

Qual é o “x” da questão com relação a dados?



RESUMO EXECUTIVO

A inovação em software continua a gerar avanços sem precedentes que transformam o mundo à nossa volta, nos empoderam como indivíduos e desenvolvem nossas economias.

Porém, o potencial pleno dessa transformação digital só pode ser realizado se explorarmos o potencial dos dados gerados por essas inovações. A verdade é que estamos vivendo uma revolução no setor de dados. O que impulsiona essa revolução, além da abundância de dados disponíveis atualmente, são as tecnologias fundamentais que alteram a forma em que coletamos, armazenamos, analisamos e transformamos as informações.

Atualmente, 90% dos líderes empresariais citam dados como um dos recursos essenciais e como um diferenciador fundamental para os negócios, no mesmo nível de recursos básicos como terra, mão de obra e capital.

Há pouco tempo, por exemplo, para coletar dados era preciso observar padrões climáticos ao longo de centenas de anos para reconhecer padrões de chuva. Isso significava ficar sentado à beira de uma estrada para registrar a velocidade do tráfego para planejar redes de transporte. Envolveria juntar milhares de quilômetros de pastas, cheias de anotações feitas à mão, para estudar o funcionamento de doenças e como elas poderiam ser curadas.

Agora, dados são gerados por sensores em milhões de dispositivos, máquinas, veículos e até mesmos postes de rua. No passado, manter essa quantidade de dados era caro e difícil, mas as capacidades de armazenamento avançaram e os custos diminuíram muito, tornando os dados armazenados um recurso renovável. Com essa capacidade de reaproveitar e adaptar dados, podemos analisá-los e transformá-los de maneiras novas para que gerem ideias valiosas para economizar tempo, dinheiro e até mesmo salvar vidas.

Alguns dos dados capturados são informações pessoais, portanto, sistemas de segurança avançados e modelos de administração responsável devem ser utilizados para garantir que as informações fiquem seguras e sejam usadas corretamente. A ampla maioria dos dados, porém, vem dos muitos dispositivos e máquinas que compartilham dados entre si e entre seus operadores. Da linha de montagem na fábrica ao avião de passageiros em voo, milhões de bytes de dados são gerados e depois analisados. Isso ajuda a melhorar o desempenho e aumentar a produtividade de maneiras que antes eram inimagináveis.

Embora os dados estejam por toda parte e sua ubiquidade e utilidade melhorem nossas vidas de tantas maneiras diferentes, muitas pessoas não entendem o que eles são, de onde vêm, como podem ser usados e como seu potencial é inerentemente gigantesco.

Este artigo descreve alguns exemplos concretos de como a inovação em dados está promovendo avanços extraordinários em alguns dos maiores desafios do mundo. Ele descreve como alterações básicas na forma em que dados são coletados, armazenados, analisados e transformados nos colocam na iminência de aproveitarmos todas as possibilidades da economia digital do Século XXI – e além. E também aborda alguns mitos que ganharam

notoriedade à medida que as pessoas continuam a trabalhar para compreender melhor o crescente setor de inovação em dados. Por fim, o artigo contém um glossário de termos que definem a linguagem da inovação em dados e que serve como um guia para quem está começando a compreender a economia dos dados.

A oportunidade que a inovação em dados oferece ao mundo é praticamente inédita. Ferramentas de software inovadoras já revolucionam nossas vidas de maneiras incríveis. Agora, essas ferramentas estão ajudando as pessoas a revelar as respostas ocultas na crescente abundância de recursos de dados. Essas novas tecnologias transformadoras estão convertendo dados em novos produtos, novas soluções e novas inovações que prometem mudar nossas vidas. Do ponto de vista econômico, uma melhor utilização dos dados poderia gerar um “dividendo de dados” de US\$ 1,6 trilhão apenas nos próximos quatro anos. Economistas estimam que ganhos de eficiência gerados por dados podem agregar quase US\$ 15 trilhões ao PIB global até 2030.

Se fizermos escolhas inteligentes hoje, essa emergente economia “centrada em dados” pode se tornar uma grande geradora de novos empregos e setores, novas descobertas e novas curas – além de estimular o crescimento econômico nas próximas décadas.

DEFINIÇÃO DE “INOVAÇÃO EM DADOS”

Muito já foi dito sobre os “quatro Vs” da inovação em dados: **V**olume, a quantidade de dados; **V**elocidade na qual eles são criados; **V**ariiedade, os tipos de dados envolvidos; e **V**eracidade, a exatidão dos dados. Contudo, menos atenção foi dada à discussão do pouco valor gerado a partir de dados brutos e da oportunidade revolucionária que compartilhamos de verdadeiramente maximizar sua utilização.

Conforme é analisado neste artigo, os dados devem ser coletados, armazenados, analisados e transformados para gerar benefícios que variam de práticos a benefícios que salvam vidas. Esses processos são essenciais para a inovação em dados – a geração de um valor imenso a partir de vastas quantidades de informação que do contrário seriam improdutivas.

SUMÁRIO

- 5** **INTRODUÇÃO**
- 7** **TRANSFORMAÇÃO DE DADOS EM RESPOSTAS**
- 14** **UMA ECONOMIA MOVIDA A DADOS**
- 17** **SEPARANDO MITOS E FATOS SOBRE DADOS**
- 28** **DISCURSO DIGITAL: COMPREENSÃO DA LINGUAGEM DOS DADOS**
- 34** **NOTAS DE FIM**
- 40** **SOBRE A BSA**



6000 A.C.

Século XV



Década de 1850

Século XXI



UM MARCO PARA O USO DE DADOS

INTRODUÇÃO

Ao longo da história humana, os marcos da civilização foram caracterizados por progressos em nossa capacidade de observar e coletar informações. Nossos ancestrais desenvolveram ferramentas para medir distância, peso, volume, temperatura, tempo e localização — todas melhoraram com o tempo e todas fundamentais para a transformação de caçadores-coletores em agricultores e, posteriormente, em habitantes de cidades.

Em 6000 a.C., utilizávamos dados sobre produção agrícola e áreas não cultivadas para aumentar a produtividade das lavouras e alimentar mais pessoas. No Século XV, utilizávamos dados dos céus para navegar pelo mundo e abrir os mares para o comércio global. Na década de 1850, utilizávamos dados para associar epidemias de cólera à má qualidade da água e salvamos vidas.

Ao longo da história moderna, até mesmo pequenas quantidades de dados nos ofereceram informações importantes na busca de soluções para alguns dos nossos maiores desafios. Registrados em pedra, rolos de papiro, volumes manuscritos ou livros impressos, os dados – e seu crescente predomínio – têm sido um dos principais motivadores do progresso humano e econômico.

No Século XXI, estamos vivendo uma rápida aceleração desse processo. Conforme dados se tornam mais abundantes e o custo do armazenamento de dados diminui, cientistas de dados podem se beneficiar de novas tecnologias e ferramentas de ponta que revelam insights valiosos a partir de vastas quantidades de dados. Conforme essas tecnologias de processamento de dados se tornam mais transformadoras, seus impactos se tornam mais profundo e as oportunidades ainda mais difundidas.

Estamos a caminho de um mundo com praticamente sem limites para informações e possibilidades. Pense em como os dados são usados para fazer previsões que melhoram nossa vida cotidiana. Dados preditivos nos ajudam a saber

com antecedência se devemos levar um guarda-chuva para o trabalho ou se devemos ir de ônibus. Dados de trânsito são usados para sincronizar semáforos, prever horários de chegada de trens e nos ajudar a encontrar o caminho mais rápido para levar nossos filhos a um ensaio. Dispositivos vestíveis nos ajudam a monitorar nosso condicionamento físico e permitem que façamos escolhas melhores, para vivermos mais e com mais saúde, e cientistas analisam terabytes de informações genéticas para descobrir novas curas e desenvolver tratamentos personalizados mais eficazes.

DADOS FAZEM A DIFERENÇA

- + **Barcelona** está aproveitando dados para construir uma cidade mais inteligente, com a capacidade de analisar os padrões de tráfego de turistas, decidir onde colocar mais estações de aluguel de bicicletas e identificar quais regiões da cidade precisam de mais caixas eletrônicas.
- + Nos **Emirados Árabes Unidos**, novas ferramentas de dados estão sendo usadas para projetar o primeiro edifício de “energia positiva” do mundo, que produz mais energia do que consome.
- + No **Quênia**, dados móveis estão sendo usados para identificar padrões de infecção por malária e definir locais de contágio que orientam os esforços de erradicação do governo.
- + Agricultores dos **Estados Unidos à Índia** estão usando dados de sementes, satélites, sensores e tratores para tomar decisões melhores sobre o que cultivar, quando plantar, como monitorar o frescor dos alimentos da fazenda ao prato e como se adaptar às mudanças climáticas.

CICLO DE VIDA DOS DADOS

ANÁLISE

TRANSFORMAÇÃO

COLETA

ARMAZENAMENTO
E CONVERSÃO

Na hora de comprar um carro, em vez de saber apenas o preço do veículo, dados nos oferecem informações sobre consumo de combustível, manutenção, seguro e segurança para que façamos escolhas mais inteligentes. Na verdade, hoje o próprio carro é um supercomputador sobre rodas. Ele tem um processador que interage com sensores de análise de desempenho para que os motoristas tenham informações sobre troca de óleo, mudança para um motor elétrico ou saibam se há crianças brincando atrás do carro antes de dar marcha a ré.

Essa crescente abundância de dados já nos ajuda a ter poder em nossas mãos colocando informações muito necessárias ao alcance de nossos dedos.

Mas o que exatamente são “dados”? Quem ou o que gera esses dados? Qual é o seu potencial de melhorar nossas vidas? Como eles devem ser utilizados para gerar o benefício máximo? E como podemos garantir que eles sejam utilizados de forma consistente com nossos valores e preocupações?

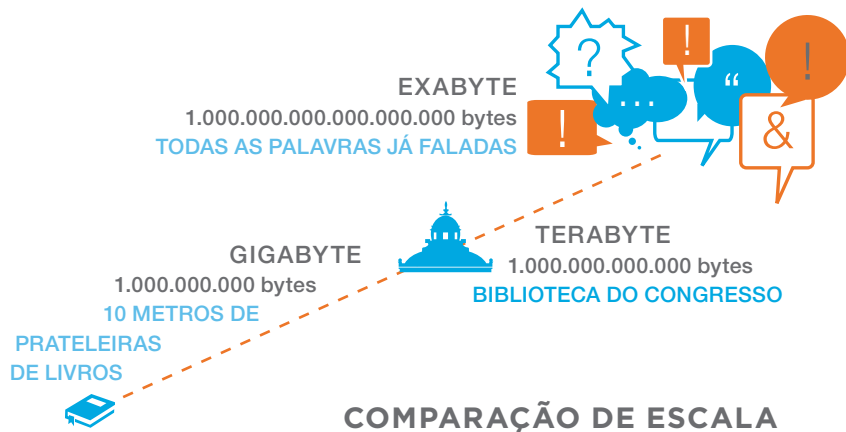
Essas são questões importantes, pois, conforme os dados deixam de ser um recurso escasso passando a ser cada vez mais abundantes, valiosos e renováveis, eles se tornam uma fonte essencial de benefícios econômicos e sociais. Historicamente, o acesso a recursos como terra, mão de obra e capital foi o diferencial econômico entre quem teve êxito e quem fracassou. Hoje, 90% dos líderes empresariais citam dados como um dos recursos essenciais e como um diferenciador fundamental para os negócios, no mesmo nível de recursos básicos como terra, mão de obra e capital.¹

Um exemplo: em uma estimativa conservadora, economistas avaliam que se um aproveitamento mais eficaz de dados gerasse pequenos ganhos, tornando os setores de atividade apenas 1% mais eficientes, isso agregaria quase US\$ 15 trilhões ao PIB global até 2030.² A “próxima grande coisa” pode se originar nas bilhões de pequenas coisas conectadas à Internet que produzem dados melhores sobre o mundo à nossa volta para gerar soluções ainda mais poderosas com base em dados.³ E já estamos encontrando respostas a perguntas que nem sabíamos que existiam.

Essa enorme mudança está em andamento. Quase tudo o que fazemos gera dados e fluxos de dados totalmente novos são criados a cada dia. Na realidade, 90% dos dados que há no mundo hoje foram criados apenas nos últimos dois anos e a cada dois anos dobramos a taxa em que os dados são produzidos. A maioria dos dados gerados não é formada por dados pessoais. Essa é uma distinção importante porque, embora seja fundamental proteger a privacidade, na maioria dos casos os dados que ajudam a melhorar nossas vidas foram gerados por um sensor conectado a uma máquina.

Nosso desafio é aproveitar os dados e colocá-los em ação, utilizando nossa engenhosidade para decifrar os valiosos ensinamentos que eles contêm. É essa capacidade de processar dados e transformar observações em ideias, e ideias em respostas, que nos permite chegar a soluções significativas para os expressivos desafios de nosso tempo.

É o momento de empresas e governos trabalharem ativamente para dar prioridade à inovação.



AS QUATRO ETAPAS ESSENCIAIS PARA TRANSFORMAR DADOS EM RESPOSTAS

A atual revolução dos dados não é impulsionada apenas pela crescente abundância de dados, mas também por tecnologias fundamentais que alteram a forma em que coletamos, armazenamos, analisamos e transformamos dados. Juntos, esses indutores nos permitem obter insights valiosos a partir dos dados e, portanto, revelar novos conhecimentos, descobrir novas conexões e fazer novas previsões.

1

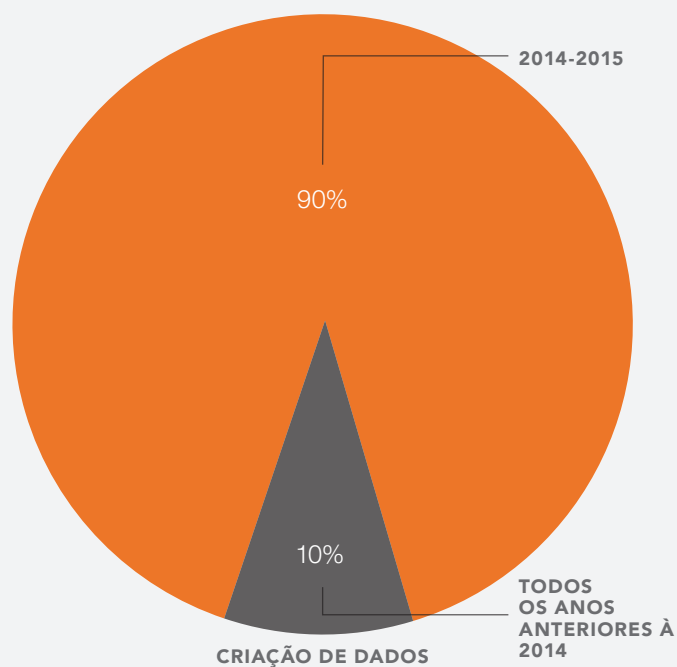
COLETA DE DADOS

Ao longo da história, sempre coletamos dados e os utilizamos para ajudar a sociedade a progredir. Frequentemente, porém, os dados eram muito escassos. Felizmente, hoje podemos nos valer de mais dados do mundo à nossa volta — dados que têm mais significado, coletados de formas mais úteis e que produzem resultados mais significativos. De algo escasso no passado, os dados se tornaram um recurso cada vez mais abundante, vital e renovável.

E não foi nossa capacidade de consolidar e extrair informações pessoais que possibilitou isso, mas sim nossa capacidade de conectar diversos dispositivos e sensores à Internet que agora gera uma fatura de novos dados que cresce exponencialmente. Consequentemente, dados são produzidos em todos os lugares — por sensores na beira de estradas utilizados para mensurar o fluxo do tráfego,

CRESCIMENTO DE DADOS NO MUNDO ATUAL

90% dos dados mundiais foram criados nos últimos dois anos.



Fonte: IBM
<http://www.ibm.com/software/data/bigdata/what-is-big-data.html>

pelos músicas e filmes digitais que criamos, pelos sensores e sistemas que controlam nossas fábricas e mercados financeiros e pelas ferramentas que usamos para projetar digitalmente a próxima novidade. Essas informações estão crescendo mais rapidamente, indo mais longe e ganhando mais importância.

Já se estima que 2,5 quintilhões de bytes de dados sejam gerados todos os dias.⁴ Para nossas mentes da era analógica é difícil até contemplar a enormidade desses dados. Para se ter uma ideia, com os dados digitais criados no mundo somente no ano passado seria possível criar uma pilha de DVDs que faria o caminho de ida e volta entre a Terra e a Lua.⁵ E o ritmo da criação de dados

A maioria dos dados não é pessoalmente identificável. Em breve, os dispositivos conectados ao redor do planeta nos ajudarão a entender melhor e melhorar o mundo à nossa volta.

também está acelerando. O volume de dados de negócios no mundo, entre todas as empresas, dobra a cada 1,2 ano.⁶ De onde vêm todos esses dados? A seguir, alguns exemplos de fontes, dentre muitas:

- + Informações digitais de hospitais, principalmente de imagens clínicas, devem alcançar 665 terabytes por dia em 2015 – ajudando a descobrir curas e salvar vidas.⁷
- + Companhias aéreas transcontinentais modernas têm tantos sensores conectados a seus motores, flaps e trens de pouso que podem gerar meio terabyte de dados por voo para melhorar o desempenho em voo,⁸ diminuir turbulências, melhorar a segurança e identificar possíveis defeitos em motores 2 mil vezes mais rápido do que antes.⁹ Multiplique isso por mais de 25.000 voos por dia e você terá uma ideia das enormes quantidades de dados úteis que estão sendo gerados apenas por jatos comerciais.
- + Satélites e observatórios climáticos, radares e outros sensores capturam mais de 2,25 bilhões de pontos de dados climáticos 15 vezes por hora, coletando 20 terabytes por dia, e possibilitando previsões climáticas mais precisas em todo o mundo.¹⁰
- + Movimentações financeiras geram de quatro a cinco terabytes de dados por dia, que são usados em análises em tempo real e na identificação de atividades comerciais problemáticas, ao mesmo tempo em que ajudam empresas a crescer e a fortalecer a economia.¹¹
- + Sensores telemáticos em dezenas de milhares de veículos de entrega monitoram o desempenho de motores, melhoram a definição de rotas e antecipam problemas. Os dados dos sensores de veículos, combinados com análises de dados de mapeamento, permitiram às empresas economizar milhões de litros de combustível e reduzir as emissões em um volume equivalente à retirada de milhares de carros das ruas durante um ano.¹²
- + O Grande Colisor de Hádrons do CERN, a Organização Europeia para a Pesquisa Nuclear, gera 40 terabytes de dados por segundo com cada experimento, fornecendo novos insights sobre os mais profundos segredos do funcionamento do universo.¹³ Da mesma forma, o Grande Telescópio de Pesquisa Sinóptica do Chile gera 30 terabytes de dados sobre o nosso universo ao observar o céu todas as noites.¹⁴
- + O sequenciamento de um único genoma de DNA gera 200 gigabytes de dados. Com o custo do sequenciamento de DNA caindo, cientistas estão criando enormes bancos de dados, com centenas de milhares dessas sequências, para encontrar as diferenças e semelhanças que se correlacionam a descobertas médicas e salvam vidas.¹⁵

E não é só a quantidade de dados que está crescendo exponencialmente, as formas em como eles são produzidos também estão. Conforme o número de dispositivos que conectam a Internet ao mundo à nossa volta aumenta, criando uma “Internet das Coisas”, um grande número de sensores cria formas inteiramente novas de dados todos os dias. A próxima grande coisa pode se basear em várias coisas pequenas, pois estima-se que 50 bilhões de dispositivos equipados com sensores poderosos estarão conectados à Internet em 2020.¹⁶

Esses dispositivos criarão dados fazendo coisas como medir a umidade do solo, o desempenho de motores, a eficiência de sistemas de energia e o local de ataques de asma. Como seres humanos, utilizamos apenas cinco sentidos para compreender o mundo à nossa volta. Em breve, dispositivos conectados ao redor do planeta detectarão uma ampla variedade de características do mundo físico para nos ajudar a compreender e melhorar o mundo em que vivemos – e, com isso, produzirão exabytes de novos dados benéficos.

Com o custo do armazenamento de dados em queda e a quantidade de dados em constante crescimento, as possibilidades de uso dos dados estão cada vez maiores.

2

ARMAZENAMENTO DE DADOS

A grande queda no custo de armazenamento está possibilitando inovações impulsionadas por dados. Em 1980, um gigabyte de armazenamento de dados era difícil de encontrar, custava centenas de milhares de dólares e exigia uma pessoa dedicada em tempo integral para gerenciar.¹⁷ Hoje, um gigabyte de armazenamento custa poucos centavos, é gerenciado facilmente e pode ser acessado a qualquer momento, em qualquer lugar.¹⁸ Desde a década de 1980, a queda no preço do armazenamento foi superior a um fator de 10 milhões.¹⁹ Para contextualizar, se o preço da gasolina tivesse diminuído na mesma medida, seria possível dar a volta ao mundo de carro cerca de 2.500 vezes gastando o mesmo que se gastava com um litro de gasolina em 1980.²⁰

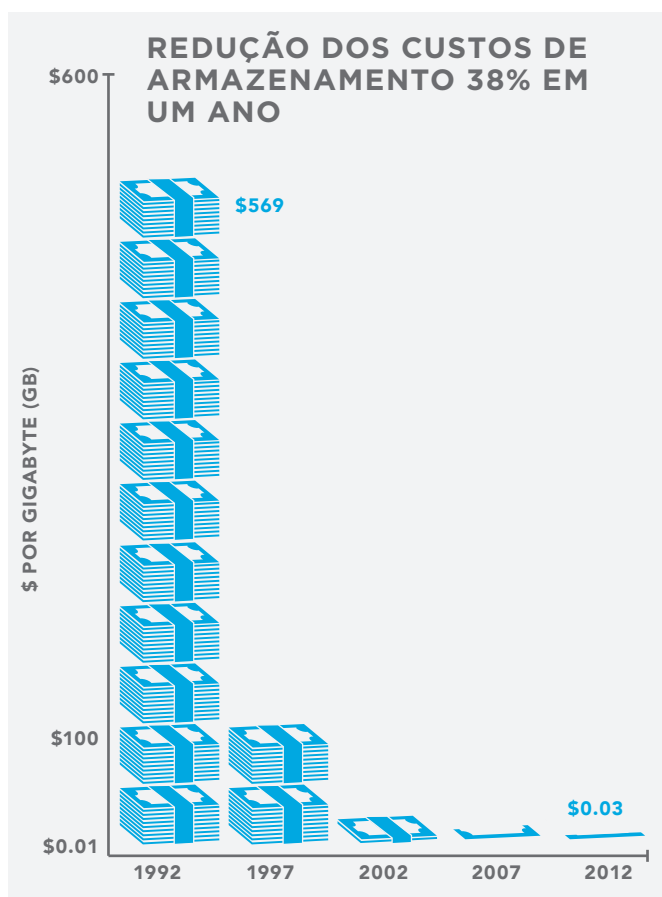
Como o custo do armazenamento continua caindo, somos capazes de armazenar quantidades cada vez maiores de dados. Em 1994, apenas 3% dos dados do mundo estavam armazenados digitalmente.²¹ Em 2007, o armazenamento digital chegava a 94%.²²

Como o custo do armazenamento de dados continua diminuindo e a quantidade de dados continua crescendo, os usos de dados continuam aumentando. Tecnologias de nuvem – que permitem que dados sejam armazenados em locais remotos e acessados por diversos dispositivos – reduziram drasticamente o preço do armazenamento de dados, fazendo com os dados não precisem mais ser excluídos após o uso inicial para dar espaço a novos dados.²³

Com isso, ao contrário de outros recursos, os dados não se esgotam conforme são usados. Dados são recursos renováveis, que podem ser combinados com outros conjuntos de dados e usados várias vezes para gerar respostas a perguntas que sequer eram esperadas quando

os dados foram criados originalmente. Por exemplo, dados climáticos não servem apenas para nos dizer se precisamos sair com um guarda-chuva, eles também podem ajudar a prever a produção de uma lavoura.

Com tecnologias de nuvem, os usuários têm um acesso melhor, mais confiável, econômico e flexível a seus dados, diminuindo a pressão que grandes quantidades de dados podem ter sobre uma infraestrutura interna de TI. Ao transformar fundamentalmente a forma em que o armazenamento de dados é comprado, vendido e fornecido – e disponibilizando dados praticamente em qualquer lugar, a qualquer momento – as tecnologias de nuvem estão se tornando as mais transformadoras da década e um dos principais motivadores de soluções impulsionadas por dados.



Fonte: Hagel III, John et al. *From Exponential Technologies to Exponential Innovation*. Deloitte University Press, 2013. Print. 2013 Shift Index Series.



DADOS EM TEMPO REAL: COMO REDUZIR O TRÁFEGO

3

ANÁLISE DE DADOS

Dados só têm valor quando podem ser compreendidos. Caso contrário, não passam de um amontoado de observações aleatórias. Compreender os insights contidos nos dados só é possível combinando engenhosidade humana e softwares inovadores.

Apesar de o mundo ser cada vez mais autônomo, ele ainda precisa de curiosidade pessoal, habilidades humanas e trabalho intenso para revelar respostas a partir de dados.

Primeiramente, dados brutos precisam ser limpos para se tornarem úteis. Cientistas de dados estimam que passam 50% a 80% de seu tempo preparando dados digitais desorganizados para que eles possam ser explorados em busca de conteúdo útil.²⁴

Em segundo lugar, é a criatividade humana que nos permite fazer as perguntas certas e depois obter respostas, classificando e reconhecendo dados ruins e interpretando os resultados de maneiras significativas. A função de cientista de dados já foi descrita como parte analista, parte artista e parte contador de histórias.²⁵ Cada dado individual é como um pixel em uma tela. Sozinho, ele fornece uma pequena quantidade de informação. Porém, ao combiná-lo com um número suficiente de pixels e na ordem correta, um cientista de dados é capaz de criar uma imagem que vale mais de mil palavras e extrair significados novos e às vezes inesperados dos dados.

Ao examinar os dados, ferramentas analíticas permitem navegar pelo aglomerado de dados para ajudar os usuários a descobrir novos padrões e tendências, obter insights inesperados a partir de dados aparentemente não relacionados e revelar relações estatisticamente interessantes de maneira automática. Utilizando bancos de dados cada vez mais ricos e algoritmos de estatística cada vez mais avançados, ferramentas de análise de software nos permitem explorar verdadeiras montanhas de dados para encontrar “pepitas de ouro de informação”.

Felizmente, a combinação das maiores capacidades de processamento nos computadores atuais com softwares inventivos fornece aos cientistas de dados ferramentas de ponta para compreender grandes quantidades de dados e obter insights valiosos deles.

Embora as redes de hoje sejam impressionantes, mover enormes quantidades de dados entre redes para um local para processá-los de uma só vez costuma ser algo economicamente inviável e logisticamente impossível. Porém, alguns dos mecanismos analíticos mais poderosos de hoje se tornaram possíveis e acessíveis por meio de enormes plataformas de computação na nuvem distribuídas paralelamente. Essas plataformas permitem que os usuários executem ferramentas de análise de dados de primeira linha em dados armazenados em diversos locais ao mesmo tempo.

O que essa análise dos dados nos permite fazer? Antes, prever o futuro parecia algo absurdo, agora, parece inevitável. Hoje, graças à moderna análise de dados, fazemos previsões confiáveis o tempo todo. Previsões climáticas para períodos de até dez dias se tornaram mais confiáveis. Gerentes de frotas podem prever quais motores precisam de reparos antes de o carro quebrar. Muitas vezes, quando dados do presente podem ser comparados com os do passado, eles podem ser usados para prever o futuro.

Economistas estão descobrindo meios para fazer previsões cada vez melhores de mercados, empregos e inflação. Por muito tempo, dados econômicos governamentais forçavam os tomadores de decisões a olhar para o passado. Estatísticas econômicas governamentais, como o crescimento do PIB, sempre analisaram meses passados para mostrar, após um longo atraso, o desempenho das economias no passado como os melhores parâmetros para o futuro. Agora, economistas estão combinando dados em tempo real, como novas vagas de emprego e produção



A IBM e a cidade de Estocolmo fizeram parceria para instalar 1.600 sistemas GPS em táxis. Os dados dos dispositivos GPS são processados por meio de um software de streaming da IBM e usados para revelar insights sobre fluxo do tráfego, tempos de trajeto e rotas favoráveis.

Reduziu a quantidade de emissões em

10%

Reduziu o trânsito na cidade em

20%

Reduziu os tempos médios de trajeto em quase

50%

Aumentou a proporção de veículos ecológicos, isentos de impostos, para

9%

Fonte: Bertolucci, Jeff. 'Dublin Points Big Data Tech At Traffic Jams'. InformationWeek 2013. Web. Nusca, Andrew. 'Stockholm Uses Real-Time GPS Data To Manage Traffic Congestion ZDNet, 2010. Web.

industrial, e comparando-os a dados históricos para terem uma ideia mais precisa da dinâmica atual e formularem políticas melhores para garantir economias mais saudáveis.

O crescimento da análise de dados em tempo real também permite uma tomada de decisões autônoma para nos ajudar, ou as máquinas que operamos, a tomar decisões mais rapidamente com mais precisão. Grandes empresas norte-americanas do setor automotivo já estão projetando novos veículos equipados com centenas de sensores, telemática e conectividade em tempo real para permitir avanços como estacionamento autônomo. Montadoras de automóveis também investem em ferramentas de análise em tempo real que permitem sistemas anticollisão autônomos e carros sem motorista. Esses progressos um dia poderão salvar vidas reagindo a situações com mais rapidez e confiabilidade do que os seres humanos.

Com a quantidade de dados em tempo real no mundo crescendo exponencialmente, quem for capaz de compreender as informações obtidas tão rapidamente quanto as obtém poderá maximizar o impacto das ferramentas de análise de dados. O poder das melhores

ferramentas de hoje em dia está na capacidade de fazer novas correlações e encontrar respostas inesperadas escondidas nos dados – mesmo quando as pessoas ainda não sabem quais são as perguntas. Ao redor do planeta, ferramentas analíticas estão encontrando correlações de alto impacto e gerando resultados inesperados. Por exemplo:

- + Rastreamento e correlacionando mais de 1000 pontos de dados por segundo, pesquisadores canadenses surpreenderam médicos ao descobrir que crianças nascidas prematuramente com sinais vitais estranhamente estáveis correlacionavam-se a casos de febre grave no dia seguinte – o que permitiu que os médicos tomassem medidas preventivas.²⁶
- + Reportagens de jornal de um período de duas décadas estão sendo usadas para prever quando epidemias de cólera ocorrerão em países como Angola.²⁷
- + Departamentos de polícia modificaram um algoritmo originalmente desenvolvido para prever terremotos e agora estão utilizando-o para prever, com precisão de 150 metros, onde crimes podem ocorrer. Roubos sofreram queda de 33% e crimes violentos de 21% nas regiões em que o software está sendo usado.²⁸
- + Utilizando análise de dados e sensores marítimos que monitoram ondas, correntes e outros dados, pesquisadores estão prevenindo tsunamis e outros desastres naturais, bem como seu impacto.²⁹
- + Dados de consultas médicas e informações sobre medicamentos prescritos revelaram como pacientes com doenças autoimunes correm risco maior de epilepsia.³⁰
- + Dados de classificação de crédito estão sendo usados para prever quais pacientes precisarão de lembretes "amigáveis" para tomar seus remédios.³¹
- + Correlacionando dados históricos de uma década de voos e padrões climáticos, passageiros de companhias aéreas podem descobrir quais voos têm menor probabilidade de sofrer atrasos.³²

A combinação das maiores capacidades de processamento nos computadores atuais com softwares inventivos fornece aos cientistas de dados ferramentas de ponta para compreender grandes quantidades de dados e obter insights valiosos deles.

Por estarmos cercados por dados, vivemos cercados por oportunidades. Quando inovadores agem com responsabilidade e criatividade, a inovação em dados pode dar respostas aos problemas do dia a dia e a alguns dos maiores desafios do mundo.

4

TRANSFORMAÇÃO E CONVERSÃO DE DADOS

Poderosas novas ferramentas de software estão nos proporcionando a capacidade de utilizar conjuntos de dados para tomar decisões melhores, com base em fatos, e não em palpites ou intuição.

Especificamente, um novo conjunto de ferramentas está ajudando a dar finalidade aos dados transformando-os de formas que nos ajudem a extrapolar, concentrar, visualizar, refletir, refinar, modelar e prever. Essas ferramentas incluem tecnologias de aprendizado de máquinas que compreendem os dados para nos ajudar a responder melhor a eles, tecnologias de modelagem e simulação que podem testar situações e transformar dados em soluções para o mundo real e ferramentas que reconhecem e convertem som, imagens ou vídeo de maneiras mais significativas.

Essas novas maneiras de transformar dados geram planos melhores, projetos de qualidade superior e decisões mais inteligentes. Por exemplo, o número de pesquisas em medicina é tão grande hoje que é praticamente impossível pra um médico se manter atualizado em relação a novos avanços, muito menos interpretar dados dos pacientes em tempo real.³³ Por isso, os hospitais estão implementando sistemas de apoio a decisões clínicas. Basicamente, são sistemas de software que analisam dados de diversas fontes para ajudar a fazer diagnósticos mais rápidos e confiáveis em um complexo ambiente de dados – comprovando-se benéfico em mais de 70% dos casos.³⁴

Outras ferramentas de software ajudam a converter dados de maneiras mais significativas. O processamento em tempo real de dados de áudio, imagens e vídeo está gerando descobertas que mudam as vidas das pessoas. Por exemplo: com o aumento da coleta de dados sobre como as pessoas falam, a tecnologia de reconhecimento de fala melhora continuamente. Isso permitiu avanços como a tradução em tempo real de conversas entre dois idiomas em diferentes continentes, o que potencialmente

pode gerar novas oportunidades para o comércio global. De maneira semelhante, com uma estimativa de que 360 milhões de pessoas sofram com problemas de audição, pesquisadores da China utilizaram reconhecimento de padrões e processamento de dados em tempo real de um sensor do 3D Kinect para desenvolver um sistema que compreende os gestos da língua de sinais e converte-os em tempo real em língua falada e língua escrita – e vice-versa.³⁵

Processamento mais rápido de imagens também está tendo um impacto profundo em áreas como detecção de câncer, computação cognitiva, neurobiologia e robótica. Por exemplo, devido à sua aparência e forma imprevisíveis, tumores cerebrais podem ser especialmente difíceis de identificar em imagens médicas. Com a ajuda da computação na nuvem e de algoritmos avançados de análise de imagens, equipes de cientistas estão competindo para descobrir os melhores algoritmos de software para identificar tumores cerebrais com mais rapidez e precisão.³⁶

A mudança de imagens 2D para 3D em mamografias está aumentando as taxas de detecção de câncer de mama. A mamografia tridimensional utiliza software para combinar muitos raios X em ângulos diferentes e criar uma imagem tridimensional, que pode aumentar a taxa de detecção de câncer de mama e diminuir o número dos tão estressantes alarmes falsos.³⁷

Na verdade, a capacidade de usar dados para criar visualizações e simulações está facilitando a compreensão e utilização de dados. Hoje, podemos modelar e simular sistemas complexos e testar projetos com mais precisão e velocidade, sem a necessidade de construí-los. Por exemplo, na década de 1980, a Boeing testou 77 protótipos do modelo 767 usando túneis de vento físicos. Em 2005, a Boeing realizou apenas 11 testes físicos para o 787, testando os protótipos usando túneis de vento virtuais



e supercomputação para economizar tempo, energia e dinheiro e salvar vidas.³⁸

Túneis de vento virtuais são um exemplo de ferramentas que trabalham com quantidades enormes de dados para que a dinâmica computacional de fluidos em 3D seja mais fácil de usar e mais rápida de implementar. Essas ferramentas permitem modelar melhor o fluxo de calor, o fluxo de fluidos, o fluxo de ar e o fluxo de processos para

gerar um desempenho melhor. Elas estão sendo usadas para modelar para onde os poluentes podem em lençóis freáticos, como aumentar o desempenho de turbinas eólicas e como projetar edifícios que possam suportar desastres naturais.

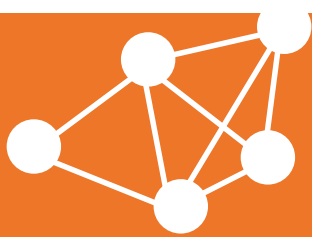
Em resumo, essas ferramentas transformam dados em soluções.

Juntos, os poderes de coletar, armazenar, analisar e transformar dados estão convergindo para revelar novas oportunidades e soluções melhores.

Na prática, porém, esses quatro motivadores costumam ser executados por pessoas diferentes, trabalhando com conjuntos de dados diferentes, armazenados em locais distribuídos. No entanto, isso faz parte do poder dessa revolução dos dados. Conjuntos de dados diferentes, que antes não tinham relação entre si, podem ser combinados e analisados mesmo quando armazenados em locais diferentes, aumentados mesmo quando os dados não forem estruturados e colocados em prática mesmo quando as partes estiverem descobrindo respostas às perguntas que os produtores de dados sequer sabiam que existiam. É impossível saber com antecedência o poder de informação que pode ser extraído de um determinado conjunto de dados – em parte porque ele pode se tornar valioso posteriormente para uma

combinação de conjuntos de dados aparentemente não relacionados e porque o algoritmo para explorar a ideia talvez ainda não tenha sido inventado.

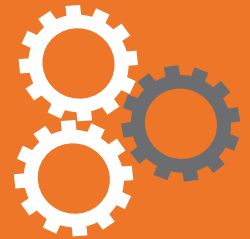
Como cercados por dados, vivemos cercados por oportunidades. Quando inovadores agem com responsabilidade e criatividade, a inovação em dados pode dar respostas aos problemas do dia a dia e a alguns dos maiores desafios do mundo. Para maximizar as oportunidades criadas pelos dados de formas que possam acelerar novas ondas de produtividade, crescimento econômico e benefícios individuais, precisamos dar prioridade à inovação e preparar o caminho para as soluções avançadas e os poderosos impactos que ela pode gerar.



Empresas de manufatura que aproveitam totalmente seus dados podem economizar

US\$ 371 bilhões

ao longo de quatro anos.



DADOS DE NEGÓCIOS: COMO REDUZIR CUSTOS

UMA ECONOMIA MOVIDA A DADOS

Dados estão emergindo como uma das forças de ganho econômico mais dinâmicas. Os impactos econômicos dos dados já estão se espalhando por diversos setores da economia, em indústrias de tecnologia avançada ou não. Nos próximos quatro anos, uma melhor utilização de dados pode gerar um “dividendo de dados” de US\$ 1,6 trilhão em todo o mundo.³⁹

A inovação em dados já mostrou sua capacidade de aumentar a produtividade. Empresas que já tomam decisões fundamentadas em dados apresentam aumentos de 5% a 6% na produtividade.⁴⁰ Se uma utilização mais eficiente dos dados gerar ganhos, ainda que pequenos, em vários setores para torná-los apenas 1% mais eficientes, os economistas estimam que pode haver um aumento de US\$ 15 trilhões no PIB global até 2030. Isso é o equivalente a uma economia dos EUA. Um aumento de 1% na produtividade pode parecer pequeno, mas, segundo Jeff Immelt, CEO da General Electric, “diga a alguém do setor petrolífero que você pode usar um software para economizar para ele 1% em alguma coisa e essa pessoa será sua amiga para o resto da vida.”⁴¹

Inovação em dados gera empregos.

A inovação em dados não está apenas incentivando o crescimento econômico. Ela também movimenta um poderoso mecanismo de geração de empregos e multiplicação da força de trabalho. A inovação em dados já gera milhares de empregos novos e bem remunerados – de analistas de dados a desenvolvedores de software e pessoas que gerenciam os enormes depósitos de dados que tornam a inovação em dados possível. Além disso, 61% dos executivos sênior nos EUA e 58% na Europa afirmam que a análise de dados é importante para os planos de suas empresas de contratar mais pessoas.⁴²

Cada emprego **RELACIONADO A DADOS**

nos EUA cria mais **3 empregos** indiretamente.

61% dos executivos americanos dizem que a **ANÁLISE DE DADOS** é importante para os **planos de contratação** ao de suas empresas.

Para cada emprego relacionado à TI que é criado, estima-se que outros três empregos sejam criados fora do setor de TI – gerando milhões de outros empregos em toda a economia.⁴³ O resultado final será de novos e amplos benefícios econômicos, de um novo boom de empregos gerados pela tecnologia.



Os edifícios inteligentes podem gerar para as empresas uma economia de energia de

US\$ 25 bilhões

por ano.



Na aviação, uma melhoria de desempenho de 1% gerada por dados pode economizar

US\$ 30 bilhões

de combustível em todo o mundo ao longo de 15 anos.



Dados estão tornando as empresas mais ágeis, reativas e competitivas.

Antes, os enormes aumentos de produtividade gerados pela TI que desenvolveram nossa economia e elevaram nossos padrões de vida estavam limitados a apenas alguns setores da economia – como o próprio setor de tecnologia. No entanto, conforme mais setores geram mais formas de dados, os dados agora devem produzir mais ganhos em setores que, tradicionalmente, ficavam para trás – da agricultura à saúde, de transportes à educação e de energia ao mercado financeiro. Na verdade, em uma pesquisa recente, 79% dos executivos sênior nos EUA e 80% na Europa disseram que a análise de dados é importante para os planos de suas empresas de atender melhor às necessidades dos clientes.⁴⁴ Além disso, 70%, dos executivos nos EUA e 80% na Europa afirmaram que a análise de dados é importante para os planos de suas empresas de criar novos produtos ou serviços.⁴⁵

ASSISTÊNCIA MÉDICA

A inovação em dados ajuda as pessoas a viver mais e com mais saúde, graças a melhores insights.

A assistência médica é um dos ambientes mais ricos em dados que temos hoje, gerando centenas de terabytes de dados por hospital todos os dias.⁴⁶ Um melhor uso desses dados pode ser a receita para um melhor atendimento e curas mais rápidas. Atualmente, graças aos dados, os médicos têm percepções inéditas da saúde de seus pacientes, o que lhes permite tomar decisões melhores. Com a ajuda de softwares analíticos, médicos podem utilizar os resultados de todos os pacientes anteriores para informar opções de tratamento para todos os pacientes futuros. Com frequência, porém, os dados de saúde são subutilizados. Se o setor de assistência médica usasse os dados com mais eficácia para aumentar a eficiência e a qualidade, estima-se que o setor poderia economizar mais de US\$ 300 bilhões por ano, reduzindo os gastos em impressionantes 8%.⁴⁷ Os maiores impactos, porém, não são apenas no dinheiro economizado, mas nas vidas salvas. Por exemplo, ao combinar dados em tempo real com o histórico médico de pacientes, pesquisadores desenvolveram um algoritmo de aprendizado de máquina capaz de prever paradas cardíacas com quatro horas de antecedência e com 66% de precisão.⁴⁸

TRANSPORTES

A inovação em dados está economizando tempo, dinheiro e combustível, além de salvar vidas.

Dados nos colocaram no caminho de um futuro com transporte mais inteligente. Em nossas vias aéreas, ferrovias e rodovias, a análise de dados em tempo real está permitindo a criação de sistemas de transporte mais inteligentes que melhoram nossa capacidade de movimentar mercadorias e pessoas com mais segurança e eficácia. Atualmente, os carros novos são equipados com tantos sensores que podem gerar até 25 gigabytes de dados por hora e contêm 10 milhões de linhas de código de software apenas para processar os dados.⁴⁹ Esses dados são usados para alimentar novos sistemas de segurança e anticollisão, cujo impacto na segurança pode ser semelhante ao do uso de cintos de segurança, reduzindo os ferimentos e mortes em até 50%.⁵⁰ Em uma época em que há cada vez mais veículos nas ruas, os dados também podem ser usados de maneiras inovadoras para reduzir congestionamentos e direcionar o trânsito mais eficientemente, economizando milhões de horas perdidas, milhares de litros de combustível e toneladas de emissões de gases do efeito estufa.

Os dados também estão decolando na aviação. Hoje, uma companhia aérea moderna pode gerar até meio terabyte de dados por voo a partir de sensores espalhados pela aeronave. Eles são utilizados para melhorar o desempenho do voo, reduzir as turbulências, melhorar a segurança e identificar defeitos em motores 2000 vezes mais rápido que antes.⁵¹ Dados de aviação também ajudam a melhorar o planejamento de rotas de voo e a informar às equipes de manutenção quais peças precisam de reposição antes de elas falharem. Os ganhos se somam. Na aviação, uma melhoria de desempenho de apenas 1% gerada por dados pode economizar US\$ 30 bilhões em combustível todo o mundo.

ENERGIA/MEIO AMBIENTE

A inovação em dados está reduzindo o consumo de energia e melhorando o meio ambiente.

Dados também estão gerando enormes economias de energia. Uma melhor utilização de dados pode promover

Se o setor de saúde fosse capaz de utilizar dados de forma mais eficiente para gerar eficácia e qualidade, estima-se que o setor economizaria mais de **US\$ 300 bilhões** todos os anos, uma redução gritante de **8%** em despesas.

economias de bilhões de dólares em energia, em diversos setores, melhorando também o meio ambiente. Insights gerados por dados permitem uma utilização de energia mais inteligente na rede elétrica, em edifícios, em nossas casas, nossas fábricas, no campo e em todo o meio ambiente. Com o uso de ferramentas de projeto de dados, por exemplo, edifícios podem ser projetados de formas que usem menos energia e equipados com sistemas inteligentes, que combinem grandes quantidades de dados de sensores com análise e atuadores para tornar as operações mais eficientes. Estima-se que a soma da economia que edifícios inteligentes podem gerar por anos para as empresas seja de US\$ 25 bilhões.⁵²

MANUFATURA

A inovação em dados está melhorando a forma como produtos são projetados, fabricados e distribuídos.

O setor de manufatura armazena mais dados do que qualquer outro.⁵³ Com isso, fabricantes têm muito a ganhar com um melhor uso de dados para aumentar a eficiência, promover qualidade e melhorar a forma como os produtos são projetados, fabricados e distribuídos. Segundo uma estimativa, uma melhor utilização de dados no setor de manufatura pode gerar uma redução de até 50% no tempo de desenvolvimento e nos custos de montagem de um produto.⁵⁴ Na realidade, a IDC (International Data Corporation) calcula que empresas de manufatura que aproveitem bem os dados podem obter um dividendo de dados de US\$ 371 bilhões ao longo de quatro anos.⁵⁵ Usando dados em tempo real, empresas também podem monitorar e gerenciar melhor suas cadeias de suprimentos globais e reduzir defeitos dos produtos.

Dados também ajudam a melhorar projetos. Projetos digitais com base em dados melhoram o processo de projetos iterativos, permitindo que projetistas testem várias ideias, ajustando-as no ambiente virtual até finalmente chegarem ao projeto ideal do produto. A Toyota, a Fiat e a Nissan reduziram o tempo de desenvolvimento de novos modelos de 30% a 50% por meio do uso colaborativo de dados e técnicas de modelagem.⁵⁶

FINANÇAS

A inovação em dados estimula a eficiência, melhora a conformidade e reduz fraudes.

No setor de serviços financeiros, o uso de crescentes quantidades de dados pode melhorar a eficiência operacional, aumentar a conformidade e identificar fraudes. Em uma pesquisa, 71% das empresas dos setores bancário e financeiro informaram que o uso de informações e análise está criando uma vantagem competitiva para suas organizações.⁵⁷ O investimento em dados pode gerar retornos enormes. A análise de dados ajudou uma empresa de cartões de crédito a identificar US\$ 2 bilhões por ano em fraudes evitáveis em cartões de crédito.⁵⁸

AGRICULTURA

A inovação em dados produz alimentos melhores e em maior quantidade utilizando menos recursos.

Dados estão brotando em fazendas em todos os lugares, ajudando a produzir alimentos mais saborosos e nutritivos para um número cada vez maior de pessoas. Graças a um conceito conhecido como agricultura de precisão, agricultores de todo o mundo podem usar dados de sementes, satélites, sensores e tratores para tomar decisões melhores visando aumentar a produtividade, diminuir custos e alimentar mais pessoas. Com um quarto da população mundial envolvida na agricultura e produção de alimentos, os ganhos obtidos com uma melhor utilização de dados podem ser enormes. Por exemplo, ao utilizar ferramentas de análise de dados, agricultores podem reduzir os custos com insumos, o uso de pesticida e produtos químicos e melhorar a produtividade em 5 a dez vezes.⁵⁹ Para produtores de laticínios, uma empresa “startup” da Croácia desenvolveu uma plataforma analítica de software com base na nuvem que fornece dados em tempo real sobre itens como impacto da qualidade da ração na produção e taxas de concepção de cada animal, o que, por sua vez, aumentou a produtividade e a eficiência em até 50% para os produtores.⁶⁰

Ao utilizar mais de 1.000 pontos de dados

por segundo, cientistas canadenses descobriram que crianças nascidas prematuramente com sinais vitais estranhamente estáveis apresentavam maior risco de desenvolver febres, o que permitiu que os médicos tomassem medidas preventivas e salvassem vidas.

CANADÁ



SEPARANDO MITOS E FATOS SOBRE DADOS

Há vários mitos em relação às recentes inovações em dados e à economia dos dados. Alguns desses mitos incluem:

- + proteção de informações e dados pessoais;
- + o impacto econômico da economia dos dados;
- + confiabilidade dos dados;
- + inovações em dados do Século XXI;
- + benefícios globais da inovação em dados; e
- + a função dos governos na regulamentação de dados.

PROTEÇÃO DE INFORMAÇÕES E DADOS PESSOAIS

MITO

A inovação em dados beneficia apenas empresas de TI, não as pessoas.

REALIDADE

A inovação em dados permite que consumidores tomem decisões melhores e que comerciantes personalizem mercadorias e serviços para nos atender melhor. Ela pode gerar grandes avanços por meio de aprendizado e medicina personalizados. Pode gerar melhores experiências para consumidores por meio de entretenimento personalizado. E pode também nos ajudar na transição de uma economia baseada na produção em massa para uma economia fundamentada na personalização em massa. De fato, há uma série de empresas que ocupam uma posição central no desenvolvimento e fornecimento de novas tecnologias inovadoras, das quais essa revolução dos dados se vale. No entanto, se fizermos escolhas inteligentes hoje, essa economia com base em dados que está surgindo pode gerar novos empregos e setores de negócios, além de colocar o mundo nas mãos dos consumidores.

MITO

Todos os dados são dados pessoais.

REALIDADE

Alguns dados podem ser informações pessoais (por exemplo, dados que geramos em nossos dispositivos móveis ou que criamos usando redes sociais). No entanto, a maioria dos dados não consiste em dados pessoais.

NA ÍNDIA, QUIOSQUES DE INTERNET
PERMITEM QUE MAIS DE

4 milhões

de agricultores acessem informações sobre
preços de produtos, clima, entre outras, em
idiomas locais.

ÍNDIA



A vasta quantidade de dados criada diariamente inclui informações como monitoramento climático por satélite, desempenho de turbinas de aviões, transações em bolsas de valores geradas por computador e sensores não relacionados a indivíduos. Mesmo quando os dados pertencem a uma pessoa, normalmente eles não são acessados por outra pessoa e são “desidentificados” – ou seja, armazenados e utilizados sem informações que revelem a identidade da pessoa envolvida.

MITO

Empresas não se preocupam com a proteção de dados pessoais.

REALIDADE

Quando dados pessoais são gerados, eles precisam ser protegidos da maneira adequada. Para expandir as oportunidades criadas pelos dados, a confiança do público e a credibilidade dos dados têm de ser altas. Empresas e organizações que usam dados devem exercer uma boa administração de tais dados. Essas práticas devem ser padronizadas, com um esforço liderado pelo próprio setor, para criar diretrizes voluntárias para a utilização responsável de dados. Muitos líderes do setor já estão agindo para deixar claro aos clientes como os dados são coletados e se eles são compartilhados. Muitas empresas seguem boas práticas que exigem que elas tornem informações pessoais anônimas sempre que isso for possível.

MITO

A inovação em dados me fará perder toda a privacidade.

REALIDADE

O sucesso da economia dos dados depende da confiança do consumidor. As pessoas devem sentir que suas informações pessoais estão seguras. Os principais desenvolvedores de software criam proteções de privacidade integradas aos seus sistemas desde o início, algo chamado de “proteção na concepção”. Além disso, desenvolvedores costumam usar ferramentas de anonimização, “desidentificação” e criptografia para minimizar o impacto de alguma possível violação dos dados. Quando os dados são agregados, impedindo que

pessoas sejam identificadas, ainda é possível analisá-los em busca de padrões de comportamento sem violar a confiança ou a privacidade dos usuários. Além disso, políticas de privacidade obrigatórias podem levar em consideração o contexto e os riscos relativos envolvidos em qualquer exposição ou mau uso dos dados, com os dados mais confidenciais (como dados financeiros ou de assistência médica) recebendo um grau de proteção de privacidade mais alto. Isso significa que dados como dados climáticos ou análises de negócios que não envolvem informações pessoais não têm a mesma exigência de proteção de dados de assistência médica específicos de pacientes.

MITO

É impossível “desidentificar” totalmente os dados. A “desidentificação” de dados é ineficaz.

REALIDADE

A “desidentificação” de dados é um processo utilizado para impedir que a identidade de uma pessoa seja associada a informações. Quando dados são “desidentificados”, eles podem ser analisados sem associação a uma pessoa. Especialistas desenvolveram técnicas de “desidentificação” que aumentam a privacidade e a qualidade dos dados.⁶¹ De acordo com especialistas, se a “desidentificação” for feita de maneira adequada, o risco de “reidentificação” de indivíduos a partir de dados anonimizados é inferior a 1%, na maioria dos casos.⁶²

MITO

Não se pode confiar em empresas que utilizam dados.

REALIDADE

O mercado ouve e leva em consideração questões de privacidade. Hoje, há sinais de uma concorrência vibrante entre grandes empresas para oferecer uma melhor proteção à privacidade. Recentemente, por exemplo, as empresas responsáveis pelos sistemas operacionais que, combinados, operam 96,4% dos smartphones do mundo anunciaram melhorias em suas configurações de privacidade. Elas estão proporcionando aos usuários

NOS EUA, GRANDES EMPRESAS DO SETOR AUTOMOTIVO estão projetando novos veículos equipados com

centenas de sensores

e ferramentas de análise para permitir avanços como estacionamento autônomo e sistema anticollisão.

EUA



controles adicionais e poder para criptografar dados e proteger a privacidade pessoal.⁶³ Empresas líderes do setor de software têm se voltado diretamente aos consumidores, divulgando seu compromisso com a proteção da privacidade.⁶⁴ As empresas estão até mesmo alterando seus serviços e políticas para aumentar a proteção da privacidade, por exemplo, não rastreando comunicações de clientes para direcionar publicidade.⁶⁵

MITO

Pessoas não têm controle sobre seus dados.

REALIDADE

Embora às vezes pareça que não temos controle sobre nossos dados, há uma série de ferramentas que talvez os consumidores não conheçam que os ajudariam a controlar melhor seus dados. Por exemplo, algumas empresas líderes no setor de software habilitaram recursos do tipo “Não rastrear” por padrão em seus navegadores da Web, portanto, os sites que você visita e as agências publicitárias que eles usam recebem solicitações de não rastreamento automaticamente. Elas informam aos sites que você não quer ser monitorado, o que pode ajudar a protegê-lo contra formas de rastreamento na Web. Além disso, alguns corretores de dados criaram páginas da Web em que os consumidores podem ver os tipos de informações que foram coletadas sobre eles, solicitar a interrupção da coleta de dados e corrigir informações incorretas.⁶⁶ Juntas, essas ferramentas oferecem aos consumidores um controle melhor da forma em que seus dados são coletados e usados ou permitem solicitar a interrupção de algumas utilizações para fins de marketing.

O IMPACTO ECONÔMICO DA ECONOMIA DOS DADOS

MITO

A inovação em dados não criará novos empregos e pode até mesmo eliminar vagas.

REALIDADE

A inovação em dados pode ser um poderoso gerador de crescimento econômico. Na realidade, 61% dos executivos sênior pesquisados nos EUA e 58% na Europa afirmam que a análise de dados é importante para os planos de suas empresas de contratar mais pessoas.⁶⁷ Contudo, a inovação em dados não gera empregos apenas no setor de TI. Estima-se que cada função relacionada a dados criará três vagas de emprego em outros setores, gerando ainda mais empregos em todos os setores da economia.

A inovação em dados ainda provocará algumas mudanças no mercado de trabalho, pois ajuda a encontrar novas maneiras de realizar tarefas antigas. No entanto, essas alterações deverão ser parecidas com a proliferação da Internet, que estima-se que tenha criado 2,6 empregos para cada emprego perdido ou que tenha deixado de existir.⁶⁸ Uma das áreas com maior potencial de crescimento de emprego é a análise de dados. Muitas vezes, só é possível compreender os insights contidos nos dados por meio da engenhosidade humana. Hoje, há uma escassez global de analistas e gerentes de dados capazes de ajudar a compreender os dados. De acordo com a McKinsey, apenas nos Estados Unidos há uma escassez de entre 140 e 190 mil pessoas com o talento analítico avançado necessário para revelar o potencial oculto nos dados e 1,5 milhão de gerentes e analistas capazes de compreender e tomar decisões com base na análise da economia dos dados.⁶⁹

Em 2014, o piso salarial médio de cientistas de dados era de US\$ 120.000 anuais e o de gerentes era de US\$ 160.000 anuais.⁷⁰ Muitas pessoas sugerem que, para extrair todo o potencial da inovação com base em dados, as empresas e as autoridades precisam tomar medidas para superar uma grande escassez de talentos que se aproxima.

NO BRASIL, um dos maiores produtores

de soja

está investindo em software e no poder da análise de dados para aumentar a eficiência das técnicas de controle de danos, reduzir custos e aprimorar as técnicas agrícolas.

BRASIL



MITO

A análise de dados elimina o discernimento humano do processo.

REALIDADE

Embora algumas perguntas possam ser respondidas com dados (por exemplo: a população da minha cidade está aumentando ou diminuindo?), muitas das respostas mais significativas não são tão claras. Nem sempre sabemos como os diversos elementos de dados se relacionam entre si. E, como talvez não saibamos com antecedência quais perguntas fazer, a análise de dados costuma ser um processo iterativo de fazer perguntas sucessivas para conseguir chegar a uma resposta. Por isso, é impossível abrir mão do discernimento e das opiniões humanas para conciliar diferenças e examinar possíveis inconsistências.

Os dados, em si, não são uma panaceia e não fazem milagres. A verdade é que, sozinhos, os dados costumam ter pouco valor. Normalmente, eles são bagunçados, sem uma organização ou estrutura coerente. O trabalho consiste em compreender os dados e encontrar o que há de relevante neles. Se os dados podem ou não resolver problemas é algo que depende da execução eficaz de uma estratégia de dados inteligente, que possa gerar soluções melhores e mais rápidas. Também depende de fazer as perguntas certas. Porém, se usarmos os dados da maneira correta, podemos ajudar a revelar respostas a alguns dos maiores desafios da sociedade, ajudar a movimentar o mecanismos da inovação e gerar uma nova poderosa onda de empregos e crescimento econômico via TI.

MITO

A inovação em dados só serve para grandes empresas – não para pequenos negócios.

REALIDADE

Com os dados cada vez mais ubíquos, os custos de armazenamento caindo e as ferramentas analíticas se tornando cada vez mais acessíveis, hoje em dia até mesmo pequenas empresas podem se beneficiar de análises avançadas de dados – ferramentas que, no passado, só estavam disponíveis às maiores empresas do mundo. Por exemplo, o recurso Trends, do QuickBooks Online da Intuit, permite que pequenas empresas se beneficiem da sabedoria coletiva de outros usuários do Intuit, podendo, assim, comparar suas receitas e despesas para revelar oportunidades. Isso permite que elas tomem decisões mais inteligentes sobre a forma em que operam. Embora o uso de soluções de analítica e inteligência de negócios não esteja muito difundido entre pequenas e médias empresas, espera-se que a adoção cresça rapidamente.⁷¹ Um estudo recente descobriu que a análise de dados é importante para 60% das pequenas empresas.⁷² Isso inclui 57% das empresas dos EUA com 50 ou menos funcionários e 62% das empresas de mesmo tamanho da Europa, de acordo com tomadores de decisão experientes. Em empresas de médio porte (com entre 51 e 500 funcionários), 87% dos executivos nos EUA e 79% na Europa afirmaram que a análise de dados é importante.

RECENTEMENTE, CIENTISTAS QUE ESTUDAM

os padrões da malária

usaram dados de telefones celulares no Quênia para especificar os locais em que a transmissão da doença estava ocorrendo, ajudando a orientar os esforços de erradicação do governo.

QUÊNIA



MITO

Os dados beneficiam apenas o setor de TI – não outros setores da economia.

REALIDADE

Dados agora estão no centro de uma grande transição tecnológica, que promete transformar e melhorar praticamente todos os setores da economia. Na realidade, muitas pessoas acreditam que a inovação em dados pode aumentar a produtividade em todos os setores econômicos. Embora as empresas que tomam decisões com base em dados informem um aumento de 5% a 6% de produtividade, se a inovação em dados gerar uma melhoria de eficácia de apenas 1%, os impactos em todos os setores da economia seriam enormes, gerando economias de energia e combustível, produzindo melhores resultados de saúde, a custos menores, e aumentando o desempenho e a vida útil de ativos físicos.⁷³ No setor de aviação comercial, por exemplo, a GE prevê que, com a capacidade de capturar dados em tempo real para melhorar a eficiência dos motores e definir rotas com mais eficácia, uma economia de combustível de apenas 1% significaria economias de US\$ 30 bilhões ao longo de 15 anos.⁷⁴

MITO

Correlação sempre implica causalidade.

REALIDADE

Essa busca pelos “desconhecidos incógnitos” é um dos principais insights que podem ser obtidos com dados. Às vezes, porém, as correlações não fazem sentido e correlações nem sempre significam causalidade. Por exemplo, apesar de a taxa de crimes violentos e assassinatos aumentar quando as vendas de sorvete aumentam, é extremamente improvável que comprar sorvete transforme as pessoas em assassinos.⁷⁵ Ainda assim, descobrir relações causais em dados está entre os insights mais valiosos que podemos descobrir nos

dados. E muitas correlações e causalidades interessantes estão sendo descobertas todos os dias, a taxas cada vez maiores, e com cada vez mais importância. A capacidade de analisar a diferença entre uma mera correlação e uma causalidade é um dos motivos pelos quais o uso de dados exige cientistas de dados qualificados, que saibam distinguir uma mera correlação de uma causalidade e eliminar resultados que não passem pelo teste de implausibilidade.

MITO

A inovação em dados requer um grande orçamento.

REALIDADE

Um bom uso dos dados não necessariamente requer grandes orçamentos. Normalmente, os investimentos necessários são pequenos e podem gerar resultados impressionantes. Prevê-se, por exemplo, que organizações que adotem uma abordagem mais holística em relação aos dados terão alguns dos maiores retornos – obtendo um “dividendo de dados” enorme, de cerca de US\$ 1,6 trilhão em receitas adicionais, custos menores e mais produtividade ao longo dos próximos quatro anos.⁷⁶ O uso de conjuntos de dados pequenos e de dados que talvez já estejam disponíveis pode ser um bom começo. Examinar uma semana de dados de transações financeiras em busca de fraudes pode ser melhor para indicar tendências do que examinar cinco anos de dados históricos. Em um estudo, empresas que empregaram análises da economia de dados com eficácia foram 26% mais lucrativas que suas concorrentes no setor, geraram 9% mais receita por meio de seus funcionários e ativos físicos e obtiveram índices de avaliação de mercado 12% superiores.⁷⁷

Nos Emirados Árabes Unidos,
novas ferramentas de dados estão
sendo usadas para projetar o primeiro edifício de

“energia positiva”

do mundo, que produz mais energia do que consome.

EMIRADOS ÁRABES UNIDOS



CONFIABILIDADE DOS DADOS

MITO

Os insights dos dados, por si só, são sempre exatos.

REALIDADE

Os insights só são exatos se os dados subjacentes forem exatos, se as ferramentas forem inteligentes e se houver um cientista de dados inteligente envolvido no desenvolvimento de um modelo para evitar resultados distorcidos. Por exemplo, o inovador aplicativo Street Bump, da cidade de Boston, EUA, usa smartphones equipados com acelerômetros e localização por GPS para encontrar buracos a partir de um grande número de usuários sempre que alguém passa por um buraco na rua. Porém, se os usuários de smartphone forem mais jovens e ricos, é possível que os dados do aplicativo não incluam dados sobre buracos de todas as regiões demográficas da cidade, portanto, a resposta da cidade aos buracos pode ser distorcida. É por isso que é frequentemente necessário o envolvimento de cientistas de dados a fim de garantir a relevância estatística e evitar distorções nos resultados dos dados.

MITO

Normalmente, decisões instintivas são as melhores decisões.

REALIDADE

A partir do momento em que nascemos, aprendemos a confiar em nossa intuição para tomar decisões. Estima-se que 19% dos gerentes globais se descrevem como “tomadores de decisão viscerais”, que confiam somente no instinto e na intuição.⁷⁸ Atualmente, porém, dados podem nos ajudar a tomar decisões mais inteligentes, com mais precisão, maior velocidade e mais impacto. De acordo com uma pesquisa feita com tomadores de decisões de TI de diversos setores, 59% afirmam que melhorar a qualidade da tomada de decisões é o principal objetivo por trás dos investimentos em tecnologias de dados.⁷⁹ Os dados não beneficiam apenas o local de trabalho.

Nós também tomamos decisões instintivas em nosso dia a dia. Em alguns casos, ferramentas podem nos ajudar a tomar decisões melhores e com mais rapidez. Por exemplo, estima-se que 93% dos acidentes automotivos sejam causados por erro humano, mas novos sensores veiculares, que coletam quantidades enormes de dados sobre seu entorno, combinados com análises em tempo real e tomada de decisões automatizada (mesmo que não seja totalmente autônoma) poderiam reduzir acidentes de carros com ferimentos ou mortes em até 50%.⁸⁰

MITO

Dados maiores sempre são dados melhores.

REALIDADE

Os conjuntos de dados nem sempre precisam ser maiores para serem melhores. O tamanho pode ser um dos fatores menos importantes. Para muitos problemas e dúvidas, até mesmo pequenas quantidades de dados, se analisadas com as ferramentas corretas, podem gerar insights interessantes. Muitas vezes, desenvolver esses insights depende mais da qualidade dos dados subjacentes e da qualidade das ferramentas usadas para analisá-los. Ainda assim, é comum a ideia de que com mais dados temos mais acesso à verdade, e de que quanto maior o conjunto de dados, mais objetivo ele é. Às vezes é o acesso a dados simples que pode ter o impacto mais imediato. Por exemplo, saber quando uma loja fecha, quantos quilômetros um carro percorreu desde a última troca de óleo ou o valor da conta telefônica do mês. Dependendo da pergunta feita, os dados contidos em páginas da Web, planilhas do Excel ou bancos de dados de CRM podem ser pequenos, mas tão capazes de fornecer respostas a perguntas quanto dados em grandes quantidades. O que importa, na verdade, é a criação de dados sólidos, o armazenamento seguro dos dados, o acesso aos dados e a capacidade de processá-los, seja qual for seu tamanho, para que eles possam ser usados quando e onde necessário para solucionar problemas.

Dados climáticos

não são usados apenas para saber se precisamos sair com um guarda-chuva. Eles podem ser usados para ajudar a prever a produção agrícola, a probabilidade de uma epidemia de gripe e ajudar a calcular a capacidade de dessalinização que precisa ser construída no Marrocos.

MARROCOS



MITO

Dados não estruturados são inúteis.

REALIDADE

Dados não estruturados podem ser úteis. Alguns dos dados mais poderosos nem sempre são compatíveis com as linhas e colunas de tabelas bem estruturadas. Eles podem ser não estruturados, como documentos de texto ou imagens de raio X. O segredo é transformar os dados em seus formatos mais úteis. Se a empresa média da lista Fortune 1000 pudesse aumentar a utilidade de seus dados em apenas 10%, ela poderia ter um aumento de mais de US\$ 2 bilhões em receitas.⁸¹

MITO

Os dados só devem ser usados para a finalidade original para a qual foram coletados.

REALIDADE

Tecnologias de nuvem reduziram drasticamente o preço do armazenamento de dados, fazendo com os dados não precisem mais ser excluídos após o uso inicial para dar espaço a novos dados.²³ Com isso, ao contrário de outros recursos, os dados não se esgotam conforme são usados. Dados são recursos renováveis, que podem ser frequentemente reutilizados, combinados com outros conjuntos de dados e usados várias vezes para gerar respostas a perguntas que sequer eram esperadas quando os dados foram criados originalmente. Por exemplo, dados climáticos não são usados apenas para saber se precisamos sair com um guarda-chuva. Eles podem ser usados para ajudar a prever a produção agrícola, a probabilidade de uma epidemia de gripe e ajudar a calcular a capacidade de dessalinização que precisa ser construída no Marrocos.

INOVAÇÕES EM DADOS DO SÉCULO XXI

MITO

Os benefícios mais profundos dos dados ainda ocorrerão no futuro.

REALIDADE

Embora a inovação em dados continuará gerando benefícios durante décadas no futuro, seus valiosos resultados já podem ser vistos em praticamente tudo à nossa volta. Diariamente, terabytes de dados são usados para fornecer previsões climáticas mais precisas. Dados de milhões de sinais de GPS são usados para calcular tempos de trajeto e a velocidade até um destino, milhões de pontos de dados de saúde são usados para identificar as possíveis causas de doenças e terabytes de dados financeiros são usados para ajudar a evitar fraudes em cartões de crédito.

Na realidade, quando a Economist Intelligence Unit pediu que os participantes de uma pesquisa descrevessem o impacto que os dados tiveram em suas organizações nos últimos cinco anos, quase 10% disseram que os dados mudaram completamente a forma de fazer negócios.⁸² 46% dos pesquisados afirmaram que os dados se tornaram um fator importante na tomada de decisões de negócios. Uma nova pesquisa sugere que, nos próximos quatro anos, uma melhor utilização dos dados pode gerar um “dividendo de dados” global de US\$ 1,6 trilhão, com dados ajudando a revelar novas oportunidades e soluções.⁸³ Em outro estudo, com foco no ano corrente, 33% dos executivos nos Estados Unidos e 24% na Europa afirmaram esperar que 10% ou mais do crescimento de suas empresas tenha relação com a análise de dados.⁸⁴ Pensando em cinco anos no futuro, 58% dos mesmos executivos nos EUA e 43% na Europa fizeram a mesma previsão.

Após o tsunami de 2004 no Sul da Ásia, pescadores da Indonésia receberam telefones celulares. Suas

receitas subiram em 30%,

pois era a primeira vez na vida que tinham dados sobre o preço de mercado dos peixes.

INDONÉSIA



MITO

Dados são supervalorizados.

REALIDADE

O uso de dados para resolver os problemas dos seres humanos não é algo novo. Utilizamos dados para tomar decisões melhores desde o início das civilizações modernas – criando técnicas de plantação em áreas sem cultivo para alimentar mais pessoas, técnicas de navegação que estimularam o comércio global e descobertas no setor de saúde que evitaram milhões de mortes por cólera. Porém, no passado, os dados eram um recurso escasso, caro para armazenar e de difícil manipulação. A diferença hoje é que os dados se tornaram mais abundantes, os custos de armazenamento caíram e as ferramentas para gerenciar dados se tornaram mais poderosas. Consequentemente, ao enfrentarmos um novo conjunto de desafios emergentes, novas poderosas tecnologias de análise de dados podem ajudar a examinar volumes crescentes de dados para permitir a descoberta de insights valiosos e soluções inesperadas para alguns dos nossos maiores desafios.

MITO

A era do crescimento econômico impulsionado por TI acabou – a inovação em dados não aumenta a produtividade.

REALIDADE

A inovação de TI e sua capacidade de estimular a economia, gerar empregos e elevar padrões de vida em todo o mundo está arraigada à sua capacidade comprovada de aumentar a produtividade – por exemplo, com aumento de produtividade de 1% a 2% na década de 1990.⁸⁵ Um importante economista da Universidade Northwestern defende que os maiores ganhos da inovação em TI já aconteceram.⁸⁶ Entretanto, a era da produtividade está viva e a todo vapor. Empresas que tomam decisões com base em dados informam um aumento de 5% a 6% na produtividade.⁸⁷ Mesmo que essa expansão de oportunidades geradas por dados aumente a produtividade nos Estados Unidos, por exemplo, em apenas 1,5% durante um período de 20 anos, o dinheiro economizado seria suficiente para aumentar a renda média do país em até 30%.⁸⁸ Se uma melhor utilização dos dados gerar pequenos ganhos em diversos setores, tornando-os apenas 1% mais eficientes, os economistas calculam um aumento de cerca de US\$ 15 trilhões no PIB global até 2030 – equivalente a uma economia dos EUA.⁸⁹

Barcelona está

utilizando dados para criar

uma cidade mais inteligente, melhorar os serviços do governo e oferecer soluções de transporte sustentáveis.

ESPAÑA



OS BENEFÍCIOS GLOBAIS DA INOVAÇÃO EM DADOS

MITO

Atualmente, apenas os Estados Unidos se beneficiam dos dados.

REALIDADE

Em todo o mundo, os dados já são usados para enfrentar desafios importantes. Por exemplo:

- + Ao rastrear mais de 1.000 pontos de dados por segundo, cientistas canadenses descobriram que crianças nascidas prematuramente com sinais vitais estáveis apresentavam maior risco de desenvolver febres – isso permitiu que os médicos tomassem medidas preventivas, salvando vidas.
- + No Brasil, um dos maiores produtores de soja está investindo em software e no poder da análise de dados para aumentar a eficiência das técnicas de controle de danos, reduzir custos e aumentar a produtividade.
- + Barcelona está utilizando dados para criar uma cidade mais inteligente, melhorar os serviços do governo e oferecer soluções de transporte sustentáveis.⁹⁰
- + Nos Emirados Árabes Unidos, novas ferramentas de dados estão sendo usadas para projetar o primeiro edifício de “energia positiva” do mundo, que produz mais energia do que consome.⁹¹
- + Agricultores na Índia estão usando dados de sementes, satélites, sensores e tratores para tomar decisões melhores sobre o que cultivar, como plantar, como monitorar o frescor dos alimentos da fazenda ao prato e como se adaptar às mudanças climáticas.⁹²

MITO

Os países em desenvolvimento ainda não estão prontos para obter benefícios com dados.

REALIDADE

A revolução dos dados e os benefícios que ela cria são um fenômeno global. Alguns dos benefícios mais importantes dos dados e algumas das maiores oportunidades se encontram no mundo em desenvolvimento, onde muitas vezes as tecnologias são obsoletas. De acordo com a IDC, os mercados emergentes no universo digital ultrapassarão os mercados maduros até 2017, crescendo de 36% para 62% da expansão do universo digital entre 2012 e 2020.⁹³ Da mesma forma, uma pesquisa com ONGs em países em desenvolvimento descobriu que 90% delas acreditam que a análise de dados seria a ferramenta mais importante para obter melhores insights a fim de ajudar seus beneficiários.⁹⁴

Há muitos exemplos da inovação em dados no mundo em desenvolvimento:

- + Recentemente, cientistas que estudam os padrões de infecção por malária usaram telefones celulares no Quênia para especificar os locais em que a transmissão da doença estava ocorrendo, ajudando a orientar os esforços de erradicação do governo.⁹⁵
- + Após o tsunami de 2004 no Sul da Ásia, pescadores da Indonésia receberam telefones celulares. Suas receitas subiram em 30%, pois era a primeira vez na vida tiveram dados sobre o preço de mercado dos peixes.⁹⁶
- + No Peru, locais históricos estão ameaçados pelo desenvolvimento. Utilizando tecnologias aéreas e softwares poderosos, que unem imagens, o Peru criou nuvens de pontos de dados tridimensionais detalhados para mapear, monitorar e proteger seus tesouros ameaçados.⁹⁷
- + Para preservar sua herança cultural, o Vietnã está usando scanners 3D e definiu uma meta de digitalizar 40.000 artefatos históricos ao longo dos próximos cinco anos.⁹⁸
- + Na Índia, quiosques de Internet permitem que mais de 4 milhões de agricultores acessem informações sobre preços de produtos, clima, entre outras, em idiomas locais. O país está usando análise avançada e tecnologias móveis para rastrear dados de fazendas individuais e, após a análise, oferecer suprimentos aos agricultores com base em suas necessidades, como fertilizantes e sementes.⁹⁹

PARA PRESERVAR SUA HERANÇA CULTURAL, o Vietnã está usando scanners 3D para digitalizar

40.000 artefatos

históricos dentro dos próximos cinco anos.

VIETNÃ



A FUNÇÃO DOS GOVERNOS NA REGULAMENTAÇÃO DOS DADOS

MITO

Os dados serão usados como uma ferramenta de exclusão, podendo reforçar as desvantagens enfrentadas por comunidades de baixa renda.

REALIDADE

A dissolução de qualquer preocupação quanto aos dados estarem sendo usados de maneiras injustas para algumas pessoas ou classes de pessoas é um desafio que deveria ser prioridade para os setores comerciais e os governos. Ao mesmo tempo, há oportunidades, que não devem ser ignoradas, de usar dados para combater a discriminação e conceder autonomia a grupos de pessoas. Na verdade, quando usados de maneira responsável, os dados podem ser uma ferramenta poderosa para ajudar a revelar as discriminações existentes, que prejudicam o acesso a empregos, financiamento, educação e oportunidades.¹⁰⁰

MITO

Governos não têm nenhuma função a desempenhar.

REALIDADE

O enorme crescimento das capacidades da computação remota e dos serviços de armazenamento, análise e software gerou uma série de novos problemas de políticas. A maioria dos tratados, leis e regulamentações não previa essas possibilidades no momento de sua elaboração. Essa ambiguidade pode ser problemática para os governos e o crescimento da economia de dados. Questões políticas antigas, como o equilíbrio entre privacidade e segurança, o fluxo comercial livre, pesquisas básicas e o desenvolvimento da força de trabalho, devem ser avaliadas novamente à luz do crescimento dessas tecnologias.

Há várias medidas concretas que as autoridades podem tomar para ajudar a alcançar todo o potencial da revolução dos dados e expandir novas ondas de produtividade, crescimento econômico e benefícios aos consumidores que estão esperando para acontecer. Elas podem investir no avanço de soluções de dados que possam lidar com alguns dos maiores desafios da sociedade — melhorar a forma como aprendemos, produzimos alimentos, vivemos, economizamos energia, viajamos e desenvolvemos nossas economias. E podem definir regras claras para cooperação internacional na imposição de leis. Podem também estimular as liberdades do mercado que afetam a capacidade das indústrias de tecnologia de investir, inovar, fazer negócios e crescer. Podem ajudar a estimular o livre fluxo de dados, sem fronteiras físicas. Podem ajudar a garantir uma força de trabalho talentosa e a superar a prevista escassez de cientistas de dados talentosos. E podem ajudar a alimentar a fogueira da inovação para acelerar os benefícios que os dados oferecem às empresas e aos consumidores. Com políticas pragmáticas, que aproveitem o talento e a tenacidade das pessoas, utilizem inovação e investimento e ampliem os recursos e capacidades, as autoridades podem ajudar a maximizar os benefícios que a economia com base em dados pode oferecer.

NO PERU, LOCAIS HISTÓRICOS ESTÃO AMEAÇADOS PELO DESENVOLVIMENTO.

Utilizando tecnologias aéreas e softwares poderosos, que unem imagens, o Peru criou nuvens

de pontos de dados tridimensionais detalhados

para mapear, monitorar e proteger seus tesouros ameaçados.

PERU



MITO

A localização de dados ajuda a proteger a privacidade e melhorar a segurança.

REALIDADE

Algumas pessoas podem acreditar que a localização dos dados — que exige que os dados sejam armazenados dentro das fronteiras de um determinado país — pode melhorar a privacidade e a segurança. No entanto, os benefícios atuais de tecnologia são possíveis graças à força global que é a Internet e alimentados por dados que cruzam o planeta entre diversos datacenters. O tráfego além fronteiras na Internet aumentou em mais de 50% desde 2005.¹⁰¹ O livre fluxo de dados entre as fronteiras permite que até mesmo as menores empresas e empreendedores se tornem “lojas de bairro” para todo o planeta, pois podem começar a vender e comprar produtos, serviços e ideias além das fronteiras de seus países. Frequentemente, porém, governos de todo o mundo consideram políticas que restringiriam o livre fluxo de dados ou exigiriam que os servidores de dados ficassem dentro de suas jurisdições como condição para atender ao mercado.¹⁰² Essas restrições prejudicam as enormes eficiências de escala, os benefícios econômicos que podem surgir com a inovação em dados e a capacidade de combinar conjuntos de dados diferentes, em locais diferentes, para fazer descobertas significativas a partir de uma crescente abundância de dados. E também podem prejudicar a segurança, impedindo o backup de dados valiosos em diferentes locais para protegê-los em caso de desastres naturais ou falhas técnicas. Para obter os benefícios que os dados podem oferecer, as leis dos países não precisam ser idênticas, mas compatíveis. Permitir o livre fluxo de dados entre fronteiras é um princípio básico para possibilitar os benefícios gerados pelos dados.

MITO

A única maneira de proteger dados é os governos exigirem que eles sejam protegidos.

REALIDADE

Regras de privacidade governamentais existentes podem ser combinadas com avanços inovadores e rigorosos de privacidade e boas práticas voluntárias do setor para garantir que os dados permaneçam seguros e que informações pessoais sejam protegidas. Em contraste, regulamentações governamentais que buscam tratar de privacidade e segurança exigindo que os dados sejam armazenados localmente podem inibir a inovação e limitar os tipos de benefícios sociais que a inovação em dados pode oferecer.



DISCURSO DIGITAL

COMPREENSÃO DA LINGUAGEM DOS DADOS

DADOS ABUNDANTES

Antes escassos, hoje há uma abundante quantidade de dados disponíveis, graças à crescente capacidade de coletar formatos de dados digitais significativos, utilizando meios totalmente novos, combinada com a redução dos custos do armazenamento de dados e novas maneiras de gerar valor a partir deles.

INTELIGÊNCIA ADAPTATIVA

Inteligência adaptativa é a inteligência computacional que não envolve apenas o processamento estatístico de dados, mas que também o combina com dados que contêm inteligência de domínio específica. Ao combinar modelos de comportamento inteligente com conhecimento avançado, os sistemas têm mais capacidade de aprender com exemplos e adaptarem-se a novas situações.

ALGORITMO

Um algoritmo é um procedimento ou série de instruções passo a passo de computação que utiliza matemática para analisar dados a fim de resolver problemas. Algoritmos são utilizados em praticamente todos os programas de software.

INTELIGÊNCIA AMBIENTAL

Inteligência ambiental é a visão de que a tecnologia se tornará cada vez mais invisível conforme crescentes quantidades de computação ubíqua e de baixo custo forem integradas ao mundo ao nosso redor. A inteligência ambiental estaria disponível sempre que precisássemos dela, pois, literalmente, tudo se tornaria conectado, inteligente e reativo.

ANÁLISE

Análise é o uso simultâneo de estatísticas e algoritmos com base em software para revelar insights, conexões e padrões significativos a partir dos dados.

DETECÇÃO DE ANOMALIAS

Detecção de anomalias é a identificação de dados, em um conjunto de dados, que não correspondem a um padrão esperado. Anomalias também são chamadas de desvios, exceções ou contaminações nos dados e podem fornecer informações úteis e importantes.

ANONIMIZAÇÃO

A anonimização de dados envolve a exclusão de todas as informações pessoalmente identificáveis que poderiam levar à identidade de uma pessoa.

DADOS INVÁLIDOS

Dados inválidos são dados ausentes ou incorretos. Pode ser algo simples, como um endereço errado, mas dados inválidos custam bilhões de dólares todos os anos às empresas da lista Fortune 1000.

BIG DATA

Big Data é um termo geral que muitas vezes se refere ao processo de aplicar análise computacional a grandes quantidades de dados para obter novos insights e melhorar a tomada de decisões. Frequentemente, o termo é usado para descrever conjuntos de dados cujo volume é tão grande, tão diversificado e que se movem com tanta velocidade que é difícil processá-los utilizando as ferramentas de processamento de dados tradicionais.



BRONTOBYTE

Um brontobyte é uma medida não oficial para uma quantidade extremamente grande de dados. Geralmente, um brontobyte é considerado equivalente a 1000 yottabytes e é representado por um 1 seguido por 27 zeros.

INTELIGÊNCIA DE NEGÓCIOS (BI)

Inteligência de negócios se refere ao conjunto de tecnologias e aplicações que transformam dados brutos em insights operacionais que podem melhorar o desempenho dos negócios e a tomada de decisões.

LIMPEZA DE DADOS

A limpeza de dados é o processo de detectar e corrigir ou remover dados inexatos, incompletos ou duplicados de um banco de dados.

ANÁLISE DE CLUSTERS

A análise de clusters é o processo de identificar objetos de dados semelhantes e agrupá-los para compreender melhor as diferenças e as semelhanças entre dados.

NUVEM

A nuvem é um termo geral que se refere a qualquer aplicativo, serviço ou dado hospedado em local remoto. Em geral, é disponibilizada por grandes grupos de servidores remotos conectados entre si para fornecer acesso de rede ubíquo e sob demanda a recursos de computação ou armazenamento.

COMPUTAÇÃO COGNITIVA

Computação cognitiva é o processo de combinar grandes quantidades de informações com técnicas de aprendizado de máquina, tecnologias de reconhecimento de padrões e, às vezes, processamento de linguagem natural para imitar a forma como o cérebro humano funciona. Muitas vezes, tais sistemas são capazes de aprender e interagir com pessoas combinando fontes de informação com contexto e insight.

DADOS GERADOS POR COMPUTADOR

Dados gerados por computador são dados gerados automaticamente por um computador, sem intervenção humana — como um arquivo de log de computador, dados de telemetria de satélites ou dados de um sensor em uma máquina industrial.

DARK DATA

Dark data (dados escuros) são dados desorganizados e inexplorados que estão sendo armazenados, mas não foram analisados nem processados, e que acredita-se que sejam negligenciados ou subutilizados de alguma forma.

DADOS

Dados são informações em formato bruto e desorganizado que podem ser manipuladas digitalmente para representar condições, objetos ou ideias. Tipos comuns de dados incluem números de vendas, resultados de pesquisas de marketing, leituras obtidas por sensores climáticos ou uma lista de cidades e suas populações. Atualmente, segundo estimativas, geramos 2,5 quintilhões de bytes de dados por dia.

AGREGAÇÃO DE DADOS

Agregação de dados é o ato de coletar dados de diversas fontes para possibilitar uma análise de nível mais avançado.

FERRAMENTAS DE AGREGAÇÃO DE DADOS

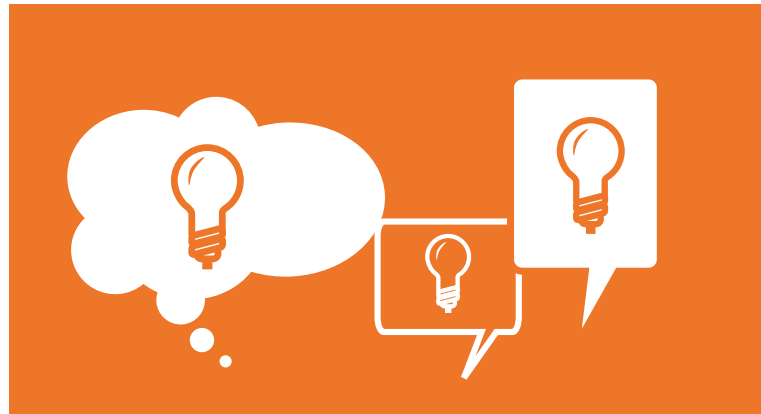
Ferramentas de agregação de dados transformam dados dispersos de diversas fontes em um único conjunto de dados.

ANÁLISE DE DADOS

Análise de dados é a aplicação de softwares como uma forma de transformar e modelar dados para obter informações, insights ou significados úteis dos dados. Normalmente, é usada para revelar padrões ocultos ou correlações desconhecidas e ajudar na tomada de decisões.

ANALISTA DE DADOS

Um analista de dados é uma pessoa responsável por preparar, limpar e processar dados.



ARQUITETURA E PROJETO DE DADOS

Geralmente, a arquitetura de dados é realizada na fase de planejamento de um novo sistema, para projetar e estruturar a forma em que os dados serão processados, armazenados, utilizados e acessados. Ao definir, no início, como dados específicos se relacionarão uns aos outros e serão colocados em movimento, é possível projetar a forma em que dados fluirão e controlar o fluxo de dados para garantir que eles sejam protegidos em todo o sistema.

BANCO DE DADOS

Um banco de dados é um conjunto grande e estruturado de dados digitais organizados, projetado para que os dados nele possam ser pesquisados, acessados e atualizados rapidamente.

DATA CENTER

Um datacenter é um local físico que abriga um grande número de servidores e depósitos de armazenamento de dados em rede usado normalmente para armazenamento e processamento remoto de grandes quantidades de dados acessíveis remotamente. Estima-se que haja meio milhão de datacenters no mundo, muitos dos quais formam a nuvem.

LIMPEZA DE DADOS

Limpeza de dados é o processo de revisar dados brutos para encontrar e excluir dados duplicados, corrigir erros, adicionar dados ausentes, remover dados corrompidos e gerar mais consistência.

TOMADA DE DECISÕES GUIADA POR DADOS

Empresas que usam a tomada de decisões guiada por dados coletam, processam e analisam dados para apoiar a tomada de decisões importantes. Uma pesquisa de Eric Brynjolfsson, economista da Sloan School of Management, do MIT, mostrou que empresas que usam tomada de decisões guiada por dados obtêm um aumento de 5% a 6% de produtividade.

DATA MINING

Data Mining (Mineração de dados) é o processo de usar algoritmos de computador de alta eficiência para encontrar padrões ou conhecimento em grandes conjuntos de dados.

QUALIDADE DOS DADOS

Qualidade dos dados é uma medida usada para definir o valor dos dados para o usuário. Refere-se à confiabilidade, à eficiência e ao valor dos dados para tomada de decisões, planejamento ou operações.

CIÊNCIA DE DADOS

A ciência de dados é uma disciplina que engloba estatísticas, visualização de dados, programação computacional, data mining (mineração de dados), aprendizado de máquina e engenharia de banco de dados para obter insights significativos que possam resolver problemas complexos.

CIENTISTA DE DADOS

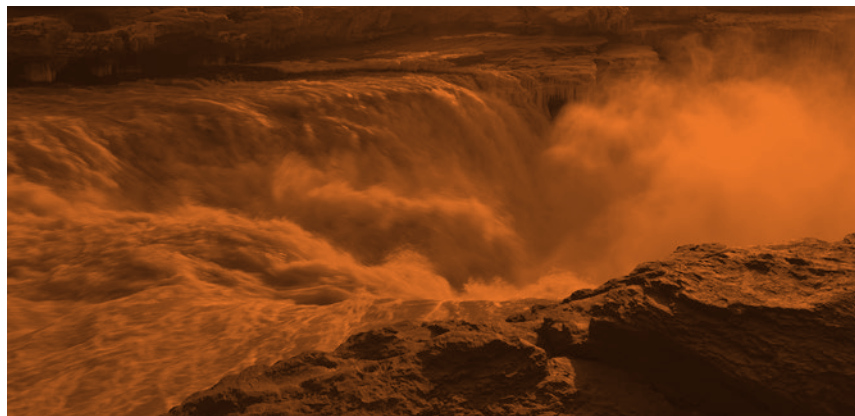
Um cientista de dados é uma pessoa capaz de combinar insights humanos, conhecimento matemático e ferramentas de tecnologia para compreender dados, por exemplo, desenvolvendo e implementando algoritmos de computador.

SEGURANÇA DE DADOS

A segurança de dados é a prática de proteger dados contra destruição, uso indevido ou acesso não autorizado. Medidas de segurança de dados adequadas podem ajudar a evitar violações de dados, garantir a integridade dos dados e proteger a privacidade. Frequentemente, envolve foco combinado em pessoas, processos e tecnologia.

CONJUNTO DE DADOS

Um conjunto de dados é uma coleção de conjuntos de informações relacionados, normalmente elementos separados, em um formato de tabela que possa ser manipulado como uma unidade.



FONTE DE DADOS

Uma fonte de dados é o principal local de origem dos dados, por exemplo, um banco de dados, uma planilha ou um fluxo de dados.

VISUALIZAÇÃO DE DADOS

Visualização de dados envolve a criação de uma representação visual dos dados para extrair significado ou comunicar informações com mais eficácia.

VIRTUALIZAÇÃO DE DADOS

Virtualização de dados é o processo de recuperar e manipular diferentes fontes de dados sem a necessidade de conhecer os detalhes técnicos sobre onde elas estão localizadas ou como estão formatadas.

“DESIDENTIFICAÇÃO”

A “desidentificação de” dados é o processo de remover informações que associem uma pessoa a uma informação específica.

MUDANÇAS INOVADORAS

Mudanças inovadoras são mudanças grandes e essenciais na sociedade e nos negócios, muitas vezes possibilitadas por novas tecnologias que criam um contexto totalmente novo para a forma como trabalhamos, vivemos, agimos e criamos valor. Frequentemente, a inovação em dados é descrita como uma tecnologia que possibilita mudanças inovadoras.

EXABYTE

Um exabyte é uma unidade de armazenamento de dados enorme — um 1 seguido por 18 zeros. Hoje em dia, é criado um exabyte por dia de novas informações.

HADOOP

Hadoop é uma plataforma de software de código aberto criada para permitir o processamento e o armazenamento de grandes quantidades de dados entre sistemas de arquivos distribuídos.

INTERNET DAS COISAS

A Internet das Coisas descreve um mundo em que dispositivos comuns se tornam muito mais inteligentes e conectados à Internet para que a revolução inteligente saia da palma de nossas mãos e ganhe o mundo à nossa volta. Como tudo o que pode ser conectado será mesmo conectado, algumas pessoas usam, apropriadamente, o termo “Internet de Tudo”. Estima-se que apenas 1% das coisas que podem ser conectadas no mundo estão conectadas atualmente. A estimativa para 2020 é de que 50 bilhões de dispositivos estarão conectados à Internet.

SISTEMA HERDADO

Um sistema herdado é um computador, aplicativo ou tecnologia que seja antiquado ou obsoleto, mas que continue sendo usado porque executa uma função necessária da maneira adequada.

APRENDIZADO DE MÁQUINA

Aprendizado de máquina é o uso de algoritmos para permitir que um computador analise dados com o objetivo de “aprender” as ações a serem tomadas quando um padrão ou evento específico ocorrer.

METADADOS

Metadados são dados sobre a dados. Podem incluir informações básicas resumidas sobre os dados, como o autor dos dados, a data de criação, o tamanho do arquivo e a data da última modificação.

DETECÇÃO DE DESVIOS

Um desvio é um dado que difere significativamente da média geral de um conjunto de dados. É numericamente distante do restante dos dados e, portanto, o desvio indica que algo está acontecendo que geralmente requer análise adicional. (Consulte também Detecção de anomalias).



RECONHECIMENTO DE PADRÕES

Reconhecimento de padrões é o processo de procurar e identificar padrões em dados. Pode ser algo simples, como identificar um conjunto de sequências repetidas em uma sequência de DNA, pode ser encontrar um padrão na forma em que dois conjuntos de dados interagem para descobrir se há um padrão que conecta um evento a outro ou, com a ajuda do aprendizado de máquina, pode ser a busca de padrões mais complexos, como encontrar caracteres numéricos em uma imagem.

PETABYTE

Um petabyte é uma medida de capacidade de armazenamento enorme, representada por um 1 seguido por 15 zeros, ou um milhão de gigabytes. Um petabyte é cerca de quatro vezes a quantidade de dados contida na Biblioteca do Congresso dos Estados Unidos.

ANÁLISE PREDITIVA

A análise preditiva envolve o uso de algoritmos de software em um ou mais conjuntos de dados para prever tendências ou eventos futuros. Muitas vezes, quando dados do presente podem ser comparados com os do passado, eles podem ser usados para prever o futuro.

MODELAGEM PREDITIVA

Modelagem preditiva é o processo de desenvolver um modelo que possa prever uma tendência, comportamento futuro ou resultado — frequentemente comparando eventos do presente com eventos do passado.

DADOS EM TEMPO REAL

Dados em tempo real são dados utilizados no momento em que são criados. Normalmente, são criados, processados, armazenados e analisados em alguns milissegundos. Dados em tempo real podem incluir tudo, de cotações de bolsas de valores à velocidade de uma roda usada no freio ABS de um carro.

MECANISMO DE RECOMENDAÇÃO

Um mecanismo de recomendação é um algoritmo de computador que faz recomendações, sugestões ou que pode personalizar algo para você com base em uma série de padrões de dados, frequentemente obtidos por meio de técnicas de aprendizado de máquina.

ANÁLISE DE REGRESSÃO

Análise de regressão é um processo estatístico de utilização de dados para estimar o relacionamento entre duas ou mais variáveis.

ANÁLISE DE RISCO

Análise de risco é o uso de ferramentas analíticas de dados de software para identificar o provável risco envolvido em um projeto, ação ou decisão. Novas ferramentas de dados podem ajudar a identificar possíveis riscos com antecedência, modelar uma matriz de situações para ajudar a reduzir o risco enfrentado por organizações e monitorar sistemas para identificar problemas se as coisas começarem a sair do rumo.

ANÁLISE DE CAUSA RAIZ

Análise de causa raiz é um método de resolução de problemas com foco na análise da relação entre causa e efeito para identificar a causa raiz de uma falha ou problema. A causa é uma causa raiz se, quando removida de uma sequência de eventos, evitar que o evento indesejado se repita.

DADOS SEMIESTRUTURADOS

Dados semiestruturados não são estruturados por um modelo de dados formal, como o utilizado em bancos de dados, mas fornecem outros meios de descrever os dados e as hierarquias. Frequentemente, dados semiestruturados usam tags ou outros marcadores de dados, no que às vezes é chamado de estrutura autodescritível.



SMALL DATA

Small Data trata-se de aproveitar até mesmo pequenas quantidades de dados, como dados contidos em uma pesquisa de consumidores, para obter resultados práticos. Geralmente, o termo se refere a dados de tamanho pequeno o suficiente para uma pessoa compreender e analisar.

DADOS ESTRUTURADOS

Dados estruturados são altamente organizados, geralmente em linhas e colunas, sendo mais fáceis de pesquisar e manipular.

TERABYTE

Um terabyte é uma medida de dados representada por um 1 seguido por 12 zeros. Discos rígidos com capacidade na casa dos terabytes já são encontrados em computadores pessoais e de trabalho ou acessados na nuvem. Um terabyte pode, por exemplo, armazenar cerca de 300 horas de vídeo em alta definição.

ANÁLISE DE TEXTO

Análise de texto é a utilização de técnicas de estatística, linguística e aprendizado de máquina em dados com base em textos para extrair significado, conceitos e insights. A análise de texto geralmente é realizada em textos em linguagem natural, como o texto de documentos, transcrições, posts na Internet, comentários ou formulários. Pode ser útil para o resumo, descoberta ou classificação de conteúdo.

DADOS TRANSACIONAIS

Dados transacionais são dados derivados de eventos específicos, como dados de movimentações financeiras, faturas, pagamentos, e expedição. Geralmente, têm um carimbo de data e hora e apoiam as operações diárias de uma organização.

DADOS NÃO ESTRUTURADOS

Dados não estruturados não possuem uma estrutura predefinida — por exemplo, anotações de uma reunião. Segundo algumas estimativas, informações não estruturadas podem responder a de 70 ou 80% de todos os dados em uma organização.

VARIEDADE

Variedade, um dos quatro “Vs” que definem a inovação em dados, representa os diversos tipos de dados, muitas vezes de diferentes fontes, que são combinados e analisados para gerar insights. Atualmente, a variedade de tipos de dados que são processados em aplicativos pode incluir bancos de dados textuais, dados transacionais, dados de streaming, imagens, áudio e vídeo.

VELOCIDADE

A velocidade, um dos quatro “Vs” que definem a inovação em dados, é a velocidade na qual os dados são criados, armazenados, analisados e visualizados. Por exemplo, grandes depósitos de dados podem receber bilhões de linhas de novas informações diariamente. Dados sensíveis ao tempo devem ser usados assim que transmitidos para maximizar seu valor.

VERACIDADE

Veracidade, um dos quatro “Vs” que definem a inovação em dados, é usada para indicar a exatidão, a confiabilidade e a precisão dos dados.

VOLUME

Volume, um dos quatro “Vs” que definem a inovação em dados, é a quantidade de dados processados — variando de megabytes a brontobytes.

YOTTABYTES

Um yottabyte é uma medida muito grande de armazenamento de dados representada por um 1 seguido por 24 zeros. Um yottabyte é equivalente à quantidade de dados armazenados em 250 trilhões de DVDs.

ZETTABYTES

Um zettabyte é uma medida de armazenamento representada por um 1 seguido por 21 zeros. Estima-se que em 2013 a World Wide Web alcançou a marca de 4 zettabytes. A projeção para 2016 é que mais de um zettabyte de dados seja transmitido por dia em nossas redes.

Notas de fim

- ¹ Economist Intelligence Unit. *The Deciding Factor: Big Data & Decision Making*. Cap Gemini, 2012. Web. Point Of View. <https://www.capgemini.com/resources/the-deciding-factor-big-data-decision-making>
- ² Um aumento de produtividade de 1 por cento pode parecer pouco, mas segundo Jeff Immelt, CEO da GE, “diga a alguém do setor petrolífero que você pode usar um software para economizar para ele 1% em alguma coisa e essa pessoa será sua amiga para o resto da vida.” Evans, Peter C., and Marco Annunziata. *Pushing the Boundaries of Minds and Machines*. GE, 2012. Web. <http://files.gereports.com/wp-content/uploads/2012/11/ge-industrial-internet-vision-paper.pdf>
- ³ A estimativa para 2020 é de que 50 bilhões de dispositivos estarão conectados à Internet, permitindo a Internet de Tudo. Evans, Dave. *The Internet Of Things How The Next Evolution Of The Internet Is Changing Everything*. Cisco Internet Business Solutions Group (IBSG), 2011. Web. White Paper. http://www.cisco.com/web/about/ac79/docs/innov/IoT_IBSG_0411FINAL.pdf
- ⁴ “Bringing big data to the enterprise.” IBM. <http://www-01.ibm.com/software/data/bigdata/what-is-big-data.html>
- ⁵ Golub, Ben. “Enough Data to Fill a Stack of DVDs to the Moon (and Back).” *ComputerWorld* 2011. Web. http://blogs.computerworld.com/18351/a_stack_of_dvds_to_the_moon_and_back
- ⁶ Corry, Will. “BIG Data / The Volume Of Business Data Worldwide, Across All Companies, Doubles Every 1.2 Years, According To Estimates.” *The Marketing Blog* 2012. Web. <http://www.themarketingblog.co.uk/2012/10/big-data-the-volume-of-business-data-worldwide-across-all-companies-doubles-every-1-2-years-according-to-estimates/>
- ⁷ “Digital Imaging in the Cloud.” *There Magazine* 2012: 16. Web. http://www.agfahealthcare.com/he/global/en/binaries/THERE_12_tcm541-95647.pdf
- ⁸ O 787 utiliza sensores de dados para reduzir o consumo de combustível, monitorar sistemas e até possui acelerômetros no nariz da aeronave para compensar turbulência. Se os sensores registrarem uma queda repentina, eles comunicam imediatamente aos flaps nas asas que eles devem fazer um ajuste (em uma questão de nanossegundos) e, fazendo isso, uma queda de 3 metros em um avião antigo pode ser reduzida para apenas 1 metro no 787, tornando o voo muito mais confortável. Humphries, Matthew. “The Boeing 787 produces over 500GB of Data during Every Flight.” *Geek.com* 2013. Web. Gosling, Kevin. “E-Enabled Capabilities of the 787 Dreamliner.” *Aero Quarterly* 2009: 22-24. http://www.boeing.com/commercial/aeromagazine/articles/qtr_01_09/pdfs/AERO_Q109_article05.pdf
- ⁹ A GE, fabricante de motores a jato para aviões, afirma que os dados dos motores permitem descobrir coisas como possíveis defeitos 2.000 vezes mais rápido do que antes. Hardy, Quentin. “What Cars Did for Today’s World, Data May Do for Tomorrow’s.” *New York Times* 2014. B7. Print. http://bits.blogs.nytimes.com/2014/08/10/g-e-creates-a-data-lake-for-new-industrial-ecosystem/?_php=true&_type=blogs&_php=true&_type=blogs&module=BlogPost-Title&version=Blog%20Main&contentCollection=Big%20Data&action=Click&pgtype=Blogs®ion=Body&r=1&
- ¹⁰ Henschen, Doug. “Bad Winter Weather Meets Big Data Prediction.” *InformationWeek* 2014. Web. <http://www.informationweek.com/big-data/software-platforms/bad-winter-weather-meets-big-data-prediction/d/d-id/1114192>
- ¹¹ Groenfeldt, Tom. “At NYSE, The Data Deluge Overwhelms Traditional Databases.” *Forbes* 2013. Web. <http://www.forbes.com/sites/tomgroenfeldt/2013/02/14/at-nyse-the-data-deluge-overwhelms-traditional-databases>
“NYSE Euronext Adapting To Market Changes with Near-Real-Time Insight into Information.” IBM Corporation, 2013. Print. NYSE Euronext. <http://www.ibmbigdatahub.com/sites/default/files/document/NYSE-Euronext-IMC14787USEN.PDF>
- ¹² Mayyasi, Alex. “Why UPS Trucks Don’t Turn Left.” *Priceonomics.com* 2014. Web. <http://priceonomics.com/why-ups-trucks-dont-turn-left/>
- ¹³ *Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update 2014–2019 White Paper*. Cisco Visual Networking Index, 2015. Print. http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/white_paper_c11-520862.html
- ¹⁴ “LSST General Public FAQs.” LSST.org 2015. <http://www.lsst.org/lsst/faq>
- ¹⁵ Quazi, Ed. “Next Generation Sequencing and the Big Data Bottleneck.” *Arkivum* 2014. Web. <http://arkivum.com/next-generation-sequencing-and-the-big-data-bottleneck/>
- ¹⁶ Valerio, Pablo. “Internet Of Things: 50 Billion Is Only The Beginning.” *EE Times* 2014. Web. http://www.eetimes.com/document.asp?doc_id=1321229

- ¹⁷ Na década de 1980, havia uma regra geral de que era necessário um administrador de dados para 1 GB de armazenamento. Naquela época, um GB de disco custava cerca de um milhão de dólares, portanto, fazia sentido ter alguém para otimizar e monitorar a utilização de espaço em disco. Gray, Jim, and Prashant Shenoy. "Rules Of Thumb in Data Engineering." Redmond, WA: Microsoft Research Advanced Technology Division, 2009. Print. Technical Report. http://research.microsoft.com/pubs/68636/ms_tr_99_100_rules_of_thumb_in_data_engineering.pdf
- ¹⁸ Wohlsen, Marcus. "Dropbox Slashes Its Price As The Cost Of A Gigabyte Nears Zero." *Wired* 2014. Web. <http://www.wired.com/2014/08/dropbox-plan-to-stay-relevant/>
- ¹⁹ De mais de US\$ 200.000 por gigabyte em 1980 (ou até um milhão de dólares) para US\$ 0,02 por gigabyte em 2013 Meeker, Mary. "Kleiner Perkins Internet Trends 2014." 2014. Presentation. <http://cryptome.org/2014/05/internet-trends-2014.pdf>
- ²⁰ Com base na eficiência média de consumo de combustível de carros de passageiros em 1980 (24,3 milhas por galão), permitir que uma pessoa compre 10 milhões a mais de capacidade pelo mesmo preço significaria 10 milhões de galões de gasolina, o que seria suficiente para um deslocamento por 243 milhões de milhas. Se a circunferência da Terra tem 24.901 milhas, então a pessoa poderia dar 9.758 voltas no planeta, ou quase 10.000 vezes. Departamento de Transporte dos EUA. Table 4-23: Average Fuel Efficiency Of U.S. Light Duty Vehicles. Washington, DC: Bureau of Transportation Statistics, 2013. Print. http://www.rita.dot.gov/bts/sites/rita.dot.gov/bts/files/publications/national_transportation_statistics/html/table_04_23.html
- ²¹ Savitz, Eric. "Big Data: The Hidden Opportunity." *Forbes* 2012. Web. <http://www.forbes.com/sites/ciocentral/2012/05/01/big-data-the-hidden-opportunity/>
- ²² A capacidade tecnológica per capita do mundo de armazenamento de informações praticamente dobrou a cada 40 meses desde a década de 1980, de acordo com pesquisa de Martin Hilbert and Priscila López. Hilbert, M., and P. Lopez. "The World's Technological Capacity to Store, Communicate, and Compute Information." *Science* 332.6025 (2011): 60-65. Web. <http://www.sciencemag.org/content/332/6025/60>
- ²³ Os custos de armazenamento caíram de US\$ 569 por gigabyte de armazenamento em 1992 para US\$ 0,02 por gigabyte em 2013 — uma taxa de queda de aproximadamente 38% por ano. Meeker, Mary. "Kleiner Perkins Internet Trends 2014." 2014. Presentation. <http://cryptome.org/2014/05/internet-trends-2014.pdf>
- ²⁴ Lohr, Steve. "For Big-Data Scientists, 'Janitor Work' Is Key Hurdle To Insights." *New York Times*. 2014: B4. Print. http://www.nytimes.com/2014/08/18/technology/for-big-data-scientists-hurdle-to-insights-is-janitor-work.html?_r=0
- ²⁵ "Data, data everywhere," *The Economist*, 25 de fevereiro de 2010. <http://www.economist.com/node/15557443>
- ²⁶ Crovitz, L. Gordon. "Why 'Big Data' Is A Big Deal." *Wall Street Journal*. 2013. Print. <http://online.wsj.com/news/articles/SB10001424127887324077704578364632408717740>
- ²⁷ Relatos de secas em Angola em 2006 acionaram um alerta sobre possíveis epidemias de cólera no país, pois eventos semelhantes anteriores mostraram que epidemias de cólera eram mais prováveis nos anos após a ocorrência de secas. Os alertas do sistema estiveram corretos entre 70% e 90% das vezes. Simonite, Tom. "Software Predicts Tomorrow's News by Analyzing Today's And Yesterday's." *MIT Technology Review* 2013. Print. <http://www.technologyreview.com/news/510191/software-predicts-tomorrows-news-by-analyzing-todays-and-yesterdays/>
- ²⁸ "Ten Big Data Case Studies in a Nutshell." *TechTarget*, 2013. Print. Essential Guide. <http://searchcio.techtarget.com/opinion/Ten-big-data-case-studies-in-a-nutshell>
- ²⁹ "Big Data to Predict Offshore Accidents, Tsunamis and Other Natural Disasters." *Predictive Analytics Today*. 2013 Web. <http://www.predictiveanalyticstoday.com/big-data-predict-shore-accidents-tsunamis-natural-disasters/>
- ³⁰ "New Developments in Big Data Visualization." *USTelecom Media* 2014. Web. <http://www.ustelecom.org/blog/new-developments-big-data-visualization#sthash.HefD5H52.dpuf>
- ³¹ Quinn, Tom. "New and Unexpected Uses for Scoring Technology." *Credit Score Blog* 2011. Web. <http://blog.credit.com/2011/06/new-and-unexpected-uses-for-scoring-technology/>
- ³² Crovitz, L. Gordon. "Why 'Big Data' Is A Big Deal." *Wall Street Journal* 2013: Print. <http://online.wsj.com/news/articles/SB10001424127887324077704578364632408717740>
- ³³ Sobrecarga de dados: Atualmente, os clínicos experientes precisam de quase 2 milhões de itens de informação para praticar medicina e os médicos assinam em média sete periódicos do setor, que representam mais de 2.500 novos artigos por ano, tornando quase impossível permanecer atualizado com as informações mais recentes sobre diagnósticos, prognósticos, terapias

- e questões relacionadas à saúde. “Clinical Decisions Support Systems: The Time Has Come...” Frost & Sullivan, 2009. Print. Market Insight. <http://www.frost.com/prod/servlet/cio/181298788>
- ³⁴ “Clinical Decisions Support Systems: The Time Has Come...” Frost & Sullivan, 2009. Print. Market Insight. <http://www.frost.com/prod/servlet/cio/181298788>
- ³⁵ “Kinect Sign Language Translator Expands Communication Possibilities.” *Microsoft Research* 2013. Web. <http://research.microsoft.com/en-us/collaboration/stories/kinect-sign-language-translator.aspx>
- ³⁶ “Brats 2012 - Multimodal Brain Tumor Segmentation Challenge.” *CodaLab*, 2012. Print. <https://www.codalab.org/competitions/191>
- ³⁷ Grady, Denise. “3-D Mammography Test Appears To Improve Breast Cancer Detection Rate.” *New York Times* 2014: p. A1 Print. http://www.nytimes.com/2014/06/25/health/breast-cancer-3d-mammography-test-x-ray.html?emc=edit_th_20140625&nl=todaysheadlines&nid=435891&r=0
- ³⁸ “The game-changing technology that’s transforming manufacturing.” *Manufacturing Weekly*, January 31, 2014. <http://web.archive.org/web/20140131233544/http://www.manufacturingweekly.com/supercomputers/>
- ³⁹ “The Return on the Data Asset in the Era of Big Data: Capturing the \$1.6 Trillion Data Dividend.” *Cloud Platform News Bytes Blog* 2015. Web. <http://blogs.technet.com/b/stbnewsbytes/archive/2014/04/15/the-return-on-the-data-asset-in-the-era-of-big-data-capturing-the-1-6-trillion-data-dividend.aspx>
- ⁴⁰ Economist Intelligence Unit. *The Deciding Factor: Big Data & Decision Making*. Cap Gemini, 2012. Web. Point Of View. <http://bigdata.pervasive.com/Solutions/Telecom-Analytics.aspx>
- ⁴¹ Um aumento de produtividade de 1 por cento pode parecer pouco, mas segundo Jeff Immelt, CEO da GE, “diga a alguém do setor petrolífero que você pode usar um software para economizar para ele 1% em alguma coisa e essa pessoa será sua amiga para o resto da vida.” Evans, Peter C., and Marco Annunziata. *Pushing the Boundaries of Minds and Machines*. GE, 2012. Web. <http://files.gereports.com/wp-content/uploads/2012/11/ge-industrial-internet-vision-paper.pdf>
- ⁴² BSA/IPSOS Global Data Analytics Poll, November 2014, www.bsa.org/datasurvey
- ⁴³ Gartner, “Gartner Says Big Data Creates Big Jobs: 4.4 Million IT Jobs Globally To Support Big Data By 2015.” 2012. Print. <http://www.gartner.com/newsroom/id/2207915>
- ⁴⁴ BSA/IPSOS Global Data Analytics Poll, November 2014, www.bsa.org/datasurvey
- ⁴⁵ BSA/IPSOS Global Data Analytics Poll, November 2014, www.bsa.org/datasurvey
- ⁴⁶ According to ESG research, data managed per hospital is expected to increase from 168 terabytes in 2010 to 6 terabytes by 2015. “Digital Imaging in the Cloud.” *There Magazine* 2012: 16. Print. http://www.agfahealthcare.com/he/global/en/binaries/THERE_12_tcm541-95647.pdf
- ⁴⁷ Manyika, James et al. *Big Data: The Next Frontier for Innovation, Competition, and Productivity*. McKinsey Global Institute, 2011. Print. http://www.mckinsey.com/insights/business_technology/big_data_the_next_frontier_for_innovation
- ⁴⁸ Pesquisadores inseriram em um algoritmo de aprendizado de máquina dados de 133.000 pacientes. O modelo ainda necessita de mais trabalho para reduzir falsos positivos. Rutkin, Aviva. “Machine Predicts Heart Attacks 4 Hours Before Doctors - New Scientist.” *New Scientist*. 2014. Web. <http://www.newscientist.com/article/mg22329814.400-machine-predicts-heart-attacks-4-hours-before-doctors.html>
- ⁴⁹ O moderno modelo híbrido Fusion da Ford gera até 25 GB de dados por hora. Hemsoth, Nicole. “How Ford Is Putting Hadoop Pedal To The Metal.” *Datanami*. 2013. Web. http://www.datanami.com/2013/03/16/how_ford_is_putting_hadoop_pedal_to_the_metal/
- O Volt da Chevrolet contém mais de 10 milhões de linhas de código de software e desenvolvedor de software é uma das profissões técnicas que mais crescem no Sudeste do Michigan, EUA, uma região conhecida há muito tempo por suas façanhas no setor de manufatura. Trop, Jaclyn. “Detroit, Embracing New Auto Technologies, Seeks App Builders.” *New York Times*. June 30, 2013. <http://www.nytimes.com/2013/07/01/technology/detroit-embracing-new-auto-technologies-seeks-app-builders.html>
- ⁵⁰ Miller, Claire Cain. “If Robots Drove, How Much Safer Would Roads Be?” *New York Times* 2014: A3. Print. <http://www.nytimes.com/2014/06/10/upshot/if-robots-drove-how-much-safer-would-roads-be.html?ref=technology&r=0>
- ⁵¹ O 787 utiliza sensores de dados para reduzir o consumo de combustível, monitorar sistemas e até possui acelerômetros no nariz da aeronave para compensar turbulência. Se os sensores registrarem uma queda

- repentina, eles comunicam imediatamente aos flaps nas asas que eles devem fazer um ajuste (em uma questão de nanossegundos) e, fazendo isso, uma queda de 3 metros em um avião antigo pode ser reduzida para apenas 1 metro no 787, tornando o voo muito mais confortável. Gosling, Kevin. "E-Enabled Capabilities of the 787 Dreamliner." *Aero Quarterly* 2009: 22-24. http://www.boeing.com/commercial/aeromagazine/articles/qtr_01_09/pdfs/AERO_Q109_article05.pdf
- ⁵² A GE, fabricante de motores a jato para aviões, afirma que os dados dos motores permitem descobrir coisas como possíveis defeitos 2.000 vezes mais rápido do que antes. Hardy, Quentin. "What Cars Did for Today's World, Data May Do for Tomorrow's?" *New York Times* 2014: B7. Print. http://bits.blogs.nytimes.com/2014/08/10/g-e-creates-a-data-lake-for-new-industrial-ecosystem/?_php=true&_type=blogs&_php=true&_type=blogs&module=BlogPost-Title&version=Blog%20Main&contentCollection=Big%20Data&action=Click&pgtype=Blogs®ion=Body&_r=1&
- ⁵³ Long, Jessica, and William Brindley. *The Role of Big Data and Analytics in the Developing World*. Accenture, 2013. Print. Accenture Development Partnerships Insights into the Role of Technology in Addressing Development Challenges. https://www.accenture.com/us-en/~/_media/Accenture/Conversion-Assets/DotCom/Documents/Global/PDF/Strategy_5/Accenture-ADP-Role-Big-Data-And-Analytics-Developing-World.pdf
- ⁵⁴ O setor de manufatura armazenou quase 2 exabytes de novos dados apenas em 2010. Manyika, James et al. *Big Data: The Next Frontier for Innovation, Competition, and Productivity*. McKinsey Global Institute, 2011. Print. http://www.mckinsey.com/~/_media/McKinsey/dotcom/Insights20and%20pubs/MGI/Research/Technology%20and%20Innovation/Big%20Data/MGI_big_data_full_report.ashx
- ⁵⁵ Manyika, James et al. *Big Data: The Next Frontier for Innovation, Competition, and Productivity*. McKinsey Global Institute, 2011. Print. http://www.mckinsey.com/~/_media/McKinsey/dotcom/Insights20and%20pubs/MGI/Research/Technology%20and%20Innovation/Big%20Data/MGI_big_data_full_report.ashx
- ⁵⁶ "Data Smart' Strategies for Customers Are Yielding 'Early But Impressive Returns.'" *Microsoft Research the Fire Hose* 2014. Web. <http://blogs.microsoft.com/firehose/2014/05/22/data-smart-strategies-for-customers-are-yielding-early-but-impressive-returns/>
- ⁵⁷ Somers, Dan. "Manufacturing 4.0 – From Industrialization to Data-Driven Product Lifecycle." *Citizenekkk*. 2013. Web. <http://citizenekkk.com/2013/11/05/manufacturing-4-0-industrialisation-data-driven-product-lifecycle/>
- ⁵⁸ Relatórios da McKinsey mostram que ao utilizar essas técnicas de projeto impulsionadas por dados, a Toyota conseguiu eliminar 80% dos defeitos antes de construir o primeiro protótipo físico. Manyika, James et al. *Big Data: The Next Frontier for Innovation, Competition, and Productivity*. McKinsey Global Institute, 2011. Print. http://www.mckinsey.com/~/_media/McKinsey/dotcom/Insights20and%20pubs/MGI/Research/Technology%20and%20Innovation/Big%20Data/MGI_big_data_full_report.ashx
- ⁵⁹ Findings of the New Intelligent Enterprise Study. IBM 2010 *New Intelligent Enterprise Global Executive Study*. 2010. Print.
- ⁶⁰ Geron, Tomio. "Cows in the Cloud: The Hot Startup Moving Farmers into the Cloud." *Forbes* 2012. Web; Helmer, Jodi. "Get Ready For Robot Farmers." Yahoo 2014. Web. <https://www.yahoo.com/tech/get-ready-for-robot-farmers-100613764059.html>
- ⁶¹ De-Identification of Personally Identifiable Information, National Institute of Science and Technology, DRAFT NISTIR 8053 (April 2015).
- ⁶² Cavoukian, Ph.D., Ann, and El Emam, Ph.D., Khaled, Dispelling the Myths Surrounding De-Identification: Anonymization Remains a Strong Tool for Protecting Privacy, Comissário de Informações e Privacidade de Ontário, (Junho de 2011); Cavoukian, Ph.D., Ann, and Daniel Castro Castro. *Big Data And Innovation, Setting The Record Straight: De-Identification Does Work*. ITIF, 2014. Print. <http://www2.itif.org/2014-big-data-deidentification.pdf>
- ⁶³ See for example, Microsoft's add on protecting privacy as their priority https://www.youtube.com/watch?feature=player_embedded&v=bt51MWII1oY
- ⁶⁴ Apple, Solicitações de Informações do Governo, observando que a empresa incorporou criptografia de ponta ao sistema operacional de seu iPhone para que "seus dados pessoais, como fotos, mensagens (incluindo anexos), e-mails, contatos, histórico de chamadas, conteúdo do iTunes, notas e lembretes, são protegidos pela sua senha", em. www.apple.com/privacy/government-information-requests/
- ⁶⁵ Hachman, Mark, "Microsoft's updated privacy policy makes it clear it's not selling ads against your words," *PCWorld*, June 11, 2014, <http://www.pcworld.com/article/2362130/microsofts-updated-privacy-policy-makes-it-clear-its-not-selling-ads-against-your-words.html>,

- Timberg, Craig. "Newest Androids Will Join iPhone In Offering Default Encryption, Blocking Police." *Washington Post* 2014: Print. <http://www.washingtonpost.com/blogs/the-switch/wp/2014/09/18/newest-androids-will-join-iphones-in-offering-default-encryption-blocking-police/>
- ⁶⁶ A Acxiom, corretora de dados, abre site de dados para consumidores, oferecendo opção de cancelamento <http://cir.ca/news/acxiom-gives-consumers-data-peek>
- ⁶⁷ BSA/IPSOS Global Data Analytics Poll, November 2014, www.bsa.org/datasurvey
- ⁶⁸ McKinsey Global Institute. *Internet Matters: The Net's Sweeping Impact On Growth, Jobs, And Prosperity*. McKinsey & Co., 2011. Print.
- ⁶⁹ Manyika, James et al. *Big Data: The Next Frontier for Innovation, Competition, and Productivity*. McKinsey Global Institute, 2011. Print. http://www.mckinsey.com/~media/McKinsey/dotcom/Insights20and%20pubs/MGI/Research/Technology%20and%20Innovation/Big%20Data/MGI_big_data_full_report.ashx
- ⁷⁰ According to Salaries of Data Scientists, an April 2014 study from Burtch Works.
- ⁷¹ Bagley, Rebecca. "How The Cloud And Big Data Are Changing Small Business." *Forbes* 2014. Web. <http://www.forbes.com/sites/rebeccabagley/2014/07/15/how-the-cloud-and-big-data-are-changing-small-business/>
- ⁷² BSA/IPSOS Global Data Analytics Poll, November 2014, www.bsa.org/datasurvey
- ⁷³ Economist Intelligence Unit. *The Deciding Factor: Big Data & Decision Making*. Cap Gemini, 2012. Web. Point Of View. <http://bigdata.pervasive.com/Solutions/Telecom-Analytics.aspx>
- ⁷⁴ Economist Intelligence Unit. *The Deciding Factor: Big Data & Decision Making*. Cap Gemini, 2012. Web. Point Of View. <http://bigdata.pervasive.com/Solutions/Telecom-Analytics.aspx>
- ⁷⁵ Gerbis, Nicholas. "10 Correlations That Are Not Causations." *HowStuffWorks*. 2015. Web. <http://science.howstuffworks.com/innovation/science-questions/10-correlations-that-are-not-causations.htm>
- ⁷⁶ Vesset, Dan, Henry D. Morris, and John F. Gantz. *Capturing the \$1.6 Trillion Data Dividend*. IDC, 2014. Print. IDC White Paper.
- ⁷⁷ Westerman, George, Didier Bonnet, and Andrew McAfee, "The Advantages of Digital Maturity." November 2012, MIT Sloan.
- ⁷⁸ "Meeting the Big Data Challenge: Don't Be Objective." *Forbes* 2013. Web. <http://www.forbes.com/sites/darden/2013/02/01/meeting-the-big-data-challenge-dont-be-objective/>
- ⁷⁹ IDG Enterprise's 2014 Big Data research. IDG. *CEOs Call for Big Data and IT Continues To Lead Investment Decisions*. 2014. Print. <http://www.idgenterprise.com/press/ceos-call-for-big-data-and-it-continues-to-lead-investment-decisions>
- ⁸⁰ Miller, Claire Cain. "If Robots Drove, How Much Safer Would Roads Be?" *New York Times* 2014: A3. Print. <http://www.nytimes.com/2014/06/10/upshot/if-robots-drove-how-much-safer-would-roads-be.html?ref=technology&r=0>
- ⁸¹ Clemens, Samuel. "7 Facts about Data Quality [Infographic]." *InsightSquared*. January 3, 2012. Web. <http://www.insightsquared.com/2012/01/7-facts-about-data-quality-infographic/>
- ⁸² Economist Intelligence Unit. *Big Data Harnessing a Game-Changing Asset*. SAS, 2011. Web. http://www.sas.com/resources/asset/SAS_BigData_final.pdf
- ⁸³ "O retorno sobre ativos de dados na era do Big Data: Captura de US\$ 1,6 trilhão em dividendos de dados." *Cloud Platform News Bytes Blog* 2015. Web. <http://blogs.technet.com/b/stbnewsbytes/archive/2014/04/15/the-return-on-the-data-asset-in-the-era-of-big-data-capturing-the-1-6-trillion-data-dividend.aspx>
- ⁸⁴ BSA/IPSOS Global Data Analytics Poll, November 2014, www.bsa.org/datasurvey
- ⁸⁵ Investimentos em TI em toda a economia dos EUA, incluindo o varejo, durante a década de alto crescimento de 1990 agregaram de 1% a 2% à taxa composta de crescimento anual de produtividade dos EUA. Farrell, Diana et al. *How IT Enables Productivity Growth*. San Francisco: McKinsey Global Institute High Tech Practice, 2002. Print. http://www.mckinsey.com/insights/business_technology/how_it_enables_productivity_growth
- ⁸⁶ National Bureau of Economic Research. NBER Working Paper No. 18315: *Is U.S. Economic Growth Over? Faltering Innovation Confronts The Six Headwinds*. 2012. Print. <http://www.nber.org/papers/w18315>
- ⁸⁷ Economist Intelligence Unit. *The Deciding Factor: Big Data & Decision Making*. Cap Gemini, 2012. Web. Point Of View. <http://bigdata.pervasive.com/Solutions/Telecom-Analytics.aspx>
- ⁸⁸ Gertner, Joey. "GE for Making the 'Internet of Things' Real." *Fast Company* 2014. Web. <http://www.fastcompany.com/most-innovative-companies/2014/ge>

- ⁸⁹ Evans, Peter C., and Marco Annunziata. *Pushing the Boundaries of Minds and Machines*. GE, 2012. Web. <http://files.gereports.com/wp-content/uploads/2012/11/ge-industrial-internet-vision-paper.pdf>
- ⁹⁰ “City Of Barcelona Realizes Vision of Innovative City Governance with Cloud, Devices, and Apps.” Customers.microsoft.com. 2014. Web. <https://customers.microsoft.com/Pages/Home.aspx>
- ⁹¹ “Autodesk the Gallery Masdar Headquarters Positive Energy Building.” *Autodesk.com*. 2015. Web. <http://www.autodesk.com/gallery/exhibits/currently-on-display/adrian-smith-gordon-gill-architecture-masdar-headquarters>
- ⁹² Bunge, Jacob. “Big Data Comes To The Farm, Sowing Mistrust.” *Wall Street Journal* 2014. Print. <http://online.wsj.com/news/articles/SB10001424052702304450904579369283869192124>
- “Supply Chain Management Solution for Fast Moving Consumer Goods & Food Industries - Farm to Fork Tech Mahindra.” *Techmahindra*. 2015. Web. http://www.techmahindra.com/en-US/wwd/solutions/Pages/Enterprises/retail_farm_fork.aspx
- ⁹³ Entre 2013 e 2020, a divisão do universo digital entre mercados maduros e emergentes (ex.: China) mudará — de 60% para mercados maduros para 60% dos dados no universo digital se originando em mercados emergentes. EMC Digital Universe. *Executive Summary Data Growth, Business Opportunities, and the IT Imperatives*. IDC, 2014. Print. <http://www.emc.com/leadership/digital-universe/2014iview/executive-summary.htm>
- ⁹⁴ Long, Jessica, and William Brindley. *The Role of Big Data and Analytics in the Developing World*. Accenture, 2013. Print. Insights de parcerias de desenvolvimento da Accenture em funções de tecnologia ao lidar com desafios de desenvolvimento. https://www.accenture.com/us-en/~media/Accenture/Conversion-Assets/DotCom/Documents/Global/PDF/Strategy_5/Accenture-ADP-Role-Big-Data-And-Analytics-Developing-World.pdf
- ⁹⁵ Long, Jessica, and William Brindley. *The Role of Big Data and Analytics in the Developing World*. Accenture, 2013. Print. Insights de parcerias de desenvolvimento da Accenture em funções de tecnologia ao lidar com desafios de desenvolvimento. https://www.accenture.com/us-en/~media/Accenture/Conversion-Assets/DotCom/Documents/Global/PDF/Strategy_5/Accenture-ADP-Role-Big-Data-And-Analytics-Developing-World.pdf
- ⁹⁶ Levy, Stephen. “Bill Gates and President Bill Clinton on the NSA, Safe Sex, and American Exceptionalism.” *Wired* 2013: Print. <http://www.wired.com/2013/11/bill-gates-bill-clinton-wired/2/>
- Chhachhar, Abdul Razaque, and Siti Zobidah Omar. “Use of Mobile Phone among Fishermen for Marketing and Weather Information.” *Archives Des Sciences* 65.8 (2012): 107-119. Print. http://www.academia.edu/4592505/Use_of_Mobile_Phone_among_Fishermen_for_Marketing_and_weather_information
- ⁹⁷ Neuman, William, and Ralph Blumenthal. “New to the Archaeologist’s Tool Kit: The Drone.” *New York Times* 2014. Print. http://mobile.nytimes.com/2014/08/14/arts/design/drones-are-used-to-patrol-endangered-archaeological-sites.html?_r=1&referrer
- ⁹⁸ “Forty Thousand Relics to Be Digitized In Five Years.” *Thanhniem News*. 2010. Web. <http://www.thanhniemnews.com/entertainment/forty-thousand-relics-to-be-digitized-in-five-years-22816.html>
- ⁹⁹ Long, Jessica, and William Brindley. *The Role of Big Data and Analytics in the Developing World*. Accenture, 2013. Print. Accenture Development Partnerships Insights into the Role of Technology in Addressing Development Challenges. https://www.accenture.com/us-en/~media/Accenture/Conversion-Assets/DotCom/Documents/Global/PDF/Strategy_5/Accenture-ADP-Role-Big-Data-And-Analytics-Developing-World.pdf
- ¹⁰⁰ Future of Privacy Forum. *Big Data: A Tool for Fighting Discrimination and Empowering Groups*. Future of Privacy Forum and Anti-Defamation League, 2014. Print. <http://www.futureofprivacy.org/wp-content/uploads/Big-Data-A-Tool-for-Fighting-Discrimination-and-Empowering-Groups-Report1.pdf>
- ¹⁰¹ Wladawsky-Berger, Irving. “The Changing Nature of Globalization in Our Hyperconnected, Knowledge-Intensive Economy.” *Wall Street Journal* 2014. Print. http://blogs.wsj.com/cio/2014/06/20/the-changing-nature-of-globalization-in-our-hyperconnected-knowledge-intensive-economy/?mod=wsj_ciohome_cioreport
- ¹⁰² Por exemplo, Argentina, Austrália, Brasil, Canadá, Chile, China, Colômbia, Costa Rica, Grécia, Hong Kong, Índia, Indonésia, Coreia do Sul, México, Peru, Rússia, Suíça e Vietnã adotaram ou propuseram regras que proíbem ou limitam significativamente a possibilidade de as empresas transferirem informações pessoais para fora de seus respectivos territórios domésticos.

SOBRE A BSA | THE SOFTWARE ALLIANCE

BSA | The Software Alliance (www.bsa.org) é a maior defensora do setor global de software perante governos e no mercado internacional. Seus membros incluem algumas das empresas mais inovadoras do mundo, que criam soluções de software que impulsionam a economia e aumentam a qualidade da vida moderna.

Com sede em Washington, DC, e operações em mais de 60 países, a BSA é pioneira em programas de conformidade que promovem o uso legal de software e defende políticas públicas que incentivem a inovação em tecnologia e impulsionem o crescimento da economia digital.



www.bsa.org

Sede Mundial da BSA
20 F Street, NW
Suite 800
Washington, DC 20001, EUA
T: +1.202.872.5500
F: +1.202.872.5501

BSA Ásia-Pacífico
300 Beach Road
#25-08 The Concourse
Cingapura 199555
T: +65.6292.2072
F: +65.6292.6369

BSA Europa, Oriente Médio e África
2 Queen Anne's Gate Buildings
Dartmouth Street
Londres, SW1H 9BP
Reino Unido
T: +44.207.340.6080
F: +44.207.340.6090