínua. A plataforma oferece relatórios abrangentes sobre movimentação de dados, linhagem de dados, significado de negócios e o impacto de mudanças e dependências nos módulos do InfoSphere Information Server para System z e ferramentas de terceiros.

O InfoSphere Information Server para System z oferece:

- Acesso a uma grande variedade de origens de informações
- Funcionalidade de integração extensiva, incluindo federação; extração, transformação e carregamento (ETL); transformações em linha; replicação e publicação de eventos
- Recursos de integração flexíveis, incluindo suporte para Arquiteturas Orientadas a Serviços (SOA), processamento acionado por eventos, processamento de lotes planejado e APIs padrão, como SQL e Java™

Tecnologias comprovadas para assegurar a qualidade dos dados

Para alcançar novos níveis de integração, velocidade e flexibilidade de informações, o InfoSphere Information Server para System z oferece recursos de transformação e qualidade dos dados capazes de aumentar a confiança nas informações, automatizar o particionamento e processar o enfileiramento – tudo em busca de escalabilidade e serviços de rápida implementação a fim de aumentar o valor. Os componentes a seguir oferecem uma funcionalidade adicional para o InfoSphere Information Server para System z.

IBM InfoSphere Discovery

Para poder implementar um programa de governança da informação ou projeto centralizado em informações, você deve saber quais dados tem, onde estão localizados e como se relacionam entre os sistemas. No caso da maioria das organizações, o processo de descoberta de dados é manual e exige meses de trabalho humano para descobrir objetos de negócios, dados sensíveis, relações de dados com diferentes origens e uma lógica transformacional. O resultado é um processo demorado e

passível de erros que desacelera o tempo de maturação, cria dúvidas em relação à precisão dos dados dentro do novo sistema e gera a possibilidade de que o novo sistema nunca entre em operação.

O IBM InfoSphere Discovery oferece uma série completa de recursos para automatizar o processo de descoberta de dados. Aborda a criação de perfis de origem única, análise sobreposta de dados de várias origens, descoberta de chaves compatíveis, criação de protótipos e testes para consolidação de dados e descoberta automatizada de transformações. O InfoSphere Discovery também utiliza algoritmos heurísticos e sofisticados que automatizam a análise para ajudar as empresas a obterem economias de tempo e custo dez vezes maiores do que quando realizam as mesmas tarefas manualmente, usando uma solução de criação de perfil².

Os recursos do InfoSphere Discovery incluem:

Criação de perfis de dados: O InfoSphere Discovery proporciona a criação de perfis de dados avançados com resultados "sob medida". Isso inclui análise de colunas, descoberta automatizada de chave primária/estrangeira e análise sobreposta simultânea de colunas com várias origens de diversas origens de dados. Tais origens podem ser simples, como arquivos de texto em um PC, ou complexas, como Virtual Storage Access Method (VSAM) no System z – ou ambas ao mesmo tempo.

Unified Schema Builder: O Unified Schema Builder pega a saída da análise sobreposta e a utiliza como entrada em um processo que auxilia o analista de dados a determinar as regras por meio das quais os dados serão consolidados, por exemplo, para migração de dados, MDM ou um data warehouse. O componente Unified Schema Builder oferece um software de automação com fluxo de trabalho integrado que ajuda a concluir seu projeto de consolidação dentro do prazo e do orçamento.

Transformation Analyzer: Esse componente foi desenvolvido para automatizar a descoberta de transformações com várias origens e regras de negócios complexas por meio da análise de valores e padrões de dados em duas origens de dados. O Transformation Analyzer é utilizado quando se sabe que duas origens de

dados estão relacionados, mas também que a relação não pode ser descrita por simples sobreposições nos valores dos dados, sendo necessário determinar como os dados são transformados entre essas origens. Quase sempre, a migração de dados, aposentadoria de aplicativos, data warehousing e MDM exigem o mapeamento e a descoberta de uma lógica transformacional complexa entre duas ou mais origens de dados. O Transformation Analyzer ajuda a acelerar esse processo mediante a automatização de grande parte da análise envolvida e a substituição do entediante trabalho manual.

O processo de análise do InfoSphere Discovery resulta na compreensão de suas origens de dados e como elas se relacionam entre si, gerando uma saída acionável que pode ser consumida imediatamente por uma grande variedade de projetos de informação, incluindo arquivamento, gerenciamento de dados de teste, privacidade de dados, integração de dados, MDM e consolidação de dados.

IBM InfoSphere Information Analyzer

O IBM InfoSphere Information Analyzer, um módulo de produto do InfoSphere Information Server, oferece funções de criação de perfis de dados e análise de regras no contexto de uma plataforma completa de integração de informações, permitindo mais acessibilidade e consistência em toda a empresa. Os metadados ativos do InfoSphere Information Server simplificam a coleta e o gerenciamento dos metadados em todo o espectro da integração. Dentro do módulo InfoSphere Information Analyzer, os resultados da criação de perfil são armazenados no repositório de metadados comum. Ao utilizar os recursos de criação de perfis de dados do InfoSphere Information Analyzer nas fases iniciais de seus projetos de integração de dados, é possível:

- Expedir a entrega de projetos acionados por dados
- Utilizar a Análise de Regras acionada por negócios com construção e aplicativos reutilizáveis em diversas origens de dados a fim de acelerar o tempo de maturação
- Ajudar a minimizar os custos e recursos de projetos de integração de dados críticos
- Eliminar o risco e o impacto da proliferação de dados incorretos e imprecisos
- Ajudar a assegurar a entrega a tempo de informações confiáveis

IBM InfoSphere Business Glossary for System z

Dificuldades em entender e interpretar dados, determinar quais dados são importantes e, em seguida, gerenciar essas informações criam bloqueios à medida que usuários de negócios e técnicos tentam colaborar visando à integração efetiva de informações. Com frequência, o problema da inconsistência na definição dos negócios em ambientes corporativos é atribuído à ausência de um dicionário de dados e de um programa de organização corporativos.

O InfoSphere Business Glossary para System z ajuda a criar, gerenciar e compartilhar um vocabulário corporativo controlado que funciona como um idioma comum entre os negócios e a TI. Este é um passo crítico para alinhar melhor a tecnologia com as metas de negócios. Além de um vocabulário controlado, os sistemas de hierarquia e classificação do InfoSphere Business Glossary fornecem um contexto de negócios adicional.

Graças à conexão ativa com os serviços de metadados do InfoSphere Information Server, o InfoSphere Business Glossary permite que organizadores de dados conectem os termos de negócios com artefatos técnicos compartilhados entre o IBM InfoSphere Data Architect, o InfoSphere Information Server ou uma solução de integração de dados de terceiros. O resultado é um conjunto comum de identificações semânticas usadas por modeladores de dados, analistas de dados, analistas de negócios, organizadores de governança, arquitetos de dados, desenvolvedores e usuários finais. Para ajudar a assegurar a alta qualidade e uma segurança avançada, somente organizadores de dados autorizados podem usar as funções administrativas do InfoSphere Business Glossary para criar e gerenciar o glossário.

A solução também funciona como um histórico de registros para ajudar a assegurar conformidade com as regras regulamentares, como a Lei Sarbanes-Oxley e o Acordo de Baliseia II. A terminologia de negócios sempre está sujeita a mudanças: A definição atual de "cliente de alto valor" pode se tornar diferente no futuro, conforme os requisitos de negócios evoluem. A possibilidade de ver o histórico daquilo que mudou, por que mudou e quem mudou é tão importante quanto a própria mudança. Tal histórico é crítico para os protocolos de governança de dados, uma vez que aumenta a confiança e a compreensão de suas informações.

O InfoSphere Business Glossary tem uma interface baseada na web que foi desenvolvida para permitir que organizadores de dados administrem os conteúdos do glossário comum. Além disso, os InfoSphere Business Glossary Packs para diferentes verticais de segmentos de mercado ajudam a acelerar a implementação do seu glossário de negócios. Com base no conhecimento adquirido ao trabalhar com mais de 400 clientes IBM e em 10 anos de experiência nos principais segmentos de mercado, como bancário, mercados financeiros, varejo, telecomunicações, saúde e seguros, esses pacotes permitem implementar, promover e adotar rapidamente o InfoSphere Business Glossary – e, portanto, obter um retorno sobre investimento mais rápido.

InfoSphere QualityStage

O IBM InfoSphere QualityStage™, parte do conjunto InfoSphere Information Server, permite que as empresas criem e mantenham uma visão exata de entidades de dados principais, como clientes, fornecedores, locais e produtos. Os principais recursos incluem investigação de dados, padronização, validação de endereço, compatibilidade probabilística e sobrevivência. O InfoSphere QualityStage pode ser implementado em aplicativos transacionais, operacionais ou analíticos e em ambientes em lote e em tempo real.

O InfoSphere QualityStage permite um processo abrangente para gerenciar e manter a qualidade dos dados. Os principais recursos dos produtos estão focados em:

- **Investigação:** Entender a natureza e a extensão das anomalias de dados e permitir uma purificação e uma compatibilidade mais efetivas
- **Padronização:** Criar uma visualização padronizada de dados de clientes, parceiros ou produtos. Este recurso também permite purificação, validação e certificação globais de endereços (para descontos postais significativos em locais selecionados) e geolocalização
- **Compatibilidade probabilística:** Oferece um dispositivo de compatibilidade líder para ajudar a assegurar os melhores resultados compatíveis possíveis; desenvolvido sobre uma plataforma que permite alta conectividade e escalabilidade
- **Sobrevivência:** Ajuda a assegurar consolidação ideal, doméstica ou visualização vinculada de informações de registro; permite uma visualização consolidada e precisa de clientes, parceiros, produtos e mais

InfoSphere Metadata Workbench

O InfoSphere Information Server foi desenvolvido para ser uma plataforma completa para a integração e enriquecimento de informações em sistemas de origens díspares. Mediante o aproveitamento de uma camada ativa e compartilhada de repositório de metadados, o InfoSphere Information Server pode dar suporte a uma série completa de atividades de integração e funções de usuário, com princípios de colaboração e reutilização. Esses artefatos incluem metadados técnicos sobre as várias origens de informação; metadados de negócios que descrevem o significado de negócios e utilização de informações; e metadados operacionais que descrevem o que acontece dentro do processo de integração.

O IBM InfoSphere Metadata Workbench oferece uma interface poderosa de gerenciamento de metadados que oferece suporte aos metadados do InfoSphere Information Server, bem como aos outros metadados principais que desempenham funções críticas nos processos de integração de dados. Uma visualização centralizada e holística de toda a paisagem dos processos de integração de dados, com visibilidade das transformações de dados que ocorrem dentro e fora do InfoSphere Information Server, oferece aos negócios informações críticas que podem resultar em melhores decisões.

Os destaques do InfoSphere Metadata Workbench incluem:

- A navegação baseada na web dos principais recursos de informações, por meio de uma interface interativa e poderosa, oferece uma maneira fácil para os funcionários acessarem informações críticas.
- A linhagem virtual de dados em várias ferramentas e plataformas permite compreender a linhagem das informações por inteiro, incluindo de onde os dados vieram e o que aconteceu a eles conforme se moviam pelos processos de integração de dados, com visibilidade estendida dos fluxos de dados corporativos fora do InfoSphere Information Server.
- A análise de impacto visual em várias ferramentas permite entender por completo o impacto de uma mudança antes que ela seja feita, inclusive quando o impacto vai além de uma única ferramenta.
- Os relatórios sobre recursos de informações, por meio de procura simples e avançada com recursos para salvar, repetir e publicar, ajudam os usuários de negócios e TI a compreender rapidamente ambientes complexos.
- Ligações automatizadas com os serviços de metadados do InfoSphere Information Server ajudam as organizações a reduzir os custos totais de TI e acelerar a produtividade.
- Metadados de colaboração e compartilhados com o InfoSphere Business Glossary promovem organização de dados, alinhamento de negócios e TI e uma melhor compreensão dos recursos de informações.
- Os vários níveis de acesso para diferentes tipos de usuários oferecem a flexibilidade necessária para customizar o gerenciamento de requisitos e impingir a segurança das informações.

System z: resiliência crítica para os negócios para governança da informação

Em função dos efeitos de longo alcance de uma iniciativa de governança da informação, é crítico que a plataforma host seja plenamente confiável e esteja sempre disponível. Pense no efeito da indisponibilidade temporária dos dados. O relatório intitulado "Business resilience: Ensuring continuity in a volatile environment", da Economist Intelligence Unit, de 2007, declarou que, segundo a U.S. National Archives and Records Administration, 25% das empresas que sofriam uma indisponibilidade de TI com dois a seis dias de duração iam à falência imediatamente³. Uma interrupção na continuidade dos negócios equivale a riscos significativos com custos diretos (perda de negócios imediatos) e custos indiretos (danos de longo prazo e credibilidade da marca). Os danos podem ir muito além do domínio financeiro, chegando a áreas principais como fidelidade do cliente, competitividade no mercado e conformidade regulamentar.

Enquanto algumas soluções oferecem tempo de atividade de semanas ou meses, a plataforma do System z proporciona tempo de atividade de anos. Fornece configurações de recuperação de desastre e foi desenvolvida para oferecer disponibilidade de aplicativos de 99,999% ao implementar a tecnologia IBM Parallel Sysplex® para criar clusters balanceados com escalabilidade quase linear. Essa configuração ajuda a evitar as desvantagens do tempo de inatividade planejado, falha de equipamentos ou perda total de um datacenter. Também utiliza algumas das tecnologias de segurança mais avançadas do

segmento de mercado – ajudando as organizações a cumprirem com requisitos regulamentares rígidos, que incluem soluções de criptografia, gerenciamento de controle de acesso e extensivos recursos de auditoria.

Como resultado, o System z é um hub para as soluções de governança da informação de muitas organizações; os maiores bancos, segurados e varejistas do mundo acreditam que a IBM os ajudará a oferecer continuidade e a proteger transações de negócios sensíveis.

Faça suas informações trabalharem mais para você

O InfoSphere Information Server para System z é uma plataforma de software totalmente integrada que cria perfis, purifica, transforma e fornece informações a partir de origens de dados de mainframe e distribuídas, oferecendo maiores insights dos negócios sem custos operacionais agregados do IBM z/OS®. Pode ajudar a obter mais valor com as informações complexas e heterogêneas espalhadas por seus sistemas – e dar suporte à sua iniciativa de governança da informação.

Com produtividade e desempenho revolucionários para purificar, transformar e mover essas informações com consistência e segurança por toda a empresa, o InfoSphere Information Server para System z permite acessar e usar as informações de maneiras diferentes a fim de promover a inovação, ajudar a aumentar a eficiência operacional e ajudar a diminuir os riscos.

Para mais informações

Para saber mais sobre a qualidade das informações e a plataforma IBM System z como parte de sua estratégia de governança da informação, entre em contato com seu representante ou Parceiro de Negócios IBM ou visite:

ibm.com/software/data/integration/info_server_system_z

ibm.com/software/data/db2imstools/solutions/data-governance.html



© Copyright IBM Corporation 2010

IBM Software Group
Route 100
Somers, NY 10589

Produzido nos Estados Unidos da América
Julho de 2010
Todos os Direitos Reservados

IBM, o logotipo da IBM, ibm.com, InfoSphere, QualityStage, System z e z/OS são marcas registradas da International Business Machines Corporation nos Estados Unidos e/ou em outros países. Se estes ou outros termos de marca registrada da IBM estiverem acompanhados, em sua primeira referência neste documento, por um símbolo de marca registrada (® ou ™), tal símbolo indica marca registrada ou de direito consuetudinário da IBM nos EUA no momento da publicação. Essas marcas registradas também podem ser marcas registradas ou de direito consuetudinário em outros países. Uma lista atualizada das marcas registradas da IBM está disponível na web em “Copyright and trademark information” em ibm.com/legal/copytrade.shtml

Java e todas as marcas e logotipos baseados em Java são marcas registradas da Sun Microsystems, Inc. nos Estados Unidos e/ou em outros países.

Outros nomes de empresas, produtos ou serviços podem ser marcas registradas ou de serviço de terceiros.

¹ Moutsos, Kim. “IMS at 40: Stronger than Ever,” IBM Database, October 2008.
www.dbmag.intelligententerprise.com/story/showArticle.jhtml?articleID=211300235

² Time and cost savings based on reports from IBM client engagements.

³ The Economist Intelligence Unit. “Business resilience: Ensuring continuity in a volatile environment,” February 2007.
http://graphics.eiu.com/files/ad_pdfs/eiu_Bus_Resilience_wp.pdf



Recycle