

A Revolução da Análise de Negócios em 2011: Otimizando Análises e Relatórios para Difundir a Inteligência

Estudo realizado para a IBM por:

David Loshin
Knowledge Integrity, Inc.
Março de 2011

Índice

1	Resumo Executivo	4
2	Volume Crescente de Dados e Decisões Mais Inteligentes	5
3	Fornecendo Inteligência	6
3.1	Estratégias que Impulsionam a Otimização Organizacional	7
3.2	Integrando Análises de Negócio e Operações	7
3.3	Acionadores Comuns de Valor	8
3.4	Exemplos nos Segmentos de Mercado	9
4	Por que Análises de Negócio Difundidas?	11
5	Entendendo as Considerações Tecnológicas	13
5.1	Disponibilidade Contínua: Sincronia de Dados e Coerência	14
5.2	Integração Coesa de Informações	16
5.3	Gerenciando o Conhecimento com a Utilização de uma Arquitetura Robusta e Extensível.	18
5.4	Técnicas de Otimização de Desempenho	19
5.4.1	Análises de Dados Volumosos	20
5.4.2	Aprimoramentos de Arquitetura: o Aplicativo de Banco de Dados Analítico	20
5.4.3	Análise Algorítmica Integrada e Complexa Utilizando Modelos Alternativos de Programação	21
5.4.4	Ajudando as Demandas de Data Warehouse com a Utilização de Compactação	22
5.4.5	Maior Desempenho através de Melhorias em Hierarquia de Memória	23
5.4.6	Computação de Fluxo	23
5.5	Serviços Analíticos Avançados e Emergentes	24
5.5.1	Mineração de Dados e Modelagem Preditiva	24
5.5.2	Modelos de Análises Preditivas e Integradas	25

5.5.3	Reconhecimento de Entidade e Extração de Entidade	25
5.5.4	Análise de texto	26
5.5.5	Análise de Opinião	26
5.6	Serviços de Entrega de Análises de Negócio	27
6	Desafios	28
7	Resumo: Direcionando os Desafios	30
	Sobre o Autor	32

1 Resumo Executivo

O reconhecimento emergente do valor das análises de negócios se opõe ao grande crescimento da quantidade de dados estruturados e não estruturados. As organizações concorrentes se desenvolvem ao adotar estratégias e métodos para a integração de inteligência de negócios e análise de negócios de modo a complementar o âmbito das decisões que são tomadas diariamente e, algumas vezes, a cada momento. As pessoas que possuem muitos dados podem sucumbir à estagnação das análises de negócios, porém inteligência confiável será fornecida aos indivíduos corretos no momento em que eles necessitarem de pequenas análises aprofundadas e, além disso, decisões confiáveis e racionais serão fornecidas.

Neste artigo, examinaremos os acionadores de valor de negócios que aceleram a introdução de análise de negócios difundida. Além disso, também revisaremos alguns exemplos gerais em que o conhecimento pode ser utilizado para desenvolver vantagem competitiva. Consideraremos o conceito de análises difundidas, nas quais origens de dados operacionais e transacionais são transmitidas em um mecanismo analítico corporativo, e os resultados são utilizados para aprimorar e, provavelmente, melhorar de modo mensurável os processos de negócios.

Este artigo, então, fornecerá uma visão geral dos componentes técnicos necessários para ativar as análises de negócio difundidas. Uma análise mais aprofundada fornecerá detalhes sobre esses princípios básicos de gerenciamento de dados:

- Sincronização contínua de dados;
- Integração coesa de informações;
- Uma arquitetura robusta para análise;
- Ajuste de desempenho para atender grandes conjuntos de dados estruturados e não estruturados;
- Computação em grande escala;
- Serviços de análíticas abrangentes, e uma
- Estrutura para relatórios, consultas ad hoc, análise dimensional e visualização e comunicação de informações.

E, finalmente, observaremos alguns dos principais desafios de sucesso e resumiremos algumas das recomendações que podem apoiar a inteligência de negócios corporativa para ser bem-sucedida e relatórios e programas de análíticas com um conjunto completo de ferramentas que dão suporte a todas as necessidades analíticas e de relatório da organização.

2 Volume Crescente de Dados e Decisões Mais Inteligentes

Há anos, analistas técnicos estão especulando o enorme crescimento de dados. Um artigo de 2010 sugere que o volume de dados continuará a expandir a uma taxa vantajosa, observando que "o tamanho do maior data warehouse... triplica aproximadamente a cada dois anos".¹ A explosão de dados é globalmente reconhecida como uma questão fundamental do setor de Tecnologia da Informação (TI). De acordo com um estudo da Gartner realizado em outubro de 2010, 47% dos entrevistados de uma pesquisa, classificaram o crescimento de dados como um de seus três principais desafios".² Como exemplo de grande crescimento de dados, a empresa varejista Wal-Mart executa mais de um milhão de transações de consumidores a cada hora, alimentando o banco de dados em uma estimativa de mais de 2.5 petabytes. A persistência também não é requisito. Um relatório de 2010 sugere que até 2013, o total de fluxo de tráfego na internet, anualmente, será de 667 exabytes.³

As informações estruturadas em banco de dados consistem em uma pequena parcela do problema. Alguns acontecimentos importantes também comprovam o enorme crescimento dos dados não estruturados: no final de 2009, estimou-se que o total de informações digitais cresceu em quase 800.000 petabytes (ou 800.000.000.000 gigabytes). E a expectativa era de que ao final de 2010 o total seria de aproximadamente 1,2 milhão de petabytes! A esta taxa, o total de dados digitais poderia crescer em 35 zettabytes (1 zettabyte = 1 trilhão de gigabytes) até 2020.⁴

Cada vez mais, espera-se que os processos complexos de negócios sejam executados através de vários sistemas interconectados. Análises e sensores integrados não apenas possibilitam continuamente a avaliação do desempenho operacional, pois a interconexão de muitos sistemas permite a comunicação rápida e a persistência dessas medições. Estima-se que 15 petabytes de novas informações estejam sendo gerados todos os dias, e 80% deles não são estruturados.⁵

A partir desses pequenos exemplos de rápida expansão do total de informação digital, podemos concluir que novos excelentes pontos de vista podem surgir de modo nunca ocorrido anteriormente, através da análise combinada de dados estruturados e não estruturados. Novos e aprimorados modos de análise de dados permitem que as organizações identifiquem novas tendências de negócios, avaliem a expansão de doenças, ou até mesmo combatam crimes, dentre muitas outras oportunidades que serão tratadas neste artigo. Parece que estamos alcançando o ponto em que a informação se torna o foco mais importante dos negócios, de modo que "os estatísticos obtêm a criação de informações dos negócios para aquisição de novas ideias".⁶ E essas novas ideias não estão somente ocultas nos bancos de dados estruturados. De certa forma, a análise também deve compreender os objetos de dados não estruturados.

¹ Merv Adrian, "Exploring the Extremes of Database Growth" IBM Data Management, Issue 1 2010

² Lucas Mearian, "Data growth remains IT's biggest challenge, Gartner says," Computerworld. 02 de novembro de 2010, (download realizado no link:

http://www.computerworld.com/s/article/9194283/Data_growth_remains_IT_s_biggest_challenge_Gartner_says)

³ The Economist, "Data, data everywhere," 21 de fevereiro de 2010 (download realizado no link:

http://www.economist.com/node/15557443?story_id=15557443)

⁴ Gantz, John and Reinsel, David, "The Digital Universe Decade – Are You Ready? The IDC 2010 Digital Universe Study," maio de 2010 (transferido por download de <http://idcdocserv.com/925>)

⁵ Bates, Pat, Biere, Mike, Weideranders, Rex, Meyer, Alan e Wong, Bill, "New Intelligence for a Smarter Planet," <ftp://ftp.software.ibm.com/common/ssi/pm/bk/n/imm14055usen/IMM14055USEN.PDF>

⁶ Op. cit., The Economist

Porém, conforme o volume de dados cresce, também aumenta a dificuldade em encontrar as informações fundamentais que são necessárias para fazer com que esses processos de negócios sejam executados de forma otimizada. Não se trata mais da necessidade de obtenção, armazenamento e gerenciamento daquele dado. O desafio, de certa forma, é necessário para extrair e fornecer o conhecimento relevante às pessoas corretas, no momento certo, a fim de aprimorar as muitas oportunidades relacionadas à tomada de decisão que ocorrem diariamente.

Fundamentalmente, isso pode ser resumido como o desejo de integrar inteligência de maneira difundida nos processos operacionais e de estratégia em todas as funções e níveis da organização. E mesmo que isso signifique o gerenciamento das oportunidades emergentes de receita, fornecendo insight em tempo real aos indicadores de desempenho corporativo, ou a reorganização por hora da equipe de reparação a fim de melhor atender as demandas no atendimento ao cliente. A capacidade de acumular, transformar e analisar as informações para fornecer análises rápidas e confiáveis às pessoas corretas, no momento certo, pode otimizar o crescimento das oportunidades e da competitividade.

3 Fornecendo Inteligência

Em praticamente todos os processos de negócios, há situações em que a informação é acumulada e implementada de modo a ajudar nas tomadas de decisões das pessoas. Algumas decisões são muito importantes, gerando estratégias corporativas globais, enquanto outras são operacionais e de pequeno escopo, como a seleção de caixas mais econômicas e de tamanho ideal para atender da melhor maneira o pedido do cliente. E para cada processo de negócios, há avaliações e desempenho, como o retorno sobre o investimento, melhores valores ou custo de materiais.

Em um cenário ideal, cada decisão seria a melhor, ou seja, a decisão na qual os resultados liderassem o melhor desempenho geral. No entanto, as pessoas que tomam decisões constantemente não recebem todas as informações necessárias para a escolha da melhor decisão. Fornecer mais informação nem sempre é a melhor resposta. Quando grandes quantidades de dados não filtrados são divulgados na organização, pessoas sobrecarregadas tendem a ficar estagnadas, ou seja, possuem compulsão em atrasar a tomada de decisões ao mesmo tempo em que esperam receber *um pouco mais de dados* que poderiam simplificar (e, talvez, justificar) a decisão iminente. Essa estagnação pode ser diminuída quando a sobrecarga de informação é reduzida, a fim de selecionar a informação específica necessária para conduzir do modo mais eficiente o processo de tomada de decisão e permitir que ações específicas sejam tomadas. Fornecendo, assim, **inteligência confiável** às pessoas certas quando elas necessitam de pequenas análises aprofundadas, favorecendo também decisões racionais e seguras.

E, no final, o desempenho geral da empresa é revisado e alternativas são consideradas para ajudar no ajuste de estratégias corporativas para geração de valor a longo prazo. Por outro lado, as atividades operacionais são aprimoradas com inteligências específicas que podem ajustar e otimizar as atividades em tempo real. No melhor cenário, o *próprio* processo de negócios incorpora a aquisição e a apresentação da inteligência de informações, monitorando os resultados das análises diretamente no processo, rastreando avaliações de desempenho e indicando quando as melhores decisões são tomadas.

Basicamente, há uma relação de proximidade entre as operações, a análise de operações, as modificações que direcionam os processos de negócios orientados por análise, as mudanças nos processos de negócios operacionais e os aplicativos correspondentes. A inteligência confiável notifica os processos operacionais e estratégicos, e a sua entrega difundida para a equipe através do cronograma completo pode facilitar uma transição a partir da reação ao que aconteceu anteriormente com o objetivo de melhorar, fazendo com que a tomada de boas decisões progrida.

3.1 Estratégias que Impulsionam a Otimização Organizacional

Conforme mais estrategistas da organização estão ficando mais conscientes sobre os recursos de relatório e análise, há um crescente desejo de utilizar análise de dados de forma a gerenciar proativamente e aprimorar as estratégias de negócios. A observação do ambiente fornece insight de como a organização funciona, assim como permite que o analista entenda as áreas com falhas, sucesso e diferencie aquelas táticas que podem ser replicadas e, conseqüentemente, utilizadas de modo a obter um melhor sucesso.

Os relatórios tradicionais em um contexto abrangente permitem que o analista observe o que aconteceu na organização e, então, revise como o indicador de desempenho pontua as tendências ao longo do tempo. A organização dos dados ao longo de diferentes âmbitos hierárquicos quantitativos e qualitativos, como localização, organização, perfil de cliente, categorias de produto, entre outros, permite que o analista estude os dados com o objetivo de encontrar oportunidades de negócios e melhorias de processo. Adicionalmente, ambientes interativos, como painéis e mash-ups, facilitam as decisões estratégicas informadas.

Além de fazer ajustes às direções estratégicas da organização, a incorporação de entrega em tempo real de métricas de desempenho permite que as partes interessadas revisem o impacto e os resultados dessas decisões estratégicas em um curto prazo, aumentando, conseqüentemente, a rapidez, e reduzindo os riscos, permitindo, assim, que a organização responda rapidamente às oportunidades emergentes de negócios. De modo mais exato, receber um insight maior sobre como a organização funciona (ou não!) pode fazer com que os executivos estratégicos sejam informados e façam mudanças significativas na maneira com que a empresa trabalha.

3.2 Integrando Análises de Negócio e Operações

Na verdade, os resultados das análises preditivas podem ser integrados diretamente nos processos operacionais, até mesmo sem o conhecimento dos consumidores de negócios. Por exemplo, mesclar os resultados das análises preditivas com os perfis dos clientes possibilita que os tomadores de decisões de todos os níveis da organização reconheçam e reajam rapidamente às oportunidades emergentes, como ajustes em tempo real de roteiros da central de atendimento, ou a nova trajetória da entrega de produtos com base em data warehouse em tempo real e estoque de dados do inventário, dados de pontos de vendas ou até dados sobre o tráfego e o tempo. As estatísticas da Web, os perfis e as preferências do cliente podem impulsionar o realinhamento de conteúdo dinâmico em Web sites a fim de melhorar a resposta do usuário final.

3.3 Acionadores Comuns de Valor

Dado o desejo comum de melhorar continuamente as interações de negócios com clientes e parceiros, podemos considerar como as inteligências de negócios e as analíticas informam os processos de tomada de decisão relacionados aos acionadores comuns de valor por todo o segmento de mercado. Atualmente, a maioria das organizações utiliza dados de duas maneiras: uso operacional/transacional (“operando os negócios”), e uso analítico (“melhorando os negócios”). Quando os resultados da análise são introduzidos no uso operacional, a organização pode explorar o conhecimento acionável descoberto para conduzir melhorias.

Uma abordagem simples para o gerenciamento das transformações de negócios utilizando análises de negócio envolve a classificação de oportunidades para impactos positivos de negócios, com um esquema de classificação simples que lista as principais categorias para melhora potencial de negócios, incluindo as seguintes áreas:

- **Oportunidades financeiras**, como redução nos custos operacionais, aumento de receita, identificação de novas oportunidades, aumento de fluxo de caixa ou diminuição de penalidades, multas e outros encargos.
- **Oportunidades de Conformidade e Risco** associadas a avaliações mais precisas de crédito, reduzindo riscos no investimento, tornando-o mais competitivo, decisões mais inteligentes relacionadas ao investimento e/ou desenvolvimento de capital, redução de fraude e vazamento de informações, conformidade auditável de acordo com os regulamentos do governo, expectativas do segmento, ou políticas autoimpostas (como as políticas de privacidade).
- **Oportunidades Confidenciais** e com base em Satisfação, como o aumento da satisfação do cliente, melhoria de condições de trabalho e na retenção de funcionários, modernização da cadeia de fornecimento, previsão mais precisa e acurada, aumento na confiança organizacional, maior consistência nos relatórios operacionais e de gerenciamento e melhores decisões.
- **Oportunidades de produtividade**, como cargas de trabalho reduzidas, aumento no rendimento, diminuição nos defeitos de fabricação e qualidade final do produto aprimorada.

Essa categorização tem o objetivo de suportar os processos de análises de negócio e ajudar no esclarecimento dos objetivos gerais, métricas e indicadores de desempenho correspondentes. A melhoria dos negócios com a utilização das análises de negócio exige mais do que a instalação e a execução de ferramentas. As principais partes interessadas devem definir as metas alcançáveis e utilizar ferramentas para informar os processos de tomada de decisão, e avaliar, medir e controlar o nível em que os objetivos estão sendo alcançados. Programas específicos podem ser projetados e desenvolvidos em relação às melhorias com qualquer uma dessas principais categorias. Considere os seguintes exemplos:

- *Geração de Receita através de Perfil do Cliente e Marketing Direcionado* – as análises do cliente podem abranger um refinamento contínuo de perfis individuais de clientes que incorporam dados de demografia, perfil e comportamento sobre cada indivíduo com o objetivo de suportar as segmentações da comunidade do cliente em vários clusters, com base na distinção de variáveis e nos conjuntos de valor correspondentes. Esses clusters diferentes de clientes e a compreensão de esquemas de classificação permitem o desenvolvimento de estratégias de micromarketing a fim de complementar campanhas que objetivam pequenos clusters de clientes com perfis similares. O marketing direcionado pode focar diretamente nas pessoas que utilizam resultados das análises de clientes.

- *Gerenciamento de Risco através da Identificação de Fraude, Abuso e Vazamento de Informações* – a fraude, que inclui atos intencionais de fraudes com conhecimento de que a ação ou representação poderia resultar em um ganho inapropriado, é frequentemente praticada através de cenários sistêmicos de utilização. A detecção de fraude é um tipo de análise que procura por padrões de transações com frequências relevantes em relação a alguns desses cenários identificados. Ao mesmo tempo, a análise abrangente dos produtos e serviços fornecidos aos clientes nos contextos de seus contratos/acordos pode evidenciar o abuso. Ambos os riscos podem ser analisados e levados ao conhecimento das autoridades adequadas internas para correção.
- *Satisfação Aprimorada do Cliente através de Perfil, Personalização e Análise de Valor de Fidelidade do Cliente* – a análise de valor da fidelidade do cliente calcula a medição da rentabilidade de um cliente em relação ao tempo de relacionamento, incorporando os custos associados com o gerenciamento da relação, assim como as receitas esperadas daquele cliente. A utilização dos resultados do perfil do cliente pode, dentre outras melhorias, aprimorar a experiência do cliente ao customizar a apresentação de material ou conteúdo. Os perfis dos clientes podem ser diretamente integrados em todas as interações com o cliente, especialmente em centros de atendimento internos, em que os perfis dos clientes podem melhorar a capacidade do responsável pelo atendimento do cliente de negociar com este, simplificar a resolução de problemas e, talvez, até mesmo aumentar as vendas de produtos e serviços.
- *Intermediação Aprimorada e Produtividade Adquirida através de Análise de Gastos* – a análise de gastos incorpora a categorização, a normatização e a coleta de compra de produtos e dados de fornecedor a fim de selecionar os fornecedores mais confiáveis, modernizar a solicitação de ofertas e processo de compras, a redução de custos, a melhoria da previsibilidade de cadeias de fornecimento de excelente valor, e aprimorar a eficiência e previsibilidade da cadeia de fornecimento.

Esse são apenas alguns de muitos exemplos em que as análises de negócio podem ser utilizadas para otimizar os acionadores comuns de valor por todos os segmentos de mercado.

3.4 Exemplos nos Segmentos de Mercado

Por um lado, as oportunidades de melhoria manifestam-se de modos diferentes, dependendo do segmento de mercado. Por outro lado, há dimensões comuns de operações que podem ser melhoradas independentemente do segmento. Utilizando as mesmas hierarquias do acionador de valor, uma empresa de um segmento de mercado específico pode beneficiar-se de relatórios e análises de negócio exclusivos para aquele segmento, como esses exemplos “verticais”:

- **Assistência Médica** – O monitoramento do desempenho de processos de negócios abrange todos os aspectos da qualidade da saúde. Por exemplo, o entendimento do motivo pelo qual alguns profissionais são mais bem sucedidos no tratamento de certas condições pode conduzir à melhoria da qualidade da saúde. As análises de negócio podem ajudar a encontrar os fatores que contribuem para o sucesso de uma abordagem e não de outras. E, além disso, também pode ajudar a observar se esses sucessos dependem de variáveis relacionadas ao controle dos profissionais ou fatores externos que estão fora do âmbito destes. Abordagens de diagnóstico melhores podem reduzir a demanda por recursos de diagnóstico de alto custo, como os equipamentos que fazem diagnósticos com base em imagens. E, melhores tratamentos podem reduzir a duração da permanência de pacientes, liberando leitos, melhorando o rendimento, e permitindo uma utilização mais eficiente dos leitos.

- **Logísticas/Cadeia de Fornecimento** – A análise integrada para o transporte e o gerenciamento de logísticas fornece insights para a avaliação de muitos aspectos de uma cadeia de fornecimento eficiente. Por exemplo, a inteligência de negócios é utilizada para analisar os padrões de uso para produtos específicos com base em uma série de dimensões geográficas, demográficas e de perfil. A previsibilidade se torna importante. Conhecer quais tipos de pessoas constituem cada tipo de área e qual a necessidade da gama de produtos em relação a períodos específicos de tempo pode ajudar (e também atender) uma demanda mais precisa e preditiva. Como resultado, o fabricante pode encaminhar a quantidade certa de produtos a fim de reduzir ou eliminar estoques esgotados. Ao mesmo tempo, compreender a demanda por região em diferentes períodos de tempo fornece planejamento mais preciso de programação, métodos e pacotes de entrega. A venda de produtos pode ser mapeada em relação à distância do ponto de origem. Se as vendas forem mais baixas em alguns locais, isto pode indicar uma falha na cadeia de fornecimento, podendo ser revisada e potencialmente corrigida em tempo real.
- **Telecomunicações** – Em um segmento que trabalha continuamente para reduzir a insatisfação do cliente, aumentar o comprometimento de negócios com o cliente contribui para a manutenção da fidelidade deste. Por exemplo, a análise do uso do telefone celular do cliente pode ajudar na identificação da rede principal de cada pessoa. Se um cliente faz chamadas para uma pequena quantidade de telefones residenciais ou celulares pessoais, ele pode ficar mais satisfeito com um plano de serviço “familiar e amigo” que reduz o custo das chamadas para os números mais frequentes. A identificação de relacionamentos domésticos em uma rede principal pode possibilitar um pacote de serviços ao consolidar as contas de telefone celular, ou ao fazer vendas complementares de serviços adicionais como serviço de linha fixa, internet e outros serviços de entretenimento. Por outro lado, se as chamadas do telefone celular do cliente são principalmente para telefones comerciais e possuem durações entre meia e uma hora, ele pode ficar mais satisfeito com um relacionamento de telefonia de negócios que agrupa chamadas com serviços de conectividade móvel adicionais.
- **Varejo** – A grande quantidade de dados de pontos de vendas torna isto um ótimo recurso para análise. Os estabelecimentos de varejo estão sempre procurando por maneiras de otimizar a reposição de produtos a fim de aumentar as vendas, ao mesmo tempo em que reduzem gastos adicionais para aumentar suas margens, especialmente quando as capacidades do mercado podem ser atribuídas diretamente às pessoas por meio de cartões de fidelidade. Compreender o relacionamento entre a localização das lojas tradicionais e os tipos de pessoas que vivem na região pode ajudar os gerentes das lojas na seleção de produtos para variedade de estoque. A posição estratégica do produto (como prateleiras centrais ou na extremidade superior) pode ser utilizada para itens que atraem lucratividade. E isto pode ser feito com base na combinação da venda de produtos por segmento do cliente, juntamente com mapas de padrões de visita do cliente aos setores da loja. A posição do produto não está limitada às localizações físicas. Muitos registros da Web podem ser analisados em relação ao comportamento do cliente para ajudar na reorganização dinâmica da posição em um Web site, bem como aumentar as vendas de produtos com base em análise de transportes desprotegidos, por meio de filtragem colaborativa, ou com base nas preferências pessoais do cliente.
- **Serviços Financeiros/Seguros** – A identificação de riscos e o gerenciamento de exposição são fundamentais para o aumento da lucratividade no setor bancário e de seguros. Os bancos que fornecem muitos serviços financeiros desenvolvem modelos precisos associados com as atividades do cliente e os perfis que identificam as variáveis adicionais de risco. Por exemplo, a análise de grandes compras com cartões de crédito

em relação às falhas de garantia pode demonstrar riscos iminentes maiores para as pessoas que fazem compras em shoppings ou comem em certos restaurantes do tipo fast food. Por sua vez, o reconhecimento de comportamentos que são indicativos de risco iminente pode ajudar o banco a antecipar eventos de risco e entrar em contato com essas pessoas, oferecendo produtos alternativos que os mantenham em casa, reduzindo o risco iminente e melhorando a previsibilidade do fluxo de caixa de empréstimo ao longo de grandes períodos de tempo.

- **Setor de Manufatura** – A análise de desempenho da fábrica é essencial para a manutenção previsível e produtividade confiável, bem como do rastreamento de desempenho de linha de produção, tempo de inatividade das máquinas, qualidade da produção, trabalho em andamento, incidentes de segurança e entrega de medições de indicadores de desempenho operacional na cadeia de escalação de gerenciamento de forma que eventos adversos possam ser encaminhados ao contexto apropriado em intervalo de tempo razoável.
- **Hotelaria** – As cadeias de hotéis avaliam os perfis dos clientes e padrões relacionados de viagem, de forma a saber que certos clientes podem estar dividindo sua reserva anual entre os competidores. Ao analisar as preferências de viagem do cliente e seus locais preferidos, a empresa pode apresentar ofertas de incentivo através do programa de fidelidade a fim de obter mais daquela reserva do cliente.

Os exemplos desses segmentos de mercado são similares, enquanto que os intervalos de análise de relatório simples dos indicadores principais de desempenho de negócios para explorar as oportunidades para obtenção de otimização no modo em que a empresa executa ou aprimora as interações com os clientes e outros parceiros de negócios. A investigação dos processos de negócios e as medições de desempenho de qualquer segmento fornecerão sugestões de maneiras de obter benefícios especificamente de relatórios e análise de negócios.

4 Por que Análises de Negócio Difundidas?

O conceito de “inteligência de negócios” constitui em ferramentas e técnicas que suportam muitas comunidades de usuários pela organização, como resultado da coleta e organização de vários (e diversos) conjuntos de dados a fim de dar suporte à tomada de decisão e ao gerenciamento nos níveis estratégicos, táticos e operacionais. Através da coleta de dados, agregação, análise e apresentação, a inteligência pode ser fornecida de maneira a melhor servir uma ampla gama de usuários alvos. As organizações que desenvolveram seus programas de data warehouse permitem que os usuários obtenham conhecimento do recurso corporativo de informações e percebam, rapidamente, o valor dos negócios.

Porém, enquanto as infraestruturas de data warehouse suportam a consulta de análise de negócios e relatórios gravados ou painéis de gerenciamento, um programa abrangente para insight de informação e inteligência pode melhorar o processo de tomada de decisão para todos os tipos de participantes da equipe em várias funções operacionais, táticas e estratégicas. Melhor ainda, a integração de informações relevantes com o contexto operacional imediato se torna um fator diferencial. A análise de clientes em modo off-line que fornece estratégias de vendas gerais é um fator, porém a inteligência acionável em tempo real pode fornecer alternativas específicas para o vendedor que conversa com um cliente específico com base no perfil de histórico de interação desse cliente, de modo a melhor servir o cliente ao mesmo tempo em que a lucratividade corporativa otimiza também a comissão do vendedor. O aumento dos benefícios gerais para todas as partes envolvidas recentemente melhora as vendas, aumenta a satisfação do funcionário e do cliente e

aprimora a taxa de resposta, enquanto reduz os custos de venda das mercadorias, e assim, todos são beneficiados.

As amplas gamas de recursos de análise de negócio ajudam a sugerir respostas às várias questões de valor progressivas:

- **O quê?** – Relatórios predefinidos fornecerão a resposta para os gerentes de operação, detalhando o que aconteceu na organização e várias maneiras de analisar os resultados das consultas a fim de compreender as características básicas da atividade de negócios (por exemplo, contas, somas, frequências, localizações, etc.). O relatório de BI tradicional fornece visão retrospectiva 20/20, ou seja, disponibiliza informações sobre o que aconteceu e pode fornecer dados agregados sobre o acontecido e até mesmo direcionar pessoas às ações específicas em relação ao ocorrido.
- **Por quê?** – Consultas ad hoc mais abrangentes, juntamente com revisões das medições e métricas em uma série de tempo, permitem uma revisão mais focada. Percorrer as dimensões relatadas permite que o cliente de negócios obtenha respostas para questões mais focadas, como as fontes de quaisquer problemas reportados, ou comparando desempenho específico pelas dimensões relevantes.
- **E se?** – Análises estatísticas mais avançadas, modelos de mineração de dados e de previsão permitem que as analíticas de negócios considerem como as diferentes ações e decisões podem impactar os resultados, possibilitando novas ideias para aprimorar os negócios.
- **E a seguir?** – Ao avaliar as diferentes opções de previsão e planejamento de modelos preditivos, os estrategistas seniores podem considerar as possibilidades e tomar decisões estratégicas.
- **Como?** – Ao considerar as abordagens para a otimização de desempenho organizacional, os gerentes seniores podem adaptar estratégias de negócios que mudam o modo em que a organização negocia.

A análise de informações possibilita a resposta a essas questões. Os processos aprimorados de tomada de decisão dependem de suporte de inteligência de negócios e recursos analíticos que aumentam em complexidade e valor por uma ampla gama para o fornecimento de conhecimento acionável (como mostrado na Figura 1). À medida que a funcionalidade analítica aumenta em sofisticação, o cliente de negócios pode obter mais insights nos mecanismos de otimização. A análise estatística ajudará a separar a causa-raiz de quaisquer problemas reportados, bem como fornecerá alguns recursos de previsão, padrões existentes e tendências que continuam sem ajustes. Os modelos preditivos que capturam padrões posteriores ajudam a preservar cenários “e se” que guiam táticas e estratégias para o excelente desempenho organizacional.

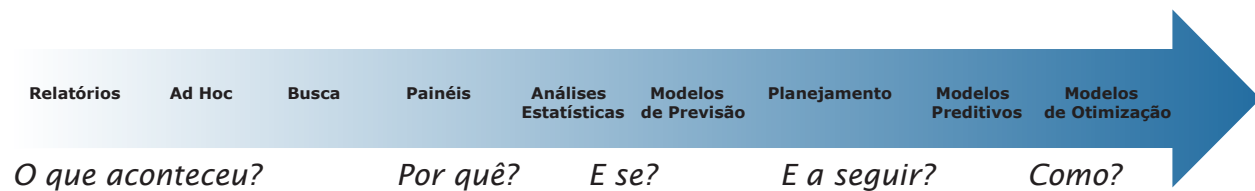


Figura 1: Uma gama de técnicas que beneficia vários clientes para a análise de negócios

A soluções de inteligência de negócios são ferramentas desenvolvidas que podem ajudar na otimização dos negócios. Isso ocorre quando as ferramentas são utilizadas para ajudar

- Executivos C-levels que revisam opções para atender os objetivos estratégicos,
- Gerentes seniores que procuram modernizar suas linhas de negócios, ou
- Tomada de decisões operacionais de uma maneira que era, anteriormente, considerada impossível.

Essas análises de negócio incorporam data warehouse, mineração de dados, análises multidimensionais, fluxos e mash-ups a fim de fornecer uma ampla visão que pode facilitar reações imediatas para oportunidades emergentes, enquanto permite a avaliação do ambiente ao longo do tempo de modo a descobrir maneiras de melhorar e expandir os negócios.

5 Entendendo as Considerações Tecnológicas

Antes, os arquitetos de data warehouse enfrentavam dificuldades com o projeto e a execução de abordagens efetivas para a extração de dados de sistemas de origem, transformando dados em formatos adequados para análise, e carregando os dados no data warehouse. Uma vez que os dados eram organizados no data warehouse, ferramentas de relatórios e análises, como o OLAP (Online Analytical Processing) e ferramentas de consulta e relatório eram utilizadas no fornecimento de dados para o consumidor de conhecimento.

Embora alguns dos principais aspectos dessas abordagens permaneçam inalterados, os detalhes da infraestrutura técnica se desenvolvem desde a metade da década de 90, com o objetivo de atender a demanda crescente para pedidos de muito mais dados disponíveis a partir de várias origens. Por sua vez, a variedade e as expectativas do conjunto de consumidores de conhecimento aumentaram as demandas por tipos, escalabilidade e mecanismos de entrega para conhecimento acionável. Atualmente, o fornecimento de informação correta às pessoas certas, no momento ideal é o ponto culminante das melhores práticas em gerenciamento de dados e organização combinados com uma infraestrutura técnica projetada para direcionar vários canais de dados constantes em uma excelente plataforma de análise de desempenho que pode fornecer resultados confiáveis em limitações de tempo real. Isso depende dessas três áreas de tecnologia:

- Sincronização contínua e dados a partir de várias origens a fim de fornecer uma visão consistente e coerente;
- Integração de informação coesa, permitindo que montantes de dados de alta qualidade sejam alinhados e uniformizados para possibilitar análise efetiva;
- Uma arquitetura robusta para coleta, ordem e gerenciamento de grandes conjuntos de dados para análise;

- Características de ajuste de desempenho que variam de melhorias em hardware a modelos de execução e programação ajustados para lidar com amplos conjuntos de dados não estruturados;
- Computação em larga escala, associando o processo e a análise de amplos conjuntos de dados acumulados a partir de muitas origens de dados e geralmente com expectativas de resultados providenciados em tempo real;
- Conjunto abrangente de análise de negócios (incluindo análise não direcionada, modelos preditivos e análises de texto) que reduz ou até mesmo elimina a necessidade de um usuário avançado para obter benefícios a partir de inteligência acionável, e uma
- Estrutura para relatórios, consultas ad hoc, análise dimensional e visualização e comunicação de conhecimento acionável.

Nesta seção, observaremos algumas das principais considerações técnicas necessárias para um ambiente de analítica moderno, e como os resultados podem ser totalmente integrados em dezenas, se não milhares, de processos de negócios diários. Com base na compreensão dos componentes tecnológicos, começaremos com a visualização de alguns desafios que surgem em ambientes com tecnologias heterogêneas agrupadas para suportar o programa de analítica.

5.1 Disponibilidade Contínua: Sincronia de Dados e Coerência

Um desafio recorrente da estrutura de data warehouse tradicional envolve o retorno de tempo prolongado entre a necessidade articulada para o relatório ou a análise e seus fornecimentos. O atraso na entrega de conhecimento acionável derivado de dados atuais dificulta as decisões oportunas, e grandes latências de dados impactarão na capacidade de avaliar os resultados dessas decisões em uma estrutura de tempo razoável. Basicamente, a sincronização de dados se tornou um componente essencial para a infraestrutura, de modo a integrar-se com as atividades de tomada de decisão operacional e estratégica. A gravidade dos dados atuais e oportunos não pode ser minimizada. Por exemplo, o aumento nas vendas por categoria de produto em certas regiões pode sugerir uma demanda crescente que necessita de realocação imediata de recursos logísticos a fim de evitar estoques indisponíveis e atender a demanda. Dependendo de informações sobre pedidos anteriores é insuficiente. Ao invés disso, plataformas atuais e informações sobre estoque, juntamente com alocação de recursos de remessa, permite decisões imediatas que mantêm um fluxo contínuo de produtos para os clientes.

Inteligência de negócios difundida exigem um excelente nível de sincronia de dados, o que significa que o ambiente deve

- Reduzir ou eliminar latência de dados;
- Manter coerência dos dados que estão sendo utilizados para análise na empresa;
- Fornecer informações atuais e oportunas, e
- Proporcionar resultados consistentes e deterministas para as solicitações semelhantes.

Manter um nível razoável de sincronia e coerência entre a variedade de origens de dados disponíveis na organização (e também fora dela) exige estratégias técnicas para disponibilidade contínua de dados que não pressionam o ambiente. Algumas dessas estratégias incluem:

- CDC (Change data capture) para replicação de dados, incluindo cópia gerenciada e sincronizada de dados a partir de uma origem para um ou mais sistemas de dados de destino. O CDC é um mecanismo orientado por evento para capturar mudanças de conjunto de dados de origem e propagar essas mudanças através de vários

canais, diretamente para bancos de dados de destino ou por fila de mensagens para processamento subsequente. *A sincronização de modificações de dados através de CDC possibilita coerência entre os sistemas operacionais e analíticos, permitindo a descoberta de oportunidades acionáveis em tempo real, enquanto mantém a consistência entre os sistemas de relatório.*

- **Federação de Dados**, que possibilita acesso transparente aos tipos de dados heterogêneos (e, geralmente, distribuídos fisicamente), plataformas e origens, e em formatos variados, sem exigir uma área temporária ou um repositório centralizado (consulte Figura 2). Federação é um modo efetivo de capturar subconjuntos de conjuntos de dados muito grandes ou muito distribuídos, e é frequentemente utilizado quando o dado é remoto, está em formato antigo ou é utilizado ocasionalmente. Por exemplo, uma estrutura de federação de dados permite que um aplicativo acesse banco de dados, dados XML, arquivos simples, ou até mesmo serviços de dados ou fluxo de dados, utilizando um mecanismo de acesso uniforme. Por sua vez, o servidor de federação acessa de modo dinâmico as origens de dados e obtém resultados sincronizados. *A federação simplifica a consolidação de dados a partir de várias origens, possibilitando "polinização cruzada" de informações a fim de melhor encontrar oportunidades e, além disso, também é muito efetiva na junção de dados não semelhantes antes da chegada ao destino.*
- **O Processamento de Fluxo de Informações**, que fornece aplicativos com acesso contínuo para origens de dados de fluxo em tempo real. *A conexão de dados de fluxo a origens de dados persistentes suporta processamento de eventos complexos e descoberta em tempo real de oportunidade com base em atividades de conhecimento emergente, como transações de mercadoria baseadas no tempo ou entregas imediatas a fim de evitar indisponibilidades no estoque.*



Figura 2: Os serviços de federação de dados fornecem acesso transparente e dinâmico às várias origens de dados.

5.2 Integração Coesa de Informações

Devido ao fato de a maioria dos dados a ser analisada ser criada por envio de dados a partir do ponto de análise para transações específicas ou propósitos operacionais, é necessário fornecer uma estrutura coesa de integração de informação a fim de facilitar a aquisição e o compartilhamento de informação. Em outras palavras, a preparação de dados corporativos para análise exige técnicas colaborativas para a aquisição de dados a partir de várias origens, gerenciamento de representações de conceitos similares de dados em diversas representações e, então, compartilhamento de informações com vários consumidores através de diferentes aplicativos.

- Integração de Dados e ETL – Do mesmo modo que a federação de dados, a integração de dados baseia-se na capacidade de acesso contínuo à dados a partir de diferentes origens e na entrega desses dados à diversos destinos, porém isso compreende muito mais ao se tratar da análise de dados extraídos, normalizando isso em formatos padrões, limpando e transformando os dados em uma representação que é adequada para o carregamento em data warehouse e quaisquer usos subsequentes para relatório e análise. *O ETL é um ativador-chave do data warehouse, fornecendo os fluxos principais de informação que permitem a integração das inteligências de negócios.*
- Master Data Management – Os objetos do “master data” (dados mestre) são os conceitos principais de negócios representados em diferentes silos de dados e utilizados em diferentes aplicativos de negócios na organização, juntamente com seus metadados associados, atributos, definições, funções, conexões, taxonomias e hierarquias. Alguns exemplos comuns incluem Clientes, Fornecedores, Partes, Produtos, Localizações, Acordos, Mecanismos de Contato. O MDM (Master Data Management) incorpora os aplicativos de negócios, métodos de gerenciamento de informação e ferramentas de gerenciamento de dados para implementar políticas, procedimentos e infraestrutura que suportem a captura, integração e subsequente uso compartilhado de “master data” completo, consistente, oportuno e adequado. O MDM deve fornecer a capacidade de gerenciar acesso às representações de identificação exclusiva para cada entidade de “master data” pela infraestrutura do aplicativo. *O fornecimento de uma “visão unificada de conceitos de dados principais agrupados a partir de várias origens ajuda a reduzir*

duplicação de clientes, produtos, fornecedores e qualquer outra classe principal de recurso de dados.

- Gerenciamento de Metadados – Somente a quantidade correta de metadados é necessária com o objetivo de dar suporte à extração de dados, integração e requisitos de análise para assegurar consistência de significado. Os metadados incorporam as definições de negócios (associadas com conceitos de dados, termos de negócios e definições e semânticas associadas), domínio de dados de referência conceitual e seus valores correspondentes, definições de elemento de dados, formato de elemento de dados, estruturas e tipos de dados, bem como modelos de entidade, relacionamentos de entidade e suporte de metadados para controle de dados como mapas de uso de informação, regras de qualidade de dados e controles de acesso. *O gerenciamento de definições de dados, tabelas de referências e mapeamento de dados possibilita a reutilização de informações e ajuda a semântica de dados, à medida que isso é integrado a partir de muitos sistemas.*
- Criação de Perfil de Dados – Como qualquer outro recurso, ele é excelente para o inventário do recurso, pois determina o que ele é, a quem pertence, onde foi utilizado e sua capacidade de manutenção. A criação de perfil de dados dá suporte a revisão de uma combinação de artefatos e análise empírica de conjuntos de origem de dados a fim de compreender as características do metadado do elemento de dados que é essencial para a análise. A criação de perfil de dados pode fornecer “estabelecimento de verdade” associado com os valores de dados atuais, bem como fornecer evidências de consistência com metadados e, além disso, proporcionará insight para a adequação das origens do candidato a fim de satisfazer as necessidades de análise de negócios. *O resultado de uma avaliação de qualidade de dados utilizando criação de perfil de dados, juntamente com uma revisão das expectativas de aplicativo de consumo, pode mostrar as regras de qualidade de dados que são gerenciadas no repositório de metadados.*
- Limpeza de Dados – A criação de perfil de dados irá expor falhas e erros potenciais nos dados, e os dados incorretos utilizados como entrada para o processamento analítico impactarão a capacidade de compreensão de seus resultados. Quando os proprietários do sistema de dados de origem encaminham projeto ou falhas de processo que introduzem problemas nos dados, então, as causas-raiz podem ser eliminadas, bem como os erros. Entretanto, quando o dado de origem está fora do controle administrativo da organização, a infraestrutura técnica e de processos deve estar adequada para analisar, normatizar, aprimorar/melhorar e limpar o dado a fim de satisfazer as necessidades analíticas de recebimento de dados. *A limpeza de dados é geralmente exigida quando o dado está sendo redirecionado, especialmente quando há novos propósitos de expectativas de qualidade de dados completos. Ela também otimiza o valor do dado, e dados consistentes e de ótima qualidade tornam confiável o processo de tomada de decisão.*
- Validação de Dados – Como alternativa, há outras maneiras de introduzir novos erros nos dados, e a identificação e eliminação antecipada de falhas potenciais de dados no fluxo de produção de informação reduzirá a variação e a inconsistência do recebimento de dados e aprimorará a eficiência operacional de modo geral. *De maior importância, a validação de dados fornecidos com regras de qualidade de dados definidas contribui com o fornecimento de dados confiáveis.*
- Resolução de Identidade – Devido ao crescimento coordenado de muitos sistemas operacionais implementados na empresa, não é comum que várias instâncias de dados, em diferentes sistemas, representem as mesmas entidades reais de diversas maneiras. De forma alternativa, há situações em que duas entidades reais compartilham os mesmos atributos de identificação, dificultando a distinção entre eles. Esses dois tipos

de problemas refletem o mesmo desafio principal: a capacidade de avaliar a semelhança entre pares de registros e determinar se eles representam o mesmo. A resolução de identidade encaminha os problemas ao calcular e pontuar o nível de semelhança entre dois registros. Quando a pontuação é acima do limite específico, presume-se que os dois registros sejam correspondentes. Se a pontuação está abaixo do limite, então é esperado que os dois registros não correspondam. A resolução de identidade é utilizada para corresponder registros na presença de variações ou atributos incompletos, ou determinar se os dois registros realmente representam entidades distintas. *Além disso, a resolução de identidade é um componente fundamental do gerenciamento de dados mestre, e o aumento na precisão da correspondência de entidade reduz a duplicação de dados e suporta análise e relatório de excelente qualidade.*

- Mecanismos de Entrega Difundida – A planilha não é mais o único modo de entrega de resultados analíticos. A autoconfiguração de relatórios e resultados de consulta de aplicativos de negócios elimina o gargalo causado pela dependência de suporte pela equipe de TI, e a entrega com base na Web simplifica a disponibilidade de resultados ao utilizar vários objetos de apresentação visual interativos e intuitivos (como gráficos e mapas de utilização). *Ao automatizar notificações orientadas por eventos que podem ser projetadas para um grupo de interfaces, que varia de áreas de trabalho a orientações de PDAs, a inteligência acionável pode ser fornecida diretamente onde ela é necessária.*

5.3 Gerenciando o Conhecimento com a Utilização de uma Arquitetura Robusta e Extensível.

Enquanto as principais partes interessadas, na maioria das organizações, estão preocupadas com os acionadores de valor de alto nível, cada área de negócios da empresa, juntamente com pressões externas específicas para suas linhas de negócios, conduz a necessidade de estrutura, enquanto encaminha melhorias para os acionadores de valor comum. E, ao mesmo tempo, informa os processos de negócios e as necessidades correspondentes de relatório e analíticas que são exclusivas para cada segmento de mercado. Isso significa que um modelo de data warehouse integrado com os domínios de dados mestre comumente utilizados (como "cliente", "produto" ou "acordo") permitem flexibilidade em modelos aprimoradas para a resolução de desafios específicos do segmento. Considere os seguintes exemplos:

- **No setor financeiro**, em que há uma necessidade de modelos que focam em um equilíbrio entre gerenciamento de risco corporativo e negligência de visibilidade de cliente e lucratividade de produto. A partir da perspectiva de risco, isso combina gerenciamento de conformidade com (e o relatório estabelecido por lei associado com) diretivas regulatórias em uma variedade de níveis governamentais, variando entre as leis Basel II, Sarbanes-Oxley, relatório financeiro internacional, bem como o gerenciamento financeiro corporativo, risco de terceiros e de crédito. Ao mesmo tempo, esses negócios ainda possuem a necessidade de relatórios variados e análise de interações com o cliente que ajudam na manutenção da fidelidade do cliente, retenção de aumento e efetividade no canal de medição.
- **No setor de seguros**, no qual há uma necessidade de modelos que encaminha desafios associados com revisão e gerenciamento de solicitações, relatórios por membro/mês, desempenho intermediário, conformidade (como os requisitos de Solvency II), gerenciamento de risco, atuarial de pensão e consumo básico, e conformidade de pensões corporativa, que ainda necessitam de visibilidade no gerenciamento e retenção de clientes, desempenho de produto e efetividade em vendas e marketing.

- **No setor de varejo**, em que a necessidade é de relatório e analíticas para o gerenciamento de cliente, de mercadorias, produtos e serviços, de operações de armazenamento, de finanças corporativas, bem como cadeia de fornecimento, gerenciamento de vários canais de vendas e também visibilidade de conformidade regulatória. Além disso, os resultados das analíticas preditivas em tempo real podem ter seu fluxo voltado para os processos de vendas e marketing operacional a fim de aprimorar a geração de receita e reduzir a insatisfação do cliente.
- **No setor de telecomunicações**, no qual há um desejo pelo monitoramento do desempenho de marketing, gerenciamento de publicidade, garantia de receita, provisionamento, classificação, faturamento, risco de crédito, dispositivo de cliente e analíticas de plano, bem como o uso de relatórios para comunicações por voz e por outros meios, retenção e insatisfação do cliente, e também desempenho de rede e sistema e utilização de análise (e chamadas reduzidas!) para informar a tomada de decisões a fim de aprimorar capitais na infraestrutura de telecomunicações.

Como pode ser visto, todos esses exemplos exigem uma arquitetura robusta que suporte a coleta, agregação, classificação e gerenciamento de grandes conjuntos de dados. E embora todos esses exemplos de segmento de mercado compartilhem vários objetivos comuns em relação à efetividade de vendas e marketing, ou gerenciamento de relacionamento com o cliente, cada um possui diferentes combinações de cargas de trabalho para relatório, análises de negócio, ou ambos podem ser beneficiados com modelos robustos de data warehouse de segmento de mercado.

5.4 Técnicas de Otimização de Desempenho

Em relação à incrível taxa de crescimento de volume de dados, a questão de manutenção de excelentes níveis de desempenho de sistema passa para uma posição dominante, especialmente com uma comunidade de inteligência de negócios e usuários de analíticas que ampliam uma ampla gama de tipos de consumidores, como

- Consumidores padrão de negócios, que usam relatórios específicos de domínio juntamente com suas próprias consultas ad hoc e análises interativas diretas;
- Consumidores ocasionais, que de modo geral revisam as métricas de desempenho operacional e funcional que simplificam os relatórios pré-projetados apresentados através de painéis e pontuações;
- Consumidores de analíticas operacionais cujos processos de negócios são informados através de resultados integrados a partir de análises operacionais;
- Consumidores corporativos estendidos, que abrangem partes externas, clientes, reguladores, análises externas de negócios, parceiros, fornecedores ou qualquer um com a necessidade de informações reportadas para a tomada de decisões tática, e
- Usuários avançados e análises sofisticadas que usam várias ferramentas e técnicas para analisar grandes quantidades de dados e cujos resultados informarão os processos de tomada de decisão.

De modo geral, os diferentes consumidores e seus usos de protótipo refletem uma combinação de demanda para relatório e análise. Consequentemente, devido aos diferentes tipos de necessidade para os resultados desses diferentes tipos de análise, a arquitetura de inteligência de negócios deve suportar uma carga de trabalho combinada de relatórios acumulados e pré-projetados, relatórios e analíticas preparados por usuários avançados, consultas ad hoc, análise interativa, bem como detalhamento interativo nos conjuntos de dados suportados.

A preparação de grandes quantidade de dados, bem como o suporte de uma carga de trabalho combinada de relatório, análise interativa e complexas significa mais do que apenas utilizar computadores potentes para a resolução do problema. Em vez disso, é preciso considerar como quaisquer aspectos do fluxo de produção de informação poderiam ser impactados como resultados de baixo desempenho, a partir do ponto em que os dados fluem no ambiente analíticos, ao ponto em que grandes quantidade de dados devem ser disponibilizados para relatório, interação com o usuário final e análise.

A necessidade de desempenho deve ser avaliada em relação às quatro áreas do fluxo de produção de informação: no ponto relacionado à integração de dados e carregamento no data warehouse, à medida que os dados são sujeitos às consultas e análise algorítmica, fornecendo resultados de modo interativo à comunidade do usuário e ao data warehouse. Podemos observar essas técnicas específicas para a melhoria do desempenho:

- Dispositivos de banco de dados analítico, que são sistemas projetados com aprimoramento de arquitetura integrada com o objetivo de melhorar as velocidades de processamento;
- Modelos de programação alternativos, como o Hadoop, para capacitar a colaboração entre a análise algorítmica e as analíticas de banco de dados. A compactação, que objetiva reduzir o armazenamento para melhorar os tempos de resposta à consulta e carregamento, e
- Aprimoramentos de hierarquia de memória, como a utilização de appliances de data warehouse sólidos e compactos em vez de discos giratórios.

5.4.1 Análises de Dados Volumosos

Os usuários finais não devem se limitar a apenas acessar resultados filtrados e reduzidos. Quando grandes quantidades de dados não estruturados contribuem para um processo de análise, o usuário final deve precisar fazer a expansão na fonte original. No entanto, a ausência de contexto pode levar à inconsistência e confusão. A manutenção de informação de histórico sobre relevância e conteúdo de cada objeto de dado coletado (como página da Web, apresentação ou vídeo) no contexto de suas conexões com os outros objetos fornece um conjunto dimensionado para análise. Além disso, aprimorar os objetos de dados com propósitos agrupados com base em hierarquias semânticas definidas fornece contextos adicionais e, conseqüentemente, dimensões adicionais para revisão.

Isso significa que o ambiente analítico deve estar capacitado com aspectos de soluções de fornecimento para “grandes” problemas. Isso envolve uma combinação de gerenciamento de acesso rápido para uma grande quantidade de artefatos de dados e a integração de resultados de analíticas computacionais (como aquelas desempenhadas com a utilização do Hadoop) juntamente com dados estruturados para o relatório e a análise no data warehouse. Reunidos, esses recursos contribuem com os tipos de serviços de analíticas emergentes avançadas que impulsionam a descoberta de insight e conhecimento acionável.

5.4.2 Aprimoramentos de Arquitetura: o Aplicativo de Banco de Dados Analítico

Em virtude do fato que o data warehouse possui necessidades extremamente limitadas para o suporte de transações, os desenvolvedores de data warehouse estão dispostos a renunciar o desempenho de transação para aumentar o desempenho de consultas complexas. Essa ideia resultou no conceito de “aplicativo de banco de dados analítico”, que é um sistema que integra recursos computacionais de alto desempenho, conexão de redes e E/S de alta velocidade, gerenciamento de banco de dados e, possivelmente, outras ferramentas pré-instaladas, configuradas e otimizadas para análise e relatório de carga de trabalho combinada.

Os aplicativos de banco de dados analíticos são, frequentemente, compreendidos por softwares otimizados estratificados em um sistema de processamento massivamente paralelo, embora a arquitetura de data warehouse subjacente possa ser configurada de modos diferentes. Três abordagens comuns são descritas com base nas configurações do data warehouse:

- Nenhum compartilhamento, em que cada unidade de processamento independente é conectada à sua própria memória e se comunica diretamente com seu sistema de disco.
- Disco compartilhado, em que cada unidade de processamento independente é conectada à sua própria memória, mas se difere da abordagem de nenhum compartilhamento, pois todas as unidades de processamento acessam um sistema de disco comum.
- Compartilhamento total, em que cada unidade de compartilhamento possui acesso aos sistemas de memória compartilhada, bem como acesso ao armazenamento de disco compartilhado.

Os aplicativos de banco de dados analíticos também são criados para a alocação otimizada de dados. Em vez de um layout de dados orientado em linha, alguns aplicativos organizam e armazenam seus dados utilizando um layout em formato de coluna. Devido ao fato de cada coluna ser armazenada separadamente, para qualquer consulta, o sistema pode avaliar quais colunas serão acessadas e, então, recuperar somente os valores solicitados por essas colunas. A indexação também pode ser simplificada, desde que os valores de cada coluna possam ser ordenados e usados para formar o índice, que não apenas modela a expansão da área de cobertura do banco de dados, como também reduz a quantidade de dados a ser transmitida a partir do data warehouse secundário, que pode melhorar consideravelmente o desempenho de consulta.

5.4.3 Análise Algorítmica Integrada e Complexa Utilizando Modelos Alternativos de Programação

Usuários avançados de análises avançadas focadas em “grandes problemas”, como processamento de grandes quantidades de dados distribuídos, análises estatísticas complexas, desenvolvimento de modelo utilizando ferramentas de mineração de dados, e redução e filtragem em grande escala, como um modo de avaliar os grandes conjuntos de dados e, possivelmente, uma combinação de dados estruturados e não estruturados, em geral com os resultados que serão integrados aos sistemas operacionais. Esses usuários provavelmente utilizarão várias técnicas de análises que operam em tarefas paralelas simples, como a combinação de data warehousing com algoritmos de mineração de dados ou aplicativos paralelos criados. Por sua vez, os resultados desses aplicativos podem ainda necessitar ser reportados através de abordagens tradicionais de inteligência de negócios, de modo que qualquer ambiente de BI possa capacitar análise complexa de algoritmos, juntamente com modelos padrões de data warehouse, principalmente em uma arquitetura de processamento massivamente paralela.

O Hadoop é um exemplo principal desse tipo de ambiente alternativo de programação que é estruturada para explorar a distribuição de dados e paralelismo massivo. Ele é uma estrutura open-source composta por dois serviços de maior importância. O primeiro é um sistema de arquivo distribuído confiável, e o segundo é um modelo de programação paralela com base em uma abordagem chamada “MapReduce”. O modelo de programação MapReduce foi apresentado e descrito pelos pesquisadores da Google para computação distribuída e paralela que abrange grandes conjuntos de dados (que variam de centenas de terabytes a petabytes). Em oposição às linguagens processuais/imperativas, como Java e C++, o modelo de programação MapReduce simula linguagens funcionais (principalmente

Lisp e APL), na maioria das vezes devido à sua dependência dessas duas operações básicas que são aplicadas aos conjuntos ou listas de pares de valores de dados:

- *O mapa* que descreve a computação ou análise aplicadas ao conjunto de pares de entrada chave/de valor para produzir um conjunto de pares-chave intermediários/de valor, e
- *A redução*, em que o conjunto de valores associados com a saída dos pares-chave intermediários/de valor pela operação de Mapa é combinado para fornecer os resultados.

Com alguns aplicativos utilizados em grandes conjuntos de dados, a teoria é que as computações aplicadas durante a fase Mapa para cada par de entradas chave/de valor são independentes umas das outras. A combinação de independência computacional e de dados significa que ambos os dados e as computações podem ser distribuídos por várias unidades de armazenamento e processamento e automaticamente feitos em paralelo. Essa paralelização permite que o programador explore os recursos de processamento escaláveis e massivamente paralelos para o aumento no desempenho e na velocidade de processamento. Assim, os resultados dos algoritmos podem ser realinhados com um modelo dimensional de data warehouse a fim de integrar análise algorítmica com relatórios tradicionais.

5.4.4 Ajudando as Demandas de Data Warehouse com a Utilização de Compactação

Embora o custo médio de armazenamento de disco continue a diminuir, a grande taxa de crescimento para todos os tipos de dados supera a diminuição relativa nos custos. Para dificultar, as necessidades de data warehouse de um ambiente de inteligência de negócios abrangem mais do que o que é armazenado no data warehouse. Devido à utilização de índices, armazéns operacionais de dados, áreas temporárias, ambientes de testes e de desenvolvimento, são necessárias maneiras de redução da área de cobertura de armazenamento sem perda de informação.

Uma técnica comum para atingir esse objetivo é a compactação. A compactação é um processo de redução da necessidade de armazenamento através do reconhecimento e reposição de padrões comuns com itens de dados de tamanho reduzido. A compactação se torna particularmente relevante ao carregar dados em seu ambiente de data warehouse. Conforme os conjuntos de dados aumentam rapidamente, a largura de banda restrita dos canais comuns de Entrada/Saída (E/S) dominam o movimento oportuno de dados na estrutura de análise. De modo alternativo, as consultas de usuário dependentes de junções de várias tabelas podem também ser dificultadas como resultado de uma largura de banda de E/S limitada. A redução da área de cobertura de armazenamento através de compactação e aplicação de layouts e filtros criativos de dados pode reduzir a demanda por movimentação de dados e, deste modo, aumenta o desempenho, os custos reduzidos e a sustentabilidade.

Existem diferentes algoritmos de compactação. Alguns exemplos abrangem codificação de execução duradoura, em que os itens de repetição de dados são representados uma vez com suas contagens correspondentes, bem como com aplicativos de compactação Lempel-Ziv juntamente com valores de dados em linhas. No exemplo anterior, os padrões comuns eram extraídos e representados pela utilização de valores numéricos compactados e, então, o padrão era substituído cada vez que isso ocorresse, de modo a reduzir a necessidade de armazenamento. As técnicas de compactação de mercadorias podem reduzir o armazenamento em até 50%, e são frequentemente utilizadas na redução de armazenamento em até 80%.

5.4.5 Maior Desempenho através de Melhorias em Hierarquia de Memória

Diferentes arquiteturas de computação geralmente variam em relação ao nível em que os vários aspectos do sistema, como CPU, cache de memória, memória principal, áreas de armazenamento em disco temporárias e armazenamento persistente em disco, contribuem para aumentar o desempenho do sistema, especialmente no contexto de cargas de trabalho de inteligência de negócios combinadas. Esses módulos de armazenamento, junto com outros, constituem o que é conhecido como "hierarquia de memória", que os arquitetos de hardware empregam em várias configurações para encontrar a combinação certa de dispositivos de memória com tamanhos e custos variados, e velocidade para fornecer ótimos resultados ao reduzir a latência para a resposta às crescentes consultas e análises complexas.

Os aplicativos orientados por desempenho tendem a utilizar a hierarquia de memória para minimizar a latência através de acesso inteligente de dados e estratégia de armazenamento em cache. O valor dessas melhorias pode ser multiplicado com o uso estratégico de hardware. Por exemplo, a resposta para diferentes tipos de consultas exige a gravação em espaço de memória temporário, utilizando, de modo mais frequente, discos giratórios como mídia. A utilização de memória flash sólida e compacta para armazenamento temporário fornece acesso mais rápido, reduz a latência de E/S, reduz o tempo de resposta de consulta e, conseqüentemente, resulta em um aumento do rendimento de consulta.

5.4.6 Computação de Fluxo

O processo de tomada de decisão é cada vez mais informado por uma variedade crescente de entradas, das quais muito poucas são baseadas em dados estruturados. De modo alternativo, as pessoas que tomam decisões utilizam uma ampla faixa de diferentes tipos de dados não estruturados a partir de diferentes origens. A varredura manual de todas essas entradas criaria um gargalo significativo para a produtividade. Em vez disso, há uma necessidade de que os métodos não apenas verifiquem de modo rápido a grande quantidade de origens de dados e analisem de modo contínuo os dados de fluxo em tempo real, de modo que os fragmentos corretos de informação possam ser filtrados, reduzidos e transformados em conhecimento acionável a fim de serem entregues às pessoas certas, no tempo correto. À medida que mais origens de informação se tornam disponíveis, esses métodos devem ser adaptados aos dados heterogêneos em diferentes origens, formatos e até mesmo tipo, incluindo texto, áudio, imagens e vídeo.

Basicamente, isso exige métodos eficientes para análise de fluxo de dados, permitindo a análise de cada fragmento de dado enviado, de modo que uma ação imediata possa ser tomada quando informações crítica são referidas. Os dados também podem ser salvos de modo seletivo para processamento subsequente. Esse tipo de análise de fluxo de informações fornece, de modo efetivo, um recurso de "federação de dados" estratificado sobre um mecanismo de consulta utilizado para a análise de grandes quantidades de dados dinâmicos a partir de muitas origens heterogêneas. A análise fornecida pela computação de fluxo fornece resposta rápida aos ambientes de mudança e eventos ao atualizar os conjuntos de resultado de busca em uma etapa de bloqueio, conforme as origens de dados são atualizadas. De modo básico, diversas (várias) origens de dados podem ser revisadas, filtradas, limpas, agregadas ou sujeitas a transformações adicionais a fim de satisfazer as várias necessidades de recebimento de dados pelo consumidor de dados.

5.5 *Serviços Analíticos Avançados e Emergentes*

Além das chamadas analíticas de “fluxo principal”, existem técnicas emergentes que estão sendo rapidamente incorporadas ao ambiente a fim de suportar inteligência operacional difundida, como:

- Mineração de dados e modelagem preditiva
- Modelos analíticos preditivos e integrados
- Extração e reconhecimento de entidade
- Análise de texto
- Análise de opinião

5.5.1 *Mineração de Dados e Modelagem Preditiva*

A mineração de dados e outras técnicas estatísticas avançadas permitem que os analistas criem modelos que, de algum modo, repliquem alguns dos processos de execução elaborados para o reconhecimento de padrões de sucesso. Os analistas de negócios aplicaram as técnicas de mineração de dados com nenhuma expectativa pré-existente a fim de identificar padrões de modo a garantir determinação adicional de valor de negócio.

Essa análise não direcionada baseia-se na não existência de condições prévias, e pode ser um ótimo modo de iniciar o desenvolvimento de modelos preditivos. Os analistas que utilizam as técnicas de mineração de dados desenvolvem modelos preditivos que podem ser refinados, treinados e, então, aplicados em um amplo conjunto de dados a fim de identificar os padrões que correspondam às oportunidades ou riscos. Por exemplo, um consumidor de dados armazenados em cluster é um processo não direcionado que reúne consumidores com base em diferenças e semelhanças encontradas. A avaliação de variáveis dependentes resultantes de um consumidor de dados armazenados em cluster possibilita a classificação direcionada de novos registros de cliente nos clusters encontrados.

Existem vários métodos de mineração de dados e técnicas que podem ser combinados para desenvolver modelos preditivos, como:

- **Armazenamento em cluster**, um processo que reúne itens de modo que cada grupo seja claramente diferente dos outros, e que os membros de cada grupo sejam reconhecivelmente semelhantes.
- **Associação**, em que um conjunto de instância de dados é analisado a fim de encontrar regras que possam ser utilizadas para prever a ocorrência de um conjunto de valores com base nas ocorrências dos outros valores em cada instância.
- **O retorno**, que é um método estatístico para a análise de um conjunto de dados de modo a ajustar um grupo de pontos de dados em uma fórmula matemática. Por sua vez, essa fórmula pode ser utilizada para a conexão dos dados a fim de possibilitar a predição de variáveis dependentes.
- **A Análise da Capacidade de Mercado**, que é um processo que procura relacionamentos de objetos que pertencem ao contexto de negócios, e seu nome é derivado do conceito de análise de conteúdos de um carrinho de supermercado para visualizar se há afinidades que “ocorrem naturalmente” e que pode ser utilizadas para vantagens de negócios.
- **Raciocínio com base em caso**, em que as situações conhecidas são utilizadas para formar um modelo para análise. As novas situações são comparadas com o modelo a fim de encontrar correspondências próximas, que podem, então, ser revisadas com o objetivo de informar decisões sobre classificação e predição.

- **A análise de decisão estruturada em árvore**, que procura um grupo de instâncias de dados e seus resultados, avalia a frequência e distribuição de valores no conjunto de variáveis, e cria um modelo de decisão em estrutura de árvore. Os nós em cada nível da árvore representam uma questão, e que possível resposta à questão é representada como uma ramificação que aponta para outro nó no próximo nível.
- **Redes neurais**, que usam métodos estatísticos e de probabilidade para analisar o treinamento de dados para a criação de um processo de “caixa preta” que coleta algumas entradas e produz saídas preditivas.

5.5.2 Modelos de Análises Preditivas e Integradas

Os modelos preditivos desenvolvidos utilizando vários dados e algoritmos de mineração de texto podem ser integrados nos processos de negócios a fim de complementar a tomada de decisão operacional, bem como a análise de estratégia, utilizando os padrões que foram revelados para prever eventos futuros ou ajudar na obtenção de metas específicas. Por exemplo, os perfis de clientes criados através da aplicação de uma análise de armazenamento em cluster podem ser utilizados para a classificação em tempo real com base em demografias específicas ou outras características. Esses perfis podem ser utilizados para recomendar oportunidades de vendas cruzadas e incrementais para o aumento de receita. Modelos preditivos integrados podem ser utilizados para encaminhar todos os nossos acionadores de valor, e são utilizados em diferentes cenários, incluindo a retenção de clientes, aquisições e compras, melhorias na cadeia de fornecimento, modelagem de fraude, melhoria na precisão de previsão, tomada de decisões com base na razão, análise de crédito e garantia automática.

5.5.3 Reconhecimento de Entidade e Extração de Entidade

Uma consequência do enorme crescimento de artefatos não estruturados é o desafio de isolar termos-chave em texto (como pessoas, lugares e coisas) e estabelecer conexões e relacionamentos entre esses conceitos. O reconhecimento de identidade em “tempo real” vem sendo, tradicionalmente, um processo em lote, porém operações em tempo crítico relacionadas aos comentários online, atendimento ao cliente, operações de centro de atendimento, ou atividades confidenciais que envolvem segurança, segredo bancário, lei anti/contra lavagem de dinheiro, ou outros aplicativos de “pessoas de interesse” se tornam significativamente mais eficientes quando as identidades dos indivíduos podem ser reconhecidas em tempo real.

O desafio é superar a procura por padrões de nome em textos sequenciais, porém ele cria conceitos de processamento de linguagem natural para expor relacionamentos que sofreram exposição (como a afinidade de uma pessoa por uma instituição de caridade específica), relacionamentos casuais (como a correlação de problemas com produtos em regiões geográficas), ou várias referências à mesma entidade (como a introdução de pronomes como “Ele” ou “Isso” que se referem a entidades nomeadas com pronomes adequados, como “George Washington”). O reconhecimento de identidade em tempo real permite a conexão rápida entre pessoas e seus atributos, características, perfis e históricos de transações relacionados e pode ser utilizado em modelos preditivos integrados em tempo real para aprimorar a tomada de decisões operacional.

5.5.4 Análise de texto

A análise de texto pode ser utilizada para isolar palavras-chave e conceitos com textos estruturados e semiestruturados, e esses artefatos de texto-chave são analisados semanticamente, modelados, e seus documentos de origem são correlacionados com base em conceitos reconhecidos. Isso implica na necessidade de taxonomias de conceito nas quais tais termos podem ser coletados e agregados em diferentes níveis de precisão, como marcas de carro, modelos, bem como versões alternativas que resultam da presença ou ausência de recursos específicos.

Os algoritmos para reconhecimento e extração de entidade e análise de textos precisam ser mais sofisticados em relação à estrutura do texto. Embora simples, as entidades não estruturadas com base em padrão (como números de telefone) podem ser examinadas utilizando técnicas como análise regular de expressão. Além disso, padrões mais complexos e técnicas confidenciais de contexto estão sendo utilizados cada vez mais. Uma norma para análises de conteúdo chamada UIMA (Unstructured Information Management Architecture) foi estabelecida em 2009 pela OASIS. Ela orienta o desenvolvimento de uma estrutura para a integração de componentes de análise de texto e regras que impulsionam o desenvolvimento de software de análise não estruturada.

Esses componentes permitem a execução de análises de termos frequentes, determina sentinelas ou termos de sinal, cria hierarquias de conceito, dicionários e normas de documento para reconhecimento de frase e extração de conceito, entre outras técnicas. Uma vez que essa análise é finalizada, a informação contida nos documentos pode ser armazenada em cluster, categorizada e organizada para suportar buscas inteligentes, e os conceitos de filtragem dos textos de fluxo ajudam a identificar artefatos de texto importantes que podem ser encaminhados diretamente às pessoas com um interesse específico no conteúdo fornecido. Depois que os conceitos foram ordenados, identificados e extraídos, eles podem ser submetidos à mineração de dados e outros tipos de análise de analíticas para ajudar o profissional do conhecimento a obter conclusões de informações acionáveis.

5.5.5 Análise de Opinião

Sites de mídia social, redes online e blogs fornecem uma grande oportunidade para uma ampla variedade de pessoas publicarem revisões pessoais, avaliações de produtos, classificação de serviço, experiências e outras opiniões, geralmente aquelas que influenciam uma ampla gama de outras pessoas na rede do autor. Em oposição aos pensamentos que não refletem o que foi apresentado, as organizações procuram tirar vantagem dessas crescentes redes influenciadoras ao encaminhar rapidamente opiniões negativas ou explorar as positivas.

A análise de opinião encaminha as análises de texto a um nível avançado através de textos não estruturados para revisar e avaliar a subjetividade na atitude no material do autor. Por exemplo, um fabricante pode analisar relatórios de centro de atendimento para os nomes de peças que aparecem frequentemente em interações negativas, como falhas em produtos. Isso pode ajudar na identificação de padrões comuns de falha, permitindo, assim, ações proativas para entrar em contato com os proprietários dos produtos antes que as peças apresentem falhas. A análise de opinião também possui outras oportunidades, como a identificação de tendências emergentes de consumidores, a identificação de preferências de clientes ou clientes insatisfeitos. Isso permite que a reputação online dos negócios seja gerenciada de modo a destacar as opiniões positivas, ao mesmo tempo em que reduz os impactos das opiniões negativas.

A análise de opinião é a soma de várias técnicas discutidas neste artigo, ou seja a análise de frequência de termos, dedução de taxonomias e hierarquias, marcação de artefatos de documento com seus rótulos de conceito correspondentes, organização de conceitos em relação aos dados estruturados alternativos, e aplicação de análises de mineração de dados para procurar padrões, associações, causalidade e outros tipos de relacionamentos.

5.6 Serviços de Entrega de Análises de Negócio

A gama de análises de negócio suporta uma variedade de consumidores de dados pela organização:

- Relatório e Consultas Ad Hoc – Relatórios estatísticos e padrões derivados de especificações de usuários fornecem um visão consistente de aspectos específicos de negócios, gerados em lote e geralmente fornecidos com base em planejamento através de interface (Web) padrão. O caráter estatístico dos relatórios impulsiona a necessidade de métodos alternativos para insight adicional. Uma abordagem é a extração dos dados reportados em planilhas para manipulação adicional de dados, ao mesmo tempo em que também permite consultas ad hoc a fim de reunir dados adicionais para análise. *Relatórios padrão podem fornecer conhecimento para uma ampla gama de consumidores, mesmo se eles necessitarem de conhecimento contextual para identificar indicadores-chave e agir. Entretanto, dado o crescimento de dados em petabytes, o relatório padrão é rapidamente condescendente ao relatório de exceções.*
- Painéis e Pontuações – Se uma pessoa capacitada for necessária para examinar as métricas de desempenho principais a partir de relatórios gravados, a simplificação da apresentação de métricas de desempenho principais pode melhor capacitar o profissional do conhecimento para a transição entre ver o que já aconteceu e entender as mudanças necessárias para a melhoria do processo de negócios. Painéis e pontuações customizam uma apresentação atualizada das métricas resumidas de desempenho, permitindo monitoramento contínuo ao longo do dia. Os mecanismos de entrega difundida podem impulsionar os painéis para uma grande variedade de canais que variam do formato tradicional com base em navegador aos dispositivos móveis portáteis. *Por meio do caráter interativo do painel, o profissional do conhecimento pode fazer detalhamento através dos indicadores principais em relação a quaisquer oportunidades emergentes, bem como agir pelos mecanismos de comunicação e fluxo de processo integrado.*
- Mash-ups – Os mash-ups impulsionam o painel para um nível superior, permitindo que os consumidores de conhecimento possuam a capacidade de identificar sua própria combinação de analíticas e relatórios com fluxos de dados externos, novas alimentações, redes sociais, e outros recursos de Web 2.0 em uma estrutura de visualização que corresponde especificamente aos seus próprios objetivos e necessidades de negócios. *A estrutura de mash-up fornece o elemento final para a integração de fluxos de dados e inteligência de negócios com aplicativos de negócios interativos.*
- OLAP (Multidimensional Analysis and Online Analytical Processing) – A análise multidimensional fornecida pelas ferramentas OLAP ajuda os analistas a examinar relacionamentos entre variáveis diferentes (em suas próprias hierarquias), como “quais são as receitas corporativas por período?” ou “qual é a disponibilidade dos produtos do fornecedor por localização?” O uso da palavra “por” sugere uma tabela dinâmica em que os dados podem ser visualizados, permitindo a observação das vendas por período de tempo e, então, por região, ou de outro modo, agrupado por região e depois por período de tempo. *O OLAP permite que o analista realize detalhamentos através das hierarquias em dimensões diferentes a fim de descobrir relacionamentos dependentes que estão escondidos nas hierarquias.*

6 Desafios

Um programa de análises e inteligência de negócios desenvolvido possuirá um suplemento completo desses componentes tecnológicos para suportar consumidores de conhecimento pela gama completa de análise. Em muitos ambientes, o programa de análises de negócio se desenvolveu de modo natural, com uma variedade de ferramentas adquiridas, soluções desenvolvidas internamente, integradas com diferentes escolhas de hardware, rede e software (como sistemas de gerenciamento de banco de dados), conduzindo a uma solução aproveitável, se não a mais eficiente.

E enquanto os vários fornecedores se esforçam para suportar a interoperabilidade (de modo que todos eles executem em conjunto), os sistemas de BI foram projetados ao longo do tempo com componentes heterogêneos fornecidos por diferentes fornecedores. O pequeno foco na complexidade da integração de componentes permite menor desempenho ou otimização. Na realidade, conforme a necessidade de rapidez aumenta, estamos reconhecendo que existem alguns desafios ocultos para as organizações, conforme eles tentam criar sua própria infraestrutura analítica abrangente, incluindo esses fatores:

- **Desenvolvimento Coordenado e Heterogeneidade** – O caráter coordenado do desenvolvimento significa que os aplicativos analíticos foram incorporados sob demanda, com nenhum plano de programa abrangente ou uma avaliação das necessidades de negócios na empresa. Isso conduz para as dependências técnicas com base nas decisões de desenvolvimento que não são especificamente relacionadas com o direcionamento das necessidades de negócio e, no momento, essas dependências podem impedir o desenvolvimento de uma solução analítica flexível e completa.
- **Flexibilidade e Extensibilidade** – Apesar das tentativas de muitos fornecedores de possibilitar interoperabilidade, cada um é limitado à suas capacidades de trabalhar bem com aquelas (geralmente liberadas) versões de produto para as quais há especificações publicadas. Na realidade, isso impõe restrições de integração rigorosas que podem evitar que o cliente ative todos os recursos disponíveis do produto. Por exemplo, se a ferramenta selecionada de limpeza de dados é somente executada com a versão 5.7 do produto ETL selecionado, o cliente deve evitar a atualização para a versão 6 deste produto até que o fornecedor de limpeza de dados capacite seu produto para suportar a atualização. Além disso, conforme as necessidades de negócios, requisitos, expectativas analíticas ou quantidade de consumidores mudam, as infraestruturas analíticas destacadas terão que ser adaptadas a essas mudanças. Isso sugere a necessidade de uma capacidade de adicionar com facilidade recursos e funcionalidade à infraestrutura de inteligência de negócios.
- **Qualidade de dados** – Embora haja uma crescente concentração em controle de dados e gerenciamento de qualidade, ainda existem possibilidades de inserção de falhas nos dados. Validação intermitente de dados, diferentes ferramentas de qualidade de dados (para análise, normatização e limpeza) e regras conflitantes ainda contribuem para falhas e inconsistência nos dados.
- **Melhores valores** – A instalação, teste e validação de vários componentes e a garantia de que eles serão executados ao mesmo tempo exigem que um investimento significativo de tempo e recursos seja feito em planejamento, projeto, implementação e disposição. A complexidade crescente da implementação e da disposição aumenta o tempo até que o sistema possa ser utilizado de modo produtivo.

- **Desempenho e escalabilidade** – Muitos sistemas de BI são prejudicados por sua excelente execução, pois conforme aumenta o número de usuários, o carregamento de consultas cresce, ou à medida que a quantidade de dados a ser analisada aumenta, a capacidade do sistema de escalar condutas apropriadas de desempenho diminui. Esse desafio na escalabilidade somente aumenta quando a interoperabilidade restringe reguladores de modo incorreto ao potencial de desempenho de quaisquer componentes integrados.
- **Soluções rápidas** – A necessidade de criar novamente os mesmos (ou semelhantes) relatórios e análises canaliza a produtividade e os recursos, e estende os melhores valores. Antes, as organizações obtinham benefícios com a incorporação de ferramentas e tecnologias para suportar aplicativos utilizados com frequência, como insight do cliente e lucratividade, análise de mercado, análise de gastos e outros acionadores comuns de valor.

7 Resumo: Direcionando os Desafios

Ao considerar todos esses desafios reunidos, surge um problema comum: as ineficiências introduzidas devido ao acúmulo progressivo de componentes tecnológicos a partir de várias origens irão, por fim, diminuir a capacidade da organização de entregar inteligência acionável às pessoas corretas, mesmo se for devido ao aumento da inteligência operacional integrada, grandes volumes de dados, aumento no número de usuários ou (mais provavelmente) uma junção dessas demandas.

E se a causa-raiz desses fatores de risco for a variedade de componentes tecnológicos, então a mitigação desses riscos exigirá a consideração das opções para a criação de uma solução completa que seja planejada para aproveitar melhor os componentes adicionais. Um conjunto completo de ferramentas que dá suporte às necessidades de análise e relatório de toda a organização direciona nossos desafios:

- **Solução completa** – Com uma arquitetura bem planejada, a equipe de programação pode articular uma estratégia que atenda as necessidades de negócios com uma solução abrangente em que todos os componentes sejam projetados para execução conjunta. A seleção de uma solução completa a partir de um único fornecedor não somente simplifica a implementação e disposição, como também simplifica o processo de compra, enquanto reduz os riscos de um ambiente heterogêneo.
- **Flexibilidade e Extensibilidade** – Um único provedor pode fornecer excelente flexibilidade, especialmente quando atualizações e liberações podem ser sincronizadas de modo que garantam que as melhorias funcionais não sejam incorretamente limitadas à versão do produto. Além disso, o cliente pode introduzir funcionalidade, conforme necessário, ao implementar novas atualizações ou módulos de modo sincronizado com o plano de programa estratégico.
- **Qualidade de dados** – A normatização da validação de dados, a limpeza, as ferramentas aprimoradas e as maneiras como essas ferramentas são utilizadas fornecem um nível previsível de consistência nos dados corporativos de uma entrada inicial (ou criação) até os vários consumidores de recebimento de dados.
- **Melhores valores** – A redução da complexidade do sistema com a unificação da plataforma de solução completa simplifica o processo de aquisição, reduz os requisitos de recursos para implementação, treinamento e disposição e, conseqüentemente, acelera a obtenção de melhores valores.
- **Soluções conjuntas** – Os fornecedores com experiência na resolução de desafios gerais dos clientes conseguem integrar as melhores práticas à sua solução conjunta de modo a alavancar seus componentes técnicos a fim de direcionar acionadores comuns de valor, como a geração de receita através de vendas cruzadas e incrementais ou redução nos custos ao utilizar análise de gastos. De modo alternativo, o tempo de trabalho com os clientes em linhas de negócios definidas ou segmentos de mercado permite que os fornecedores customizem suas ofertas de soluções de modo a obter vantagem dos especialistas no assunto para criar soluções que atendam às necessidades de negócios para segmentos específicos, como o posicionamento da localização para o setor de varejo, o encaminhamento de otimização para logísticas, ou avaliação de áreas difíceis para o setor de seguros.

- **Desempenho e Escalabilidade** – Quando uma solução completa é projetada para ser executada em um hardware específico, os desenvolvedores conseguem obter vantagem de várias otimizações integradas diretamente nas plataformas de software e hardware, como o gerenciamento de fluxo de trabalho, planejamento de tarefas e processo, balanceamento de carga, canais paralelos de E/S ou alta disponibilidade. Os serviços de gerenciamento de banco de dados analíticos e otimizados permitem o data warehousing de alto desempenho, suportados por integração de dados feita em paralelo, além de serviços de federação em alta velocidade. A crescente quantidade de consultas pode ser transferida para unidades alternativas de processamento ou encaminhadas para banco de dados em memória, reduzindo as cargas de DBMS, ao mesmo tempo em que aumenta as taxas de resposta e rendimento.

Com a transição de uma evolução coordenada de inteligência de negócios corporativa, e o ambiente analítico criado em uma gama de componentes tecnológicos para uma solução completa estrategicamente planejada, sua empresa poderá obter melhores valores em tempo real e analíticas integradas, resultando na entrega de inteligência vantajosa às pessoas adequadas que tomam decisões e no momento ideal.

Sobre o Autor

David Loshin, presidente da Knowledge Integrity, Inc, (www.knowledge-integrity.com), é um líder de ideias reconhecido e consultor especialista nas áreas de qualidade de dados, gerenciamento de dados principais e inteligência de negócios. Ele escreve muito sobre as melhores práticas de BI, através de um canal especialista em www.b-eye-network.com e possui vários livros e artigos sobre qualidade de dados e BI. Seu livro, "Business Intelligence: The Savvy Manager's Guide" (junho de 2003) foi reconhecido como um recurso que possibilita que os leitores "obtenham entendimento sobre inteligência de negócios, instrução sobre gerenciamento de negócios, data warehouse e o modo como todas essas capacidades são executadas em conjunto" Ele é o autor de "Master Data Management", que foi aprovado pelos líderes do segmento de mercado de gerenciamento, e lançou recentemente o "The Practitioner's Guide to Data Quality Improvement", que foca em processos práticos para a melhoria dos utilitários de informação. Acesse <http://dataqualitybook.com> para obter mais insights sobre qualidade de dados, e <http://mdmbook.com> para ideias sobre o gerenciamento de dados principais.

David pode ser contatado em loshin@knowledge-integrity.com