



A ERA DOS DADOS PARA O SETOR PÚBLICO

UMA NOVA CULTURA ORGANIZACIONAL ANALÍTICA



SOCIAL
GOOD
BRASIL



iris

Laboratório de
Inovação e Dados



CEARÁ
GOVERNO DO ESTADO

ÍRIS | Laboratório de Inovação e Dados do Governo do Ceará
AWS Institute
Social Good Brasil

A Era dos Dados para o setor público:
uma nova cultura organizacional analítica

Coleção Inovar | Transformar | Inspirar
Tema: Alfabetização em Dados

Coordenação editorial

Jessika Moreira | ÍRIS
Ana Janaina Nelson | AWS Institute

Autoria

Carol Andrade | Social Good Brasil
Fernanda Campagnucci | Open Knowledge Brasil
José Borbolla Neto | Branded Brain
José Macedo | ÍRIS | Programa Cientista Chefe do Governo do Ceará
Marianna Gonçalves | ÍRIS
Mariana Zonari | ÍRIS
Silvana Paula Martins de Melo | ÍRIS
Ticiania Linhares | ÍRIS | Programa Cientista Chefe do Governo do Ceará

Preparação e revisão

Amélia Gomes | ÍRIS
Isabel Ferreira Lima | ÍRIS

Projeto gráfico e diagramação

Mariel Maffessoni Ramos | Social Good Brasil
Angélica Freitas | ÍRIS

Brasil, 2021

SUMÁRIO



	04	
Sobre esta edição	05	
Prefácio	06	Capítulo 1 Bem-vindos à Era dos Dados
	15	Capítulo 2 “Solucionismo tecnológico”
Capítulo 3 Alfabetização em dados	20	
Capítulo 4 A jornada de dados	32	
Capítulo 5 Gestor público, lidere a mudança de cultura	43	
Capítulo 6 Transparência e dados abertos	48	
Capítulo 7 LGPD no setor público	53	
	58	Capítulo 8 Dimensões éticas e dilemas da atualidade
	64	Capítulo 9 Uso de dados na gestão pública: casos e aplicações

SOBRE ESTA EDIÇÃO

Esta edição foi idealizada pelo ÍRIS | Laboratório de Inovação e Dados do Governo do Ceará em parceria com a AWS Institute e o Social Good Brasil.

Ela objetiva orientar a gestão pública sobre o uso de dados e evidências nos atuais contextos tecnológico e humano, motivando gestoras e gestores a liderar, de forma estratégica, uma mudança de cultura no campo analítico.

De caráter introdutório, a edição aborda a alfabetização em dados como uma competência imprescindível para o setor público. Mostra, com base em ferramentas práticas, o caminho que o gestor deverá seguir para promover o aculturação analítico da gestão.

Também são apresentados projetos que confirmam como a alfabetização e a transparência de dados impactam de forma positiva a gestão pública.

Na Era dos Dados, o desafio para avançar está nos modelos mentais, na cultura, nos aspectos humanos, e não apenas na tecnologia. Por isso, para liderar e fazer a mudança, é fundamental posicionar o ser humano no centro, garantindo diálogo aberto e sentimento de propósito.

A Era da Informação e a 4ª Revolução Industrial trouxeram novos parâmetros para aferir o progresso de uma sociedade. Capacidade de conexão, produção, armazenamento e processamento de dados. Transformação de dados em informação e em inteligência: produtiva, governamental, social e coletiva.

Sociedades marcadas por relações de poder e hierarquia abrem espaço para modelos cooperativos, cocriadores e horizontais, nas quais a inteligência coletiva ocupa local de destaque nos processos participativos e nas tomadas de decisões públicas. As novas tecnologias, portanto, surgem como verdadeiros portais que nos levam a caminhos, avanços e, por que não dizer, progressos em escala ainda não vivenciada pela humanidade, em ciclos transformativos tão curtos.

É nesse contexto que nasce esta obra. Ela destaca o poder e o papel transformador dos dados, sem se deixar seduzir pelo “solucionismo tecnológico”, ou se perder em suas virtudes, nem sem descurar dos desafios do nosso tempo.

A obra enfatiza aspectos humanos, tecnológicos, normativos e de gestão que perpassam os dilemas e as atribuições dos gestores do nosso tempo. Com corações e mentes abertas para aceitar o novo, os autores nos convidam a percorrer os trilhos do progresso com a velocidade em que circulam os dados, tendo como nortes o cuidado e a preservação com o que há de mais humano em nós: a capacidade de nos conectarmos para resolver problemas de forma conjunta.

Talvez esse tenha sido o grande e imemorial tesouro descoberto pela humanidade, que, de tão óbvio, muitas vezes é esquecido. Esta obra nos inspira a percorrer esse caminho milenar: o de unidos elegermos necessidades comuns, que nos unem e preservam a todos e à “grande mãe, Gaia”. Não existe caminho solitário na estrada do progresso.

Boa leitura!

ANA CARLA BLIACHERIENE

— Professora doutora da
Universidade de São Paulo | USP

#1

BEM-VINDOS À ERA DOS DADOS

por José Borbolla



Como chegamos até aqui?

“Cientista de dados: a profissão mais *sexy* do século XXI”¹, profetizaram, em 2012, os professores Davenport e Patil em artigo publicado na *Harvard Business Review*.

De lá para cá, as áreas de dados cresceram rapidamente, espalhando-se por diversas disciplinas, segmentos e indústrias, e prometem, em dez anos, acrescentar até U\$ 15 trilhões na economia global. É o que mostra pesquisa publicada em 2017 pela PricewaterhouseCoopers Brasil (PwC Brasil)².

Por todos os lados e âmbitos da nossa vida, seja na dimensão individual, seja na coletiva, parece haver um consenso de que o *big data* e a inteligência artificial (IA) transformarão o mundo como o conhecemos e resolverão “definitivamente” os problemas da sociedade.

A resposta não é simples nem óbvia, mas talvez seja determinante para o futuro da humanidade. Para explorá-la,

vamos voltar um pouco no tempo e entender como chegamos até aqui.

Não é novidade que as tecnologias sempre evoluíram, se transformaram e transformaram a nossa maneira de interagir com o mundo ao nosso redor. Ou seja, falar sobre evolução tecnológica é falar também sobre a evolução da nossa espécie.



Para entendermos o tema central desta publicação, vamos partir de uma perspectiva sociológica e histórica apoiada em três premissas detalhadas a seguir.

1 Confira o artigo completo em: <<https://hbr.org/2012/10/data-scientist-the-sexiest-job-of-the-21st-century>>.

2 Confira a pesquisa em: <<https://www.pwc.com/gx/en/issues/data-and-analytics/publications/artificial-intelligence-study.html>>.

PREMISSA 1

VELOCIDADE DA MUDANÇA

Sim, as técnicas e ferramentas sempre mudaram ao longo do tempo. Mas, do início do século XX aos dias atuais, a velocidade dessas mudanças está mais acelerada (Figura 1). Esse ritmo, cada vez mais frenético, traz consigo desafios para a sociedade e para as instituições, nas esferas pública e privada.

PREMISSA 2

AVANÇO DAS TECNOLOGIAS DE COMUNICAÇÃO

Ainda que possamos conceber essa premissa como um componente da anterior, é importante lançar luz sobre a maneira como as diferentes tecnologias de comunicação transformaram, e seguem transformando, a forma e a velocidade da informação.

PREMISSA 3

PADRÃO E VELOCIDADE DE CIRCULAÇÃO DA INFORMAÇÃO

Essa terceira premissa é também um desdobramento lógico da anterior. Sempre que uma nova tecnologia de comunicação foi criada e “implementada”, ela acabou por transformar, radicalmente, as estruturas sociais de suas respectivas épocas.

• • •

Neste início de século XXI, estamos em meio a um processo de transformação em que essas três premissas são protagonistas. Compreendê-las nos tornará mais aptos a enfrentar os novos desafios que elas nos oferecem.

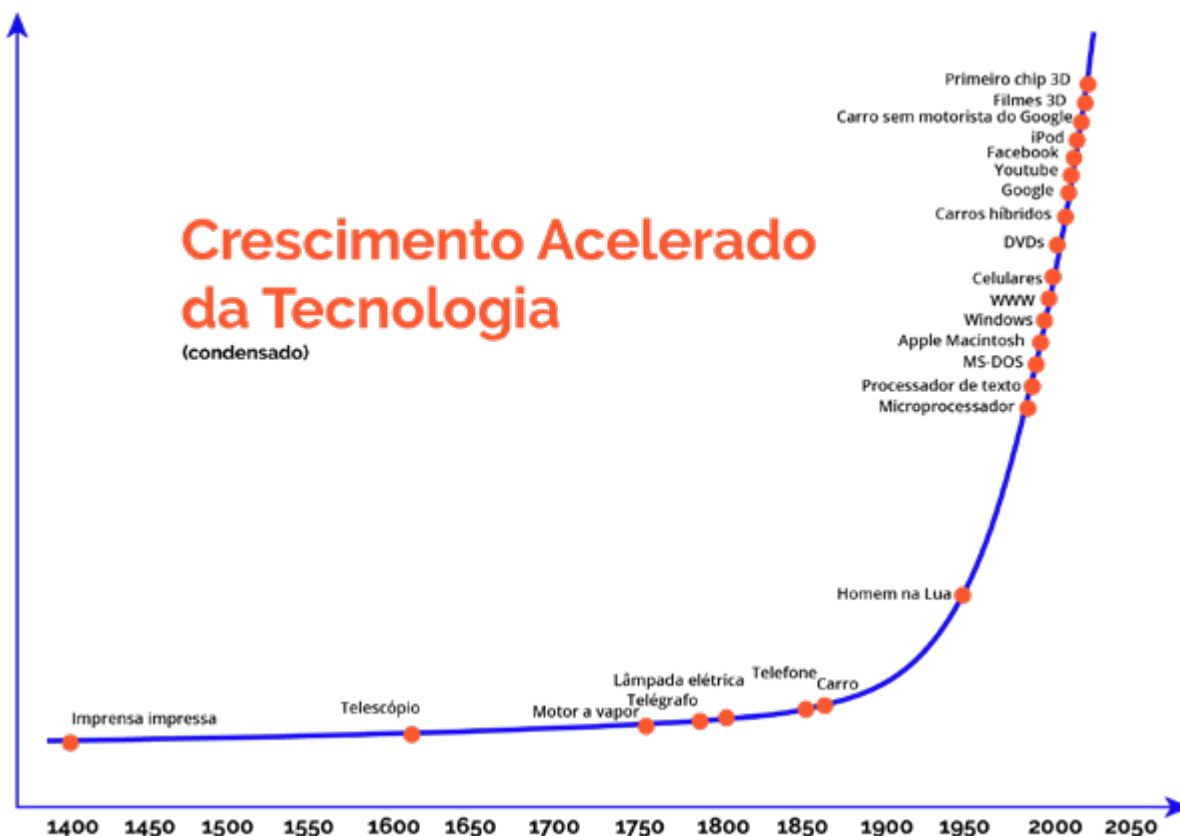


Figura 1. Curva referente ao crescimento acelerado das tecnologias

Efeitos da velocidade do avanço tecnológico

O avanço alcançado pelas sociedades ao longo dos últimos séculos é indiscutível. Segundo o historiador e filósofo israelense Yuval Harari, em seu livro *Sapiens: uma breve história da humanidade*, nenhuma outra espécie que já habitou este planeta ocupou praticamente todos os cantos do globo em tão pouco tempo. Esse processo se deu, obviamente, pela capacidade humana de colaborar em larga escala e de produzir ferramentas e técnicas que iam transformando padrões de interação e impacto no ambiente ao nosso redor.

Na medida em que vamos chegando a épocas mais recentes, o que parece ficar claro, ao observarmos a história dessas invenções, é uma aceleração na frequência e na ocorrência de inovações e novidades envolvendo técnicas e ferramentas.

Portanto, quanto mais acelerado for o avanço, mais frequentes serão os impactos e efeitos sobre a sociedade humana, ainda que nossas capacidades de compreensão e adaptação não tenham evoluído no mesmo compasso.

A Figura 2 traz a ideia de exponencialidade e ilustra como cada revolução técnica se dá sobre as bases da anterior. Ela é baseada no livro *Digital technology and social change: the digital transformation of society from a historical perspective*, de Martin Hilbert.

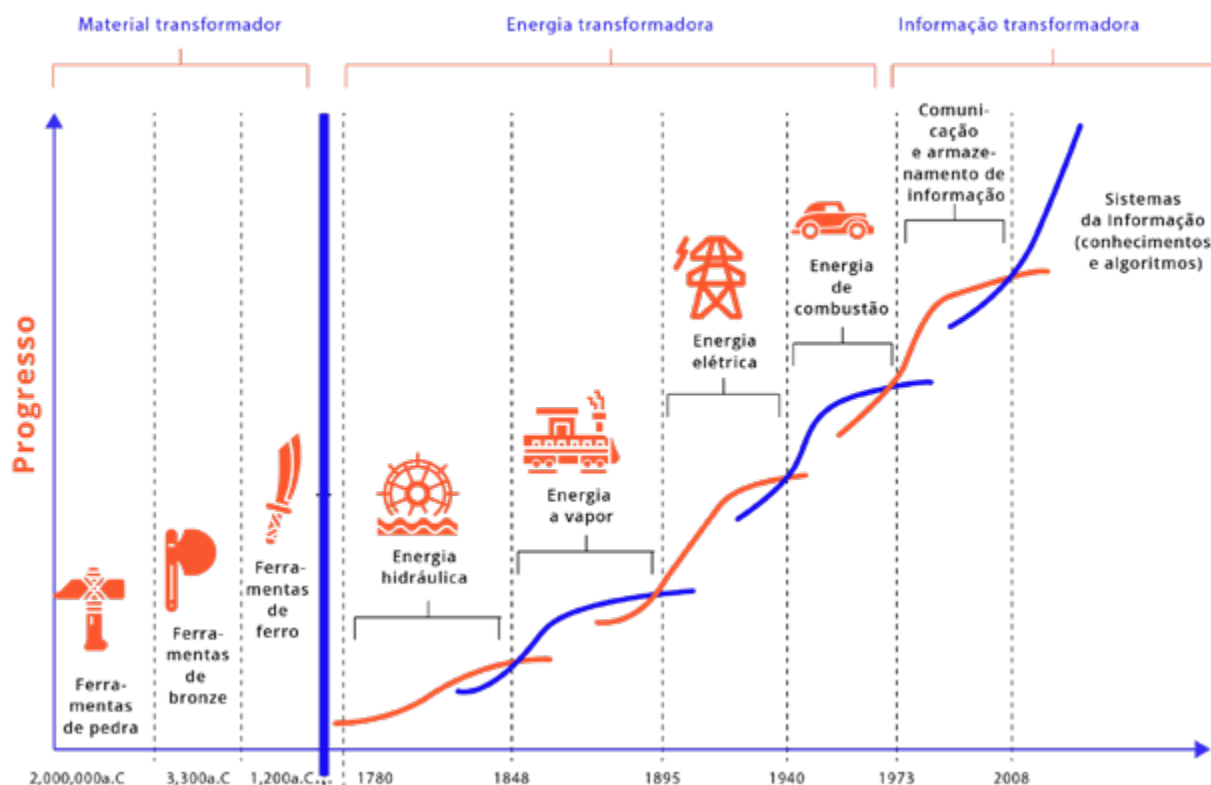


Figura 2. Revoluções técnicas

Explosão informacional

Todas as tecnologias de comunicação inventadas ao longo do tempo também transformaram, radicalmente, nossas relações e padrões organizacionais. Para ilustrar, vamos olhar para a prensa mecânica de tipos móveis, criada por Gutenberg na segunda metade do século XV, na Europa.

Por essa época, a produção de informação era restrita devido a dois fatores: velocidade de confecção de um livro manuscrito e quantidade de pessoas alfabetizadas e letradas, com tempo e engenhosidade necessários para esse tipo de tarefa.

A prensa mecânica, quando surge, impacta diretamente esses dois fatores. Em poucas décadas, espalha-se pela Europa. Panfletos, literatura de cordel e livros passam a ser produzidos em escala e com linguagens mais acessíveis. Foi amor à primeira vista.

Mas esse amor trouxe impactos, da Reforma Protestante à Contra-Reforma, do Iluminismo às Revoluções Francesa e Americana. As novas possibilidades e velocidades de produção e circulação de informação tiveram grande efeito sobre os principais eventos que esculpiram o mundo no formato que hoje conhecemos.

Se todas essas mudanças se deram ao longo de três ou quatro séculos, o que está diante de nós hoje, em 2021, é algo de outra magnitude. Daí sua importância e relevância.

Inspirados no livro *A praça e a torre*, de Niall Ferguson, podemos dizer que a internet e as telecomunicações disponíveis atualmente seriam, de maneira similar e com liberdade poética, a prensa mecânica do século XXI em relação ao potencial de impacto e transformação que elas possuem.

Mas essa comparação só faz sentido se acrescentarmos a velocidade e entendermos como essa combinação muda radicalmente o padrão de circulação das informações.

Imagine todas essas mudanças que a Europa enfrentou ao longo de 300 ou 400 anos acontecendo hoje no mundo, só que em poucas décadas.

Sempre que o padrão mudou, o mundo mudou junto

Da década de 1970 para cá, a evolução dos computadores e o surgimento da internet aceleraram a níveis incomparáveis a quantidade de informação produzida (em formato digital), alterando profundamente os padrões de circulação, acesso e consumo.

Se durante os últimos 100 anos estruturamos nossas vidas e sociedades a partir de tecnologias de *broadcast* (transmissão ampla de um emissor para muitos receptores, em via de mão única), a internet e as tecnologias de comunicação móvel transformaram radicalmente esse padrão. Antes, nosso problema era a dificuldade e o custo para acessar e obter informações; hoje, vivemos em uma realidade diametralmente oposta.

É esse mesmo contexto histórico recente de explosão informacional que cria as condições para o (re)surgimento de conceitos como inteligência artificial ou aprendizado de máquina, ideias que já vinham sendo gestadas nas esferas acadêmicas desde a década de 1940.

Enquanto, por um lado, a capacidade de processamento cresce exponencialmente, por outro, o aumento da adoção e do uso de computadores contribuiu de maneira definitiva para acelerar o processo de digitalização de informações de todos os tipos.

Se estamos aqui, poucas décadas depois, discutindo temas relacionados às áreas e tecnologias de dados (como as definimos atualmente), esse é apenas o capítulo mais recente de um processo histórico mais amplo e antigo. **A novidade, portanto, não é o fato de haver transformação tecnológica mas quão rápido ela ocorre, uma vez que nossa capacidade de adaptação não avança no mesmo ritmo.**

Por isso, hoje, vivemos em um mundo em que coexistem cartórios (exemplos perfeitos do impacto que a prensa mecânica e o papel nos deixaram de recordação) e carros que dirigem sozinhos a partir de algoritmos e sistemas avançados de processamento e análise de dados.

Os tipos de problemas e desafios que instituições como os cartórios resolvem são típicos de épocas anteriores e já estão dando seus últimos sinais de vida. Ao mesmo tempo, ainda desconhecemos os problemas e desafios produzidos por um carro que dirige e toma decisões sozinho.



Figura 3. Manchete da revista *Wired*, 2014. "Dados são o novo petróleo da economia digital".

Aspectos culturais e organizacionais

Vamos imaginar que o tipo e/ou modelo de instituição e o paradigma de gestão de uma determinada época estão intimamente conectados à estrutura e à tecnologia comunicacional, os quais, por sua vez, determinam o padrão e a velocidade de circulação da informação.

O exemplo do cartório, citado há pouco, ilustra bem essa relação. Em uma sociedade que tem o papel como seu principal meio de armazenamento de dados, é natural que haja uma instituição para validar a qualidade e a veracidade de seus registros. Usaremos essa mesma lógica para pensar nas aceleradas e profundas mudanças que testemunhamos nos anos recentes.

O cartório, representante do paradigma anterior, é capaz de resolver o problema da qualidade ou da veracidade da informação digital? Sabemos que a resposta é óbvia.

E se reformularmos a mesma pergunta substituindo o cartório por empresa ou instituição pública? Será que todos esses modelos, concebidos em outras épocas e contextos, estão preparados para navegar nesse mundo da explosão informacional?

As estruturas organizacionais e os modelos de gestão com os quais estamos acostumados no momento foram gestados e esculpidos ao longo do século XX, em uma realidade ainda distante da atual, de comunicação instantânea e carros autônomos.

Vamos resgatar duas premissas mencionadas no início deste capítulo — avanço das tecnologias de comunicação e padrão e velocidade de circulação da informação — para explorar a ideia que o economista Ronald Coase publicou em seu artigo “The nature of the firm”³, em 1937.

Coase destaca, entre outros fatores, que a dificuldade e o custo de se obter informações e dados obrigavam as organizações a assumirem determinadas formas, como estruturas piramidais ou centralizadas, áreas e silos muito especializados, paradigma de comando-e-controle.

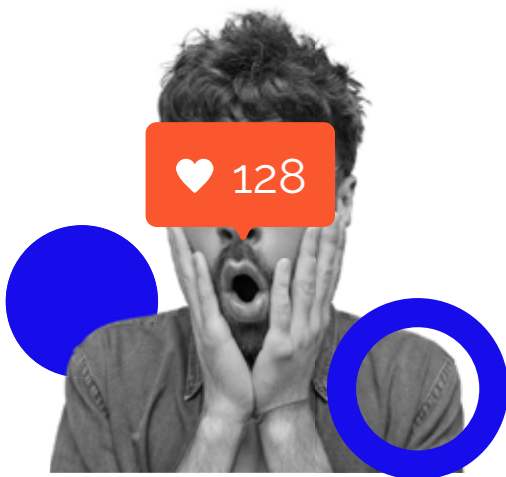
Toda essa ideia que ainda hoje sobrevive, nos mais diversos segmentos e indústrias, nas esferas pública e privada, nasceu, portanto, em um mundo escasso de informações. E foi preciso desenvolver estratégias para lidar com isso.

Já avançamos quase 90 anos desde a publicação de Ronald Coase, e vivemos agora um contexto oposto: se, antigamente, o acesso e o custo para obter informação eram obstáculos, hoje o obstáculo é o excesso de informações.

3 Confira o artigo completo: <<https://onlinelibrary.wiley.com/doi/full/10.1111/j.1468-0335.1937.tb00002.x>>.

Mas, assim como no caso do cartório, toda a lógica por trás do funcionamento de uma empresa tradicional ou de uma repartição pública⁴ foi concebida para lidar com problemas e desafios de um mundo que não existe mais. Não bastaria, portanto, e ao contrário do que afirma uma visão muito popular da nossa época, adquirir tecnologias e ferramentas mais modernas para fazermos melhor e de maneira mais eficiente o que sempre fizemos, do jeito que sempre fizemos.

Novos problemas demandam novas ferramentas. E, para concebê-las, é fundamental que entendamos a natureza desses novos desafios.



Estamos preparados para destravar os imensos potenciais de novas tecnologias?

Toda transformação carrega consigo potenciais aplicações e usos de novas tecnologias, capazes de nos ajudar a tomar melhores decisões e a superar os desafios da sociedade, como nunca houve no passado.

Enquanto nos preparamos e nos protegemos contra possíveis riscos e efeitos colaterais das novas tecnologias, estamos diante de complexas escolhas para criar condições necessárias de forma que o potencial positivo se concretize. De maneira prática, quando pensamos nesse universo de dados e nos papéis e responsabilidades dos gestores públicos, precisamos considerar dois fatores: alfabetização em dados e como nosso cérebro toma decisões. Eles serão detalhados a seguir.

⁴ Referência, de maneira mais ampla, a todo paradigma tradicional de gestão, estrutura organizacional e liderança.

ALFABETIZAÇÃO EM DADOS

É crucial que todas as pessoas que trabalham com dados saibam não apenas operar ferramentas, seja ela um *software* de *business intelligence* (BI) ou uma nova linguagem de programação, mas também que desenvolvam e ampliem suas capacidades de compreender, analisar, interpretar e traduzir dados.

COMO NOSSO CÉREBRO TOMA DECISÕES

Se pudéssemos resumir o poder que as tecnologias nos oferecem, isso certamente estaria relacionado a como podemos usar uma quantidade enorme de dados para tomar melhores decisões. Mas essa relação entre quantidade de dados e qualidade de decisão, como veremos, não acontece de maneira linear e direta.

...

Nosso cérebro não é uma máquina de processar dados, uma linha de produção, que na medida em que vai recebendo novas informações vai produzindo decisões melhores. Talvez esse seja o grande sonho da nossa espécie nos tempos atuais, mas não descreve bem nossa natureza e como nosso cérebro funciona.

As decisões que tomamos são influenciadas por uma grande quantidade de elementos internos, externos, conscientes e inconscientes.

A racionalidade é, sim, uma capacidade cognitiva humana. No entanto, ela está, de maneira desordenada, conectada

a diversos outros sistemas cerebrais e aos contextos histórico, cultural e social, dentro dos quais estamos inseridos. Somos mulheres e homens da nossa época, afinal.

A cultura institucional dentro da qual exercemos nossas atividades profissionais tem um peso importante na maneira como interpretamos dados, atribuímos valores e tomamos decisões. Toda essa força bruta de processamento, todas essas técnicas com nomes difíceis, ferramentas poderosas, quantidade inimaginável de dados e outros fatores são capazes de nos ajudar. Mas, ainda assim, não nos fornecem a resposta definitiva.

E o que isso significa, na prática? Significa que nossa expectativa ao utilizarmos essas tecnologias cada vez mais avançadas, capazes de processar uma quantidade de dados que já ultrapassa os quintilhões de *bytes* diários, não deve ser alocada na esperança de obtermos uma “verdade absoluta”.

A maior parte dos grandes desafios da nossa sociedade são problemas de natureza complexa, que não costumam admitir respostas simples e únicas. A inteligência artificial e o *big data* não são as “bolas de cristal do século XXI”, mas sim mecanismos que amplificam nossa capacidade de analisar cenários e pensar criticamente sobre a realidade.

Quanto mais a tecnologia avança, mais fundamental o aspecto humano se torna.

#2

“SOLUCIONISMO TECNOLÓGICO”

por José Borbolla



Novas tecnologias e seus efeitos na sociedade

A narrativa parecia perfeita. Tecnologias avançando rapidamente, nossa vida sendo transformada por dispositivos mais modernos, portáteis, poderosos, multifuncionais e com conexão mais estável e rápida.

Foi a década de 1990 que gestou a crença de que havíamos alcançado o ápice do nosso desenvolvimento social e político e que todos os problemas, a partir dali, seriam endereçados pela tecnologia⁵.

No campo da Educação, por exemplo, ganhou força a ideia de que computadores nas escolas e o acesso à internet, oferecendo conhecimento gratuito, acessível e instantâneo para todos, resolveria, de forma definitiva, as limitações e os desafios da educação tradicional.

Como compara o historiador Yuval Harari, criamos, inclusive, nosso próprio Vaticano do século XXI no Vale do Silício californiano. De lá vieram as principais profecias sobre esse futuro incrível e exponencial. Bastaria que não se colocassem obstáculos ao avanço tecnológico e este, por si

só, daria conta das imperfeições do nosso mundo, como se todos os problemas da sociedade humana fossem causados por má gestão e falta de poder de processamento...

Em particular, a gestão pública sofre de maneira mais intensa pressões e cobranças por essa crença equivocada e simplória de que a má gestão é a causa central de todos os males.

A ideia de que todos os tipos de serviços públicos podem ser digitalizados, em formato mais eficiente e menos oneroso, protagoniza diversas esferas de discussões sobre o avanço tecnológico. Entre elas, a que diz respeito ao impacto e à transformação do paradigma atual de como as instituições públicas devem ser operadas.

Se, por um lado, novas tecnologias nos ajudam a resolver problemas, por outro, trazem consigo novos desafios. E isso não é de hoje.

⁵ Francis Fukuyama, cientista político e economista norte-americano, explorou ideias análogas em seu livro *O fim da história e o último homem*, de 1991.

Inúmeras aplicações trouxeram soluções incríveis às quais estão atrelados riscos e dificuldades.

Nos momentos de lançamento das novas tecnologias, o mundo todo se enche de esperança. Mas bastam alguns anos para ficar evidente que a questão era um pouco mais complexa do que parecia no início.

Vamos pensar em exemplos concretos, como a eletricidade ou a energia nuclear. Olhando para trás, parece óbvio que enfrentaríamos todos os problemas e catástrofes que essas tecnologias traziam consigo, na medida em que avançavam e se espalhavam pelo planeta. Mas isso não era óbvio quando elas nasceram. E talvez o otimismo tenha nos dominado quando os primeiros experimentos com eletricidade ou energia nuclear demonstraram que elas eram capazes de resolver problemas antigos de maneira aparentemente simples.

Cultivar a simplória e perigosa ideia de que mais tecnologia é sempre melhor e que os problemas e desafios atuais da nossa sociedade só existem porque ainda não desenvolvemos a tecnologia apropriada para resolvê-los é o que alguns autores chamam de **“solucionismo tecnológico”**.

A energia nuclear trouxe a promessa de eletricidade barata e acessível para o planeta. Pouco ou nada se falava, no início, sobre outros usos possíveis. Em teoria, se os governos decidissem combinar uma política global para utilizar a energia nuclear, seria possível oferecer eletricidade para o planeta sem queimar combustí-

veis fósseis ou esgotar recursos naturais escassos? Há quem diga que sim.

O fato de a tecnologia dispor de um determinado atributo ou potencial técnico não diz nada sobre como ela será utilizada na vida real.

É absolutamente crucial e indispensável entender essa limitação e superá-la, não apenas para um profissional que ocupe uma posição estritamente ligada à análise, ciência ou engenharia de dados, mas também a todas as outras áreas e disciplinas.

Enquanto adotarmos o “solucionismo tecnológico” como uma perspectiva única, seguirão altas as chances dos nossos sonhos utópicos se tornarem pesadelos reais.

E o que significa superar a visão solucionista?

Significa entender que os atributos técnicos de uma tecnologia qualquer sugerem potenciais aplicações, mas nada informam ou determinam sobre o uso que aquela tecnologia receberá.

Quando deixamos de olhar para a inteligência artificial como provedora de “verdades absolutas”, exercitamos a capacidade crítica e o dever cidadão de entender as tecnologias como mecanismos que só se realizam por meio do nosso uso e das nossas decisões. E, portanto,

são capazes de produzir efeitos positivos e negativos na sociedade.

Enquanto você, gestor público, preenche sua planilha, escolhe suas fontes de dados, escreve linhas de código em seu algoritmo e analisa seus gráficos, há outra combinação de elementos atuando de maneira simultânea e indissociável.

Dispomos de grande quantidade e variedade de dados, aliados a tecnologias com poder crescente de processamento. Essas tecnologias parecem, a cada ano, mais capazes de lidar com quantidades de informações cada vez maiores. Isso, por si só, deveria estar produzindo as melhores decisões em todos os âmbitos.

O que determina nossas escolhas e decisões sobre o que faremos com o conhecimento que o método científico pode nos conferir não pode ser calculado, muito menos determinado, matematicamente ou racionalmente.

Definições sobre certo ou errado, bom ou ruim são construções culturais, sociais e históricas, que dizem respeito às dimensões morais e éticas da nossa existência. Não há quantidade de cálculo que possa ser feito para que haja uma resposta definitiva sobre questões dessa natureza.

Ainda que o resultado de uma operação seja matemática e metodologicamente perfeito, a forma como as pessoas interpretarão esses números e como suas decisões imediatas serão ou não influenciadas não dependem apenas da exatidão estatística.

Precisamos entender que vivemos num mundo complexo, de relações complexas, desde as partículas mais básicas da física quântica às estruturas mais massivas do universo.

A notícia positiva é que entender as limitações das nossas técnicas e ferramentas pode, na verdade, ampliar nossa capacidade de utilizá-las para resolver problemas e responder perguntas. E é justamente aí que as dimensões técnicas e o pensamento crítico se encontram.

Em outras palavras: estamos falando da combinação entre os elementos importantes para operacionalizarmos nossos modelos mentais, da escolha das fontes de dados e das técnicas que utilizaremos enquanto, simultaneamente, atuamos com um olhar crítico “sempre ligado”. É esse olhar crítico que nos ajudará a entender a complexidade, as nuances, os contextos histórico e social, bem como a natureza do problema que precisamos resolver.

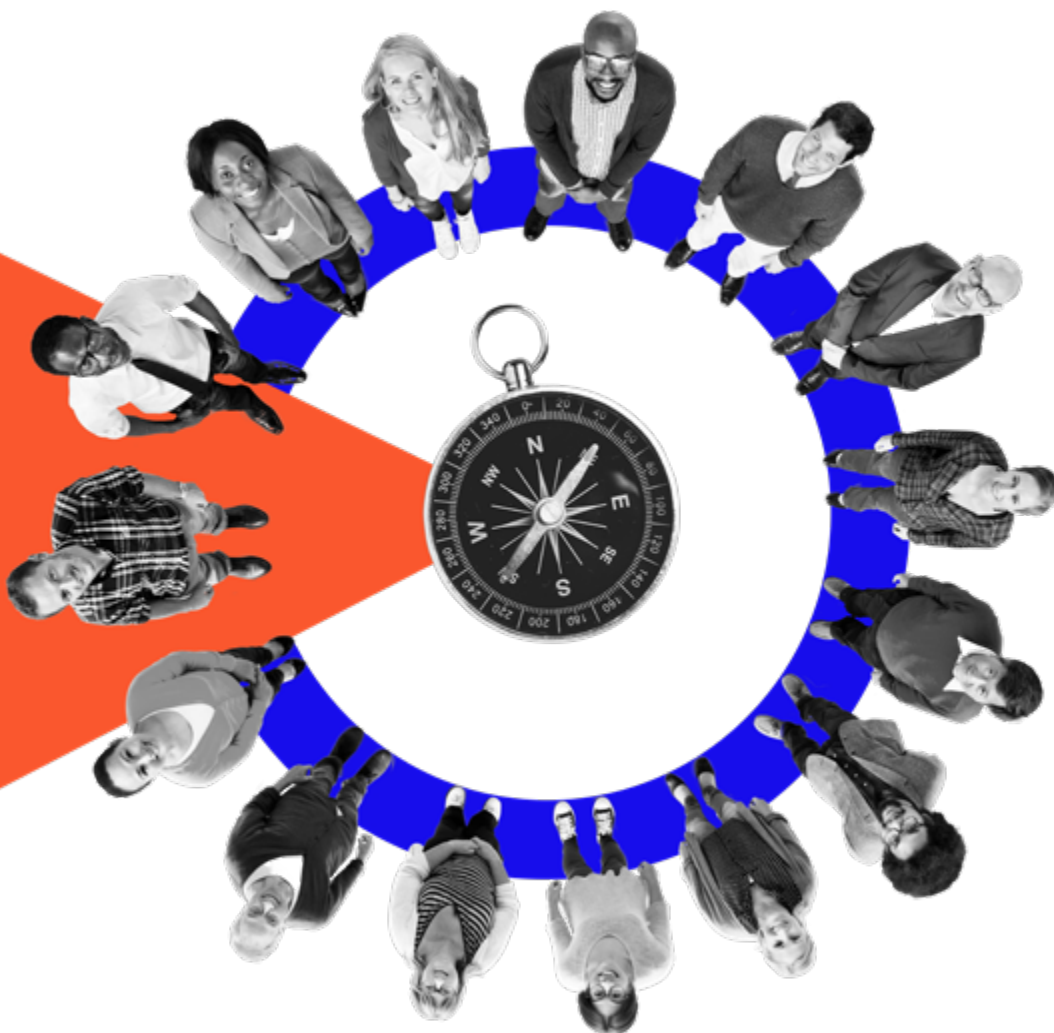
Na escolha de modelos, importa estarmos cientes de suas vantagens e desvantagens, interpretar os resultados, compreender e analisar criticamente as possíveis relações encontradas entre variáveis e, especialmente, traduzir esses aprendizados de volta para o mundo real.

É a capacidade de análise crítica, interpretação e tradução que permite transformar os dados em melhores decisões. A matemática e as ferramentas são apenas parte dessa jornada. O aspecto mais importante da tecnologia é o humano.

#3

ALFABETIZAÇÃO EM DADOS

por Carol Andrade



Alfabetização em dados no mundo

Neste capítulo, entenderemos o que é alfabetização em dados e por que ela é uma das competências fundamentais do gestor público do futuro.

Para ilustrar a magnitude da alfabetização em dados, apresentamos algumas pesquisas e publicações recentes sobre esse tema.

A PricewaterhouseCoopers Brasil (PwC Brasil)⁶ estimou que *analytics* e inteligência artificial contribuirão com cerca de 15,7 trilhões de dólares na economia global até 2030, o que pode ser traduzido em um impacto de até 26% no Produto Interno Bruto (PIB) da China e 14,5% no dos Estados Unidos.

É difícil encontrar na história recente um exemplo similar ocorrendo em um intervalo de tempo tão curto.

Diferente de outras fases da inovação, em que substituímos ferramentas tradicionais por versões mais modernas e eficientes, na Era dos Dados tem ficado cada vez mais claro que o desafio

demandará mudanças organizacionais, estruturais, culturais e no paradigma de liderança.

Em todos os segmentos, nas esferas pública e privada, apenas as organizações capazes de elevar seus níveis de alfabetização em dados, e incorporar essa nova cultura, sobreviverão.

Criar condições para fazer crescer uma visão mais integrada e aumentar a capacidade analítica da organização como um todo a torna mais apta a entender seu ecossistema e as necessidades de seus usuários.

Estudo publicado na edição de agosto de 2019 da *Harvard Business Review*⁷ revela que, até aquela data, apenas 8% das companhias estavam obtendo sucesso na implementação dessas mudanças em larga escala.

6 Confira o estudo completo em: <<https://www.pwc.com/gx/en/issues/data-and-analytics/publications/artificial-intelligence-study.html>>.

7 Confira a pesquisa completa em: <<https://hbr.org/2019/07/building-the-ai-powered-organization>>.

Conceito, pesquisas e competências

A tradução de *data literacy* é “alfabetização em dados”. Outra tradução muito utilizada no Brasil é “fluência em dados”, que, na Era dos Dados, **é como um segundo idioma que todos nós devemos aprender e ser fluentes.**

Em 2015, os pesquisadores do Instituto de Tecnologia de Massachusetts (MIT)⁸ Catherine D’Ignazio e Rahul Bhargava descreveram em seu artigo “Abordagens para desenvolver a alfabetização em *big data*”⁹ que a alfabetização em dados é a capacidade de ler, trabalhar, analisar e se comunicar com dados.

A partir desse estudo, o conceito ganhou popularidade internacional e criou-se um consenso acerca de quatro *competências essenciais*:

LER DADOS

TRABALHAR COM DADOS

ANALISAR DADOS

ARGUMENTAR COM DADOS

Na sequência, Jordan Morrow, chefe de alfabetização em dados da empresa Qlik¹⁰, ganhou destaque com o tema “Por que todos deveriam ser fluentes em dados”¹¹, no TEDxBoise¹². **Ele enfatizou que a alfabetização em dados não é apenas para especialistas ou cientistas de dados, mas para todas as pessoas.**

A Qlik capacita funcionários de todos os níveis para torná-los capazes de fazer as “perguntas certas” sobre dados, desenvolver conhecimentos, tomar decisões e comunicar o significado a outras pessoas.

Em 2018, essa mesma empresa lançou a pesquisa “Liderar com dados: como impulsionar a alfabetização em dados na empresa”¹³, em que foram consultados mais de 7 mil tomadores de decisão em distintas organizações, nos cargos de gerência e supervisão, na Europa, na Ásia e nos Estados Unidos.

A pesquisa evidencia que o nível de alfabetização em dados nas empresas ainda é baixo, porém os próprios colaboradores consideram essa competência necessária para melhorar sua performance e credibilidade profissional, estando dispostos a investir tempo e energia no desenvolvimento de tais habilidades.

8 O MIT é uma universidade privada de pesquisa referência em tecnologia e inovação localizada em Cambridge, Massachusetts, Estados Unidos. <<https://www.mit.edu/>>.

9 Consulte o artigo completo em: <https://dam-prod.media.mit.edu/x/2016/10/20/Edu_D%27Ignazio_52.pdf>.

10 Qlik é uma empresa que oferece produtos de análise e integração de dados e serviços complementares com ênfase em treinamentos de *data literacy*. Conheça mais em: <<https://www.qlik.com/pt-br/bi/data-literacy>>.

11 Assista à palestra em: <https://www.youtube.com/watch?v=8ovyQZ_Z8Xs>.

12 Conheça o TEDx Boise em: <<http://tedxboise.org/>>.

13 Confira a pesquisa completa em: <https://www.qlik.com/us/bi/-/media/08F37D711A58406E83BA8418EB-1D58C9.ashx?ga-link=datlitreport_resource-library>.

Vejamos aqui alguns dados que justificam as descobertas da pesquisa:

» Apenas **24%** dos tomadores de decisão pesquisados estão totalmente confiantes em sua capacidade de ler, trabalhar, analisar e discutir os dados.

» Como apenas **32%** dos executivos de alto nível são vistos como alfabetizados em dados, a força de trabalho das empresas, ou seja, os líderes seniores, não são encorajados a usar dados.

» **94%** dos entrevistados que usam dados em sua função concordam que os dados os ajudam a fazer um trabalho melhor. Eles também acreditam que uma maior alfabetização em dados lhes daria mais credibilidade (**82%**).

» **78%** dos tomadores de decisão disseram que estariam dispostos a investir mais tempo e energia para melhorar suas habilidades analíticas.

A pesquisa indica ainda caminhos para uma mudança de cultura organizacional, por meio do desenvolvimento e do apoio às lideranças e equipes, uma vez que a baixa alfabetização em dados impede o crescimento de iniciativas de transformação digital, assim como leva à deficiência generalizada na confiança dos dados.

Sem a alfabetização em dados, os líderes não conseguem prosperar na atual economia analítica ou impulsionar qualquer tipo de transformação organizacional ou social.

Outra publicação de relevância internacional é a *Global Data Literacy Benchmark*¹⁴, lançada em 2020 pela iniciativa *Data to the people*¹⁵. Trata-se da primeira medição abrangente de alfabetização em dados, realizada com mais de 5 mil funcionários na Austrália, nos Estados Unidos, na Índia, no Reino Unido e no Canadá.

O diferencial dessa publicação, que é mais do que uma pesquisa, é ser um modelo e uma ferramenta de avaliação de competências de alfabetização em dados, a Databilities®.

A estrutura descreve *15 competências essenciais* nas dimensões de leitura, escrita e compreensão, sendo que esta última equivale às habilidades de análise e argumentação. Veja a Figura 4 na página a seguir.

14 A tradução literal da publicação é "Comparativo de mercado global de alfabetização de dados". Acesse para conhecê-la: <https://bd2aff5a-3466-4f31-9830-5ad0042d9703.filesusr.com/ugd/1ff4ae_3e3b4a50ead44939b50713c3c9af4b66.pdf>.

15 *Data to the People* é uma iniciativa que reúne reconhecidos especialistas globais e líderes da indústria na construção e promoção da alfabetização em dados. Conheça mais em: <<https://www.datatothepeople.org/>>.



LEITURA

- 1 Descoberta de dados
- 2 Avaliação e garantia da qualidade dos dados



ESCRITA

- 3 Coleção de dados
- 4 Gestão e organização de dados
- 5 Manipulação de dados
- 6 Curadoria e reutilização de dados
- 7 Criação e uso de metadados
- 8 Conversão de dados (formato)



COMPREENSÃO

- 9 Análise de dados
- 10 Interpretação de dados (compreensão)
- 11 Identificação de problemas usando dados
- 12 Visualização de dados
- 13 Apresentação de dados (verbalmente)
- 14 Tomada de decisão baseada em dados
- 15 Avaliação de decisões/conclusões com base em dados

Figura 4. Databilities®: 15 competências essenciais de alfabetização em dados



Se você quiser receber uma avaliação gratuita das suas habilidades de alfabetização em dados nessas competências e níveis, pode acessar o [questionário da Databilities®](#).

No Brasil, o Social Good Brasil lançou em português uma [autoavaliação](#) que, além de medir o grau de alfabetização em dados, também avalia o uso ético e a promoção do bem-estar social com dados.

Ainda em 2020, foi lançada a pesquisa “O impacto humano da alfabetização em dados”¹⁶, realizada pelo *Data Literacy Project*¹⁷ e envolvendo 9 mil funcionários de setores e cargos de alto nível a nível básico, em países da América do Norte, da Europa e da Ásia-Pacífico.

O estudo aponta que a democratização dos dados dentro das organizações é fundamental para acelerar a mudança de cultura. O desafio é que muitos funcionários ainda não possuem as habilidades necessárias para trabalhar com ferramentas de dados de forma confortável e confiante.

Em relação à confiança, apenas **21%** da força de trabalho global está totalmente confiante sobre suas habilidades de alfabetização em dados.

16 Conheça a pesquisa completa “The human impact of data literacy” em: <https://www.accenture.com/_acn-media/PDF-118/Accenture-The-Human-Impact-Data-Literacy.pdf#zoom=50>.

17 O *Data Literacy Project* é uma iniciativa que une organizações e líderes acadêmicos para discutir e desenvolver as ferramentas necessárias para uma sociedade alfabetizada em dados. Conheça mais em: <<https://thedataliteracyproject.org/>>.

Um efeito negativo disso é que apenas **32%** dos executivos entrevistados disseram ser capazes de criar valor mensurável a partir dos dados, enquanto **27%** disseram que dados e projetos analíticos produzem descobertas que levam a ações concretas.

Outro efeito também negativo é que quando as organizações não apoiam ativamente seus funcionários para que eles adotem novas tecnologias e práticas de trabalho isso acaba por afetar seu bem-estar.

O estudo aponta que 6 em cada 10 entrevistados (**61%**) relataram se sentir sobrecarregados pelos dados e mais estressados no local de trabalho. Isso ocorre quando as organizações não apoiam ativamente seus funcionários na adoção de novas tecnologias e práticas de trabalho.

Como pudemos ver, o Brasil ainda ficou fora das pesquisas apresentadas. Elas se concentraram no setor privado e na língua inglesa.

Alfabetização em dados na gestão pública

Em 2017, a Organização para a Cooperação e Desenvolvimento Econômico

(OCDE)¹⁸ lançou a publicação “Habilidades essenciais para inovação no setor público”¹⁹, abordando a necessidade do desenvolvimento de seis competências essenciais para inovar na gestão pública. Uma dessas competências é a alfabetização em dados.

A publicação defende que aumentar o número de cientistas ou especialistas em dados no governo não garante o sucesso na condução para uma gestão orientada a dados, devendo-se, portanto, estimular o seguinte: **que todos os agentes públicos sejam alfabetizados em dados e valorizem a importância do uso de dados.**

Também é imprescindível que a análise de dados não seja um “pensamento posterior”. A alfabetização em dados se refere à tomada de decisão pautada em dados e evidências, e não em “achismos”, e os gestores devem entender seu caráter estratégico.

O estudo aponta ainda que um gestor público quando alfabetizado em dados:

- » usa dados, em vez de suposições, para tomar decisões;
- » gerencia serviços públicos com base em dados;
- » trabalha com especialistas em dados;
- » comunica-se efetivamente com não especialistas sobre dados, análises e resultados.

18 OCDE é uma organização econômica intergovernamental com 38 países membros, fundada em 1961 para estimular o progresso econômico e o comércio mundial: <<https://www.oecd.org/latin-america/countries/brazil/brasil.htm>>.

19 Confira o estudo “Core skills for public sector innovation”: <https://www.oecd.org/media/oecdorg/satellite-sites/opsi/contents/files/OECD_OPSI-core_skills_for_public_sector_innovation-201704.pdf>.

Os resultados gerados quando gestores públicos são fluentes em dados são visíveis em três esferas:

- 1** Soluções de dados criadas para melhorar a eficiência da gestão, otimizando e economizando recursos.
- 2** Políticas públicas baseadas em evidências, reconhecendo-se e aumentando o impacto positivo de iniciativas públicas que deram certo.
- 3** Transparência dos dados públicos, permitindo uma cidadania plena a partir do acesso à informação e a um real engajamento da população e de outros setores nas questões públicas.

Alfabetização em dados para a população e os impactos sociais

Alfabetizar as pessoas em dados, preparando-as para navegar nesse atual oceano de informações, é uma necessidade social. Isso permite que elas estejam aptas a mais oportunidades de trabalho e de negócios, em um mundo cada vez mais digital.

Além de inclusão produtiva e melhoria da renda, a alfabetização em dados dá condições para que as pessoas tomem melhores decisões nos campos da saúde, das finanças, da política e em outros aspectos de sua vida pessoal.

Nesse sentido, é fundamental refletirmos sobre:

(a) **a necessidade de a Ciência de Dados posicionar o ser humano no centro de seus processos e, assim, prosperar nas organizações públicas.** Além do desenvolvimento tecnológico, é fundamental considerar o desenvolvimento humano caminhando junto da implementação de novas tecnologias. Caso contrário, pode-se investir muito recurso em uma solução engavetada e subutilizada.

(b) **os efeitos negativos das tecnologias no mundo atual.** É preciso ir além da alfabetização em dados, democratizando uma educação em dados que inclua reflexões acerca de uma visão crítica dessa nova Era dos Dados e nosso papel ético na sociedade.

No próximo tópico, vamos abordar quatro desses efeitos.

Efeitos negativos das tecnologias no mundo atual

O **primeiro efeito** podemos chamar de ditadura dos dados. Segundo Yuval Harari, “Dados são uma das mais importantes questões políticas no mundo, pois estão se tornando o capital mais valioso”, e hoje a maioria dos dados está sob posse de poucas organizações.

O **segundo** é a “pandemia das *fakes news*”. Ainda que a desinformação sempre tenha existido, a novidade agora é o ritmo que ela ganhou. Pessoas deixam de lado a capacidade de pensar de forma crítica e, por conta dos algoritmos, são induzidas a uma opinião restrita a sua pequena bolha.

O **terceiro efeito** se relaciona aos ciberataques, acelerando crimes digitais e vazamentos de dados pessoais sensíveis. O Brasil é o segundo país do mundo que mais sofre ciberataques, segundo o *Mapa em tempo real de ameaças cibernéticas*²⁰ criado pela empresa de segurança virtual Kaspersky.

O **quarto efeito** negativo é a possibilidade de violação e a utilização imprópria de dados pessoais. No Brasil, por exemplo, não há geração de dados confiáveis sobre determinados locais e populações mais vulneráveis, isto é, os dados são retratados com viés discriminatório ou sequer são retratados. Como consequência, localidades e populações podem ser excluídas de políticas públicas ou receber menos investimentos sociais.

Os desenvolvimentos humano e tecnológico devem caminhar juntos.

Uso de dados para o bem comum

Globalmente, há uma série de iniciativas de uso de dados para o bem comum, um campo de inovação conhecido como *Data for Good*, ou, em português, *Dados para o Bem*, do qual o Social Good Brasil faz parte.

Jake Porway, pesquisador da *Data.org*²¹, cofundador e diretor executivo da *DataKind*²², está mapeando iniciativas ao redor do mundo nesse campo. Uma versão preliminar desse mapeamento já foi publicada, está no estudo “Traçando o cenário Dados para o Bem”²³.

Destacamos a seguir quatro categorias de iniciativas do campo *Data for Good*.

COLETAS E NARRATIVAS DE DADOS SOBRE GÊNERO E RAÇA

Como exemplo no mundo, citamos o *Data2x*²⁴. É uma iniciativa da Organização das Nações Unidas (ONU) para melhorar a disponibilidade, a qualidade e o uso de dados de gênero.

Já no Brasil, a empresa social *Gênero e Número*²⁵ produz e distribui jornalismo orientado por dados e análises sobre questões urgentes de gênero e raça, vi-

20 Acesse o Cyber Threat Real Time Map em: <<https://cybermap.kaspersky.com/>>.

21 Conheça a iniciativa global: <<https://www.data.org/>>.

22 Conheça a organização: <<https://www.datakind.org/>>.

23 Confira o artigo na íntegra: <<https://www.data.org/charting-the-data-for-good-landscape/>>.

24 Conheça mais o Data2x em: <<https://data2x.org/>>.

25 Conheça mais a Gênero e Número em: <<https://www.generonumero.media/>>.

sando qualificar debates rumo à equidade. A partir de linguagem gráfica, conteúdo audiovisual, pesquisas, relatórios e reportagens multimídia alcançam e informam uma audiência interessada no assunto. Outra iniciativa nacional interessante é o Data Labe²⁶, um laboratório de dados e narrativas na favela da Maré, no Rio de Janeiro.

DADOS ABERTOS, O PODER DA TRANSPARÊNCIA PARA O EXERCÍCIO DA CIDADANIA

No mundo, há várias iniciativas que oferecem *datasets* (em português, conjuntos de dados) para que qualquer cidadão possa consultá-los. O objetivo é promover a transparência e cada vez mais pesquisas e projetos baseados em dados confiáveis.

A plataforma Dados Abertos do Banco Mundial²⁷ é uma dessas iniciativas. Ela disponibiliza uma coleção de dados referentes a desenvolvimento sustentável, indicadores financeiros, de desenvolvimento, estatísticas das dívidas externas dos países, entre outros dados no plano global.

No Brasil, o Open Data Index²⁸, da Open Knowledge Brasil, acompanha como o poder público do país está atuando de acordo com princípios de transparência. Além disso, ajuda a avaliar políticas, identifica gargalos e orienta os muni-

cípios a aprimorarem suas políticas de dados abertos.

APOIO À CRIAÇÃO DE SOLUÇÕES DE DADOS

O objetivo dessa iniciativa é aproveitar o melhor da inteligência artificial e de dados para resolver problemas sociais complexos e, dessa forma, impactar a vida de milhares de pessoas e conectar diversas instituições.

No mundo, o DataKind²⁹ se propõe a colocar a Ciência de Dados a serviço da humanidade, disseminando o campo Data for Good.

No Brasil, as iniciativas do Laboratório de Dados do Social Good Brasil e do ÍRIS | Laboratório de Inovação e Dados do Governo do Ceará são pioneiras e se destacam nessa jornada.

EDUCAÇÃO E TREINAMENTO

Já são muitas as iniciativas globais de educação em dados voltadas para cientistas de dados, novos talentos e todo e qualquer cidadão.

Destacamos a OCDE Learning Compass 2030³⁰, uma estrutura de aprendizagem em evolução que fornece pontos orientadores, como bem-estar individual e coletivo.

26 Conheça mais o Data Labe em: <<https://datalabe.org/>>.

27 Conheça mais o World Bank Open Data em: <<https://data.worldbank.org/>>.

28 Conheça mais o Open Data Index em: <<https://ok.org.br/projetos/open-data-index/>>.

29 Conheça mais o DataKind em: <<https://www.datakind.org/>>.

30 Conheça mais sobre o OCDE Learning Compass 2030 em: <<https://www.oecd.org/education/2030-project/teaching-and-learning/learning/learning-compass-2030/>>.

Outra instituição referência global é a Data Pop Alliance³¹, que tem como objetivo principal escalar a alfabetização na Era dos Dados para ajudar indivíduos e instituições a se tornarem agentes de mudança social positiva. Isso se dá por meio do uso de dados de maneira a refletir os princípios éticos fundamentais de equidade e empoderamento.

No contexto nacional o Social Good Brasil encabeça programas de educação que incorporam a alfabetização em dados como forma de democratizar o conhecimento. Uma dessas iniciativas — on-line e gratuita — é a *Habilidades*³², primeira série sobre alfabetização em dados no Brasil. Cada episódio aborda uma ou mais competências com linguagem acessível e didática.



31 Conheça a Data Pop Alliance: <<https://datapopalliance.org/mobilize/>>.

32 Acompanhe a série Habilidades: <<https://www.youtube.com/user/SocialGoodBrasil/>>.

Estar preparado para a Era dos Dados não significa apenas ter um uma vasta base de dados ou dispor de um robô que atenda seus usuários de maneira digital e automatizada. Estamos falando de um outro código-fonte, de um sistema operacional distinto. Será necessário redesenhar estruturas e processos decisórios, de forma que todas as decisões da organização estejam apoiadas, em menor ou maior grau, por dados.

#4

A JORNADA DE DADOS

por Carol Andrade



Como fazer avançar uma gestão pública orientada a dados

Em qualquer instituição, o desafio para avançar na Era dos Dados não é a tecnologia, mas os modelos mentais, a cultura e os aspectos humanos. Por isso, para que haja a implementação de novas tecnologias e uma efetiva mudança de cultura, é fundamental que o ser humano esteja no centro dos processos e seja garantido um diálogo aberto e o sentimento de propósito.

A dificuldade de integração e o compartilhamento de bancos de dados entre áreas distintas é um desafio recorrente para uma análise de dados mais sofisticada e com potencial de diagnóstico de valor para a gestão de serviços e políticas públicas.

Segundo artigo de Anthony Spadafora³³, mais da metade (**55%**) das organizações concordam que a fragmentação de dados em distintos bancos de dados traz lentidão ao processo de maturidade analítica da organização.

Outro desafio comum é a crença de que trabalhar com dados é apenas para pessoas da área de Exatas e/ou cientistas

de dados. Como consequência, não há capacitação e engajamento das lideranças e demais pessoas que deveriam tomar decisões com base em evidências.

Também é um desafio para as instituições não descentralizar a informação e o conhecimento dos dados, principalmente em modelos hierárquicos mais rígidos e burocráticos.

Como sociedade, avançamos em tecnologias e inovações, mas nem tanto em modelos mentais. Predominam os modelos tradicionais de escassez e competição, em vez de abundância e colaboração, assim como o apego a *status* e cargos.

Esses são os reais problemas, assim como a resistência das pessoas em adotar novidades por resistirem à mudança. Por isso, como parte do processo, é fundamental ouvir melhor as pessoas envolvidas, entender seus receios e captar suas motivações.

33 Leia o artigo completo em: <<https://www.itproportal.com/news/many-companies-dont-know-where-their-critical-data-is-kept/>>.

Toda mudança e inovação exige sair da “zona de conforto”, daí ser fundamental o envolvimento das lideranças. Líderes que compreendem que a Era dos Dados apenas começou — e é o nosso futuro — são capazes de compartilhar uma visão inspiradora e gerar um sentimento de engajamento rumo a uma nova cultura analítica.

Ferramentas para apoiar decisões baseadas em dados

O gestor público fluente em dados é capaz de definir estratégias, pautar suas decisões com base em evidências, e não apenas em percepções pessoais e vieses³⁴.

A seguir, apresentamos duas ferramentas para apoiar gestores na tomada de decisão: o *Canvas de Dados* e o *Fluxo de Dados*, ambas criadas pelo Social Good Brasil.

O propósito dessas ferramentas é orientar a gestão a navegar no pensamento analítico, a partir de objetivos e problemas bem definidos, para pautar projetos e decisões e dar os primeiros passos para criar uma solução de dados.

Canvas de Dados

Para apresentar o Canvas de Dados, é necessário primeiro entendermos a diferença entre **pensamento convencional** e **pensamento analítico**.

Na lógica convencional, de acordo com a Figura 5, na página a seguir, começamos o raciocínio pela grande quantidade de dados disponíveis e nos perguntamos: “O que eu consigo extrair disso?”. Nos passos seguintes, identificamos e testamos hipóteses. Essa lógica convencional tem três principais problemas:

- 1** Muitas vezes, não chegamos a bons resultados por não termos um objetivo claro em relação ao projeto. Instituições contratam cientistas de dados, mas não têm êxito em suas jornadas. E por que isso acontece? De acordo com estudo da Gartner³⁵, **85%** dos projetos de dados falham, e uma das principais razões para isso é a falta de um objetivo claro.
- 2** Ficamos imersos em grandes quantidades de dados e não avançamos ou avançamos lentamente.
- 3** Quando finalmente chegamos a uma conclusão, não sabemos se ela foi a mais efetiva entre as opções, porque não havia, desde o início, um objetivo claro em relação ao final do projeto.

34 Vieses cognitivos ou inconscientes são associações automáticas, suposições, julgamentos e atitudes em relação a outras pessoas, situações e contextos.

35 Leia artigo que cita o estudo da Gartner: <<https://towardsdatascience.com/why-data-science-projects-fail-revisited-85fe242c3931>>.

Já pela lógica analítica, apresentada na Figura 6, iniciamos e partimos do propósito, ou seja, do objetivo ou do problema que pretendemos resolver. Esse objetivo será o foco em um universo de dados que ultrapassa nossa capacidade humana de consumo.

Os dados são identificados apenas na terceira etapa do processo. Só depois de termos um objetivo claro e as perguntas para tomada de decisão, é que definimos os dados relevantes.

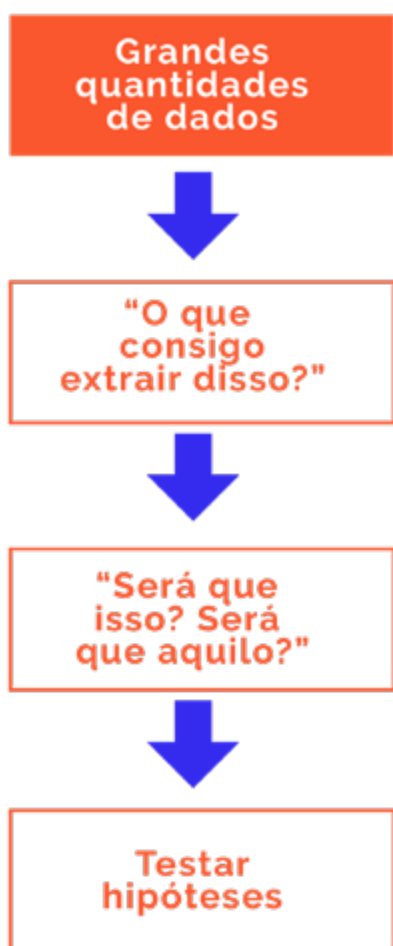


Figura 5. Lógica convencional

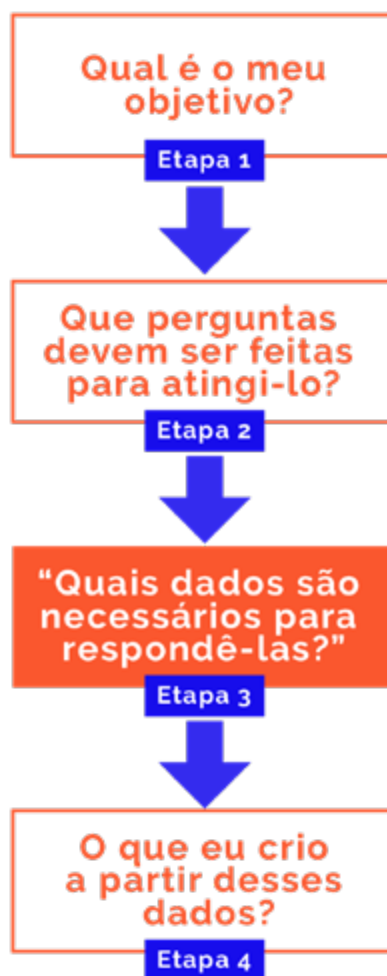


Figura 6. Lógica analítica



CANVAS DE DADOS

ORGANIZAÇÃO: _____

OBJETIVOS / RESULTADOS Os principais objetivos ou resultados de curto ou médio prazo que a organização quer alcançar. Normalmente encontrados em ferramentas estratégicas como Planejamentos ou Teoria da Mudança. Qual o principal objetivo ou resultado de curto ou médio prazo que sua organização quer alcançar? Pode ser sua meta de 2021 ou a de 2022, por exemplo.		
PERGUNTAS ESTRATÉGICAS As perguntas/decisões estratégicas que a organização deveria fazer para entender e qualificar melhor os objetivos ou resultados. Classificar as perguntas de acordo com o tipo de análise esperada (descritiva, diagnóstica, preditiva ou prescritiva).	DADOS NECESSÁRIOS Dados necessários para responder as perguntas estratégicas. Classificar de acordo com origem (primário ou secundário) e frequência de coleta.	PRODUTOS DE DADOS Possíveis produtos de dados que ajudam a responder as perguntas estratégicas a partir dos dados necessários. Podem ser relatórios, estudos, painéis de controle etc. Classificar de acordo com o tipo de análise esperada (descritiva, diagnóstica, preditiva ou prescritiva).

Esta obra foi desenvolvida pelo Social Good Brasil sob licença creative commons que permite sua replicação atribuindo a fonte e sem fins comerciais.

Figura 7. Ferramenta Canvas de Dados Social Good Brasil

Portanto, o Canvas de Dados (Figura 7) é tanto uma ferramenta estratégica como prática, que pode ser utilizada diariamente pelo gestor público. Ela o orienta a navegar pelo pensamento analítico e estratégico.

A seguir, explicamos como utilizar essa ferramenta, passo a passo.

ETAPA 1 OBJETIVOS/RESULTADOS

Defina e alinhe objetivos ou resultados esperados

Além da importância da lógica analítica, que considera a clareza de objetivos para iniciar um projeto de dados, é fundamental o alinhamento entre lideranças, coordenadores, gestores e técnicos.

Garantir que toda a equipe envolvida compreenda o que esperamos do projeto diminui as chances de desgastes e retrabalhos. Fazer bem essa primeira etapa garante o sucesso das seguintes e traz qualidade ao resultado final. Além disso, evita potenciais desentendimentos e lentidão ao longo do caminho, ou mesmo o fracasso do projeto.

Os objetivos podem ser de níveis estratégicos ou táticos, de impacto de médio ou curto prazo. Também podem ser definidos a partir de problemas a resolver, como as questões internas do dia a dia da instituição ou questões relativas à sociedade.

Qualquer que seja a natureza do objetivo ele deve ser minimamente SMART. Esse método é usado para construir objetivos em planejamentos estratégicos.



Para alcançar um objetivo SMART³⁶, você deve considerar os fatores acima.

Como exemplo de um objetivo SMART citamos o caso da Portabilis³⁷, empresa social que atua com a gestão pública municipal através do uso de tecnologia e dados nas áreas de educação e assistência social. Ela definiu um objetivo SMART a partir de um Canvas de Dados da seguinte maneira: reduzir em 50% as taxas de abandono escolar em todas as escolas municipais da Grande Florianópolis até 2030.

ETAPA 2 PERGUNTAS ESTRATÉGICAS

Refleta e defina perguntas para tomar decisões estratégicas

Na segunda etapa do Canvas de Dados, vamos refletir e definir perguntas que a organização deve fazer para entender e qualificar objetivos e/ou resultados.

No caso da Portabilis, por exemplo, as perguntas foram:

- » Qual o histórico de evasão escolar no município?
- » Em qual período do ano letivo mais ocorre abandono?
- » Em quais localidades do município o abandono é mais predominante?
- » Qual é a faixa de renda média das famílias dos alunos que evadem?
- » Quais outros indicadores não são usualmente relacionados à evasão escolar?
- » É possível identificar uma relação entre faixa de renda familiar e outros indicadores não usualmente relacionados com a evasão escolar?
- » A partir do histórico, qual perfil de aluno tem maior tendência à evasão escolar?

Atenção! A qualidade das perguntas influencia a qualidade dos produtos gerados ao final do Canvas.

36 SMART é a sigla de um método de definição de metas baseada em 5 fatores: S (specific/específico), M (measurable/mensurável), A (achievable/alcançável), R (relevant/relevante) e T (time-bound/temporal).

37 Conheça a Portabilis: <<https://portabilis.com.br/>>.

Nessa segunda etapa, faz diferença conhecer a complexidade de análises dos dados. A Figura 8 apresenta os quatro níveis de complexidade de análise de dados. A análise é uma das competências de alfabetização em dados.

Se aproveitarmos esse conhecimento para elaborar perguntas do Canvas de Dados, elas poderiam ser formatadas assim:

ANÁLISE DESCRITIVA

O que aconteceu?

Quais perguntas levariam você a entender melhor o contexto atual em relação ao seu objetivo?

ANÁLISE DIAGNÓSTICA

Por que aconteceu?

Quais perguntas levariam você a entender a razão pela qual as coisas aconteceram de determinada maneira em relação ao seu objetivo?

ANÁLISE PREDITIVA

O que vai acontecer?

Quais perguntas ajudariam você a entender o que pode acontecer daqui para frente em relação ao seu objetivo?

ANÁLISE PRESCRITIVA

Que ação tomar?

Quais perguntas ajudariam você a decidir como reagir ou responder aos acontecimentos futuros em relação ao seu objetivo?

Às vezes, um time que deseja se destacar como mais inovador, ao saber de potenciais análises mais avançadas, pode perder o foco do seu objetivo, gerando um produto mais inovador em termos de tecnologia, porém menos usado na prática para a tomada de decisão.

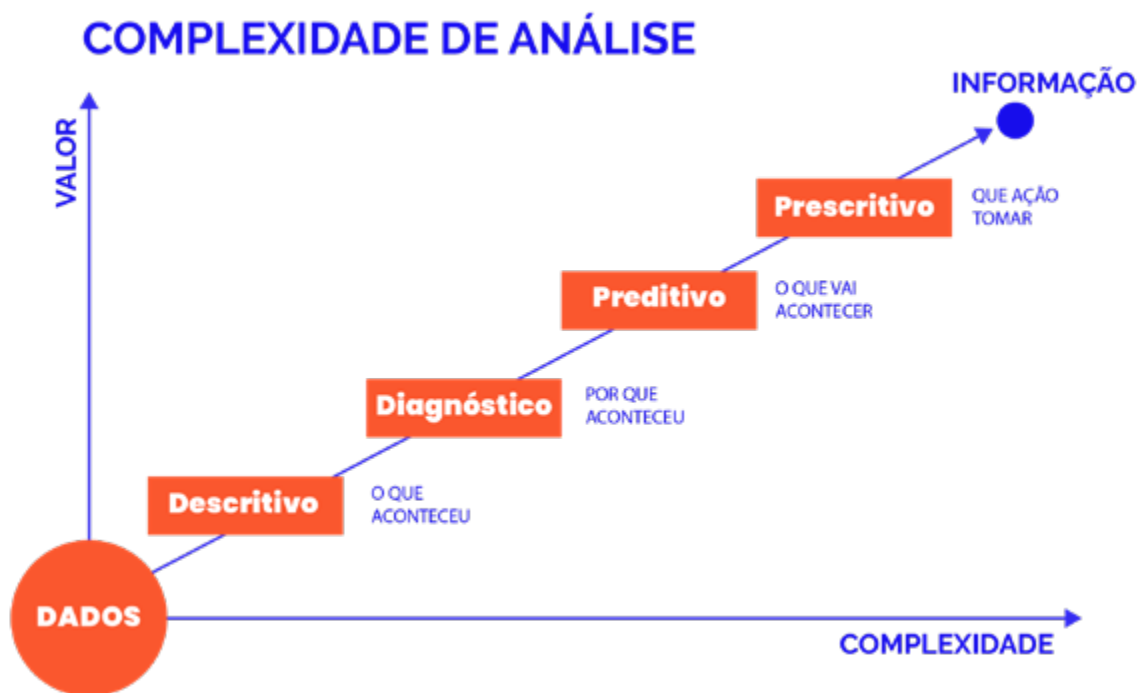


Figura 8. Níveis de complexidade de análise de dados

ETAPA 3 DADOS NECESSÁRIOS

Enumere dados e informações necessárias para cada pergunta

Quando as perguntas estiverem bem definidas, levantaremos as informações e dados necessários para respondê-las.

Os dados podem ser evidências e informações de todos os tipos, quantitativos e qualitativos, primários ou secundários, dados que sua instituição possui internamente ou até que precise coletar. Não se restrinja no momento. Inicie listando tudo o que for necessário.

No exemplo da Portabilis, eles enumeraram como dados necessários:

- » Dados de alunos e suas matrículas no sistema de gestão escolar.
- » Dados das famílias no sistema de gestão de assistência social.

ETAPA 4 PRODUTOS DE DADOS

Desenhe soluções ou produtos de dados que reúnam tais evidências

Na última etapa do Canvas de Dados, desenhamos os produtos de dados, que também chamamos de **soluções**.

Esses produtos irão conter os dados enumerados anteriormente e necessários para a tomada de decisão. Podem ser relatórios, estudos e análises descritivas e diagnósticas, painéis de controle com atualização em tempo real (*dashboards*), indicadores-chave de desempenho (KPIs), entre outros.

Um mesmo produto de dados pode conter dados primários e secundários e cruzar as informações para análises de mais de uma fonte de dados. Ao final, dependendo do objetivo, pode ser necessário desenhar mais de um produto de dados.

Nesta última etapa, além de clareza de objetivos, perguntas e dados necessários, é fundamental também, para a tomada de decisão, ter clareza do público-alvo que irá consumir o produto.

O perfil desse público é necessário para decidir o tipo de visualização e as narrativas de dados que serão adotadas. Como já mencionamos, importa construir uma solução de dados que seja de fato usada para a tomada de decisão, e não uma solução com tecnologia sofisticada, mas que pode gerar uma barreira.

A única regra aqui é considerar o público-alvo com muita empatia. Antes de realizar o projeto e investir recursos no desenvolvimento, é importante prototipar e validar a solução desenvolvida.

Vale enfatizar a importância de desenhar um plano de implementação do produto de dados que engaje todos os envolvidos, desde o time de execução até o público-alvo.

No caso da Portabilis, eles desenharam e criaram dois produtos de dados:

- » Índice Municipal de Evasão Escolar (IEV): produto com análise preditiva, que prediz, na matrícula do aluno e de acordo com seu perfil, sua tendência à evasão escolar.

» Evasão Zero: painel de monitoramento intersetorial de frequência, abandono e evasão escolar.

Para garantir a adoção de tais produtos pela gestão escolar, o time da Portabilis se envolveu diretamente com as equipes das escolas municipais, apresentando os produtos e seus benefícios.

Fluxo de Dados

O Fluxo de Dados é outra ferramenta criada pelo Social Good Brasil para apoiar uma instituição na tarefa de desenhar o processo da jornada dos dados no dia a dia.

Compreende etapas da jornada dos dados, desde sua coleta até a sua visualização, como mostra a Figura 9.

Em cada etapa, devem ser descritos: a) como funciona o produto de dados; b) quais as implicações da Lei Geral de Proteção de Dados (LGPD); c) quais ferramentas podem ser usadas.

ETAPA 1 COLETA

A primeira etapa do fluxo permite definir a estratégia, as ferramentas e os dados a serem coletados para o sucesso de um produto de dados recém-desenhado.

Se esse fluxo for usado como ferramenta para mapear o processo de dados já coletados e tratados em diversos projetos, é possível considerá-lo para refletir sobre dados coletados desnecessários e sobre o que é necessário, mas não se coleta.

FLUXO DE DADOS

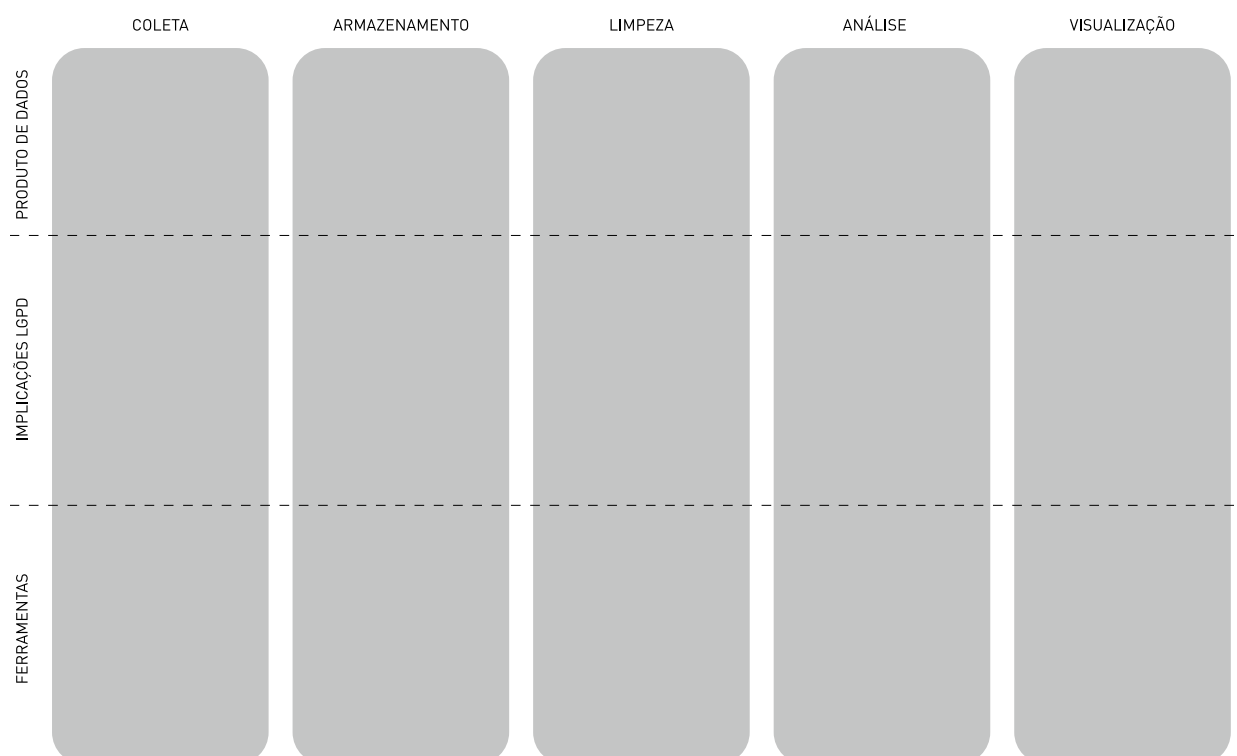


Figura 9. Ferramenta Fluxo de Dados Social Good Brasil

Os dados são a matéria-prima, e a qualidade da matéria-prima coletada refletirá diretamente na qualidade do resultado.

Uma diretriz, de acordo com a LGPD, é coletar apenas os dados que forem extremamente necessários e, além disso, justificar o porquê da coleta de cada dado, assim como a sua base para tratamento.

ETAPA 2 ARMAZENAMENTO

Na segunda etapa, devemos considerar o armazenamento dos dados de forma organizada e consultável. Por isso, é fundamental prever uma estrutura lógica que dê suporte a uma organização do tipo “guardar para encontrar”.

Diretrizes de LGPD incluem a verificação de fornecedores de armazenamento de dados com o certificado “EU-US Privacy Shield” ou similar, para garantir a proteção dos dados coletados.

ETAPA 3 LIMPEZA

Quanto melhor a qualidade da coleta de dados, melhor será a consistência dos registros para que as análises estejam corretas. Mesmo assim, deve-se considerar uma etapa de limpeza de dados, que serve como uma higienização de dados, e identifica possíveis falhas na migração entre sistemas.

De qualquer forma, no caso de um vazamento de dados, uma base suja apresenta um elevado grau de risco para a organização. Portanto, ter bases limpas também é uma diretriz da LGPD.

ETAPA 4 ANÁLISE

Nesta etapa, os dados se transformam em informações relevantes. As perguntas já estarão mapeadas previamente no Canvas de Dados, e o trabalho agora será identificar as variáveis necessárias para respondê-las e, então, fazer análises.

Se esse fluxo de dados for usado para mapear o tratamento que a instituição já faz no seu dia a dia, esta é a etapa na qual os tipos de análises realizadas, e as análises ainda inexploradas são identificados. De acordo com o nível de complexidade, as análises podem ser descritivas, diagnósticas, preditivas ou prescritivas.

Outra diretriz da LGPD é realizar o máximo possível de análises em formato de *clusters* e anonimizados, prezando pela privacidade dos dados dos seus usuários.

ETAPA 5 VISUALIZAÇÃO

Esta etapa normalmente usa a mesma ferramenta da etapa de análise de dados, e a somatória das duas é o produto de dados em si. São exemplos: painéis de controle (*dashboards*), artigos, pesquisas, planilhas, infográficos e relatórios impressos.

É a etapa em que dados são traduzidos em informações visuais para melhor compreensão, usando elementos visuais (linhas, figuras, recortes, contrastes) para representar o comportamento dos dados.

Um livro referência é *Storytelling com dados*, de Cole Nussbaumer Knaflitz.

O gestor poderá identificar oportunidades de somar a visualização dos dados a narrativas, que transformem os dados em histórias que engajam o público para a tomada de decisão ou relacionado ao objetivo inicial do projeto de dados.



#5

GESTOR PÚBLICO, LIDERE A MUDANÇA DE CULTURA

por Carol Andrade



Mais do que usar dados é necessário mudar uma cultura

Passar de uma cultura tradicional para uma cultura analítica significa menos percepção e mais *insights* a partir de dados.

No contexto desta publicação, cultura significa o conjunto de crenças e valores de uma organização ou de um grupo, podendo ser entendida também como “o jeito que se faz as coisas”.

Em 2018, Groysberg, Lee, Price & Cheng descreveram no artigo “O guia do líder para a cultura corporativa” que as normas culturais definem aquilo que é encorajado, desencorajado, aceito ou rejeitado pelo grupo.

Algumas décadas antes, em 1962, Everett Rogers publicou seu livro *Difusão de inovações*. Nele, tratou sobre como gerar e espalhar inovações na sociedade, enfatizando que tais inovações não são percebidas por todos os consumidores da mesma forma. Esse é um estudo para difundir a inovação dentro das instituições e promover a mudança de cultura.

O gestor público inovador e que lidera a mudança de cultura deve engajar outras pessoas dentro da organização e assim criar uma “legião” de embaixadores.

Esses embaixadores são fundamentais para ajudar a influenciar a **maioria inicial**, que adotarão a novidade quando tiverem clareza dos benefícios e resultados. Se esse perfil participar, é possível alcançar a **maioria tardia**, formada pelos resistentes à mudança, o grupo de conservadores. Por último, ficam os **retardatários**, que só participam quando há uma nova regra, uma nova lei os obrigando a se adequar.

A Figura 10, na página a seguir, ilustra a Lei da Difusão da Inovação, teoria proposta por Everett Rogers, no livro *Difusão de inovações*, citado no tópico anterior. Essa teoria procura explicar por que e em que proporção novas ideias e tecnologias se disseminam.

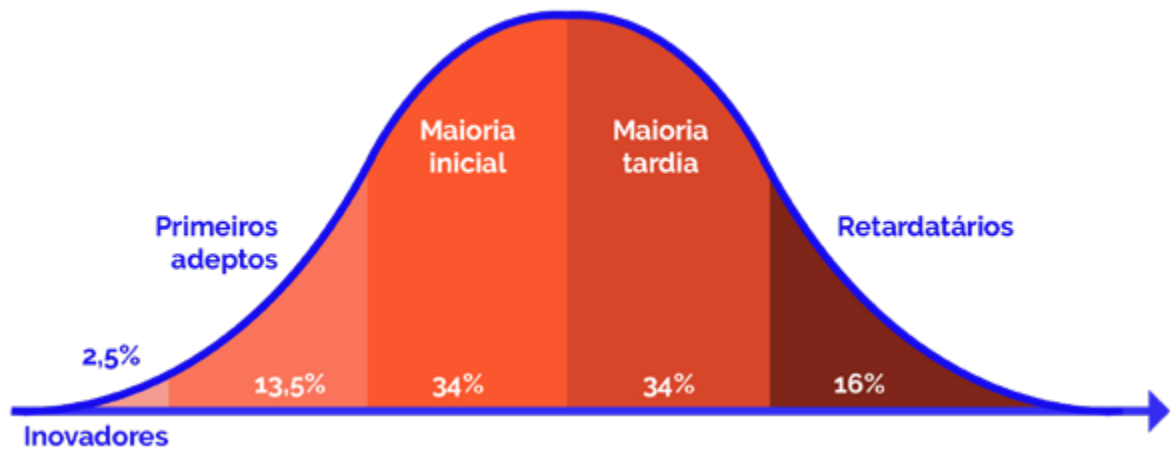


Figura 10. Lei da Difusão da Inovação segundo Rogers (1962)

Como implementar um plano para a mudança de cultura

Para implementar um bom plano de mudança de cultura, a gestão deverá contemplar pelo menos 4 passos:

PASSO 1 DIAGNÓSTICO

Quais são as características da cultura vigente? Essa pergunta é que vai nortear o diagnóstico das condições e do contexto social e cultural para a mudança.

Nesse primeiro passo, é importante compreender as especificidades da cultura, seja em uma área, secretaria ou instituição, e identificar as motivações para a mudança.

Por exemplo, quais são as potenciais **pressões externas**, como a Lei de Acesso à Informação e a Lei Geral de Proteção de Dados (LGPD)?

Da mesma forma, deve-se diagnosticar as **pressões internas** — necessidade de medir resultados, desenvolvimento de políticas públicas mais efetivas, excesso de burocracia e outras.

PASSO 2 MAPEAMENTO

No segundo passo, é feito o mapeamento dos perfis presentes na instituição, de acordo com a Lei da Difusão da Inovação. A partir daí, uma estratégia de engajamento é desenhada para cada perfil: primeiros adeptos, maioria inicial, maioria tardia e retardatários.

Busque engajar os primeiros adeptos como embaixadores, tornando-os multiplicadores analíticos dentro da instituição.

Apresente resultados ao grupo da maioria inicial, que pode se comprometer ao ver resultados e benefícios do uso de dados.

Para apresentar resultados que engajem o perfil da maioria inicial — os pragmáticos —, faça pilotos simples que gerem resultados de curto prazo e por meio dos quais seja possível visualizar benefícios tangíveis do uso de dados.

Já para envolver os resistentes, ative os embaixadores para conversar com esse grupo, e assim compreender os medos e desejos para a mudança.

Por fim, é importante pensar estratégias para envolver também os que estão na linha de frente coletando dados críticos para o sucesso dos produtos de dados.

PASSO 3 ESTRATÉGIA E PROCESSO DE COMUNICAÇÃO

Esse é o passo para engajar os perfis mencionados, tanto por meio de interações individuais quanto coletivas. Nesse momento, é fundamental criar uma narrativa baseada em dados que seja atraente e convincente.

PASSO 4 MUDANÇAS EM PROCESSOS, PRÁTICAS, RITUAIS E SISTEMAS DE RECONHECIMENTO

No quarto e último passo para implementar um plano de mudança de cultura, deve-se refletir sobre:

- a) ferramentas em uso e novas ferramentas a serem adotadas pela instituição;
- b) mudanças necessárias no processo decisório atual e em reuniões;
- c) reconhecimento e avaliação de quem usa os dados.



O desafio para avançar na Era dos Dados está nos modelos mentais, na cultura, nos aspectos humanos, e não na tecnologia. Por isso, para liderar e fazer avançar o processo de mudança, é fundamental posicionar o ser humano no centro, garantindo diálogo aberto e sentimento de propósito.

#6

TRANSPARÊNCIA E DADOS ABERTOS

por Fernanda Campagnucci



Transparência e dados abertos são fundamentais

O trabalho cotidiano na gestão pública é desafiador: geralmente, as equipes são reduzidas, a substituição de tecnologias defasadas é lenta e as cobranças e os desafios são crescentes e cada vez mais complexos.

Muitos dos problemas enfrentados pelo setor público são chamados de *wicked problems* — em português, “problemas complexos”. São exemplos os fenômenos decorrentes das mudanças climáticas, da desigualdade e da saúde pública.

Onde encontrar soluções para esses problemas? Em geral, as melhores experiências de inovação pública são abertas à inteligência coletiva e à participação social. Para isso, transparência e dados abertos são fundamentais.

A transparência é um fator de prevenção da corrupção, e dados abertos estimulam inovação dentro e fora do Estado.

Além disso, o acesso à informação pública é um direito fundamental, previsto em convenções e tratados internacionais reconhecidos pelo Brasil, tanto na Constituição Federal de 1988 quanto em normas como a Lei de Acesso à Informação (Lei Federal nº 12.527/2011) e a Lei de Governo Digital (Lei nº 14.129/2021).

O “emaranhado” dos dados

O mais importante passo para uma política efetiva e sustentável de transparência e dados abertos é estabelecer uma *estrutura de governança de dados*, definindo padrões e quem zela por eles.

Em todos os níveis da organização e em diálogo com a sociedade, essa estrutura ajudará a traçar planos, metas e prioridades de abertura e melhorias.

É preciso compreender que dados são construções humanas. Alguém, em algum lugar, decide o que medir ou não, o que coletar ou não, e como fazer isso. Os dados passam por tratamentos e transformações, às vezes com aportes de mais de um departamento, e a falta de documentação de como isso deve ser feito gera conflitos de interpretação e comunicação “dentro de casa” — imagine para quem está fora.

Uma boa governança ajuda a garantir a integridade, a qualidade e o acesso aos dados, evitando que eles sejam apenas um “emaranhado” burocrático de onde pouco se extrai valor.

Por onde começar?

O processo de publicação de dados pode ser pensado como uma “trilha”³⁸, em que algumas etapas precisam acontecer de forma contínua, e não necessariamente linear. Conheceremos melhor essas etapas a seguir.

ETAPA 1 COMPREENDER

Transparência e dados abertos são ideias diferentes e complementares. Uma **política de transparência** visa tornar informações acessíveis e compreensíveis ao público em geral. Seus elementos incluem painéis de consulta, gráficos, relatórios e outros documentos.

Já uma **política de dados abertos** visa publicar bases de dados seguindo padrões técnicos correntes³⁹. Esses dados podem ser acessados e reutilizados livremente, inclusive por máquinas, e os sistemas podem trocar informações com mais facilidade (são “interoperáveis”). Isso coopera para que a própria sociedade crie soluções inovadoras com esses dados, sem onerar o orçamento público.

As duas políticas devem ser feitas de forma conjunta. Uma base de dados não é necessariamente acessível para todos os públicos. Um gráfico, embora possa ser mais direto, não permite checar as informações que ele mostra. É fundamental, portanto, oferecer capacitação permanente, para que tanto a equipe da gestão quanto quem usar os dados possam lidar com essas duas frentes.

Essa formação não deve se limitar a ferramentas ou linguagens de análise de dados. O mais crucial e desafiador é capacitar para a adoção da cultura de uso de dados, bem como localizar benefícios e oportunidades de gerar valor. É importante ter visão da infraestrutura de dados (em inglês, *data infrastructure literacy*⁴⁰), o que permite criar e intervir na maneira como são gerados e extraídos os dados.

ETAPA 2 PLANEJAR

O planejamento de políticas de dados abertos depende do mapeamento de todas as bases existentes, que podem ser organizadas em uma espécie de “inventário” ou catálogo de dados⁴¹. Esse instrumento reúne os “metadados” sobre as bases, que

38 A Open Knowledge Brasil propõe essas etapas no e-book *Publicadores de dados: da gestão estratégica à abertura* (2021). Disponível em: <<https://ok.org.br/publicacoes/>>.

39 O Ceweb.br possui diversas publicações que detalham esses padrões. Veja, por exemplo, o livro *Dados abertos conectados* (2015), de Seiji Isotani e Ig Ibert Bittencourt, em: <<https://ceweb.br/publicacao/livro-dados-abertos/>>.

40 Consulte o artigo “Data infrastructure literacy”, de Jonathan Gray, Carolin Gerlitz e Liliana Bounegru (2018), em: <<https://doi.org/10.1177/2053951718786316>>.

41 Conheça o exemplo do Catálogo Municipal de Bases de Dados (CMDDB), da Prefeitura de São Paulo, criado em 2014 e que, desde então, passa por ciclos de atualização semestrais. Acesse: <<http://bit.ly/cmbd-sp>>.

descrevem o conteúdo: quem são os responsáveis, se contêm ou não variáveis restritas por lei etc.

Um trabalho assim traz revelações surpreendentes sobre informações coletadas dentro da gestão. É essa visão do todo que permite priorizar e direcionar o esforço de abertura. A orientação é começar com o possível e, gradualmente, aumentar e melhorar — esse processo é permanente e subsidia também a governança de dados da organização.

Fazem parte dessa etapa os planos de dados abertos (PDAs)⁴². Para refletirem de fato a demanda interna e externa por dados, eles devem ser feitos de forma colaborativa. Os PDAs ajudam a sociedade a monitorar o esforço de abertura, além de servirem de roteiro para priorizar o que será aberto.

ETAPA 3 DESENVOLVER

Para que os processos de gestão e abertura de dados sejam sustentáveis, o “encanamento” deve ser sólido, bem planejado e de preferência automatizado, para depender o menos possível de intervenções manuais.

Há ferramentas do tipo *open source*, de uso livre, para criar os processos ETL — *Extract, Treat and Load*, ou,

em português, extração, tratamento e carga dos dados. Também aqui é possível pensar em um processo gradual: não é preciso ter tudo pronto e perfeito para começar a abrir dados que já podem gerar alto valor para a gestão e os usuários⁴³.

ETAPA 4 ABRIR

Esta é a etapa em que os dados finalmente “dão as caras” ao mundo: a publicação. Para que eles sejam úteis e compreensíveis, além de seguirem padrões de formato aberto, precisam estar bem documentados.

Há várias opções para torná-los públicos. Painéis (*dashboards*) são ferramentas atraentes para dar transparência a um assunto dinâmico e complexo, pois reúnem formas de visualização e consulta. É recomendável que as bases de dados que geram os painéis estejam disponíveis, de forma que cumpram o propósito de uma política de dados abertos.

Os dados também podem estar publicados em repositórios — como o CKAN, ferramenta aberta usada por governos e organizações ao redor do mundo⁴⁴ para construir portais que disponibilizam *downloads* de dados abertos.

42 O manual *Elaboração de planos de dados abertos*, da Controladoria Geral da União (CGU), traz uma proposta de passo a passo. Conheça em: <<http://bit.ly/manual-plano>>

43 Veja mais detalhes nos capítulos 5 e 6 da edição *Publicadores de dados: da gestão estratégica à abertura*: <<https://ok.org.br/publicacoes/>>.

44 Para conhecer casos interessantes, acesse: <<https://ckan.org/showcase>>. No Brasil, além do Governo Federal (<http://dados.gov.br>), diversos estados e municípios também utilizam essa ferramenta. Ver, por exemplo, o Portal de Dados Abertos de Santa Catarina: <<http://dados.sc.gov.br/>>.

Em um nível mais alto de abertura, os dados podem ser acessados pelas pessoas usuárias por meio de uma API (do inglês, *Application Programming Interface*), método que permite que programas de computador ou aplicativos façam consultas pontuais e acessem dados diretamente.

ETAPA 5 CONECTAR

O uso e reúso dos dados é o propósito. Para que sejam úteis, eles precisam estar sempre atualizados, bem documentados e acessíveis pelos meios mais adequados.

Mas quem são essas pessoas usuárias? Você é uma, por exemplo. Usam-se os dados dentro ou fora da organização, com vários níveis de acesso, entendimento e de capacidades técnicas.

Essas pessoas fazem parte de um ecossistema, e integrá-las em todas as etapas precedentes é o caminho para o sucesso das políticas de transparência e dados abertos.

Por isso, além de mapear essas partes interessadas (os *stakeholders*), é preciso colocar em prática estratégias para que elas possam opinar sobre os dados e o planejamento de abertura.

Finalmente, o uso desses dados deve ser incentivado por meio de desafios, competições (*datathons, hackathons*)⁴⁵ ou outros processos colaborativos, internos e externos, para que os dados possam gerar valor para a gestão e a sociedade com a inovação aberta. O uso dos dados aumenta a demanda pela qualidade deles e aproxima o público da gestão que os abre.

45 *Hackathons* são maratonas nas quais se desenvolvem soluções criativas para problemas propostos por uma organização. Nesses esforços concentrados, que podem durar um ou mais dias, equipes costumam competir por um prêmio ou trabalhar de forma colaborativa. Bases de dados são insumos importantes em um *hackathon*. Já um *datathon* — em português, maratona de dados — é uma versão mais específica desse tipo de evento, girando apenas em torno de análise, visualização e outros produtos derivados de dados.

#7

LGPD NO SETOR PÚBLICO

por Silvana Melo e Mariana Zonari



Privacidade e proteção de dados pessoais

O cenário da Era Digital, aqui apresentado, fundamenta-se na necessidade de compartilhamento de conhecimento e acesso – quase que ilimitado – a informações. Contudo, atrelado a este processo, temos uma grande preocupação com o mau uso de dados e com a privacidade das pessoas. Essa preocupação é ainda maior no setor público, uma vez que o Estado possui um dever inerente de respeito, zelo e transparência com os cidadãos. Nesse sentido, é necessário entendermos que privacidade e proteção de dados pessoais não são expressões sinônimas.

No Brasil, essas duas expressões estão inseridas no panorama dos direitos de personalidade⁴⁶. No entanto, existem diferenças significativas entre elas, principalmente relacionadas a origem, objeto, tipo de esfera de proteção que oferecem ao titular, comunicação com outros direitos e previsão no ordenamento jurídico⁴⁷.

A noção de privacidade, em síntese, encontra-se prevista na Constituição Federal de 1988 (art. 5º, X e XII) e busca preservar o direito à intimidade, à liberdade de decisão sobre a divulgação de informações e a autonomia de cada cidadão⁴⁸. É um conceito complexo, que muitas vezes apresenta dimensões concorrentes e contraditórias, mas se relaciona diretamente com as percepções de escolhas, controle, transparência e segurança.

O direito à proteção de dados pessoais, por sua vez, tem origem posterior ao direito à privacidade e é resultado da sociedade da informação. Com o surgimento de computadores e, em seguida, dos bancos de dados, o controle sobre a informação, especialmente relacionada aos dados pessoais, passou a ser visto como uma forma de poder⁴⁹.

Indispensável citar que esse movimento não é apenas do Brasil e que uma

46 Os direitos da personalidade são aqueles relacionados à pessoa e à sua dignidade. São exemplos: vida/integridade física, honra, imagem, nome e intimidade.

47, 48, 49 Veja o artigo “Análise comparativa entre direito à privacidade e direito à proteção de dados pessoais...”, de Gabriela Machado Vergili, em: <<https://dataprivacy.com.br/analise-comparativa-entre-direito-a-privacidade-e-direito-a-protecao-de-dados-pessoais-e-relacao-com-o-regime-de-dados-publicos-previsto-na-lei-geral-de-protecao-de-dados-2/>>.

maior preocupação com a privacidade e a proteção dos dados pessoais ganha força, a nível global, após inúmeras violações no decorrer da história.

O desconhecimento, em geral, desses conceitos é sintoma da ausência de uma cultura forte de privacidade e proteção de dados no Brasil. Como consequência, o setor público tem dificuldade para se adequar à Lei Geral de Proteção de Dados (LGPD), ainda que tenha havido crescente preocupação com a proteção de dados pessoais nos últimos anos.

• • •

A LGPD é a lei brasileira que regulamenta o tratamento de dados pessoais, seja realizado por pessoa natural, seja por pessoa jurídica de direito público ou privado, desde que ofereçam produtos ou serviços no Brasil ou tratem dados de pessoas localizadas em território nacional. Esse tratamento pode ocorrer tanto por meios físicos quanto digitais.

Ela não é uma legislação específica sobre sigilo ou confidencialidade, tratando da proteção de dados pessoais, isto é, de informações que identifiquem uma pessoa natural ou a tornem identificável.

Essa legislação foi criada para permitir e facilitar o tratamento de dados pessoais e proporcionar o desenvolvimento tecnológico, econômico e da inovação. Por ser uma legislação recente, existem mi-

tos que a cercam, como, por exemplo, o de que é necessário o consentimento do titular para qualquer tratamento de dados ou o de que a LGPD veio para proibir o uso de dados pessoais.

Portanto, é cada vez mais importante incentivar uma mudança cultural sobre privacidade e proteção de dados pessoais. A ideia é observar nessa regulamentação uma “janela de oportunidade”, refletindo sobre o quanto as instituições públicas poderão ganhar e se tornar mais eficientes com a adequação à nova lei⁵⁰.

Conceito e tratamento de dado pessoal

A LGPD é uma lei sobre o tratamento de *dados pessoais*. O conceito desse tipo de dado está em seu artigo 5º: informação relacionada a pessoa natural identificada ou identificável.

No Brasil, foi adotado o conceito expansionista de dado pessoal, ou seja, não apenas a informação relativa à pessoa diretamente identificada estará protegida pela lei, como também aquela informação que possa tornar a pessoa identificável⁵¹.

A lei também traz a noção de dado pessoal sensível: informações sobre origem

50 Veja o artigo “Regulação de dados é uma janela de oportunidade”, de Bruno Bioni, em: <<https://brunobio-ni.com.br/blog/2019/03/29/regulacao-de-dados-e-uma-janela-de-oportunidade/>>.

51 Veja o artigo “The PII Problem...”, dos professores George Schwartz e Daniel Slove, em: <https://www.law.berkeley.edu/files/bclt_Schwartz-Solove_NYU_Final_Print.pdf>.

racial ou étnica, convicção religiosa, opinião política, filiação a sindicato ou a organização de caráter religioso, filosófico ou político, dado referente à saúde ou à vida sexual, dado genético ou biométrico, quando vinculado a uma pessoa natural.

De acordo com a LGPD, *tratamento* é toda operação realizada com dados pessoais, como as que se referem a coleta, produção, recepção, classificação, utilização, acesso, reprodução, transmissão, distribuição, processamento, arquivamento, armazenamento, eliminação, avaliação ou controle da informação, modificação, comunicação, transferência, difusão ou extração.

Essa definição de tratamento de dados pessoais é abrangente, indo desde a coleta até a eliminação. Engloba todas as possibilidades de manuseio dos dados, independente do meio utilizado. Dessa forma, os atos de receber, acessar, arquivar ou armazenar dados pessoais estão contidos no conceito de tratamento, tornando indispensável manter os registros dessas operações⁵².

...

Depois de conhecermos os conceitos de dados pessoais e tratamento, é importante agora entendermos o que a LGPD dispõe sobre titular, controlador, operador e encarregado, além da relação entre esses atores.

TITULAR

Pessoa natural a quem se referem os dados pessoais que são objeto de tratamento. É o núcleo de existência da LGPD. Sempre que pensamos em dados pessoais, devemos ter os nossos holofotes voltados ao titular de dados pessoais. Esse é um princípio máximo da autodeterminação informativa trazido pela lei.

CONTROLADOR E OPERADOR

São os agentes de tratamento. O controlador é pessoa natural ou jurídica, de direito público ou privado, a quem competem as decisões referentes ao tratamento de dados pessoais. O operador é pessoa natural ou jurídica, de direito público ou privado, que realiza o tratamento de dados pessoais em nome do controlador. A LGPD impõe seu maior peso jurídico sobre o controlador, pois ele é o responsável pela tomada de decisões, enquanto o operador apenas trata os dados pessoais de acordo com as determinações do controlador.

ENCARREGADO

Também conhecido como *data protection officer* (DPO), atua como canal de comunicação entre o controlador, os titulares dos dados e a Autoridade Nacional de Proteção de Dados (ANPD). Ele é o responsável pela gestão interna do programa de privacidade.

52 Veja o artigo “The PII Problem: Privacy and a New Concept of Personally Identifiable Information”, dos professores George Schwartz e Daniel Slove, disponível em: <https://www.law.berkeley.edu/files/bclt_Schwartz-Solove_NYU_Final_Print.pdf>.

Desafios do setor público para se adequar à LGPD

A LGPD trouxe para o setor público a necessidade de mudança na cultura da privacidade e proteção de dados pessoais.

A urgência de adequação à LGPD não diz respeito apenas a penalidades administrativas, mas implica também em questionamentos sobre a melhor forma para torná-la efetiva.

Ponderar sobre a necessidade de criação de um comitê de privacidade interno ou sobre a contratação de empresa especializada em Direito Digital são exemplos desses questionamentos⁵³.

O regulamento não responde essas questões. Entretanto, em seu capítulo IV, disciplina a base fundamental para o correto tratamento de dados pessoais no setor público, com destaque para a ideia de interesse público.

A observância do interesse público é o que norteia as principais regras a serem seguidas pela esfera pública no exercício do tratamento e do uso compartilhado de dados, duas atividades necessárias para executar políticas públicas.

O setor público, por meio de suas organizações, coleta e mantém grande volume de dados, inclusive, de natureza sensível.

Isso torna obrigatória a adoção de medidas preventivas para evitar o vazamento de dados.

Outro aspecto relevante é que, muitas vezes, as bases de dados são sobrecarregadas com informações desatualizadas e desnecessárias, tornando as tarefas da adequação mais desafiadoras.

É importante, por isso, avaliar impactos e planejar ações preventivas antes, durante e depois do período de estruturação, sempre considerando os aspectos legais.

Hoje, os principais desafios para implementar a LGPD no setor público se concentram nas seguintes ações:

- » Realizar um diagnóstico com registro de dados e ambiente normativo adequado.
- » Construir um programa público de governança em privacidade e proteção de dados que estimule a criação de cultura sobre o tema.
- » Ajustar os processos de contratação, conciliando-os com outras normas, considerando a responsabilidade solidária dos agentes de tratamento.

Outros desafios também se impõem: visão setorializada, falta de integração e interoperabilidade dos dados, diversidade tecnológica e limitações orçamentárias.

53 Veja o artigo "A LGPD e seus efeitos no setor público", de Rodrigo Otero, em: <<https://www.serpro.gov.br/lgpd/noticias/2019/lgpd-setor-publico-efeitos>>.

#8

DIMENSÕES ÉTICAS E DILEMAS DA ATUALIDADE

por José Borbolla



O fim da privacidade como a conhecemos

A produção de novas legislações e políticas para lidar com as transformações provocadas pelo rápido avanço tecnológico, como é o caso da privacidade, demonstra, na prática, os efeitos que essa aceleração produz em nosso dia a dia.

Esse período de profundas e velozes transições nos obriga a voltar para a “mesa de discussão”. Precisamos redesenhar consensos e definições neste mundo cada vez mais hiperconectado e hiperdigital.

Nesse contexto, o papel do poder público, em um sentido mais amplo, e o papel do gestor, mais especificamente, são cruciais e indispensáveis.

Isso significa que a gestão pública deve ser capaz de usar, interpretar e analisar criticamente dados internos e externos em todas as esferas de atividade, independente da área. Essa capacidade é condição indispensável para responder e endereçar desafios causados por mudanças contínuas.

Uma vez explicado que a utilização de técnicas e ferramentas — por mais modernas e avançadas que sejam e por maior quantidade de dados que sejam capazes de processar — não resolvem nossos problemas nem respondem nossas perguntas sozinhas, é fundamental destacar, mais uma vez, que o pensamento crítico é o ingrediente que falta para dar “liga” nessa receita.

Ele é obrigatório na construção da “ponte” que conecta todo o potencial oferecido por determinada tecnologia para sua aplicação na vida real. Daí precisamos falar aqui sobre o que parece ser o aspecto mais importante quando falamos sobre dados e inteligência artificial: a Ética.

É imprescindível olhar para a tecnologia como um meio que pode nos ajudar a resolver desafios, e o que determina o sucesso ou o insucesso é o uso que nós, humanos, decidimos dar para ela.

Se a Matemática pode nos dizer se a variável A causa um efeito B, é na dimensão ética que decidimos o que fazer com essa informação e como aplicá-la no mundo real.

Ainda que pareça óbvio, quando olhamos para a história das sociedades percebemos que a racionalidade científica pura⁵⁴ e sua dimensão ética nem sempre estão de mãos dadas.

Dois aspectos centrais podem ampliar nossa compreensão sobre a natureza desse desafio e os riscos que o avanço tecnológico atual pode trazer.

O primeiro deles é o conceito de viés cognitivo, fruto da complexidade do nosso cérebro e da maneira como a humanidade evoluiu. O outro aspecto é como decidimos o que é certo, errado, bom e ruim.

Nosso cérebro é resultado de milhões de anos de evolução, e parece ser o mesmo, estruturalmente, há pelo menos 100 mil anos. Ao longo do tempo, ele foi sofrendo diferentes pressões evolutivas e novas estruturas e peculiaridades em seu funcionamento foram surgindo.

De maneira resumida, segundo Daniel Kahneman, no livro *Rápido e devagar: duas formas de pensar*, a evolução da nossa espécie produziu um cérebro humano composto por dois sistemas diferentes: um que toma decisão de maneira automática em níveis predo-

minantemente inconscientes; outro que delibera racionalmente para então decidir.

O primeiro é muito mais eficiente do ponto de vista energético. O segundo gasta mais “combustível metabólico”. O primeiro é rápido. O segundo pode demorar.

Há milhares de anos, a combinação entre esses sistemas nos ajudou a encontrar comida, sobreviver a extremos climáticos, escapar de predadores e desenvolver estratégias e relações sociais cada vez mais complexas.

No mundo de hoje, os “efeitos colaterais” que esses dois sistemas cerebrais indissociáveis produzem são denominados, de maneira genérica, de vieses cognitivos, ou, em outras palavras, tendências e padrões que nosso cérebro apresenta em determinadas circunstâncias.

Talvez o mais conhecido entre tantos vieses já mapeados seja o viés de confirmação, ou a tendência que nosso cérebro possui de buscar informações para confirmar crenças anteriores, enquanto resiste e produz desconforto quando apresentado a fatos que contradizem as opiniões que já possuímos.

É esse viés que nos ajudará a conectar todos os pontos que discutimos até aqui.

54 Essa racionalidade está na base do mito do “solucionismo tecnológico”, discutido no capítulo 2.

Solucionismo, racionalidade, viés cognitivo e ética: juntos, em um exemplo real e atual

Vamos imaginar um mundo ou uma época qualquer em que dispomos, totalmente digitalizado, na íntegra e sem nenhum tipo de erro ou problema nos dados, de todo o histórico penal do Brasil (infrações, condenações, prisões, tudo...), desde 1889 (portanto, pós-abolição).

Trabalhamos juntos em um órgão de segurança pública de um futuro não muito distante e contratamos uma *startup* que desenvolveu uma solução com inteligência artificial incrivelmente poderosa, capaz de usar o histórico de ocorrências para prever o risco de crimes futuros, de maneira precisa, geolocalizada e dentro de uma janela de tempo ideal.

Qual tipo de previsão você acha que essa inteligência artificial seria capaz de produzir?

Alguns poderiam dizer que, pelo fato de esse histórico criminal estar disponível na íntegra e seguir sendo alimentado por dados atuais, isso bastaria para que tal sistema pudesse atingir níveis de precisão muito superiores ao que fizemos até aqui.

A lógica parece simples: olhamos para todos os tipos de ocorrências, quando e onde ocorreram, para o perfil das pessoas envolvidas e cruzamos isso com uma infinidade de outras variáveis. A previsão será produzida, portanto, a partir de uma quantidade de dados gigantesca.

Mas que tipo de previsão essa inteligência artificial está produzindo?

O exemplo brasileiro é bom para explorarmos a complexidade dessa pergunta. Após a invasão europeia no início do século XVI, passamos cerca de **70%** do nosso tempo de existência sob um regime de trabalho escravista, tendo como atividade econômica central o tráfico negreiro. Fomos o país que mais recebeu seres humanos de diversos locais da África.

Durante boa parte desses quase 350 anos, as pessoas pretas não foram consideradas pessoas. Eram mercadoria. E assim também eram tratadas em nosso conjunto de leis. A primeira vez que o preto é humanizado em nosso regramento jurídico é no Código Penal de 1830, especificamente para que punições contra o seu corpo pudessem ser bem estabelecidas e operadas pelo Estado.

A abolição não foi um processo que se preocupou com a integração dessas pessoas, que sempre estiveram à margem das estruturas de poder enquanto indivíduos.

No final do século XIX e início do XX, dos diversos absurdos que ocorreram,

destacamos dois para finalizarmos nosso exemplo: o primeiro é a criação de regras e tipos penais, como o conceito de “vadiagem”, para sustentar e manter as profundas separações e diferenças socioeconômicas entre brancos e pretos, criando as bases legais para criminalizar praticamente tudo que negros recém-libertos faziam.

O outro elemento são as ideias de “eugenia” e “racismo científico”, que forneceram as bases racionais para que todas as crueldades e violências possíveis seguissem sendo cometidas contra as mesmas pessoas. Se o regime escravista havia acabado, nossa sociedade inventou um outro regime, com outras regras e outros nomes para continuarem fazendo a mesma coisa.

Vamos retomar algo que já aprendemos juntos até aqui e que este exemplo nos ajuda de maneira fundamental a compreender de maneira ainda mais objetiva: **se a correta aplicação metodológica da matemática e da estatística nada nos dizem sobre o que é “certo” e “errado”, “bom” e “ruim”, esse tipo de decisão ocorre em uma outra esfera, a partir de outras premissas e mediada pelas especificidades cognitivas do nosso cérebro.**

E por que isso acontece?

Porque analisar histórico de criminalidade apenas nos diz quem ultrapassou a linha do “certo”, do “permitido” e do

“aceitável”, mas nada nos dizem sobre como as linhas que separam “legal” e “ilegal” foram estabelecidas.

As sociedades continuaram a traçar linhas e criar regras que mantivessem a segregação racial mesmo depois da abolição da escravidão. Isso significa que qualquer sistema ou “inteligência”, por mais avançado que seja, nunca será capaz de prever o “próximo crime”. A única coisa que esse sistema fará é replicar os padrões do passado, só que numa velocidade e numa escala inimagináveis.

É aqui que precisamos conectar a dimensão ética da nossa atuação como profissionais de dados. Quando olharmos, por exemplo, para o Atlas de Violência⁵⁵, aqueles números podem receber diferentes interpretações e explicações.

Um gestor descuidado, mal informado sobre a história do seu próprio país e partidário do “solucionismo tecnológico” pode interpretar que o percentual de pessoas presas indica como a violência se distribui na sociedade. Se este indivíduo criar um algoritmo preditivo, ele o fará a partir dessa visão de mundo e interpretará os resultados a partir dessa perspectiva.

Mas podemos olhar para os mesmos números e chegar a conclusões diferentes. Se a realidade da violência urbana e do sistema prisional está desse

55 O Atlas da Violência é um portal que reúne, organiza e disponibiliza informações sobre violência no Brasil. Foi criado em 2016 e é gerido pelo Instituto de Pesquisa Econômica Aplicada (Ipea). Acesse: <<https://www.ipea.gov.br/atlasviolencia/>>.

jeito, isso revela como todos esses elementos históricos ainda ecoam e reverberam na nossa época, significa que ainda não fomos capazes, enquanto sociedade, de equacionar essas questões estruturais históricas.

Ao contrário: de diversas formas, tudo o que fizemos até aqui só contribuiu para manter e aprofundar essas diferenças. Uma pessoa que tenha essa visão sobre o processo, ao criar um algoritmo preditivo, chegaria a conclusões totalmente diferentes.

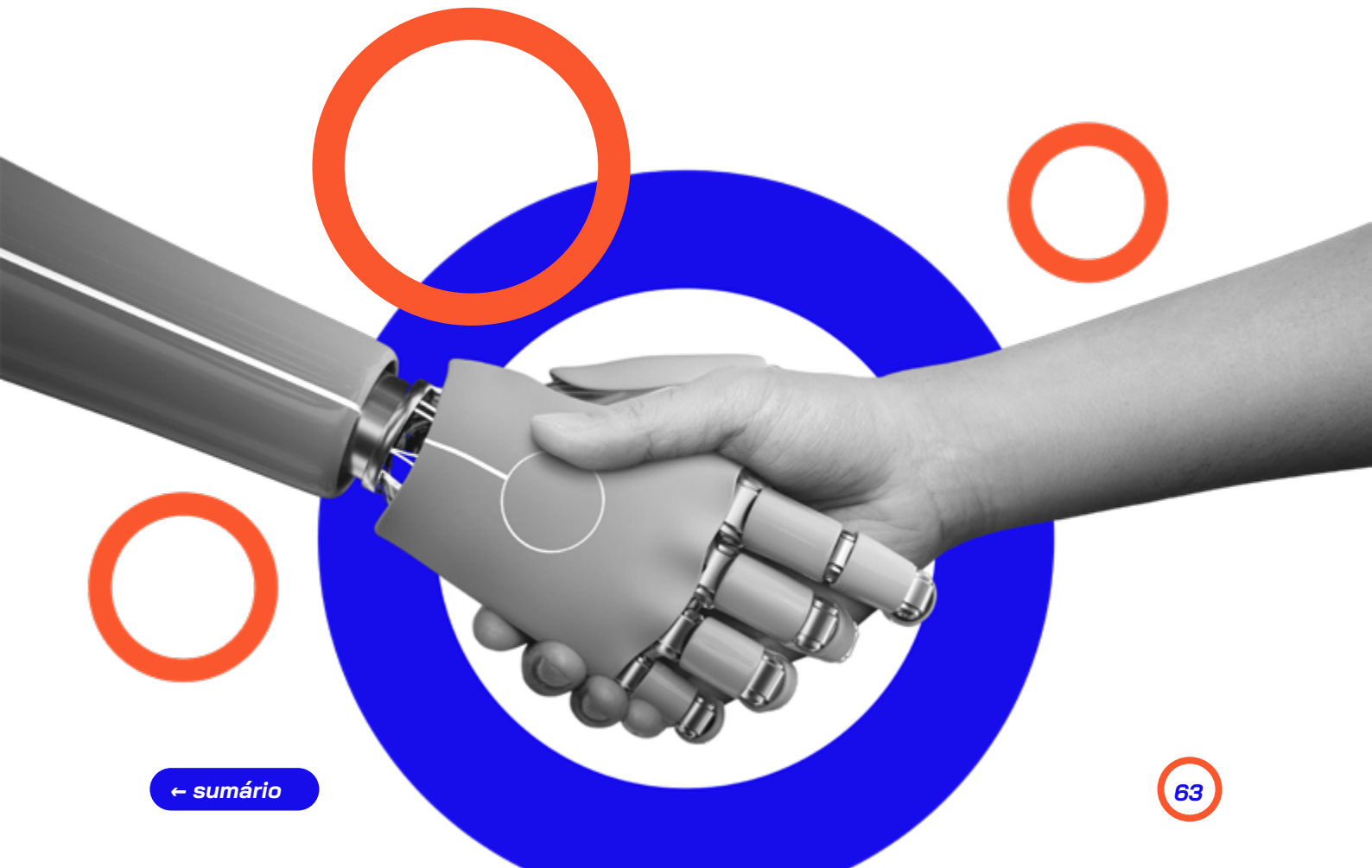
Quem está mais certo? Essa é uma questão em aberto em 2021, cuja resposta não é única, óbvia, simples ou “solucionista”. E cujos efeitos sentiremos nos próximos anos e décadas.

Sempre que lidamos com problemas complexos, como as questões sociais, por exemplo, precisamos utilizar um

modelo. Um modelo será, necessariamente, uma simplificação da realidade.

Se não tivermos clareza e consciência sobre o que os nossos modelos incluem ou excluem, sobre a representatividade das bases de dados e sobre como elementos históricos e culturais estão presentes, seja na origem do dado, seja em nossos vieses cognitivos, correremos o risco de ampliar e aprofundar desigualdades e outros problemas estruturais da nossa sociedade.

E, por fim, mas não menos importante, lembremos: qualquer modelo de inteligência artificial tem seus limites, por mais avançado que seja e por maior que seja sua capacidade de processar volumes enormes de dados.



#9

USO DE DADOS NA GESTÃO PÚBLICA: CASOS E APLICAÇÕES

por José Macedo, Ticiano Linhares e Marianna Gonçalves



CASO 1

Rastreador de sintomas de COVID-19

PROBLEMA

A pandemia de COVID-19 exigiu dos governos medidas eficientes e ágeis para mitigar os efeitos causados pela doença.

O Plantão Coronavírus foi uma das soluções desenvolvidas e disponibilizadas para a população no Estado do Ceará. É um projeto da Secretaria da Saúde (Sesa) acelerado pelo ÍRIS. Trata-se de uma plataforma com tecnologia de chatbot, em que o usuário mantém um bate-papo com um sistema de inteligência artificial.

Primeiro, é feita uma triagem a partir dos sintomas informados pelo usuário, a fim de classificar seu estado de saúde em uma das três categorias: leve, moderado e grave. Em seguida, a depender da classificação do seu estado de saúde, o usuário poderá ser encaminhado para o atendimento humano, na mesma plataforma.

Por ser uma doença nova e pela gravidade da pandemia, era fundamental conhecer naquele primeiro momento, via Plantão Coronavírus, os sintomas relatados pela população. No entanto, seria impossível para os profissionais rastrear nas interações textuais todos os sintomas, dos mais frequentes aos mais raros.

SOLUÇÃO

Para acompanhar as descobertas de novos sintomas — frequentes ou raros — presentes em pessoas infectadas pelo novo coronavírus, o ÍRIS, em cooperação com o Programa Cientista Chefe do Governo do Ceará, criou a Sintomatic, uma rede neural profunda (*deep learning*) que processa textos em linguagem natural.

A identificação de sintomas em texto foi mapeada como um problema de reconhecimento de entidade (em inglês, *Named Entity Recognition* – NER). NER corresponde à capacidade de identificar as entidades nomeadas nos documentos e rotulá-las em classes definidas de acordo com cada tipo. De forma geral, o robô de captura de sintomas possui uma rede neural capaz de reconhecer entidades. Neste caso, uma entidade é um sintoma.

Devido a possíveis mutações do vírus e consequente aparecimento de novos sintomas, esse modelo proporcionou avanços para entender a doença. Por causa de sua capacidade de reconhecer novos padrões, identificou uma alta frequência de comportamentos psicológicos alterados, como ansiedade, angústia e tristeza, em usuários positivos ou não para COVID-19.

A ferramenta mostrou, portanto, a necessidade de o Estado ampliar o atendimento através do canal para o cuidado com a saúde mental da população.

RESULTADOS E IMPACTOS

- » Monitoramento dos sintomas frequentes e raros associados à COVID-19, de abril a outubro de 2020. Esse período compreende o começo da pandemia e a primeira onda de novos casos e mortes, quando pouco se sabia sobre a doença.
- » Automatização do processo de análise dos dados gerados pelo Plantão Coronavírus. Seria inviável aos profissionais de saúde mapear novos sintomas, diante do volume de textos gerados pela Plataforma. No dia 7 de maio de 2020, por exemplo, houve um pico de 8.700 acessos.
- » Identificação de novos sintomas e comportamentos psicológicos alterados.
- » Acompanhamento da evolução temporal dos sintomas para entender quais deles mais afetavam a população cearense.

CASO 2

Plataforma Big Data Social

PROBLEMA

A Secretaria de Proteção Social, Justiça, Cidadania, Mulheres e Direitos Humanos (SPS) precisa identificar e acompanhar todas as famílias em situação de vulnerabilidade social no estado do Ceará.

A existência de bases de dados descentralizadas e não estruturadas dificulta o desenvolvimento de políticas públicas mais direcionadas e efetivas, bem como o acompanhamento de diversos programas gerenciados pelo órgão.

Nesse contexto, o ÍRIS, apoiado pelo Programa Cientista Chefe do Governo do Ceará e em parceria com a SPS, está desenvolvendo a plataforma Big Data Social.

SOLUÇÃO

A plataforma Big Data Social integra dados de diversas bases da SPS. Inclui também dados das demais pastas ligadas à Proteção Social, como as secretarias de Educação e da Saúde e, ainda, de ministérios do Governo Federal e organizações não governamentais (ONGs).

Sustentada por tecnologias de *big data* e de inteligência artificial, ela permitirá integrar, analisar e avaliar as políticas públicas de assistência e proteção social no Ceará, com o objetivo maior de tornar as políticas públicas mais efetivas.

A principal entrega dessa plataforma é a visão integrada de um conjunto de dados e indicadores confiáveis sobre políticas públicas de assistência social, para subsidiar gestores, pesquisadores da assistência social, organizações não governamentais, entre outros.

Órgãos públicos, como os Centros de Referência da Assistência Social (CRAS), por exemplo, poderão identificar, de forma mais eficiente, os serviços acessados por cada família e, assim, direcio-

nar a assistência de acordo com o perfil e a necessidade das pessoas.

RESULTADOS ESPERADOS

O uso de *big data* e inteligência artificial são fundamentais para analisar e interpretar grandes volumes de informações, o que permite tomadas de decisões mais assertivas pelos gestores públicos.

Além disso, a combinação dessas tecnologias torna possível compreender e desenhar políticas públicas para atender às complexidades do sistema de assistência e proteção social.

Com a integração dos dados e a transparência dos resultados, a população se beneficia através de políticas mais assertivas, da avaliação e intervenção dos problemas identificados por meio de indicadores, da facilidade na tomada de decisões estratégicas, da promoção, do acesso a números atualizados e da acurácia (precisão) na aplicação dos recursos destinados a programas sociais nas diversas instâncias.

Com a implementação total dessa plataforma, os seguintes resultados são esperados:

- » Aumento da capacidade dos gestores para desenvolver políticas públicas orientadas às necessidades reais da população.
- » Aumento da efetividade na aplicação dos recursos públicos envolvidos nos programas sociais.
- » Diminuição da desigualdade social no estado do Ceará.

» Aproximação entre o governo e a população.

CASO 3

Transparência na Saúde Pública

PROBLEMA

A Secretaria da Saúde do Governo do Ceará (Sesa) não possuía um sistema de monitoramento em saúde integrado capaz de analisar dados epidemiológicos para, em seguida, fornecer informações e orientar de forma mais assertiva a tomada de decisão por parte da gestão. O gerenciamento de informações encontrava-se fragmentado ou impreciso, gerando fragilidade nos dados ou tomadas de decisões ineficientes.

A descentralização dessas informações provocava consequências críticas na gestão da saúde pública, visto que havia um tempo acima do desejado entre receber, tratar e processar para ter a informação apta ao uso.

SOLUÇÃO

O IntegraSUS é uma plataforma de transparência de dados da saúde pública do Estado do Ceará. Ela integra, em tempo oportuno, as informações do Sistema Único de Saúde do Ceará, ao implementar e operacionalizar, tratar, compilar, analisar e disponibilizar informações.

O monitoramento e o gerenciamento dos indicadores, por meio da integração dos sistemas utilizados pela Sesa, contemplam informações dos 184 municípios cearenses, envolvendo as áreas administrativa, financeira e de planejamento, bem como a integração com sistemas de outros órgãos.

RESULTADOS E IMPACTOS

Com a integração dos sistemas e a transparência das informações em saúde, a Sesa fomenta a melhoria da gestão entre as secretarias de saúde municipais e as unidades gestoras/prestadoras de serviço para os usuários do SUS Ceará.

O ambiente tecnológico contribui para ações estratégicas dos gestores, subsidiando, de forma antecipada, as respostas para as crises e otimizando a comunicação com a imprensa. Além disso, é possível:

- » Acompanhar os contratos de gestão (monitoramento e avaliação das metas pactuadas).
- » Aumentar o número de atendimentos, qualificar os profissionais e os processos de trabalho e reduzir custos.
- » Possibilitar convênios de cooperação entre a Sesa e as secretarias municipais do Estado.
- » Monitorar em tempo oportuno indicadores de saúde da Sesa.
- » Monitorar e avaliar dados de vigilância sanitária.
- » Garantir o nível de excelência da prestação de serviço, assim como a qualidade e a humanização no atendimento.
- » Aplicar recursos disponíveis de maneira mais assertiva, para qualificar a

assistência, aumentar a produção e a produtividade hospitalar.

- » Economizar o licenciamento de *softwares*, considerando o uso de ferramentas *open sources* e o desenvolvimento da biblioteca de visualização.
- » Otimizar os indicadores estratégicos monitorados pelo IntegraSUS.

O IntegraSUS torna públicos dados relevantes, que passaram a ser melhor geridos. Além disso, proporciona também:

- » Transparência da informação.
- » Maior acessibilidade de pacientes.
- » Monitoramento em tempo real das filas dos hospitais.
- » Acompanhamento de processos/convênios.
- » Agilidade na comunicação paciente/instituição.

CASO 4

Plataforma Big Data e Inteligência Artificial para Segurança Pública

PROBLEMA

De acordo com os últimos 30 anos de pesquisa sobre policiamento, há um consenso de que, para que as abordagens de policiamento sejam efetivas, elas devem ser focadas e usadas de modo sistemático através da abordagem de solução de problemas.

Há diversas abordagens utilizadas para policiamento, algumas não muito efetivas no que se refere à redução de crimes. As abordagens consideradas mais bem-sucedidas ou mais promissoras, nas quais a análise visual de crimes constitui um elemento fundamental são: policiamento baseado em *hotspots* e policiamento orientado para o problema (POP).

Hotspots são regiões geográficas que possuem desproporcionalmente mais crimes do que outras regiões em uma dada jurisdição. O policiamento baseado em *hotspots* consiste em priorizar a alocação de recursos nessas regiões. A base de evidências para o policiamento baseado em *hotspot* é uma das mais sólidas.

Os resultados mostram que quando a polícia se concentra em pequenas regiões geográficas com alto crime, consegue-se reduzir o crime nessas regiões. Em torno de **50%** das chamadas ou dos incidentes são tipicamente concentrados em menos de **5%** dos lugares (por exemplo, endereços ou segmentos de rua) em uma cidade.

O policiamento orientado para o problema consiste na ideia de que a polícia deve assumir um papel proativo, em vez de reativo, ao identificar, entender e responder a problemas (não apenas incidentes individuais) em suas comunidades. Vários estudos científicos concluíram que o essa abordagem possui um impacto, em geral modesto, mas estatisticamente significativo no enfrentamento do crime e da desordem. Outros estudos menos rigorosos também encontraram resultados positivos na redução do crime.

SOLUÇÃO

Considerando esses estudos científicos, a Secretaria da Segurança Pública e Defesa Social do Governo do Ceará (SSPDS) desenvolveu o projeto “Inteligência Científica e Tecnológica na Segurança Pública”.

Esse projeto resultou na criação de uma plataforma para segurança pública baseada nas tecnologias de *big data* e inteligência artificial. Essa plataforma é composta de várias soluções tecnológicas que serviram para a implementar estratégias de prevenção, investigação e policiamento ostensivo mais efetivas.

As plataformas são compostas por diversas ferramentas. Primeiramente, o Cerebrum, uma ferramenta de busca que integra mais de 60 bases de dados da segurança pública, oferece um ponto de acesso único para realizar análises avançadas sobre os dados integrados.

Como exemplo de análises, é possível acompanhar em tempo real a movimentação de viaturas policiais, acessar câmeras de monitoramento, geocodificar todos os eventos criminosos sobre o mapa do Estado do Ceará, entre outras análises. Segundo, a ferramenta CrimeWatcher (atualmente Status), cuja função é calcular os *hotspots* de eventos criminosos (manchas criminais) e fornecer diversas análises espaço-temporais desses eventos.

Outra ferramenta importante é o HumanNerd, que permite a coleta automática de dados a partir da leitura de boletins de ocorrência, usando um algoritmo sofisticado de inteligência artificial. Além disso, foram desenvolvidos diversos modelos de inteligência artificial para interpretar imagens de veículos e pessoas, fundamentais para a pronta resposta aos eventos criminosos.

dos homicídios e **63%** no roubo a veículos entre os anos de 2018 e 2020.

Em 2021, o número de homicídios no Ceará caiu **28%** no primeiro semestre deste ano em comparação a igual período de 2020. Essa foi a maior redução de homicídios do país no período, conforme os dados do Monitor da Violência, índice nacional de homicídios criado pelo Portal G1, com base de dados oficiais dos 26 estados e do Distrito Federal⁵⁶.

Outro resultado bastante relevante foi a capacidade de coletar e processar um grande volume de informações da segurança pública em tempo real. Dessa forma, os atores vinculados à segurança pública do Ceará podem ter acesso simplificado a uma grande quantidade de informações que ajudam na tomada de decisão, incluindo o próprio governador, que possui em sua sala um painel que reflète a situação da segurança pública.

O uso dessa plataforma e de novas estratégias de segurança pública estão permitindo a implementação de uma gestão baseada em evidências, que auxilia na tomada de decisões mais assertivas e orientada a dados.

RESULTADOS E IMPACTOS

O principal impacto alcançado pelo uso dessa plataforma, aliada às novas estratégias da SSPDS, foi a redução de **54%**

56 Confira a matéria "Ceará tem a maior queda no número de homicídios no Brasil no primeiro semestre de 2021": <<https://g1.globo.com/ce/ceara/noticia/2021/08/20/ceara-tem-a-maior-queda-no-numero-de-homicidios-no-brasil-no-primeiro-semester-de-2021.ghtml>>.

