

# SUPER INTERESSANTE

MARCO CIVIL  
DA INTERNET

A legislação inédita pode ser uma solução para a segurança digital



A dense grid of binary code (0s and 1s) arranged to form the shape of a question mark. The pattern is composed of numerous vertical columns of binary digits, with the arrangement creating the outline and interior of a question mark. The background is dark, making the white binary digits stand out.



# Bem-vindo à era do BIG DATA

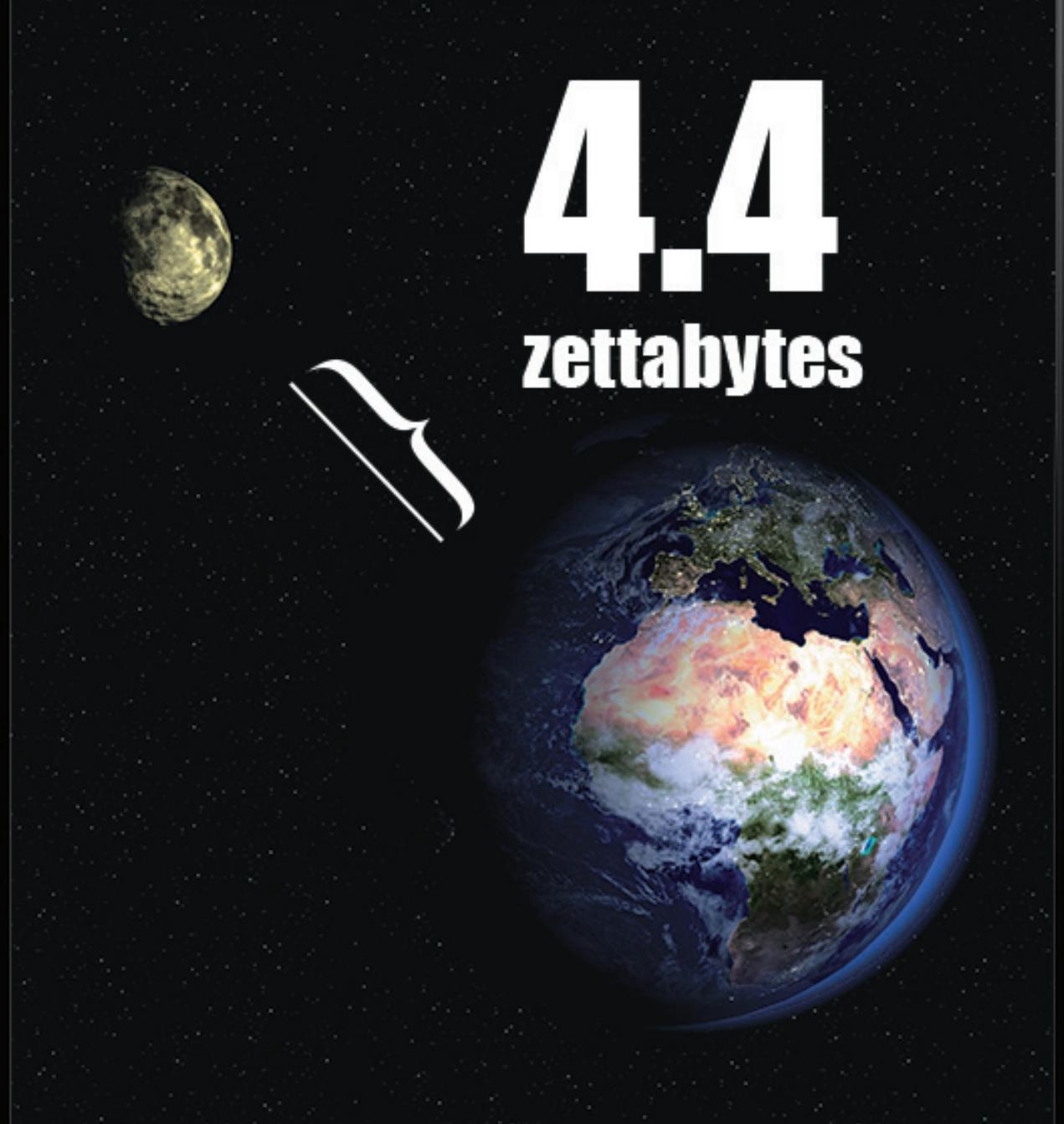
Para o mundo de hoje, unidades de medida como o gigabyte já não comportam mais os dados e será cada vez mais comum ouvirmos sobre petabytes, exabytes, zettabytes e yottabytes. É nessa explosão de números que encontramos o Big Data.

**TEXTO, EDIÇÃO E DESIGN**

*Bianca Yuki Enomura*

# IT/A





**4.4**  
**zettabytes**

**Se todos os dados digitais produzidos em 2013 fossem armazenados em tablets de 128Gb, formariam uma pilha de 253,704 quilômetros, dois terços da distância entre a Terra e a Lua.**

**Fonte:** DC e EMC

# D

efinir o termo “Big Data” é quase tão complicado quanto prever seu futuro. A primeira definição da maneira que conhecemos hoje surgiu em 2001, com Doug Laney, da empresa Gartner Group, os famosos 3 V’s: Volume, Velocidade e Variedade. São bancos

de dados tão grandes, que precisam de formas inovadoras de processamento de informação para uma melhor percepção e tomada de decisão. Na prática, significa que ferramentas poderosas estão minerando informações com mais eficiência para fazer coisas como: mapear o trânsito para sugerir melhores rotas, monitorar interesses para oferecer serviços e produtos personalizados, e até acelerar a decodificação do genoma humano.

Nesse contexto, os dados podem ser divididos em estruturados e não estruturados. O primeiro sendo informações fornecidas organizadamente, como um levantamento de perfil dos clientes ou uma pesquisa eleitoral do Ibope; já os não estruturados são aqueles coletados num “caos organizado”, como comentários no Facebook, movimentações do cartão de crédito e pesquisas no Google.

O modelo que existe hoje se desenvolveu nos últimos anos, mas a análise de dados acontece há séculos. John Graunt pode ser considerado o primeiro minerador de dados, graças ao seu trabalho para encontrar padrões que explicassem as causas da Peste Bubônica (ou Peste Negra), que matou aproximadamente 75 milhões de pessoas, um terço da população europeia no século XIV. Por meio dos boletins sobre mortalidade publicados em Londres, Graunt também fez análises sobre o tamanho da população londrina na época e o número de nascimentos de homens e mulheres.

A diferença entre o modo de pesquisas, estatísticas e dados anteriores ao universo “www”

reside justamente no superlativo dos 3 V’s. Se antes a dificuldade era encontrar a informação, hoje o desafio e o grande trunfo é o processamento correto para entendê-la. E isso, é claro, está diretamente ligado às novas tecnologias, e em particular, às possibilidades de armazenamento e manipulação de dados de código aberto, como o Hadoop.

Trabalhar efetivamente com grandes volumes de dados significa interpretar corretamente as correlações entre os números. E, para isso, além de manter os bancos de dados atualizados com as informações necessárias, é preciso utilizar ou desenvolver ferramentas para decodificá-los e traçar estratégias para colocá-los em prática.

## GRANDE POTENCIAL

As projeções e expectativas de empresas do setor de Tecnologia de Informação para o Big Data são grandiosas, assim como o volume de dados que o definem. A EMC Corporation, multinacional da área de tecnologia da informação, anunciou, em abril, o único estudo que quantifica e prevê a quantidade de dados produzidos anualmente. O estudo deste ano intitulado “O Universo Digital de Oportunidades: Dados valiosos e o valor crescente da Internet das Coisas”, com pesquisas e análises da IDC, revela como o surgimento de tecnologias sem fio, produtos inteligentes e empresas definidas por softwares estão desempenhando um papel central no aumento do volume de dados do mundo. Devido, em parte, a esta internet das coisas, o universo digital está dobrando de tamanho a cada dois anos e vai multiplicar 10 vezes entre 2013 e 2020 – de 44 trilhões de gigabytes para 44 trilhões.

Uma outra pesquisa, divulgada em 2013 pelo IDC, aponta que o mercado para Big Data vai atingir US\$ 16,1 bilhões em 2014, incluindo infraestrutura, serviços e softwares. E a razão de ser ➤

» um mercado com potencial lucrativo tão alto são os resultados obtidos até o momento.

Você pode não ver, mas o Big Data não está presente somente no dia a dia das grandes empresas. Cada ligação, mensagem, pesquisa no Google, viagem ou compra no cartão de crédito é mais informação adicionada ao seu perfil. Toda ação do seu cotidiano pode ser transformada em dados, mas nem tudo é necessariamente armazenado. É claro que nem todas essas informações têm utilidade no universo digital, apenas 22% foram consideradas como “dados úteis” e menos de 5% dos dados úteis realmente foram analisados, de acordo com o estudo.

Embora dois terços dos dados digitais sejam produzidos pelos consumidores, as empresas têm obrigação ou responsabilidade por 85% do universo digital. As informações que você posta e seus dados de navegação no Facebook, por exemplo, enriquecem as bases de dados corporativos com hábitos de consumo e comportamento. E a partir disso, os softwares podem descobrir seus interesses e direcionar a publicidade mais adequada e aumentar a probabilidade de que a compra seja efetuada.

Cada solução atende à necessidade da companhia que o contrata, e pode variar no tipo de informações coletadas e nas aplicações dadas ao resultado. Na rede de lojas Renner, o sistema cruza em tempo real a posição dos caminhões com os produtos por GPS e as quantidades dos seus estoques. O programa ainda monitora redes sociais para identificar a aceitação dos produtos divulgados nos comerciais, um recurso que vem sendo utilizado pelas empresas como uma maneira rápida de obter feedback e analisar se as tomadas de decisão estão corretas ou precisam ser modificadas.

Em alguns casos, o objetivo não é aumentar a porcentagem de vendas, mas melhorar a efetividade da empresa. Em 2010, a Central

Nacional Unimed, operadora de planos de saúde, decidiu implantar um sistema para automatizar tarefas e eliminar o fluxo manual de documentos. A empresa investiu aproximadamente R\$ 1 milhão em uma solução de gestão desenvolvida pela IBM e implantada pela CPM Braxis Capgemini, que permite capturar as contas médicas, distribuir as informações com segurança no repositório de conteúdo e a partir do cruzamento dos dados, a Unimed pode monitorar os indicadores em tempo real. Na prática, o recurso gerou um aumento de produtividade de 20%, incluindo redução de cerca de 2,5 milhões de folhas de papel em decorrência da utilização do novo sistema.

Se você mora numa cidade grande, provavelmente já gastou muito tempo no trânsito. Na cidade de São Paulo, o recorde de congestionamento registrado no dia 23 de maio deste ano foi de 344 km, de acordo com informações da Companhia de Engenharia de Tráfego (CET). No mesmo dia, 834 km foram registrados pela MapLink, empresa que calcula o volume real do tráfego com mais abrangência e precisão do que a CET.

A MapLink identificou um nicho de mercado no problema de mobilidade de mais de 11 milhões de habitantes só na capital paulista. ▶

## A EXPLOSÃO DE NÚMEROS ATÉ 2020

### 318 SMARTPHONES

de 32Gb será a produção de dados de uma família média. Quase 5 vezes mais do que hoje.

### 40%

dos dados serão arquivados nas nuvens. O dobro do que foi armazenado em 2013.

### 32 BILHÕES

de aparelhos serão conectados à internet (carros, brinquedos, coleiras de cães com softwares).

# a cada **MINUTO** *do dia*

Usuários do  
**PINTEREST**

marcam

**3.472** imagens

**8.333** vídeos são  
compartilhados no

**VINE**

Usuários do  
**SKYPE**  
se conectam  
por 23.300  
horas

**26.380**  
recomendações são  
postadas no  
**YELP**

**48 mil**  
aplicativos são  
baixados na  
**Apple Store**

Usuários do  
**PANDORA**  
escutam 61.141  
horas de música

Usuários do  
**YOUTUBE**  
publicam  
**72 H** de novos  
vídeos

**204**  
milhões de  
**E-MAILS**  
são enviados

**Google**  
recebe  
**4 milhões**  
de buscas

Usuários do  
**FACEBOOK**  
publicam 4.460.000  
postagens

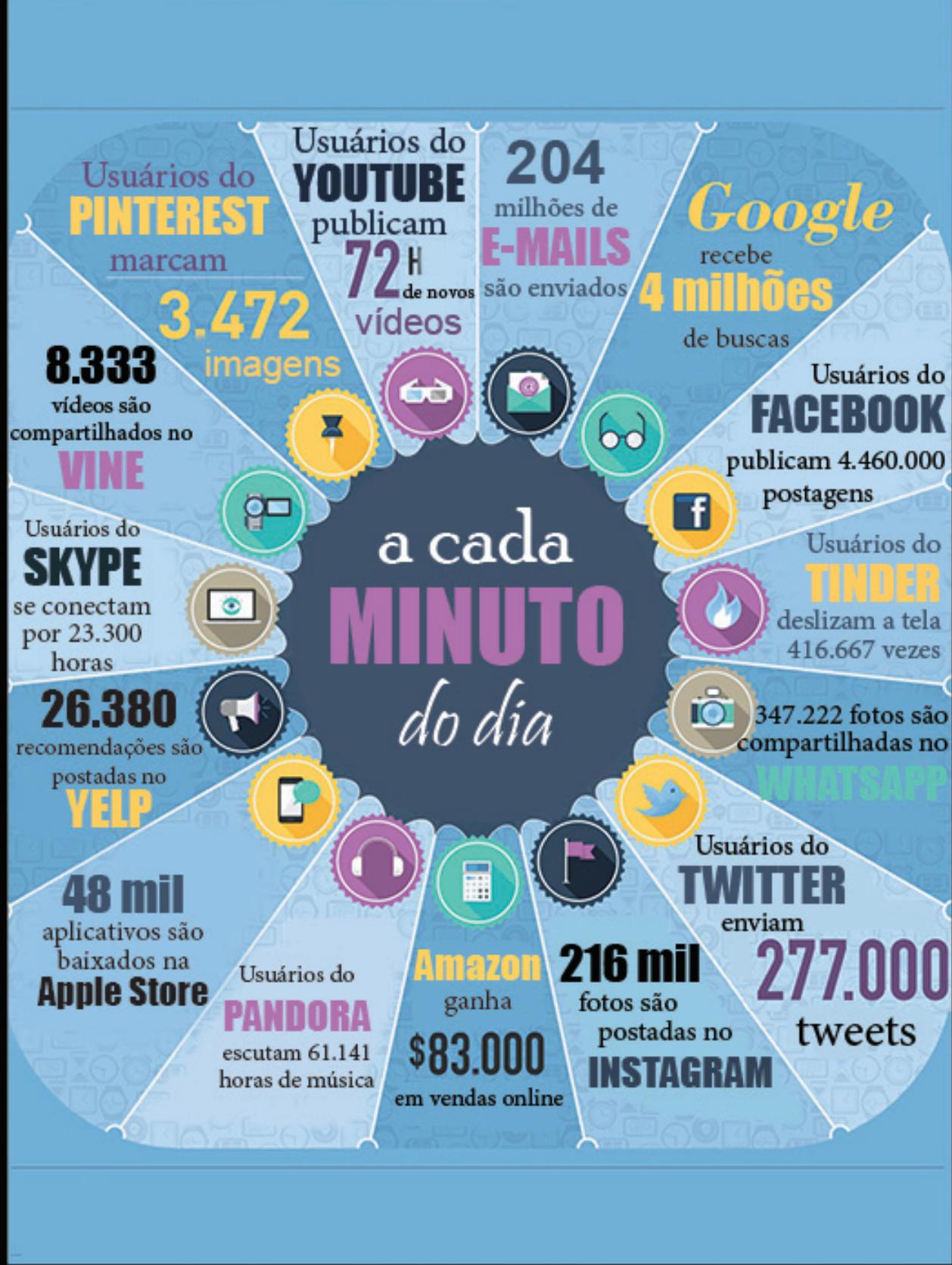
Usuários do  
**TINDER**  
deslizam a tela  
416.667 vezes

347.222 fotos são  
compartilhadas no

**WHATSAPP**

Usuários do  
**TWITTER**  
enviam

**216 mil**  
fotos são  
postadas no  
**INSTAGRAM**  
**277.000** tweets



# DADOS EM PERSPECTIVA

Você já parou pra pensar na quantidade de dados produzidos? O volume de informações já é tão grande que fica difícil entender sem dimensionar com outros exemplos.

Nota: os exemplos são baseados no sistema decimal e não nos tamanhos binários dos arquivos.  
Logo, 1MB = 1.000 KB e não 1.024KB.

## YOTTABYTES

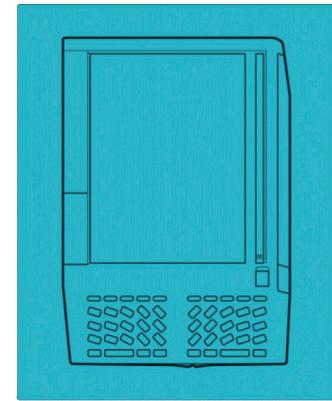
1 YOTTABYTE = aproximadamente 1.000.000.000.000.000.000 megabytes



**10.000 YB**  
Todos os micróbios do planeta

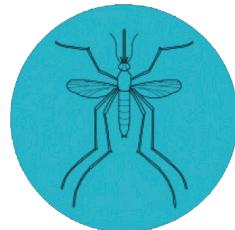
## MEGABYTES

1 MEGABYTE = aproximadamente 1000 kilobytes



**256 MB**  
Memória da primeira geração do Kindle

**66 MB**  
Genoma do mosquito da malária



## ZETTABYTES

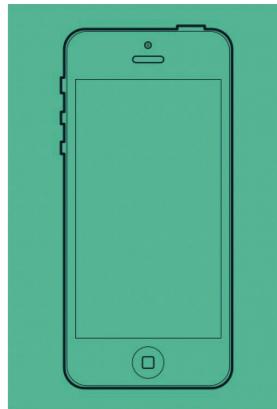
1 ZETTABYTE = aproximadamente 1.000.000.000.000.000.000 megabytes

A large grid of binary code (0s and 1s) representing 44 ZB of data. The grid is approximately 20 columns by 20 rows of binary digits.

**44 ZB**  
Previsão da quantidade de dados no mundo em 2020

## GIGABYTES ➤➤➤➤➤

1 GIGABYTE = aproximadamente 1000 megabytes

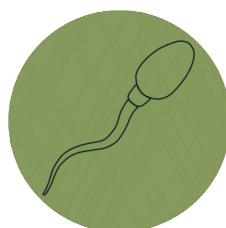


**32 GB**

Capacidade de armazenamento total do iPhone 5

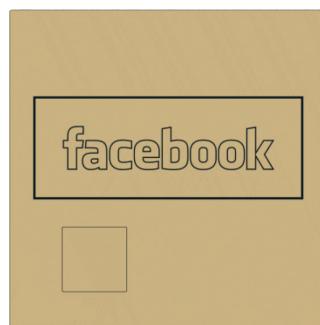
**1.760 GB**

Esperma humano:  
DNA criado por  
homem, por segundo



## TERABYTES

1 TERABYTE = aproximadamente 1.000.000 megabytes



**274 TB**

Fotos e vídeos gerados por dia no Facebook, em 2012

**12 TB**

Todo o tráfego de internet do ano 1990



## EXABYTES ➤➤➤➤➤

### EXABYTES

1 EXABYTE = aproximadamente 1.000.000.000.000 megabytes



**330 EB**

Tráfego de internet em todo o ano de 2011

## PETABYTES ➤➤➤➤➤

### PETABYTES

1 PETABYTE = aproximadamente 1.000.000.000 megabytes

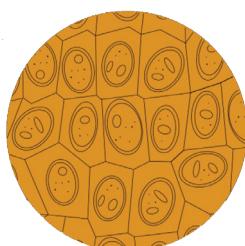


**20 PB**

Dados processados por dia no Google, em 2008

**2.400 PB**

Células da pele humana eliminadas em um mês



# Fluxo de dados deve inverter até 2020



O Brasil será  
responsável por 4%  
do universo digital



Brasil, China, Índia,  
México e Rússia vão  
representar 60% dos  
dados do mundo



EUA, Europa Ocidental,  
Japão, Canadá, Austrália  
e Nova Zelândia cairão  
para 40%

## DEMOCRATIZAÇÃO DO BIG DATA

Assim como o acesso aos computadores está mais democratizado, a utilização de Big Data não deve ficar restrita às grandes corporações. O preço para armazenar dados está diminuindo e estão surgindo ferramentas baratas ou gratuitas para lidar com eles.

## Hadoop

é um dos exemplos. A plataforma em Java faz o processamento de grandes volumes de dados. Foi construída e é atualizada por vários colaboradores e organizações.

informações de GPS instalados em mais de 500 mil veículos no Brasil e em alguns países da América Latina.

De acordo com Mariana Lucas, assessora de imprensa, a MapLink possui parcerias com empresas de rastreamento de frota e seguradoras de automóveis, que costumam usar esses equipamentos de segurança para monitorar seus veículos. Mas a captação de dados se restringe às posições geográficas de veículos, já que não é possível identificar quem fornece as informações.

## CENTRO DE OPORTUNIDADES

Com o mercado em expansão e a explosão de oportunidades, companhias do setor já sinalizam a falta de cientistas de dados especializados. De acordo com projeções do Instituto Garner, a área demandará a contratação de 44 milhões de

ta, segundo dados do IBGE. A ferramenta desenvolvida pela empresa prevê as regiões com maior fluxo de automóveis e oferecem alternativas menos movimentadas. Enquanto aplicativos concorrentes utilizam o sinal emitido pelos smartphones de forma colaborativa para mapear a movimentação de veículos e a CET faz a coleta de informações de trânsito basicamente de forma visual, através de câmeras ou agentes, o software da MapLink coleta e cruza

profissionais em todo o mundo até 2015 sendo 500 mil vagas só no Brasil.

Para preencher esses cargos, muitas empresas recorrem a profissionais com outras formações familiarizadas com números, estatísticas e bancos de dados.

Em maio, a EMC inaugurou o Centro de Pesquisa e Desenvolvimento (Centro de P & D) Richard e Maureen E. Egan, voltado para a área de Big Data. A construção do centro, no Parque Tecnológico da Universidade Federal do Rio de Janeiro (UFRJ), foi feita por meio de um memorando de intenções com o Ministério de Ciência, Tecnologia e Inovação (MCTI). Outros centros globais de Pesquisa e Desenvolvimento serão das empresas SAP Labs Latin America, Microsoft e Intel.

O investimento previsto para os próximos cinco anos é de US\$ 100 milhões, aplicados na construção do centro e para as pesquisas aplicadas em Big Data, com foco inicial na indústria de petróleo e gás, relacionados com a aquisição, a mobilidade, análise, colaboração e visualização dos dados geofísicos, geológicos, de engenharia e de negócios utilizados através do ciclo de vida do setor.

## PÉ ATRÁS

Na contramão do otimismo que impulsiona o mercado, alguns especialistas mais cautelosos acham precipitado deixar tudo por conta do Big Data. O caso mais conhecido é o Google Flu Trends, lançado em 2008, que prometia prever surtos de doenças mais rápida e precisamente do que os sistemas dos governos. O algoritmo buscou cinco anos de registros da internet, com centenas de bilhões de buscas e criou um modelo de previsão utilizando 45 termos procurados.

A ideia, apontada como um caso de sucesso por Viktor Mayer-Schönberger e Kenneth Cukier, no livro “Big Data – Como extrair »

»volume, variedade, velocidade e valor da avalanche de informação”, não funcionou como se esperava. Em 2009 o sistema não foi capaz de prever a gripe A (ou gripe suína). O surto global do vírus da Influenza A (H1N1) foi registrado pela primeira vez em março de 2009 no México e se espalhou por mais de 75 países nos três meses seguintes. De acordo com dados da Organização Mundial da Saúde (OMS), foram registradas 18.500 mortes provocadas pela gripe A entre abril de 2009 e agosto de 2010.

Tattiana Tozzi, pesquisadora de tendências e comportamentos sociais digitais na Universidade de São Paulo (USP), não considera que o sistema tenha sido falho. “Atualmente todos os softwares de mineração e análise de dados estão sofrendo constantes melhorias. O caso do Flu Trends pode ser atribuído aos diversos fatores interdependentes que toda a construção de um cenário sofre diariamente, que podem provocar falhas como nos demais sistemas”, afirma.

O impacto das previsões incorretas não foi catastrófico para a população, pois a ferramenta não substituiu o controle de doenças tradicional. Mas o Google Flu Trends deixa claro que as correlações feitas pelos logaritmos nem sempre são significativas ou verdadeiras.

Um dos desafios dos analistas de dados é desenvolver programas e interpretar padrões que identifiquem a existência de causalidade entre as informações analisadas, pois a ligação entre dois acontecimentos não implica que um tenha causa direta sobre o outro. No exemplo do Flu Trends, a pesquisa por termos de sintomas de gripe não garante que o usuário possua a doença.

Outra ressalva é o perigo de restringir as tomadas de decisões aos dados e não considerar outros fatores. O site norte-americano FiveThirtyEight, pertencente à rede de TV ESPN, desenvolveu uma ferramenta interativa para prever quais equipes têm maior probabilidade

de avançar e vencer a Copa do Mundo 2014, sediada no Brasil.

De acordo com texto divulgado no site, as previsões são baseadas no Soccer Power Index (SPI), algoritmo desenvolvido por Nate Silver, fundador e editor-chefe, em conjunto com a ESPN em 2010. O SPI é um modelo complexo que requer mais tempo que outros projetos para preparar e limpar os dados. Os princípios por trás do programa envolvem mais previsões do que retrospectiva: uma escala de importância é atribuída aos jogos (às vezes jogos amistosos são levados a sério, se a partida for contra um rival histórico), os setores de ataque e defesa também recebem classificações (sendo que, a pontuação da defesa tende a ser mais precisa em jogos mais competitivos, como a maioria da Copa do Mundo). Por fim, ainda são equacionados dados de ligas de clubes internacionais (Inglaterra, Espanha, Alemanha, Itália e França) e competições como a Liga dos Campeões e Liga Europa para avaliar os jogadores. Ainda foi considerada a vantagem por jogar em casa e a distância a ser viajada pelas equipes.

Assim como indicam grandes casas de apostas, a seleção do Brasil é a favorita a vencer a competição, com 45% da probabilidade. Seguida pelos times da Argentina (13%), Alemanha (11%) e Espanha (8%). No entanto, ainda na primeira fase, o sistema apresentou uma falha grave: a equipe espanhola foi eliminada, contrariando as chances de 80% de chegar às oitavas de final. O que mostra que embora o algoritmo considere inúmeras variáveis, o fator humano é difícil de ser mensurado por números e, portanto, foi excluído da equação. Questões como preparação psicológica, entrosamento do grupo e disposição física afetam a performance dos atletas e podem ser definitivas para o resultado de uma competição esportiva. **S**



# A Face Humana do Big Data

*The Human Face of Big Data* é um livro de fotografias e ilustrações de informações digitais em tempo real, que conta como histórias são transformadas pelo Big Data. O projeto foi criado pelo jornalista Rick Nolan e Jennifer Erwitt, produzido pela Against All Odds e com o patrocínio da EMC.

Um time de designers e cientistas de dados da EMC e do Tableau analisou um bilhão de mensagens no Twitter e outros dados e definiu 15 conjuntos de filtros usando palavras-chave e hashtags com base nos seis grandes temas do projeto ou com outras histórias específicas.

Após a coleta e interpretação, os conjuntos de dados foram transferidos com o Aspera para o software do Tableau fazer a análise e visualização.

Uma das áreas exploradas no projeto é a tecnologia no esporte. O beisebol é apontado como um campo de uso tradicional de estatísticas, em parte por ser uma série de eventos que podem ser contabilizados em um jogador: número de arremessos bem sucedidos, bolas rebatidas, arremessos capturados.

E o exemplo de maior repercussão é o *best-seller* que se tornou hit em Hollywood, *Moneyball*.

## CONEXÕES

### DO BIG BROTHER AO BIG DATA

#### **Big Brother**

Reality show criado na Holanda, em 1999, reúne pessoas para conviverem dentro de uma casa, vigiada 24 horas por câmeras e sem contato com o mundo exterior. No Brasil, o programa teve 14 edições e oferece o prêmio de 1,5 milhão de reais. O nome do programa foi inspirado no livro...

#### **1984**

Livro clássico do autor Eric Arthur Blair, conhecido pelo pseudônimo George Orwell. Publicado em 1949, aborda o cotidiano de um regime político totalitário e repressivo. Em 2005, foi listado como uma das 100 melhores obras de língua inglesa publicadas desde 1923, pela revista...

#### **Time Magazine**

Publicação de notícias semanais, distribuída nos Estados Unidos. Há mais de 80 anos faz a nomeação de Pessoa do Ano, que reconhece a pessoa ou grupo que teve o maior destaque do ano. Em 2011, o ator Brad Pitt foi escolhido como melhor performance pelo filme...

#### **Moneyball**

O longa-metragem é baseado na história de Billy Beane, gerente geral do time de beisebol Oakland Athletics, que criou uma equipe competitiva com um orçamento muito inferior ao dos rivais, com ajuda de análise estatística de desempenho dos jogadores, baseada em...

#### **Big Data**

O termo 'Big Data', refere-se à tecnologia de armazenamento e análise de grandes volumes e variedade de dados com maior velocidade. Esse recurso se fez necessário devido à explosão de dados vivenciado de uns anos para cá: a quantidade de bits é similar ao de estrelas conhecidas no céu.

# Big Data ou Big Brother?

Segurança da informação e direito à privacidade são pontos que levantam dúvidas quanto à mineração de dados na internet.



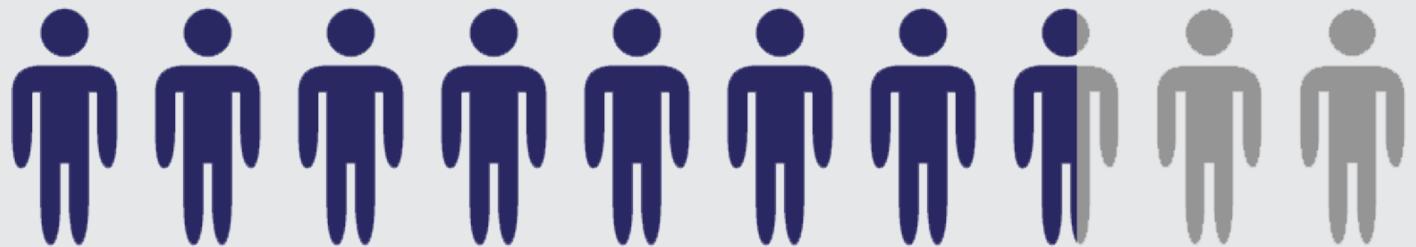
**E**m março de 2012, o governo dos Estados Unidos anunciou a “Big Data Research and Development Initiative” (Iniciativa de Pesquisa e Desenvolvimento de Big Data, em tradução livre). O programa reúne seis departamentos e agências federais e investimento de 200 milhões de dólares para melhorar ferramentas e técnicas necessárias para acessar, organizar e recolher as descobertas de grandes volumes de dados digitais. A intenção, segundo nota divulgada, é acelerar o ritmo de descoberta em ciência e engenharia, fortalecer a segurança nacional e transformar o ensino e a aprendizagem através da extração de conhecimento dos bancos de dados.

No ano seguinte o ex-agente da Agência

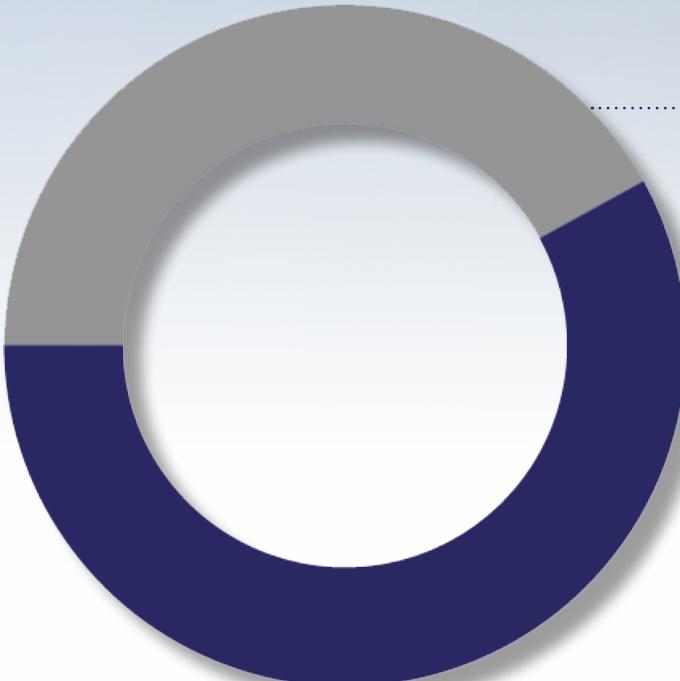
de Segurança Nacional dos Estados Unidos (NSA, em inglês), Edward Snowden, revelou que o governo norte-americano havia coletado 97 bilhões de dados em todo o mundo. Nesse montante constavam desde materiais de espionagem em empresas até violações de comunicação de chefes de Estado, como a Presidente Dilma Rousseff.

E não são apenas os governos que estão monitorando e captando informações na internet, sem o conhecimento dos usuários. O caso relatado no livro “O Poder do Hábito”, do jornalista Charles Duhigg, é um exemplo clássico para explicar esse assunto: a rede de departamentos Target, dos Estados Unidos, coletou dados pessoais do cadastro dos clientes e analisou os padrões de consumo, com ➤

# Segurança de dados no Brasil



**72%** dos consumidores já passaram por violação de dados, mas muitos não tomam as medidas necessárias para proteger sua privacidade



**42% dados não protegidos**

Números financeiros corporativos, informações de identificação pessoal, registros médicos e dados de conta dos usuários são alguns tipos que não estão totalmente protegidos.

» o objetivo de traçar o perfil do comprador e direcionar publicidade com mais precisão. Certa vez, um homem entrou numa das lojas da Target, reclamando aos gritos por sua filha adolescente ter recebido cupons de desconto com itens relacionados à gravidez e acusando a rede de estimular a garota a engravidar. Preocupado com a possibilidade de perder o cliente, o gerente da loja ligou para se desculpar pela confusão e, no fim, descobriu que as projeções estatísticas estavam corretas e a adolescente daria à luz em alguns meses.

Para muitos, pode ser inconveniente e até irritante que os softwares saibam tanto da sua vida, façam sugestões de produtos para serem adquiridos e presumam comportamentos a partir de cálculos pré-definidos. Janet Vertesi, professora assistente do curso de sociologia da Universidade de Princeton, nos Estados Unidos, quis fazer o caminho inverso ao do caso da Target e evitar a invasão de privacidade. Ela utilizou métodos para despistar o rastreamento de dados e, assim, ficar livre do bombardeamento de propaganda de itens para o enxoval do bebê. Além de evitar postagens em redes sociais, Janet utilizou e-mails pagos, recorreu a um navegador que não rastreasse seus registros e fez compras por meio de vale-presentes (para não ter o produto identificado pelo cartão de crédito). Esse último recurso alertou instituições financeiras do governo, por ser considerado um “comportamento suspeito”, pois essas ações para tentar evitar o monitoramento poderiam ser indicativos de atividade criminosa.

## MARCO CIVIL

As proteções legais para evitar abusos ainda são um problema, pois os meios de regulação não acompanharam o ritmo de evolução da tecnologia e não atendem totalmente às neces-

sidades atuais. O Brasil foi pioneiro ao discutir e aprovar a Lei 12.965 conhecida como Marco Civil da Internet, um passo importante para preservar a neutralidade da rede, estabelecer direitos de consumidores e deveres das empresas de telecomunicações e provedores de conteúdo, como Google e Facebook.

A construção do texto foi feita de forma participativa, com mais de duas mil colaborações e através da própria internet. E também foi uma resposta ao projeto de lei do então senador Eduardo Azeredo (PSDB-MG), apelidado de “AI5 digital”, que propunha uma regulamentação penal e criminalizava condutas genéricas praticadas na internet.

A reviravolta que fez o projeto tramitar em regime de urgência, a pedido da presidente Dilma Rousseff, foi quando Edward Snowden começou a divulgar o monitoramento ilegal feito pela Agência Nacional de Segurança dos Estados Unidos. O caso de espionagem fez com que a defesa da privacidade de informações aparecesse como destaque para o governo brasileiro.

Após ajustes que agradaram e desagradyaram defensores da liberdade na internet e as poderosas empresas de telecomunicações, o projeto foi aprovado na Câmara dos Deputados e no Senado e sancionado pela presidência em 23 de abril de 2014.

Ainda não está claro o impacto que a decisão terá sobre o universo do Big Data, mas o Art. 7º abre espaço para discussões ao garantir a inviolabilidade e sigilo do fluxo de comunicações pela internet e comunicações privadas armazenadas, só sendo autorizada com ordem judicial. Outros pontos do artigo ainda exigem que a coleta, uso, armazenamento, tratamento e proteção de dados pessoais sejam informados com clareza e sejam feitos apenas com consentimento expresso do ➤

» usuário.

Nesse período de consolidação das práticas na internet e suas regulações, é quase impossível escolher livremente que dados fornecer às plataformas. Assinalar a caixa “Li e aceito os termos e condições” muitas vezes significa ceder informações valiosas para as empresas. Mas a importância desse contrato parece não abalar a maioria das pessoas, já que, de acordo com pesquisa da Universidade

de Stanford, 97% dos usuários vão direto para o “concordo”.

Quando se trata do impacto direto do Big Data, a ameaça à privacidade do indivíduo é um tópico que causa receio. É aceitável que uma empresa disponibilize os dados pessoais de seus usuários para terceiros ou os utilize para fins comerciais? O direcionamento dos softwares seria uma forma de privação da liberdade de escolha? **S**

# 4 pontos para entender o Marco Civil da Internet



## NEUTRALIDADE DA REDE

O projeto regulamenta que o tráfego de pacotes de dados não pode favorecer certos tipos de serviço. Não pode haver limite de velocidade e nem censura de conteúdo.

## PRIVACIDADE

Fotos e textos de redes sociais inativas deverão ser efetivamente excluídos, o monitoramento massivo torna-se ilegal e será exigida uma autorização prévia para a venda de informações.

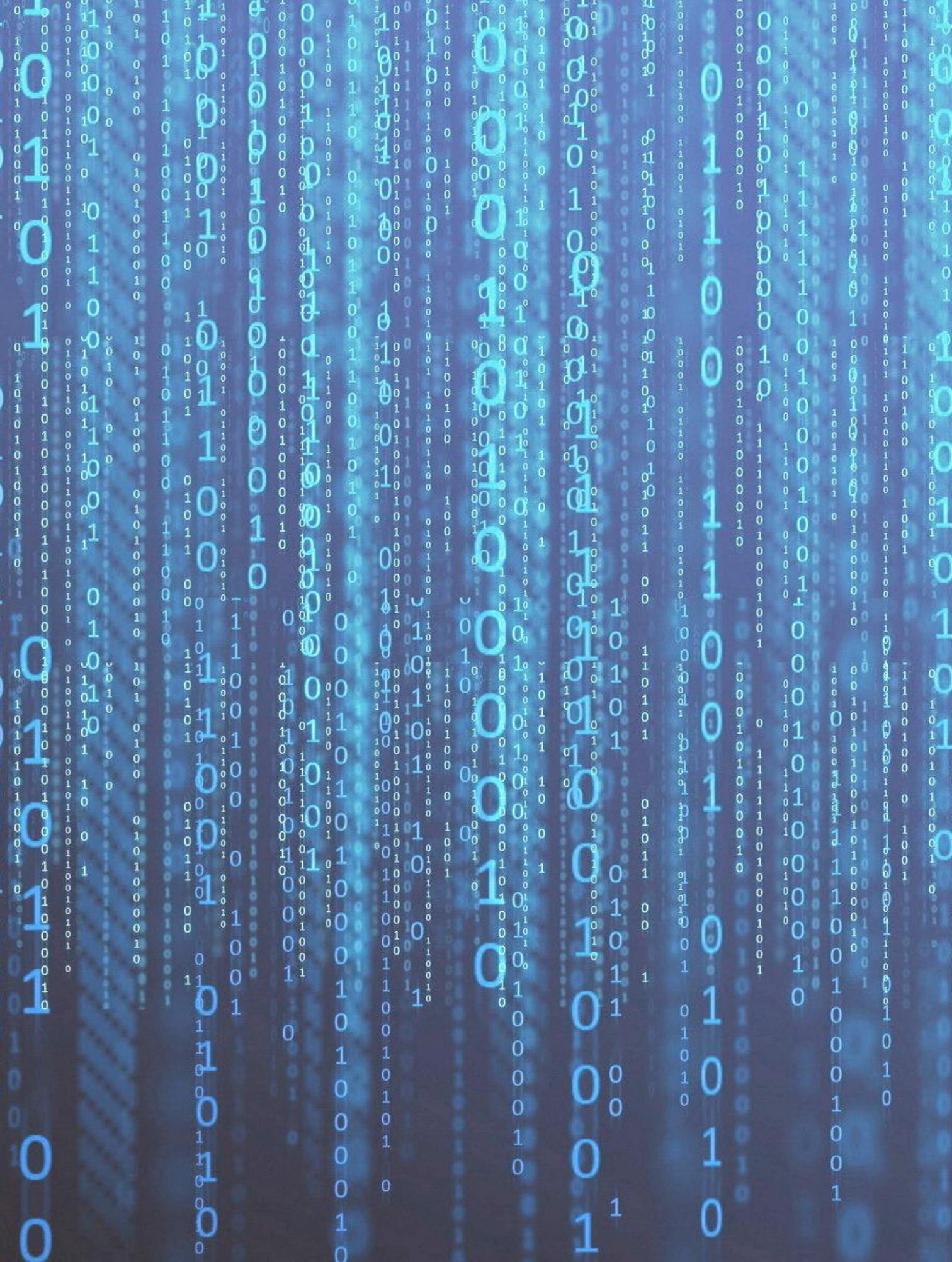
## RESPONSABILIZAÇÃO CIVIL

O provedor não será mais responsável por qualquer conteúdo publicado pelo internauta. Entretanto, para removê-lo, será necessária uma ação judicial.

## PUNIÇÕES

Estabelece punição criminal, civil e administrativa para o caso de quebra de sigilo, além de indenizar usuários cujos direitos sejam violados.

A dense grid of binary code (0s and 1s) on a dark background. The pattern forms the letters 'b' and '1'. The letter 'b' is located in the upper right quadrant, and the letter '1' is located in the lower left quadrant. The binary digits are arranged in a grid where each row and column contains a sequence of 0s and 1s.



Universidade Federal de Santa Catarina  
Centro de Comunicação e Expressão  
Departamento de Jornalismo  
Curso de Jornalismo

Trabalho de Conclusão de Curso apresentado em julho de 2014.

Aluna: Bianca Yuki Enomura  
Orientadora: Gislene Silva

Banca avaliadora:  
Rita Paulino  
Rogério Christofoletti