ANNUAL REVIEWS

Revisão Anual de Direito e Ciência Social

Algoritmos e Tomada de Decisão no Setor Público

Karen Levy, Kyla E. Chasalow e Sarah Riley

Departamento de Ciência da Informação, Universidade Cornell, Ithaca, Nova York 14853, EUA; e-mail: karen.levy@cornell.edu

Baixado de www.annualreviews.org. Convidado (guest) IP: 191.177.164.206 Em: Ter, 17 Mar 2026 21:02:23

ANNUAL REVIEWS CONNECT

www.annualreviews.org

- Download figures
- Navigate cited references
- Keyword search
- Explore related articles
- Share via email or social media

Rev. Anual de Direito e Ciênc. Soc. 2021. 17:309–34

Publicado pela primeira vez como Revisão Antecipada em 27 de julho de 2021

A Revisão Anual de Direito e Ciência Social está online em lawsocsci.annualreviews.org

https://doi.org/10.1146/annurev-lawsocsci-041221-023808

Copyright © 2021 por Annual Reviews. Todos os direitos reservados

Palavras-chave

algoritmos, tecnologia, governo, política, previsão, responsabilidade

Resumo

Este artigo examina o uso de sistemas algorítmicos para apoiar a tomada de decisão no setor público. Os governos adotam, adquirem e utilizam sistemas algorítmicos para apoiar suas funções em diversos contextos — incluindo justiça criminal, educação e concessão de benefícios — com consequências importantes para a responsabilidade, privacidade, desigualdade social e participação pública na tomada de decisão. Exploramos as implicações sociais dos sistemas algorítmicos municipais em várias etapas, incluindo formulação de problemas, aquisição de tecnologia, implantação e avaliação. Destacamos várias questões em aberto que exigem mais pesquisas empíricas.

1. INTRODUÇÃO

Notícias de destaque sobre algoritmos frequentemente vêm do setor privado. Algumas reportagens são elogiosas e otimistas, descrevendo o potencial dos algoritmos para transformar a sociedade para o bem; outras são menos generosas, focando na substituição de empregos, ameaças à privacidade e autonomia, ou no papel dos algoritmos na promoção de desinformação e polarização nas redes sociais. Esses fenômenos capturam a atenção pública e influenciam como as pessoas percebem os algoritmos, moldando políticas, legislações e agendas de pesquisa. Os sistemas algorítmicos do setor público ficam um pouco fora desse foco, especialmente no nível local. Esses sistemas influenciam funções governamentais que vão desde as rotineiras (coleta de lixo) até as de alto risco (decisões judiciais pré-julgamento). No entanto, apesar de seu alcance e impacto, o trabalho empírico sobre o impacto e a eficácia dos sistemas algorítmicos do setor público ainda é incipiente.

Ao sintetizar trabalhos empíricos e analíticos existentes sobre essas tecnologias controversas, buscamos fornecer aos pesquisadores um roteiro para orientar a investigação empírica sobre o uso público de sistemas algorítmicos—particularmente no nível local e com foco nos Estados Unidos. Destacamos vários locais estratégicos de pesquisa para investigar algoritmos do setor público, organizando-os por estágios no ciclo de vida do desenvolvimento e implantação do algoritmo. Na prática, o desenvolvimento e a adoção de sistemas algorítmicos são processos iterativos, improváveis de ocorrer em um conjunto de etapas distintas e ordenadas. A abordagem de ciclo de vida que propomos é uma ferramenta analítica, e não uma representação fiel da realidade. Dito isso, os estágios que descrevemos destacam temas comuns e questões em aberto, ao mesmo tempo em que identificam decisões e suposições que provavelmente fornecerão detalhes empíricos e vantagem analítica para o cientista social.

Prosseguimos da seguinte forma. Na Seção 2, introduzimos nossa investigação, observando ambiguidades de definição e escopo que complicam as discussões sobre algoritmos e fornecendo um panorama sobre algoritmos no setor público. Nas Seções 3 a 8, destacamos tensões, restrições e dinâmicas-chave que surgem em diferentes estágios da vida do algoritmo do setor público. A Seção 3 trata da formulação do problema—a identificação de um problema social a ser abordado e sua tradução em termos algorítmicos. A Seção 4 considera o papel dos dados—sua aquisição, integração e papel tanto como recurso quanto como restrição. Na Seção 5, abordamos como os sistemas são adquiridos por meio de compras de fornecedores comerciais ou outros meios. A Seção 6 considera abordagens para a governança de algoritmos do setor público, desde requisitos de transparência até proibições totais. Na Seção 7, tratamos das interseções com o trabalho de linha de frente e respostas de comunidades e organizações cívicas. A Seção 8 considera como os sistemas algorítmicos são mantidos, avaliados, reconsiderados, aprimorados ou desmantelados ao longo do tempo. A Seção 9 conclui.

2. O ALGORITMO DO SETOR PÚBLICO

2.1. Algoritmos e Sistemas Algorítmicos

"Algoritmo" é uma palavra carregada hoje em dia, usada para invocar uma tecnologia poderosa, às vezes misteriosa, que orienta decisões e molda nossas experiências on-line e off-line (Kitchin 2017, Ziewitz 2016). Em um nível, um algoritmo é simplesmente um "conjunto de regras que definem precisamente uma sequência de operações" (Stone 1971, p. 4). Por exemplo, o aprendizado de máquina, tema de muitas discussões contemporâneas sobre algoritmos, depende de algoritmos de treinamento para aprender um padrão entre entradas e saídas de interesse. Na verdade, dois algoritmos estão envolvidos aqui: o algoritmo de treinamento, que especifica um procedimento para aprender a partir dos dados, e um modelo final aprendido, que recebe uma entrada (por exemplo, detalhes médicos, padrões de tráfego, dados de prisões), aplica uma regra e produz saídas (frequentemente, uma previsão) (Kleinberg et al. 2018, Lum & Chowdhury 2021). "Algoritmo" às vezes é usado de forma ampla para se referir a um ou ambos.

Outros termos complicam ainda mais a questão, incluindo inteligência artificial; aprendizado de máquina; automação; e variantes de análise preditiva, ciência de dados e estatística. Embora esteja fora de

Baixado de www.annualreviews.org. Convidado (guest) IP: 191.177.164.206 Em: Ter, 17 Mar 2026 21:02:23

embora não haja espaço aqui para aprofundar as distinções mais sutis entre esses termos, vale notar que eles são usados de forma um tanto intercambiável no contexto de tecnologias intensivas em dados—ainda que, por vezes, com conotações muito diferentes. Inteligência artificial, ou IA, é um termo guarda-chuva para várias tecnologias computacionais que visam realizar tarefas entendidas como exigindo algum tipo de cognição complexa. Às vezes, a expressão pode evocar imagens futuristas de sistemas com habilidades semelhantes às humanas, talvez à custa de atenção às tecnologias atuais que são incapazes de tal autonomia, mas ainda assim poderosas (Krafft et al. 2019). Da mesma forma, aprendizado de máquina, por mais impressionante que possa soar, descreve o que são fundamentalmente técnicas estatísticas para ajustar (às vezes modelos muito complicados) aos dados. No entanto, como nem todos os algoritmos podem ser classificados como aprendizado de máquina, e até mesmo algoritmos relativamente simples—including fórmulas explícitas baseadas em regras (Citron 2008)—podem desempenhar papéis significativos em contextos de políticas públicas, alguns preferem falar em sistemas automatizados de tomada de decisão. Este termo foca mais na função política que esses sistemas desempenham—auxiliar ou substituir julgamentos humanos—do que nos procedimentos técnicos que os fundamentam (Molnar & Gill 2018, Spielkamp 2019).

Baixado de www.annualreviews.org. Convidado (guest) IP: 191.177.164.206 Em: Ter, 17 Mar 2026 21:02:23

Decisões de definição são mais do que disputas semânticas. Elas afetam a comunicação entre pesquisadores e formuladores de políticas, moldam a compreensão pública e determinam quais sistemas estão sujeitos à regulação (Krafft et al. 2019, Young et al. 2019b). Ou seja, a terminologia pode ser, em si, uma questão de política substantiva e divisiva, não apenas uma questão de escopo processual. Por exemplo, em 2018, a cidade de Nova York convocou uma Força-Tarefa de Sistemas Automatizados de Decisão—um painel de especialistas da academia, indústria, sociedade civil e agências municipais encarregado de fornecer recomendações ao governo da cidade sobre o uso e a governança desses sistemas. O painel passou seus primeiros meses sem conseguir chegar a um consenso sobre quão amplamente deveria definir seu objeto de estudo (Richardson 2019), questionando se (por exemplo) o uso de uma planilha Excel qualificava como um “sistema automatizado de decisão”. Esse esforço de definição foi ainda mais complicado pela relutância da cidade em fornecer dados sobre sistemas algorítmicos atualmente em uso, tornando os debates sobre terminologia e escopo necessariamente abstratos (Cahn 2019). Os desafios do painel destacam tanto a dificuldade de construir consenso sobre a definição e o escopo de sistemas algorítmicos quanto as barreiras potenciais para fundamentar tais discussões na prática atual.

Além disso, a terminologia do algoritmo traz consigo a capacidade de criar uma distância retórica entre os formuladores de políticas e os resultados das políticas—uma distância que pode ser usada para desviar a responsabilidade por decisões impopulares. Como descrevem Lum & Chowdhury (2021), funcionários do Stanford Medical Center foram rápidos em culpar um “algoritmo muito complexo” que “claramente não funcionou direito” quando seu sistema de alocação de vacinas contra a Covid-19 priorizou administradores de alto escalão em vez de trabalhadores da linha de frente. O algoritmo, na verdade, funcionou corretamente—no sentido de que implementou fielmente as regras determinadas pelos tomadores de decisão. Quando os resultados das políticas encontram desaprovação pública, desviar a culpa para o algoritmo serve a uma função política, permitindo que os tomadores de decisão salvem (ao menos em parte) sua imagem.

Aqui, usamos algoritmo ou algorítmico como abreviação geral para tecnologias que dependem de técnicas de aprendizado de máquina ou regras explicitamente programadas para informar ou executar ações governamentais. Como muitos estudiosos enfatizam, esses processos técnicos operam dentro de contextos sociais e institucionais, compondo sistemas sociotécnicos (Kitchin 2017, Selbst et al. 2019, Spielkamp 2019). Nesse sentido, a pesquisa em ciências sociais sobre algoritmos no setor público geralmente foca nas dimensões políticas, legais, econômicas e sociais dos sistemas algorítmicos, em vez de apenas em seus componentes computacionais. Seguimos essa abordagem aqui.

2.2. Raízes e Crescimento dos Sistemas Algorítmicos no Setor Público

Focamos nossa análise nos governos locais dos Estados Unidos devido ao escopo de suas responsabilidades e ao seu alto grau de influência na vida pública. Os governos locais têm sido

chamados de “motores do federalismo”, pois “lideram o gasto de fundos para fins domésticos e empregam o maior número de pessoas para alcançar esses objetivos” (Advis.Comm.Int. Relat. 1982, p. 11). São o principal provedor de serviços ao público, supervisionando segurança pública, saneamento, estacionamento e serviços educacionais dos quais muitos dependem diariamente. Dado o volume e a abrangência dos serviços administrados pelas localidades, elas são terreno fértil para sistemas automatizados emergentes. Se os estados oferecem 50 laboratórios de federalismo, os municípios oferecem mais de 90.000 (Hogue 2013).

Embora o funcionamento diário do governo local afete a todos, os constituintes mais vulneráveis são os mais frequentemente expostos e diretamente impactados por sistemas algorítmicos municipais, muitas vezes sem aviso prévio ou opção de exclusão. Indivíduos com menos acesso a recursos materiais, especialmente pessoas negras, têm maior probabilidade de interagir com agências governamentais — tanto porque seu status marginalizado lhes oferece menos opções privadas quanto porque o Estado controla essas populações de forma proativa e desproporcional (Barabas 2020, Bridges 2017, Eubanks 2018, Madden et al. 2017). Assim, os sistemas algorítmicos municipais têm influência desproporcional sobre a vida dos mais vulneráveis.

Embora foquemos no recente crescimento dos sistemas algorítmicos do setor público, esses sistemas não surgiram do nada. Eles têm raízes em uma longa história de coleta de dados governamentais; análise estatística; e o papel crescente dessas ferramentas na administração pública, formulação de políticas, prestação de serviços e gestão populacional. Seus afluentes incluem infraestruturas de dados fundamentais — sistemas de registro civil, censos, pesquisas, administração burocrática — bem como métodos estatísticos e científicos sociais frequentemente desenvolvidos em estreita ligação com o planejamento governamental e a prestação de serviços (Bouk 2015, Breckenridge & Szreter 2012, Didier 2020, Salsburg 2001, Scott 1998).

De modo geral, os sistemas algorítmicos do setor público emergem de um zeitgeist de formalização, racionalização e automação no governo que está presente pelo menos desde a Segunda Guerra Mundial. O ideal de tomada de decisão governamental sistematizada e quantificada tem raízes na pesquisa operacional, com sua ambição de desenvolver técnicas neutras, científicas e objetivas para analisar situações e determinar cursos de ação ideais (Harcourt 2018). Um dos primeiros atores foi a RAND Corporation, um think tank fundado em 1948 para fornecer pesquisas à Força Aérea dos EUA. A RAND posteriormente aplicou seus métodos em cidades de todo o país para “reestruturar a burocracia municipal em linhas racionais e científicas” em vez da “confusão reacionária” que caracterizava as respostas típicas dos governos municipais aos problemas (Flood 2010, pp. 111, 116). A metade do século XX também viu formas iniciais de automação sob a forma de sistemas especialistas, destinados a automatizar a tomada de decisão codificando a expertise do domínio em sistemas de regras lógicas (Hadden & Feinstein 1989). Ideias relacionadas se refletem no impulso por políticas e tomadas de decisão baseadas em evidências (Engstrom et al. 2020, Stevenson 2018) e no crescimento de práticas atuariais que dependem de previsão estatística em vez de julgamento clínico desenvolvido por experiência e treinamento (Dawes et al. 1989, Simon 1988).

Se o impulso por uma tomada de decisão formalizada e orientada por dados no governo não é totalmente novo, o que, então, distingue os algoritmos atuais? Uma resposta simples é sua onipresença. Novas ferramentas algorítmicas estão mais “profundamente incorporadas” aos processos governamentais do que esforços anteriores, “levando as novas ferramentas de governança algorítmica ao centro do poder coercitivo e (re)distributivo do Estado” (Engstrom et al. 2020, p. 11). O que possibilita essas ferramentas é o aumento da coleta e integração de dados, que apoia uma capacidade crescente de aproveitar a “inteligência coletiva” sobre a prestação de serviços públicos e a vida dos cidadãos (Vogl et al. 2020). Esses dados também são processados de novas maneiras: técnicas de aprendizado de máquina se caracterizam pelo foco no desenvolvimento de modelos para previsão em vez de explicação (Breiman 2001, Hofman et al. 2017, Kleinberg et al. 2015) e, ao fazer isso sem codificar regras explicitamente, tornam-se mais difíceis de interpretar e explicar (Caruana et al. 2015, Lipton 2018, Selbst & Barocas 2018).

Baixado de www.annualreviews.org. Convidado (guest) IP: 191.177.164.206 Em: Ter, 17 Mar 2026 21:02:23

Outras duas características notáveis frequentemente atribuídas aos sistemas algorítmicos atualmente são a escalabilidade e a personalização. Escalabilidade refere-se à forma como os algoritmos podem ser usados para influenciar “decisões de maneira uniforme e abrangente”, replicando sequências de tomada de decisão (e quaisquer suposições ou falhas que incorporam) em diferentes contextos e ao longo do tempo (Brauneis & Goodman 2018, p. 129). Personalização implica uma expansão dos usos tradicionais de dados para o planejamento de políticas amplas com base em dados agregados e para usos de dados que “categorizam cidadãos, alocam serviços e prevêem comportamentos” de forma individual (Dencik et al. 2018, p. 3). A expectativa é que isso permita uma melhor alocação de recursos, menores custos e mais intervenções preventivas.

Embora as razões pelas quais os governos locais adotam esses sistemas estejam situadas em histórias mais amplas e tendências atuais, os governos locais ocupam uma posição distinta, com restrições e necessidades particulares que tornam os sistemas algorítmicos especialmente atraentes. Embora os governos locais variem consideravelmente em termos de financiamento, estrutura organizacional e prioridades políticas, algumas de suas características comuns ajudam a explicar a adoção de sistemas algorítmicos no nível local.

Primeiramente, os governos locais têm menos poder financeiro em relação a outros níveis de governo. Eles dependem de um mosaico complexo e em constante mudança de financiamentos dos governos estaduais e federais — complexo em parte porque os valores mudam ao longo do tempo, mas também porque diferentes categorias de financiamento (por exemplo, subsídios categóricos, subsídios em bloco) podem vir com condições complicadas. A proporção da receita local proveniente de transferências intergovernamentais tem diminuído constantemente há décadas (Berman 2019, p. 111; Wildasin 2010). Além disso, muitos estados restringem como e quanto de receita os governos locais podem arrecadar por conta própria, por meio de medidas como limitações de impostos e despesas. Essas medidas, como a Proposição 13 da Califórnia, reduzem os orçamentos locais e forçam as localidades a depender de métodos alternativos para gerar receita, muitas vezes por meio de multas e taxas impopulares politicamente. Assim, apesar de suas muitas responsabilidades, os governos locais enfrentam a dupla dificuldade de reduções de auxílio e condições onerosas de subsídios. Isso torna altamente desejáveis os sistemas modernos que prometem realizar mais com menos recursos.

Além dos fatores financeiros, os governos locais enfrentam a pressão de fornecer serviços públicos de alto impacto e alta visibilidade. As localidades frequentemente tomam decisões críticas, pelas quais os cidadãos podem atribuir responsabilidade a um único órgão ou até mesmo a um indivíduo: juízes decidem quem liberar ou deter antes do julgamento, os bombeiros priorizam edifícios para inspeção de violações de código e as autoridades de habitação distribuem assistência de aluguel limitada. As agências precisam tomar decisões difíceis no melhor interesse de seus cidadãos, enquanto navegam por restrições orçamentárias e políticas. Incorporar o apoio à tomada de decisão algorítmica pode, portanto, ter um apelo especial como uma forma de as localidades tentarem lidar com essas tarefas de maneira objetiva e precisa.

3. FORMULANDO O PROBLEMA

3.1. Definição do Problema

Soluções algorítmicas são, aparentemente, aplicadas a problemas. Mas como os governos passam a identificar problemas como necessitando de soluções algorítmicas? O ato de reconhecer, definir e especificar o problema que um algoritmo pretende resolver é, em si, uma questão profundamente política.

Estudiosos de políticas públicas há muito caracterizam certos problemas como "espinhosos" — complexos e confusos, sem limites claros ou soluções óbvias, e sujeitos a interesses conflitantes. Pobreza e mudança climática são exemplos clássicos. Os formuladores de políticas frequentemente "recortam" componentes tratáveis desses problemas sistêmicos que podem ser "resolvidos" (Churchman 1967). Isso envolve simplificar o problema espinhoso em algo facilmente mensurável e politicamente aceitável. É aí que os algoritmos ganham destaque. Como Kaplan (1964, p. 303) afirmou: "Tendemos a formular nossos problemas

Baixado de www.annualreviews.org. Convidado (guest) IP: 191.177.164.206 Em: Ter, 17 Mar 2026 21:02:23

de tal forma que pareça que as soluções para esses problemas exigem exatamente o que já temos à disposição."

O processo de converter um problema sistêmico em um problema algorítmico tem importantes implicações para quais problemas são abordados e como são compreendidos. Alguns problemas não se prestam facilmente a soluções algorítmicas — ou, ao forçá-los a se encaixar em tal solução, a natureza do problema muda consideravelmente. A parte de um problema que pode ser tratada de forma algorítmica pode roubar o oxigênio político de reformas não algorítmicas ou aliviar preocupações de que "algo precisa ser feito", mas sem abordar as causas raízes (Abebe et al. 2020a) — o que Churchman (1967) chamou de "domar o rugido" de um problema perverso. Por exemplo, Eubanks (2018, p. 197) observa que, antes de cidades e estados implementarem sistemas algorítmicos para gerenciar programas de combate à pobreza, "a nação estava fazendo perguntas difíceis: Qual é a nossa obrigação uns com os outros em condições de desigualdade? Como recompensamos o cuidado?" Ela aponta para os algoritmos como "reformulando esses grandes dilemas políticos como questões mundanas de eficiência e engenharia de sistemas: Como melhor combinamos necessidade e recurso? Como eliminamos fraudes e desviamos os inelegíveis? Como fazemos o máximo com o mínimo de dinheiro?" (Eubanks 2018, pp. 197–98). Os sistemas algorítmicos não só falharam em abordar as causas raízes da pobreza, como também restringiram o que era considerado politicamente possível e alteraram os valores usados para avaliar essas opções.

Não surpreendentemente, a definição de um problema depende em grande parte de quem está fazendo essa definição. Os atores da comunidade podem conceituar um problema através da lente de uma questão social subjacente (pobreza, falta de moradia, transporte público insuficiente). Dentro do governo, os problemas podem ser formulados como ineficiências operacionais, como redução de fraudes ou economia de custos. Enquadramentos alternativos podem ou não ser articulados explicitamente. O impulso para soluções algorítmicas no nível local também é, às vezes, moldado por outros níveis de governo, uma dinâmica que Solow-Niederman et al. (2019) chamam de "federalismo algorítmico". Por exemplo, após os ataques de 11 de setembro, o governo federal expandiu significativamente o apoio financeiro à aplicação da lei local em um esforço para combater o terrorismo, estimulando o desenvolvimento de programas de vigilância por departamentos de polícia locais (Crump 2016). Os problemas podem se formar em parte como uma função dos recursos disponíveis para enfrentá-los.

3.2. Formalização do Problema

Abordar problemas sociais por meio de métodos algorítmicos exige um passo adicional de tradução, no qual os conceitos são operacionalizados em termos formais e mensuráveis. Transformar um problema social em um problema algorítmico requer a formalização explícita dos insumos e dos objetivos desejados. Ao fazer isso, cria-se necessariamente "uma lente particular para entender o que é o problema" (Abebe et al. 2020a, p. 254).

O trabalho de tradução da formulação do problema "faz a violência necessária ao mundo que [ele] tenta modelar, mas também fornece maneiras acionáveis e inovadoras de abordar problemas complexos" (Passi & Barocas 2019, p. 47). A necessidade de formalizar um problema em termos algorítmicos implica em concessões: embora possa reduzir e simplificar excessivamente problemas complexos, também pode ajudar na articulação precisa de objetivos gerais de políticas e tornar claras diferentes compreensões de um conceito. Com transparência e deliberação suficientes, os interessados podem aproveitar a formalização exigida por sistemas algorítmicos como um meio de perguntar (por exemplo) o que exatamente queremos dizer com risco, ou esclarecer exatamente quais fatores devem determinar a ordem de alguma alocação. A formalização pode, em seu melhor cenário, tornar julgamentos políticos vagos em algo concreto e fornecer pontos de contato para uma defesa legal e política focada (Abebe et al. 2020a). No entanto, o processo de formalizar um problema em termos passíveis de medição, modelagem estatística e dados disponíveis envolve decisões que, se não forem consideradas cuidadosamente, podem mudar a natureza do problema e os resultados que se seguem.

Baixado de www.annualreviews.org. Convidado (guest) IP: 191.177.164.206 Em: Ter, 17 Mar 2026 21:02:23

Considere o desenvolvimento da Ferramenta de Triagem do Vale do Silício no Condado de Santa Clara, Califórnia. A ferramenta utiliza dados de sete agências do condado para prever quais pessoas em situação de rua provavelmente se tornarão "usuários de alto custo dos serviços públicos no próximo ano" (Toros & Flaming 2018, p. 119; Early 2016). O objetivo é priorizar essas pessoas para moradia permanente e, assim, reduzir custos. Crucialmente, a escolha de usar o risco de incorrer em altos custos como resultado para essa ferramenta é uma decisão ética. Pode ser uma decisão justificável: alto custo pode estar correlacionado com alta necessidade, e a redução de custos pode liberar recursos para outras pessoas e programas. Mas a ferramenta pode subpriorizar jovens que custam menos agora, mas podem custar e sofrer mais a longo prazo — e priorizar altos custos pode gerar resultados desiguais se mulheres, imigrantes e minorias tendem a ter custos mais baixos, inclusive por razões não relacionadas ao seu nível de necessidade (Harkinson 2016, Obermeyer et al. 2019). Embora descrita como "uma ferramenta justa e objetiva para triagem" por seus criadores, ela só pode ser assim, na melhor das hipóteses, dado um conjunto de decisões sobre quais resultados (objetivamente) prever (Toros & Flaming 2018, p. 118). Por outro lado, se essas decisões não estiverem bem estabelecidas antes de construir tal modelo, então os criadores do modelo implementam implicitamente uma escolha de política por meio da escolha da variável alvo. (Para um modelo diferente de alocação de assistência habitacional — usando a probabilidade de retorno ao sistema de abrigos como resultado — veja Kube et al. 2019.)

Preocupações semelhantes se aplicam à escolha das variáveis a serem usadas como entradas para uma tarefa de previsão ou classificação. Embora a precisão preditiva seja uma preocupação primária aqui, houve muita discussão recente sobre como lidar com o potencial impacto discriminatório do uso de certas características (mesmo quando contribuem para a precisão) e como responder e mitigar essas questões (Barocas & Selbst 2016). Uma preocupação relacionada é se as escolhas das variáveis de entrada foram feitas sob uma "abordagem baseada no déficit" que se baseia principalmente em dados sobre fatores negativos — indicações de alto risco — sem também incorporar indicadores positivos de menor risco (Brown et al. 2019).

A abordagem do déficit e as preocupações sobre impacto desigual podem ser descritas como um tipo de viés. Muitas vezes se alerta que dados e modelos podem ser tendenciosos — mas a grande questão é: de que maneira? Todos os dados são não representativos de alguma forma (por exemplo, em relação a alguns alvos ou para algumas variáveis); todos os dados têm limitações (Chasalow & Levy 2021). A tarefa passa a ser tornar essas limitações explícitas e investigar se elas importam para o objetivo em questão. Por exemplo, é importante distinguir a questão do viés de seleção daquela do viés histórico (Suresh & Guttag 2020). Enquanto o primeiro pode ser resolvido por um melhor esquema de coleta de dados, o segundo surge mesmo com medição perfeita, quando os dados capturam realidades passadas que são, de alguma forma, indesejáveis. Esses dados são tendenciosos no sentido de que previsões feitas a partir deles tenderão a replicar essas realidades passadas. A própria escolha de usar dados históricos (e, de certo modo, todos os dados são históricos) pode ser vista como uma escolha de política, pois expressa um desejo ou ao menos uma disposição de replicar aspectos do passado para alcançar objetivos de política. Para possibilidades futuras e sistemas não testados, não temos dados.

As escolhas de política se estendem ainda ao desenvolvimento do modelo. Mulligan & Bamberger (2019) dão como exemplo o uso de modelos sismológicos para policiamento preditivo (a intuição sendo que o crime se propaga como réplicas de terremotos), mas nem toda modelagem se baseia em tais analogias intuitivas. Abordagens baseadas em regressão e modelos de aprendizado de máquina mais complexos podem ter bom desempenho, embora sejam difíceis de interpretar (Lipton 2018). Em qualquer caso, aprender um modelo a partir de dados exige escolhas sobre como medir o ajuste do modelo ou o desempenho preditivo. É bem conhecido que tais escolhas envolvem trade-offs — por exemplo, entre falsos positivos e falsos negativos (Kleinberg et al. 2017). Uma literatura crescente sobre justiça algorítmica explora maneiras de impor restrições que exigem que modelos atinjam diferentes critérios de justiça em nível de grupo ou individual (Chouldechova & Roth 2018, Mitchell et al. 2021, Mulligan et al. 2019). Em última análise, projetistas e governos também devem fazer uma escolha tão básica que pode ser negligenciada: usar ou não um modelo preditivo. Às vezes, mesmo o melhor modelo entre os desenvolvidos pode ter um nível inaceitavelmente alto

Baixado de www.annualreviews.org. Convidado (guest) IP: 191.177.164.206 Em: Ter, 17 Mar 2026 21:02:23

nível absoluto de erro (Salganik et al. 2020). Além disso, em alguns casos, pequenas diferenças na formalização do problema podem levar a resultados de políticas muito diferentes. Um exemplo surge em Abebe et al. (2020b), explorando possíveis alocações de subsídios destinados a prevenir que interrupções financeiras súbitas (choques de renda) causem ruína financeira para famílias. Embora o objetivo geral da política (prevenir resultados financeiros adversos para famílias) e o mecanismo da política (fornecimento de subsídios) sejam consistentes, formalizações plausíveis ligeiramente diferentes do problema podem resultar em ordens completamente invertidas de quem deve receber assistência.

Problemas de alocação, em particular, levantam outra distinção: embora muitos sistemas algorítmicos tenham como alvo uma tarefa de previsão ou classificação onde faz sentido considerar medidas de erro, nem todos os algoritmos são formados dessa maneira. Por exemplo, Lee et al. (2019) desenvolveram um algoritmo para alocar doações para uma organização sem fins lucrativos de distribuição de alimentos, agregando as preferências de uma variedade de partes interessadas. Nesse caso, o algoritmo é construído para refletir as escolhas das partes interessadas sobre quais fatores priorizar, e não há uma alocação correta verdadeira para compará-lo. Ou seja, um algoritmo pode tentar prever com precisão um resultado medido (por exemplo, custos incorridos por pessoas em situação de rua), mas também pode implementar diretamente um conjunto de preferências de políticas (Lum & Chowdhury 2021).

Baixado de www.annualreviews.org. Convidado (guest) IP: 191.177.164.206 Em: Ter, 17 Mar 2026 21:02:23

4. OBTENÇÃO DE DADOS

4.1. Dados como Recurso e Restrição

O ciclo de vida do algoritmo governamental começa com um problema ou tarefa. Também começa com dados. Os governos locais são grandes geradores de dados sobre a vida diária de indivíduos e comunidades. Esses dados assumem muitas formas (estruturados versus não estruturados, gerados por humanos versus máquinas, entre outros), podem residir em diferentes sistemas tecnológicos e podem ou não ser produzidos com análise posterior em mente. A produção de dados é moldada de dentro para fora e de fora para dentro—por infraestruturas tecnológicas, procedimentos organizacionais, leis e políticas, além do julgamento dos funcionários e contingências externas (Ku & Gil-Garcia 2018). Fundamentalmente, os dados que os governos coletam ou podem acessar moldam e refletem os tipos de problemas que os governos estão dispostos e aptos a enfrentar. Os dados possibilitam, mas também restringem, e essa dinâmica se manifesta de várias maneiras.

A definição do problema frequentemente levanta uma tensão entre a mensurabilidade de uma variável e sua relação com o verdadeiro resultado de interesse. Mensurabilidade e validade de construto são preocupações comuns em toda a ciência social empírica, onde as variáveis que podemos medir ou acessar são proxies para os verdadeiros resultados de interesse (Jacobs & Wallach 2021, Singleton & Straits 2005). Essas considerações são agudas no contexto de políticas públicas, onde programas implementados e as previsões que fazem podem determinar resultados de vida para pessoas reais. Algumas das medidas que mais nos importam em termos de resultados de vida são difíceis de medir e difíceis de isolar causalmente, tornando a avaliação de políticas difícil. Mesmo quando há dados extensos e de alta qualidade sobre a vida das pessoas, os resultados podem ser difíceis de prever no nível individual (Garip 2020, Salganik et al. 2020). Isso deve nos fazer refletir sobre políticas que centralizam a previsão em nível individual.

Dada a complexidade do mundo social, os construtores de modelos podem ser inclinados a medir resultados operacionais internos menos sujeitos ao acaso e ao ruído, mas mais distantes da consequência que realmente importa. Um exemplo surge no estudo de Flood (2010) sobre o fechamento de departamentos de bombeiros na cidade de Nova York: como medidas mais diretas de sucesso ou fracasso (fatalidades, ferimentos, danos à propriedade) são relativamente raras e impactadas por fatores fora do controle do departamento, os modeladores decidiram usar o tempo de resposta—o tempo entre o acionamento do alarme e a chegada do caminhão de bombeiros, que poderia ser medido dando cronômetros aos tenentes—como proxy (Flood 2010). Essa era, de muitas maneiras, uma medida imperfeita: a chegada do caminhão de bombeiros não corresponde perfeitamente ao sucesso no combate a incêndios, e bombeiros não colaborativos estavam dispostos a manipular os dados. Usando

uma medida de resultado interna simplificou o modelo e o tornou analiticamente tratável, mas ao custo de depender de suposições questionáveis sobre a validade do resultado.

Outro exemplo surge no estudo de Eubanks (2018) sobre a Ferramenta de Triagem Familiar do Condado de Allegheny, um modelo usado para prever o risco de abuso e negligência infantil a fim de priorizar intervenções por assistentes sociais. Aqui, o verdadeiro resultado de interesse é o mau-trato infantil, que é difícil de medir diretamente. Em vez disso, o Condado de Allegheny usou duas variáveis proxy como resultados: reencaminhamento comunitário (se o condado recebeu outro relatório sobre a criança) e colocação em família adotiva (que resultou de ações da própria agência e dos tribunais). Como resultado, a Ferramenta de Triagem Familiar previu comportamentos futuros da comunidade—cujos relatos de suspeita de abuso visavam desproporcionalmente famílias negras e birraciais—e da própria agência, em vez do resultado direto de interesse. Esses tipos de simplificações são comuns e muitas vezes necessários para conseguir construir um modelo, mas podem introduzir viés, espaço para manipulação e circularidade nos sistemas preditivos.

Baixado de www.annualreviews.org. Convidado (guest) IP: 191.177.164.206 Em: Ter, 17 Mar 2026 21:02:23

Mesmo que medidas relativamente diretas dos resultados de interesse estejam disponíveis, elas podem não estar disponíveis para tudo ou todos. Os dados coletados pelo governo geralmente estão sujeitos a viés de seleção: governos coletam mais dados sobre pessoas que têm maior grau de contato com programas e instituições governamentais (Chouldechova et al. 2018, Dettlaff et al. 2011). A Ferramenta de Triagem Familiar do Condado de Allegheny baseou-se em dados obtidos de programas públicos: escritórios de assistência social, Medicaid, serviços de saúde mental e a cadeia do condado, entre outros. No entanto, o condado não tinha acesso a dados sobre residentes que não tiveram contato com esses programas e instituições—residentes que poderiam, por exemplo, pagar por serviços de saúde privados. O resultado, nas palavras de Eubanks, é que “o modelo confunde criar filhos sendo pobre com criar filhos de forma inadequada” e é relativamente cego para alguns segmentos da população (Eubanks 2018, p. 158). Um modelo preditivo construído com dados históricos pode, em alguns casos, melhorar a tomada de decisão—mas também corre o risco de reproduzir ou agravar as desigualdades estruturais embutidas no sistema.

4.2. Integração de Dados e Dinâmicas Interagências

Sistemas algorítmicos frequentemente envolvem coordenação entre múltiplas agências públicas e geralmente são acompanhados por um impulso para integração de dados—ou seja, para o vínculo de dados de departamentos e serviços dentro ou entre municípios, ou de dados do setor público e privado, para criar uma visão mais integrada das operações governamentais (Malomo & Sena 2017). A integração de dados pode ser vista como um meio para previsões mais precisas; medidas preventivas mais precoces e direcionadas; ou melhor compreensão do funcionamento dos serviços públicos (Dencik et al. 2018, Toros & Flaming 2018, Young et al. 2019a). Esforços relacionados podem ser encontrados em propostas para cidades inteligentes, viabilizadas pelo uso de sensores (Green 2019) e portais de dados abertos que permitem que membros do público e do setor privado acessem dados do governo (Schrock 2016). A visão é de um governo mais eficiente, eficaz e conectado.

A integração de dados não é fácil. Os dados do setor público são frequentemente fragmentados, limitados por infraestruturas organizacionais e tecnológicas que restringem a capacidade de compartilhar dados entre organizações ou serviços. Calcular um índice de risco pré-julgamento, por exemplo, requer informações pessoais e demográficas (por exemplo, idade, emprego e situação habitacional), informações sobre acusações e histórico criminal, que podem vir de diferentes agências e bancos de dados. Avaliações de risco, portanto, exigem coordenação e cooperação contínuas para serem implementadas e podem ser dificultadas por agências com diferentes restrições e infraestruturas não projetadas para tais usos. Malomo & Sena (2017) descrevem como os “silos estruturais de dados” resultantes da fragmentação da atividade do setor público e da falta de padrões comuns de dados significam que o compartilhamento de dados pode exigir considerável esforço legal e tecnológico. No caso de parcerias público-privadas, a resistência à integração pode vir do medo das empresas de perder vantagem competitiva ao compartilhar dados proprietários (Young et al. 2019a).

Mas essas barreiras também podem servir como proteções (Hartzog & Selinger 2015, Surden 2007). A integração de dados levanta a perspectiva de vigilância de grande alcance e violações de privacidade, de sistemas de tomada de decisão coordenados que conectam áreas anteriormente distintas da vida das pessoas, e de um esbatimento entre o setor público e privado (Nissenbaum 2010, Ohm 2009). Por exemplo, no Condado de Pasco, Flórida, o Conselho Escolar do Condado de Pasco e o Departamento Estadual de Crianças e Famílias coletam dados sobre os alunos para apoiar objetivos educacionais e o bem-estar das crianças e famílias. Esses dados também foram, no entanto, combinados pelo Gabinete do Xerife de Pasco, juntamente com seus próprios registros, para "identificar jovens em risco de se tornarem infratores prolíficos" e manter uma lista de alunos que acredita serem "os mais em risco de cair em uma vida de crime" (Bedi & McGrory 2020). Neste caso, a integração de dados permitiu usos controversos bastante diferentes daqueles para os quais os dados foram originalmente coletados.

A integração de dados é ainda mais preocupante, dadas as desigualdades na disponibilidade de dados. Anteriormente, no contexto do caso de bem-estar infantil de Eubanks, referimo-nos a danos que decorrem da super-representação em dados administrativos, mas os benefícios e danos dos dados integrados do setor público não acompanham claramente a exclusão e inclusão em conjuntos de dados (Lerman 2013). A inclusão pode resultar em vigilância e serviço—muitas vezes ao mesmo tempo. A exclusão pode refletir e reforçar poder e recursos (por exemplo, famílias que podem pagar por serviços de saúde privados recebem menos escrutínio das agências de assistência social) ou desvantagem e marginalização (por exemplo, pessoas com menos meios de relatar problemas ou defender suas necessidades) (Crawford 2013). Pesquisadores também enfrentam essas dinâmicas. Como argumentam Johnson & Rostain (2020), os dados do setor público podem servir tanto como "vigilância" quanto como "holofote", sendo este último uma metáfora para o potencial de usá-los para monitorar instituições e expor desigualdades e violações de direitos.

5. ADQUIRINDO O SISTEMA

Definição do problema, aquisição de dados, construção do sistema—muitas vezes isso não acontece totalmente internamente. Em vez disso, muitas ferramentas algorítmicas são compradas de fornecedores privados que as desenvolvem e licenciam para os municípios. A aquisição é uma parte rotineira de como os governos obtêm tecnologias e muitas vezes é necessária, dada a limitação de acesso a especialistas e recursos internos. Normalmente, as agências emitem um pedido de proposta para solicitar ofertas de fornecedores. Em teoria, os pedidos de proposta permitem que as agências selecionem a melhor solução para um determinado problema e aumentem a transparência do governo, embora isso nem sempre aconteça na prática. Na verdade, os processos de aquisição apresentam desafios para a capacidade dos governos de manter o controle sobre como prestam serviços e governam. A aquisição de algoritmos levanta questões sobre responsabilidade (De quem é?), discricionariedade no design e na tomada de decisões (Quem tem?), e formulação de políticas (Quem faz? Quem sabe disso?). Em jogo, escrevem Brauneis & Goodman (2018, p. 109), está o potencial de uma "captura corporativa do poder público", deixando "o governo, que sozinho é responsável perante o público... esvaziado, burro e às escuras."

5.1. A Mentalidade de Aquisição

Nos Estados Unidos, as agências governamentais locais, estaduais e federais têm certo grau de liberdade em como implementam suas funções legalmente mandatadas. Quando as ações do governo atingem o nível de decisões políticas substantivas, as leis administrativas geralmente exigem que sigam procedimentos destinados a promover valores como transparência, responsabilidade, envolvimento público e tomada de decisão fundamentada (Adams 2018, Davidson 2016, Mulligan & Bamberger 2019). Tais procedimentos incluem exigir que as agências mantenham registros das deliberações e das evidências usadas para justificar decisões políticas, registros que podem posteriormente estar sujeitos a revisão judicial. A aquisição governamental, embora sujeita a regulamentação destinada a garantir concorrência justa entre fornecedores, muitas vezes é isenta de

Baixado de www.annualreviews.org. Convidado (guest) IP: 191.177.164.206 Em: Ter, 17 Mar 2026 21:02:23

esses requisitos administrativos. Mulligan & Bamberger (2019) alertam que as práticas atuais de aquisição de algoritmos criam o potencial para que escolhas políticas substantivas sejam incorporadas em decisões contratuais e deixadas para desenvolvedores terceirizados. A "mentalidade de aquisição", argumentam eles, posiciona a aquisição de algoritmos como uma questão de gestão interna. Ela vê os sistemas algorítmicos como fornecendo um novo mecanismo para cumprir uma "função bem definida" existente, em vez de métodos que possam mudar como "uma instituição entende e executa sua missão" (Mulligan & Bamberger 2019, p. 787).

Lembre-se do exemplo da força-tarefa da cidade de Nova York, que ficou presa à questão de saber se até mesmo o Excel se enquadrava na categoria de "sistemas de decisão automatizados". De fato, alguns sistemas governamentais que dependem de algoritmos fazem realmente parte da gestão interna—um sistema de cibersegurança, por exemplo—e a mentalidade de aquisição pode ser apropriada nesse caso. Isso não isenta tais sistemas de escrutínio, pois até decisões aparentemente banais sobre infraestrutura técnica podem ser impactantes (Kelion 2020). Mas a mentalidade de aquisição pode ser especialmente problemática para sistemas voltados ao público que afetam direitos e obrigações das pessoas, restringem a discricionariedade das agências e envolvem escolhas sobre como traduzir mandatos legais gerais em implementações e regras de aplicação específicas. Uma tarefa fundamental, então, passa a ser determinar quais sistemas são voltados para dentro e quais são suficientemente voltados ao público para que exijam mais transparência e justificativa pública (Mulligan & Bamberger 2019). Não traçamos essa linha aqui, mas o primeiro passo é delinear os tipos de escolhas políticas que podem tornar um sistema algorítmico mais do que apenas uma questão de implementar objetivos e políticas preexistentes.

A crítica à mentalidade de aquisição implica que, em contraste com decisões que apenas implementam um procedimento bem definido, decisões políticas afetam direitos, valores e resultados sociais, envolvem dilemas éticos ou determinam a alocação de recursos valiosos. A aquisição levanta questões importantes sobre quais partes interessadas têm poder para definir e formalizar um problema algorítmico e quais processos moldam essa participação. Uma onda de pesquisas recentes explorou qual papel processos participativos de tomada de decisão podem desempenhar para garantir que as prioridades públicas estejam bem refletidas na formulação do problema algorítmico desde o início (Lee et al. 2019, Sloane et al. 2020, Young et al. 2019c). No entanto, a aquisição pode representar uma barreira a essa participação. Quando o processo de construção de um sistema algorítmico é terceirizado para fornecedores ou analistas profissionais, os formuladores de políticas podem passar a ver a formalização de um modelo em termos algorítmicos como domínio de especialistas quantitativos. Mas, uma vez que a autoridade é cedida a especialistas quantitativos, esses especialistas podem não ver seu papel como avaliar se a definição do problema é apropriada, se as suposições que ela reflete se sustentam na realidade ou quais componentes não técnicos de implementação (como proteções de devido processo) precisam acompanhar a ferramenta técnica (Goodman 2019). Como escreve Flood (2010, p. 219), embora os especialistas quantitativos da RAND

> fossem, em última análise, responsáveis por suas suposições falhas [refletidas em seu modelo de fechamento de quartéis de bombeiros], é um pouco como esperar que um motorista de caminhão de cimento reflita se construir novas estradas é a melhor maneira de aliviar os problemas de transporte do país—pensar sobre isso não é o trabalho dele, e pensar em voz alta pode realmente fazê-lo ser demitido.

5.2. Caminhos Alternativos para Aquisição

Os processos de aquisição variam. Os governos podem estar envolvidos em diferentes graus, com diferentes tipos de vínculos contratuais com empresas e com sistemas mais ou menos prontos ou personalizados. Embora frequentemente desenvolvidos por empresas privadas, algoritmos também podem ser adquiridos por meio de competição aberta. Em 2017, as Escolas Públicas de Boston anunciaram o Desafio de Transporte de US$ 15.000 para projetar um algoritmo para realocar suas rotas de ônibus e horários de início das aulas enquanto atendiam a um conjunto de

Baixado de www.annualreviews.org. Convidado (guest) IP: 191.177.164.206 Em: Ter, 17 Mar 2026 21:02:23

objetivos declarados (Goodman 2019, Whittaker et al. 2018). Esses objetivos eram para que o algoritmo melhorasse a eficiência das rotas e corrigisse as desigualdades na distribuição existente dos horários de início das aulas, que prejudicavam famílias minoritárias e mais pobres, ao mesmo tempo em que atendia a um conjunto de restrições motivadas por considerações de saúde e segurança. O algoritmo construído pelos vencedores, uma equipe de pesquisadores do Instituto de Tecnologia de Massachusetts, atingiu esses objetivos (Bertsimas et al. 2019). No entanto, isso não foi suficiente para evitar uma forte reação da comunidade, especialmente contra a disruptividade do algoritmo nas rotas e horários existentes, levando o distrito a abandonar parcialmente o projeto.

Goodman (2019) argumenta que os problemas que levaram ao fracasso parcial do algoritmo decorrem do envolvimento insuficiente entre o público e o modelo, e que essa questão já era evidente na forma como a competição foi estruturada, sem qualquer ênfase na implementação sociopolítica ou na comunicação com o público sobre como o algoritmo funcionava. Embora o distrito escolar tenha reconhecido as questões de equidade em jogo e envolvido o público sobre "o que eles queriam do transporte escolar e dos horários de início das aulas de forma abstrata", ela escreve, "quase não houve envolvimento com o próprio modelo, transparência insuficiente sobre as compensações do algoritmo e nenhuma oportunidade de ajustá-lo" (Goodman 2019, p. 3). Uma comunicação mais direta sobre as compensações entre aumentar a equidade e minimizar a disrupção poderia ter levado a uma maior aceitação pública ou a um algoritmo ajustado que também priorizasse a estabilidade. Embora a contratação por concurso delegue inerentemente algumas decisões aos concorrentes, os poderes de definição de tarefas, avaliação, adoção e apresentação pública permanecem nas mãos do governo.

Apesar de seus fracassos, o algoritmo de transporte escolar de Boston oferece um exemplo em que um governo obteve um algoritmo para ajudar em uma tarefa complexa, declarando abertamente seus objetivos de política, fazendo algumas tentativas de engajamento público e mantendo controle próximo sobre a avaliação e implementação do algoritmo. É exatamente isso que muitas vezes falta nos processos de contratação, especialmente quando ocorrem por meio de fornecedores privados—parte de um conjunto maior de preocupações sobre o equilíbrio de poder em evolução entre governos e empresas de tecnologia (Pasquale 2015, Zuboff 2019). Recentemente, ao descrever o caso do uso do ShotSpotter pelo Departamento de Polícia de Oakland (uma tecnologia amplamente utilizada para detectar sons de tiros), Mulligan (citado em Kaye 2021) pintou um quadro sombrio:

> Eles [o governo da cidade] na verdade não recebem os dados sobre os padrões de movimento em sua própria cidade... e assim, de certa forma, acabam se tornando reféns de empresas que agora sabem mais sobre a infraestrutura urbana do que a própria cidade, porque possuem os dados brutos... e a cidade fica apenas com um conjunto de relatórios.

É esse tipo de dinâmica que Brauneis & Goodman (2018, p. 109) têm em mente quando escrevem que a contratação ameaça deixar os governos "esvaziados, cegos e no escuro".

6. GOVERNANDO A TECNOLOGIA

6.1. Transparência e Avaliação

A imagem do algoritmo como uma caixa-preta, ou da contratação sem escrutínio público, gera apelos por abertura como caminho para avaliação crítica, processo democrático e responsabilização. No entanto, a transparência significativa tem se mostrado difícil de definir e realizar. A transparência pode ter muitos objetos, incluindo código-fonte e pressupostos de modelagem, detalhes do conjunto de dados, deliberações internas, justificativas para ações efetivamente tomadas, contratos e orçamentos. Faz sentido pensar na transparência como um objetivo multidimensional, medido em graus, para o qual diferentes dimensões levantam diferentes questões técnicas, legais e sociais (Burrell 2016, Coglianese & Lehr 2019, Fink 2018, Levy & Johns 2016, Pozen 2018, Pozen & Schudson 2018).

Baixado de www.annualreviews.org. Convidado (guest) IP: 191.177.164.206 Em: Ter, 17 Mar 2026 21:02:23

Em meio a essa multiplicidade, uma distinção ampla, porém útil, é entre formas mais proativas e reativas de transparência—a primeira ocorrendo mais por design, obrigação legal e divulgação voluntária, a segunda em resposta a solicitações ad hoc de informações (ben-Aaron et al. 2017, Bloch-Wehba 2020). A transparência reativa pode, às vezes, ser insuficiente e tardia. Mulligan & Bamberger (2019) enfatizam a necessidade de intervenções no nível do design e do procedimento e observam que a transparência ex post é insuficiente para desafiar decisões políticas incorporadas em sistemas já em uso. Quando governos adquirem sistemas algorítmicos de fornecedores, às vezes ficam no escuro sobre os detalhes de como os sistemas funcionam. Diante de solicitações de registros abertos, os governos podem não conseguir atendê-las porque não possuem as informações. Os fornecedores podem alegar isenções de segredo comercial para não compartilhar dados, softwares e processos de design, e contratos governamentais podem permitir isso (Wexler 2018). As próprias agências também podem ter interesse no sigilo e, às vezes, interpretam acordos de confidencialidade e isenções de segredo comercial de forma ampla (Brauneis & Goodman 2018).

Baixado de www.annualreviews.org. Convidado (guest) IP: 191.177.164.206 Em: Ter, 17 Mar 2026 21:02:23

Solicitações de informações, embora reativas, continuam sendo um mecanismo importante de exposição pública e uma forma de estudiosos analisarem empiricamente a transparência. Ben-Aaron et al. (2017) utilizam solicitações de informações a governos locais para estudar se os governos têm mais probabilidade de atender aos pedidos quando informados de que outros condados já cumpriram. Brauneis & Goodman (2018) usam solicitações de registros abertos para catalogar algoritmos usados em avaliações de risco pré-julgamento, bem-estar infantil, policiamento preditivo e avaliação de professores, revelando assim lacunas no conhecimento dos governos sobre seus próprios sistemas. Esse tipo de pesquisa sobre transparência fornece não apenas informações sobre algoritmos e procedimentos governamentais, mas também uma indicação de quão dispostos ou capazes os governos estão de fornecer essas informações em primeiro lugar.

Alternativamente, esforços de transparência proativa concentram-se em exigir que os governos produzam ou divulguem informações sobre o design do sistema, aquisição, desempenho e impactos projetados. Nesse sentido, alguns formuladores de políticas buscaram inspiração em requisitos de documentação na elaboração de políticas ambientais. Muitas propostas para avaliações de impacto algorítmico (ou AIAs) exigem uma análise obrigatória a ser concluída antes que um sistema seja adquirido ou implantado. Essas avaliações podem incluir períodos para comentários e respostas do público; divulgação de detalhes-chave sobre o propósito e a implementação; discussão de meios alternativos de alcançar objetivos políticos; e avaliação prospectiva de possíveis preocupações quanto à precisão, viés e impacto sobre grupos de interesse (Reisman et al. 2018, Selbst 2017). Da mesma forma, após um sistema estar em uso, ele pode ser submetido a auditorias obrigatórias periódicas ou outras formas de avaliação e testes, realizados internamente ou por pesquisadores externos ou conselhos de supervisão (Ada Lovelace Inst. 2020). Essas avaliações contínuas concentram-se principalmente na detecção de viés, embora também possam inspecionar outras qualidades, como confiabilidade ou conformidade com outros mandatos políticos (Kroll et al. 2016).

Dados empíricos sobre a eficácia das AIAs na prática são relativamente escassos, e políticas que as tornam obrigatórias estão apenas começando a surgir (por exemplo, McKelvey & MacDonald 2019). Na melhor das hipóteses, as AIAs podem estimular uma reflexão proativa sobre questões-chave, fortalecer a capacidade dos governos locais e proporcionar uma forma de participação das comunidades afetadas no processo político. Na pior das hipóteses, podem se tornar pouco mais do que um item de verificação—burocracia pro forma que pouco contribui para promover a responsabilização (Metcalf et al. 2021, Waldman 2019).

6.2. Orçamento

Os processos orçamentários também podem desempenhar um papel na promoção da tomada de decisões democráticas sobre sistemas algorítmicos. Na maioria das localidades, itens orçamentários na escala de sistemas algorítmicos municipais exigem aprovação do conselho municipal. Em teoria, o orçamento atua como um controle, garantindo que os fundos públicos sejam alocados adequadamente e dando aos membros do conselho e aos constituintes a oportunidade de fazer perguntas

sobre a tecnologia e seus impactos esperados. Dito isso, orçamentos não são uma salvaguarda garantida. Por exemplo, quando localidades recebem fundos de outras fontes (ou seja, agências estaduais ou federais), às vezes conseguem passar pelo processo orçamentário ou até mesmo evitá-lo completamente. Em outros casos, os períodos de aviso público ou de comentários são tão breves e mal divulgados que não servem para nenhum propósito prático.

Ambas essas dinâmicas ocorreram em Oakland, Califórnia. Em 2010, o conselho municipal aprovou pedidos dos departamentos de polícia e bombeiros para construir um Centro de Consciência de Domínio (DAC) usando fundos federais originalmente destinados a esforços antiterrorismo. Os pedidos atraíram pouca atenção pública. O DAC agregou fontes de dados existentes, criando um "centro para analisar feeds de vídeo e sensores" (Wood 2013) com o objetivo de manter o Porto de Oakland e a área ao redor seguros. No entanto, em 2013, quando os departamentos de polícia e bombeiros solicitaram permissão para receber uma segunda bolsa federal para expandir enormemente o DAC, o público percebeu (Winston 2014). Desta vez, uma investigação do grupo sem fins lucrativos Oakland Privacy Working Group descobriu que a aquisição do DAC violava uma proibição de contratar com entidades envolvidas em trabalhos com armas nucleares. Separadamente, uma conversa orçamentária rotineira revelou todo o escopo do programa e chamou a atenção pública. Juntos, esses eventos catalisaram um movimento em um momento em que os moradores estavam sensibilizados pelos recentes vazamentos de Edward Snowden para reagir a programas de vigilância com ceticismo e preocupação. Em resposta à pressão da comunidade, o conselho municipal aprovou uma resolução reduzindo o programa e criou uma força-tarefa cidadã (posteriormente tornada permanente) para aconselhar sobre tecnologias de vigilância. O caso demonstra como os processos orçamentários podem — mas nem sempre — facilitar a supervisão pública.

6.3. Mandatos Estaduais

Nos Estados Unidos, mandatos em nível estadual também regem o uso local de sistemas algorítmicos. Os detalhes de implementação desses mandatos podem ter consequências profundas na vida dos cidadãos e muitas vezes são pouco examinados. Esses mandatos podem ceder ou exercer controle sobre diferentes aspectos do processo de implementação, escolhas que introduzem uma tensão entre entendimentos globais e locais de justiça, precisão e viés, e entre administração de políticas centralizada e descentralizada (Solow-Niederman et al. 2019).

A Lei de Reforma da Fiança Monetária da Califórnia (SB10) e a Lei de Segurança Pública e Responsabilidade de Infratores de Kentucky (HB463), ambas destinadas a reger práticas de avaliação de risco na justiça criminal, ilustram essas dinâmicas. Kentucky determinou que investigações pré-sentença incluíssem resultados de avaliação de risco e que juízes incorporassem esses resultados na tomada de decisão pré-julgamento. Ao instituir uma avaliação estadual, Kentucky buscou uma versão de justiça semelhante à padronização: um indivíduo poderia viajar para qualquer lugar do estado e esperar o mesmo resultado do instrumento de risco. No entanto, resultados uniformes podem ser considerados injustos se diferentes fatores de risco (por exemplo, idade, histórico criminal) tiverem diferentes graus de poder preditivo entre jurisdições. Em contraste, a Califórnia tentou reconciliar essa tensão permitindo que as localidades escolhessem sua própria avaliação de risco, ao mesmo tempo em que concedia autoridade de supervisão ao Conselho Judicial do estado. Em um esforço para buscar equilíbrio, a Califórnia deu ao conselho e às localidades o poder (e a responsabilidade) de validar cada avaliação de risco, mas criou confusão e carga administrativa adicional (Solow-Niederman et al. 2019). Comparar esses dois mandatos demonstra como os estados buscam um conjunto complexo, e às vezes incompatível, de objetivos enquanto equilibram restrições políticas e orçamentárias. Sem esclarecer as tensões e compensações entre estado e município que surgem sob os mandatos, os estados "correm o risco de codificar esses entendimentos implicitamente, de maneiras opacas e que podem resistir à responsabilidade democrática" (Solow-Niederman et al. 2019, p. 724).

Baixado de www.annualreviews.org. Convidado (guest) IP: 191.177.164.206 Em: Ter, 17 Mar 2026 21:02:23

6.4. Restrições, Moratórias e Proibições

Além dos requisitos de transparência, avaliações de impacto, medidas orçamentárias e mandatos estaduais, os governos locais podem regular diretamente sistemas algorítmicos específicos. Por exemplo, Spivack & Garvie (2020) categorizam as medidas que estados e localidades tomaram em relação às tecnologias de reconhecimento facial. As medidas mais rigorosas proíbem totalmente a aquisição e o uso desses sistemas pelo governo — ou até mesmo por governos e empresas privadas. Vários governos locais adotaram tais medidas (São Francisco, Berkeley e Boston, entre outros). Outras cidades abordaram o problema por meio de moratórias que suspendem a aquisição e o uso por um determinado período ou até que estudos adicionais sejam concluídos. Ainda outros governos adotam abordagens mais limitadas, permitindo o uso da tecnologia, mas impondo condições para seu uso. Como detalham Spivack & Garvie (2020) no contexto do reconhecimento facial, essas restrições variam desde a proibição de integração com câmeras corporais até a limitação de acesso (por exemplo, permitindo acesso apenas para investigação de crimes violentos). O equilíbrio entre restrições baseadas em propósito e uso, por um lado, e rejeições totais da tecnologia, por outro, é politicamente controverso: os defensores divergem em suas avaliações sobre a aceitabilidade política e a eficácia de cada abordagem. Há poucas evidências empíricas sobre os impactos de longo prazo das diferentes estratégias.

Baixado de www.annualreviews.org. Convidado (guest) IP: 191.177.164.206 Em: Ter, 17 Mar 2026 21:02:23

7. USANDO E RESPONDENDO AO SISTEMA

Sistemas algorítmicos municipais operam em ambientes complexos e dinâmicos. Eles influenciam e são influenciados por múltiplas agências e atores, cada um com seus próprios objetivos e interesses, tornando difícil prever os impactos de um sistema antes da implementação. Uma vez implantados, outros atores entram em cena, moldando como o algoritmo funciona na prática — e às vezes resistindo.

É possível, até certo ponto, antecipar essas interações entre humanos e algoritmos e fazer escolhas cuidadosas de design sobre como os algoritmos são inseridos nas burocracias e ambientes de trabalho (Bansal et al. 2020, Dietvorst et al. 2016, Veale et al. 2018, Yang et al. 2020). Um aspecto importante do design do sistema envolve não apenas o que o algoritmo prevê, mas como essas previsões são apresentadas às pessoas que as utilizam. Os sistemas são frequentemente projetados para traduzir probabilidades em escalas, pontuações e categorias amigáveis ao usuário, mas isso traz riscos. Se, por exemplo, juízes que usam previsões de risco de fuga pré-julgamento para tomar decisões de fiança recebem apenas pontuações de risco, sem serem informados sobre a probabilidade real de comparecimento ao tribunal que uma pontuação alta ou baixa representa, eles podem fazer suposições incorretas sobre o significado das pontuações. Em um estudo de validação de tal algoritmo, aproximadamente 90% dos réus com a menor pontuação de risco (1) e 70% daqueles com a maior pontuação (6) compareceram ao tribunal (Brauneis & Goodman 2018). Aqui, as pontuações 1 e 6 parecem apresentar uma diferença mais extrema do que esses percentuais realmente indicam — um risco comum ao usar pontuações e classificações (Espeland & Sauder 2007). Dessa forma, as decisões de interpretação e apresentação incorporadas ao design de um sistema algorítmico moldam sua eficácia e impacto.

7.1. Respostas dos Trabalhadores da Linha de Frente

Talvez uma decisão ainda mais fundamental no design de sistemas algorítmicos seja se e como envolver os trabalhadores da linha de frente desde o início. Uma distinção comum aqui é entre automação total e sistemas de recomendação (Coglianese & Lehr 2019, Veale & Brass 2019, Yeung 2017). Os sistemas de decisão algorítmica são frequentemente retratados como substituindo os tomadores de decisão humanos, mas muitos sistemas não têm o objetivo de substituir totalmente os trabalhadores (Kellogg et al. 2020). Em vez disso, eles

fornecem suporte à decisão (recomendações, previsões, pontuações), enquanto os trabalhadores mantêm a capacidade de ajustar ou substituir as decisões finais sugeridas pelo sistema. Por outro lado, algumas pesquisas demonstraram o viés de automação, no qual os trabalhadores tendem a se submeter à saída do sistema mesmo quando aparentemente têm margem para discricionariedade (Citron 2008, Skitka et al. 1999).

A pesquisa sobre a discricionariedade dos trabalhadores do governo—se ela existe, as formas que assume, suas vantagens e seus perigos—tem uma longa história que antecede os sistemas de decisão automatizados (Bovens & Zouridis 2002). Essa discricionariedade surge porque "os burocratas de nível de rua frequentemente trabalham em situações muito complicadas para serem reduzidas a formatos programáticos...[eles] têm discricionariedade porque as definições aceitas de suas tarefas exigem observação sensível e julgamento, que não são reduzíveis a formatos programados" (Lipsky 1980, p. 15). Os avanços tecnológicos não eliminaram essa característica do trabalho público, e é importante entender como trabalhadores e máquinas interagem para gerar uma decisão, em vez de apenas o resultado estatístico do sistema algorítmico (Guay & Parent 2018, Veale & Brass 2019, Vogl 2020, Vogl et al. 2020). No contexto da justiça criminal, por exemplo, algumas pesquisas demonstram heterogeneidade significativa na conformidade com recomendações de avaliação de risco em dimensões como idade (Stevenson & Doleac 2019), raça (Albright 2019) e renda (Skeem et al. 2020). Trabalhadores da linha de frente também podem ativamente mitigar o viés do sistema ao substituir decisões da máquina (Albright 2019, Stevenson & Doleac 2019). Por esse motivo, alguns pesquisadores alertaram contra a automação total, demonstrando que humanos são preditores superiores em alguns contextos (por exemplo, De-Arteaga et al. 2020). Embora as agências implantem sistemas algorítmicos em parte por sua precisão e objetividade, ter um humano no processo pode ter efeitos positivos e negativos.

Além das decisões de sobreposição, os sistemas algorítmicos municipais mudam as experiências dos trabalhadores—para melhor e para pior—e as formas como os trabalhadores respondem a essas mudanças podem influenciar o desempenho do sistema e levar a consequências imprevistas. Brayne & Christin (2020) exploram esse fenômeno em departamentos de polícia e tribunais criminais. Eles documentam "estratégias de resistência" que policiais e juízes empregaram em resposta à vigilância e à desqualificação. Juízes, por exemplo, temiam que administradores e advogados de defesa usassem avaliações de risco para avaliar produtividade e desempenho, comparando diretamente suas decisões de sentença e taxas de encarceramento, colocando assim em risco sua autonomia profissional. Eles também argumentaram que as pontuações de risco não forneciam insights além do que sua experiência profissional já proporcionava. Em resposta a ameaças como essas, juízes e policiais empregaram duas estratégias: ignorar recomendações e ofuscar dados. Policiais, por exemplo, às vezes desconsideravam as recomendações geográficas geradas por tecnologias preditivas ou até mesmo usavam a tecnologia contra si mesmos—deliberadamente adulterando antenas de carros para interromper equipamentos de gravação e impedir que a administração ouvisse suas conversas. Outras pesquisas sobre o uso policial de sistemas algorítmicos também demonstram que policiais usam sistemas de muitas maneiras não previstas, o que pode prejudicar o funcionamento do sistema e violar pressupostos sobre precisão (Ehrenkranz 2019, Garvie 2019). A discricionariedade, portanto, é uma faca de dois gumes: continua sendo parte crucial de como os sistemas algorítmicos são implementados, mas viés, resistência e até adulteração do sistema podem moldar como os sistemas funcionam.

7.2. Respostas Públicas

Pesquisas emergentes de opinião pública apresentam um quadro nebuloso e um tanto fragmentado das atitudes dos americanos em relação aos sistemas algorítmicos. Dados de pesquisas indicam uma divisão na opinião pública sobre se decisões automatizadas de alto risco são eficazes, justas ou aceitáveis (Smith 2018). Também há um desacordo significativo sobre onde delegar a responsabilidade de supervisão dos sistemas automatizados. Em uma pesquisa nacional, apenas 26% dos entrevistados tinham confiança na capacidade dos governos estaduais de gerenciar IA. A confiança no governo federal era aproximadamente a mesma. Em comparação, 41% tinham confiança na capacidade das empresas de tecnologia de gerenciar IA (Zhang & Dafoe 2019).

Baixado de www.annualreviews.org. Convidado (guest) IP: 191.177.164.206 Em: Ter, 17 Mar 2026 21:02:23

Talvez a confiança do público dependa da aplicação: Em outra pesquisa, 56% dos adultos dos EUA confiaram que as agências de aplicação da lei usariam softwares de reconhecimento facial de forma responsável, enquanto apenas 36% confiaram que as empresas de tecnologia fariam o mesmo (Smith 2019).

No entanto, as tendências gerais da opinião pública podem desempenhar apenas um papel limitado na orientação das ações políticas (Brown et al. 2019). Seja qual for a opinião pública em relação aos sistemas algorítmicos em geral, um número crescente de pessoas e organizações está envolvido em esforços para combater a ameaça potencial que esses sistemas representam para suas comunidades. Seus objetivos e métodos são diversos. Advogados legais contestam sistemas de tomada de decisão algorítmica com base em vários fundamentos relacionados ao devido processo, tanto em contextos criminais quanto civis—contestando, por exemplo, o uso de avaliação de risco na sentença sem transparência suficiente para permitir que o réu questione sua validade (Wexler 2018) e a privação de moradia e benefícios públicos com base em decisões automatizadas (Gilman 2020, Richardson et al. 2019).

Alguns grupos atuam dentro do governo local, desempenhando um papel consultivo e de supervisão. A Oakland Privacy Commission, por exemplo, é um órgão liderado por cidadãos que aconselha a cidade sobre como defender os direitos de privacidade da comunidade, e Pittsburgh criou recentemente sua própria força-tarefa destinada a “garantir a responsabilidade algorítmica e a equidade para todos os residentes” (Inst. Cyber Law Policy Secur. 2021). Outros grupos exercem pressão sobre os governos locais de fora. Exemplos incluem Eye on Surveillance, uma organização de base em Nova Orleans que trabalha para impedir a expansão das ferramentas de vigilância da cidade organizando protestos públicos e fazendo campanha por leis anti-vigilância, e MediaJustice, uma organização sem fins lucrativos de Oakland que combate a vigilância digital e o policiamento e prisões de alta tecnologia. Essas diversas respostas públicas refletem a longa variedade de abordagens para a reforma de políticas. Os impactos e compensações de diferentes abordagens para a reforma e oposição algorítmica permanecem pouco explorados, e há poucas evidências empíricas sobre a eficácia relativa das várias abordagens.

8. AVALIANDO, MANTENDO OU DESMONTANDO O SISTEMA

O momento em que um sistema é introduzido é quando ele tem mais chances de atrair atenção. Mas o escrutínio contínuo é mais difícil. Com o tempo, algumas ferramentas algorítmicas tendem a se tornar partes assumidas da infraestrutura burocrática. No entanto, várias dinâmicas sociais importantes só surgem depois que um sistema está em funcionamento há algum tempo; essas também fornecem arenas importantes de investigação para o pesquisador social. Entre elas, destaca-se como um sistema é avaliado, mantido ou corrigido ao longo do tempo—se esses processos ocorrem de fato—e como os sistemas são usados para novos fins além dos originalmente previstos.

8.1. Avaliação, Manutenção e Mudança

Avaliar se um sistema algorítmico está alcançando seus efeitos pretendidos é desafiador e específico para cada contexto. Embora avaliar o desempenho do modelo em dados de teste reservados seja uma parte rotineira do desenvolvimento do modelo, avaliar se um modelo continua a ter bom desempenho em novas entradas após ser implantado traz dificuldades adicionais. Avaliar se esse modelo está ajudando um governo a atingir seus objetivos de política amplia ainda mais o escopo. Isso exige lidar não apenas com a precisão preditiva, mas também com o funcionamento prático do sistema.

Algoritmos às vezes são promovidos como um caminho para políticas mais adaptativas, evitando que fiquem obsoletas e ultrapassadas (Mormann 2021). Embora atualmente muitos algoritmos do setor público não sejam sofisticados o suficiente para se adaptar em tempo real aos dados recebidos, existe pelo menos o potencial de re-treinar modelos periodicamente com dados novos. Mas, adaptativos ou não, algoritmos não são

Baixado de www.annualreviews.org. Convidado (guest) IP: 191.177.164.206 Em: Ter, 17 Mar 2026 21:02:23

autossustentáveis e existem em ambientes dinâmicos. A avaliação ao longo do tempo torna-se crítica, para que os próprios algoritmos não fiquem obsoletos e desatualizados.

Dentro do aprendizado de máquina, a questão de sistemas desenvolvidos em um contexto, mas aplicados em outro, é frequentemente discutida sob o termo "shift" (Quiñonero-Candela 2009, Rabanser et al. 2019). O shift ocorre quando as condições do mundo real sob as quais um modelo foi desenvolvido são diferentes das condições em que ele é implantado. De forma simples, o shift pode ocorrer à medida que o mundo muda ao longo do tempo. Essa mudança pode refletir um desvio gradual, choques externos repentinos (a pandemia de Covid-19 é um exemplo), ou até mesmo ações dos próprios governos. Por exemplo, um município pode adotar políticas que afetam resultados que também está tentando prever. Se não atualizar seu modelo com dados do novo ambiente, suas previsões podem permanecer cegas aos impactos dessas mudanças, criando "previsões zumbis" (Koepke & Robinson 2018). Questões semelhantes podem ocorrer quando um sistema é treinado com dados de uma população, mas depois é exportado para outra população (por exemplo, para outro município) em que as condições diferem (Buolamwini & Gebru 2018, Gama et al. 2014, Hamilton 2015, Koh et al. 2020).

O shift não é apenas uma questão técnica; é também uma questão legal e burocrática. Sem procedimentos e recursos para apoiar estudos de validação que testem viés e comparem previsões com resultados reais, o shift pode passar despercebido e não ser tratado (Koh et al. 2020). Em contextos não algorítmicos, agências governamentais já estão acostumadas com a necessidade de avaliações repetidas. Uma abordagem legal comum para tais questões é o uso de cláusulas de expiração ou datas de validade, pelas quais um programa deve ser revisitado ou reautorizado (Mormann 2021). Esses tipos de mecanismos criam atrito proposital e podem fornecer pontos de corte naturais para garantir reavaliação periódica e supervisão pública. No contexto algorítmico, procedimentos para criar esse tipo de atrito são menos desenvolvidos, mas as apostas são altas. Ferramentas automatizadas de fiscalização "raramente são sistemas prontos para uso[.]...Se dados baseados em padrões históricos de fiscalização forem usados sem reflexão para treinar modelos e os esforços para atualizar esses modelos forem ad hoc, as ações de fiscalização correm o risco de...lutar a última guerra em vez de enfrentar novos [problemas]" (Engstrom 2020, pp. 25–26).

8.2. Feedback

Outro conjunto de complicações para avaliar algoritmos surge quando a mudança emerge do próprio algoritmo. Pesquisas das ciências sociais sobre performatividade fornecem um análogo útil aqui: a própria existência de um sistema destinado a modelar e medir um determinado fenômeno pode, por si só, moldar esse fenômeno, alinhando o comportamento humano ao que o sistema preveria (Callon 1998, MacKenzie 2008).

Ciclos de feedback ocorrem quando os resultados do algoritmo são usados para tomar decisões e os resultados dessas decisões são realimentados nos dados de treinamento subsequentes, reforçando padrões e vieses existentes. O viés nos dados de prisões usados para treinar algoritmos de policiamento preditivo tem esse efeito: dados de prisões por delitos de drogas super-representam bairros negros, levando as previsões a direcionar esses bairros para patrulhas, resultando em mais prisões e, assim, mais super-representação (Ensign et al. 2018, Lum & Isaac 2016). Por outro lado, o modelo pode não ser informado sobre o uso de drogas em outros bairros onde a polícia não faz prisões e deixar de direcionar essas áreas, contribuindo para ciclos de policiamento desequilibrado. Outras dinâmicas de feedback incluem as formas como os comportamentos humanos podem mudar em resposta à existência de um modelo. Como propõe a Lei de Goodhart, as medidas de resultado tornam-se menos úteis quando os atores começam a direcioná-las (Abebe et al. 2020a, Bambauer & Zarsky 2018): escolas ensinam para o teste, faculdades manipulam rankings (Espeland & Sauder 2007) e candidatos a emprego aprendem quais sinais incluir em seus currículos para avançar em triagens automatizadas (Weed 2021). Essas formas de feedback podem se tornar pontos de disputa à medida que os afetados por um sistema tentam se adaptar e antecipá-lo, enquanto os criadores do sistema tentam garantir que ele funcione conforme o esperado.

Baixado de www.annualreviews.org. Convidado (guest) IP: 191.177.164.206 Em: Ter, 17 Mar 2026 21:02:23

8.3. Falha e Expansão Gradual

Como os sistemas algorítmicos falham? Já aludimos a vários fenômenos que podem ser considerados formas de falha: proibições, moratórias, desafios legais, mudanças, ciclos de feedback e reação pública negativa. Um algoritmo também pode simplesmente não alcançar os efeitos pretendidos (por exemplo, O'Brien et al. 2018). Quando sistemas são adquiridos de fornecedores privados, a possibilidade de os fornecedores "falirem, mudarem seu foco ou alterarem seus modelos de negócio" cria caminhos adicionais para falhas, como ocorreu no desenvolvimento abortado pela Alphabet de uma orla "inteligente" em Toronto (Valverde & Flynn 2020). Tudo isso sugere potenciais locais para estudar como as coisas podem dar errado, quem é responsável quando isso acontece e o que ocorre em seguida.

O contraponto à falha é, de certo modo, a expansão. O ciclo de vida de um algoritmo não começa nem termina necessariamente com uma única tarefa bem definida. O conceito de expansão gradual (function creep) é polissêmico e frequentemente usado para descrever como o uso de tecnologias do setor público pode se expandir além de seus propósitos originais, especialmente quando essas expansões ocorrem sem debate aberto ou reconhecimento (Koops 2021). A introdução de uma tecnologia pode estimular novos projetos não planejados originalmente ou diminuir a barreira para implementar usos adicionais de dados, ferramentas, financiamento e outros recursos burocráticos agora disponíveis. Claro, tal mudança e expansão podem ser uma forma de inovação. Mas falar em expansão gradual geralmente é expressar preocupação e chamar atenção para mudanças ocultas e sem responsabilidade. Analiticamente, a ressalva que a expansão gradual impõe à nossa visão do ciclo de vida algorítmico é que a tarefa, problema ou contexto inicial para o qual um conjunto de dados ou algoritmo foi criado e implementado pode não ser tudo para o que ele é usado.

9. CONCLUSÃO: ALGORITMO COMO DESAFIO E OPORTUNIDADE

Sistemas algorítmicos do setor público levantam questões desafiadoras sobre como garantir transparência, responsabilidade e controle do governo sobre seus próprios sistemas — mas também oferecem oportunidades para transparência e responsabilidade, bem como para um design mais cuidadoso e formulação explícita de problemas. Lembre-se do caso do algoritmo de transporte escolar de Boston, criado para melhorar a eficiência e a equidade, mas que foi contestado por causar perturbações nos horários existentes. Esse algoritmo acabou se tornando um alvo mobilizador para reação pública, gerando slogans como "famílias acima dos algoritmos" (Goodman 2019). Mas não está claro se o algoritmo era realmente o problema. Goodman o descreve como um caso de "bode expiatório algorítmico", no qual o algoritmo representava questões substantivas de equidade e mudanças disruptivas que estavam realmente em jogo (embora potencialmente mais polêmicas de discutir) e que poderiam estar em jogo mesmo sem um algoritmo na equação. A tragédia do caso é que o algoritmo poderia ter proporcionado flexibilidade para envolver o público na escolha entre múltiplos trade-offs. Se implementado, poderia ter criado um sistema mais equitativo do que o que existia originalmente.

A dualidade entre desafio e oportunidade é um tema recorrente no ciclo de vida algorítmico. Neste artigo, traçamos esse tema desde a formulação do problema, onde algoritmos podem tornar objetivos e trade-offs explícitos, mas também podem suavizar problemas complexos sem abordar as causas raízes; passando pela aquisição e contratação de dados, onde a integração de dados e parcerias público-privadas expandem os recursos do governo, mas trazem perigos; e finalmente até a governança, implementação e avaliação, onde algoritmos geram temores legítimos, mas também podem se tornar bodes expiatórios. O ciclo de vida algorítmico que apresentamos aqui representa não tanto uma única linha de vida, mas um mosaico formado por muitos campos, desenrolando-se entre os "trabalhadores do federalismo". Ele destaca locais importantes para pesquisa social. Para os pesquisadores, o desafio é compreender esse espaço em meio a escopo ambíguo, muita variedade e altos riscos — especialmente no nível local, onde muitos dos algoritmos que impactam diretamente a vida pública estão presentes. A oportunidade é entender melhor não apenas uma tecnologia, mas a sociedade e seu futuro mediado por algoritmos.

Baixado de www.annualreviews.org. Convidado (guest) IP: 191.177.164.206 Em: Ter, 17 Mar 2026 21:02:23

DECLARAÇÃO DE DIVULGAÇÃO

Os autores não têm conhecimento de quaisquer afiliações, associações, financiamentos ou participações financeiras que possam ser percebidas como afetando a objetividade desta revisão.

AGRADECIMENTOS

Agradecemos o apoio da Fundação John D. e Catherine T. MacArthur (todos os autores) e da Bolsa Microsoft Ada Lovelace (S.R.).

LITERATURA CITADA

Abebe R, Barocas S, Kleinberg J, Levy K, Raghavan M, Robinson DG. 2020a. Papéis da computação na mudança social mudança. In Proceedings of the 2020 ACM Conference on Fairness, Accountability and Transparency, pp. 252–60. Nova York: Assoc. Comput. Mach.

Abebe R, Kleinberg J, Weinberg SM. 2020b. Alocações de subsídios na presença de choques de renda. Proc. AAAI Conf. Artif. Intel. 34(5):7032–39

Ada Lovelace Inst. 2020. Examinando a caixa preta: ferramentas para avaliar sistemas algorítmicos. Relatório, Ada Lovelace Inst., Londres. **https://www.adalovelaceinstitute.org/report/examining-the-black-box-tools-assessing-algorithmic-systems/**

Adams C. 2018. Regras domésticas: o caso do procedimento administrativo local. Fordham Law Rev. 87:629–69

Advis. Comm. Intergov. Relat. 1982. Papéis estaduais e locais no sistema federal. Relatório, Advis. Comm. Intergov. **Relat., Washington, DC. https://library.unt.edu/gpo/acir/Reports/brief/B-6.pdf**

Albright A. 2019. Se você der a um juiz uma pontuação de risco: evidências das decisões de fiança do Kentucky. Discuss. Pap. Ser. 85, John M. Olin Cent. Law Econ. Bus., Harvard Law School, Cambridge, MA

Bambauer J, Zarsky T. 2018. O jogo do algoritmo. Notre Dame Law Rev. 94:1–47

Bansal G, Nushi B, Kamar E, Horvitz E, Weld DS. 2020. A IA mais precisa é o melhor colega de equipe? Otimi- **zando a IA para trabalho em equipe. arXiv. https://arxiv.org/abs/2004.13102**

Barabas C. 2020. Além do viés: reimaginando os termos da 'IA ética' no direito penal. Georgetown J. Law Mod. Crit. Race Perspect. 12:83–111

Barocas S, Selbst AD. 2016. O impacto desigual do big data. Calif. Law Rev. 104(3):671–732

Bedi N, McGrory K. 2020. O xerife de Pasco usa notas e históricos de abuso para rotular crianças escolares como potenciais criminosos. As crianças e seus pais não sabem. Tampa Bay Times, 19 de nov. **https://projects.tampabay.com/projects/2020/investigations/police-pasco-sheriff-targeted/school-data/**

ben-Aaron J, Denny M, Desmarais B, Wallach H. 2017. Transparência por conformidade: um experimento de campo avali- ando a abertura em governos locais. Public Admin. Rev. 77(1):68–77

Berman DR. 2019. Governo Local e os Estados: Autonomia, Política e Política Pública. Abingdon, Reino Unido: Routledge

Bertsimas D, Delarue A, Martin S. 2019. Otimizando os horários de início das escolas e rotas de ônibus. PNAS 116(13):5943–948

Bloch-Wehba H. 2020. Acesso a algoritmos. Fordham Law Rev. 88:1265–314

Bouk D. 2015. Como Nossos Dias Foram Numerados: Risco e a Ascensão do Indivíduo Estatístico. Chicago: Univ. Chicago Press

Bovens M, Zouridis S. 2002. Da burocracia de nível de rua ao nível de sistema: como a informação e a comuni- cação tecnológica está transformando a discricionariedade administrativa e o controle constitucional. Public Adm. Rev. 62:174–84

Brauneis R, Goodman EP. 2018. Transparência algorítmica para a cidade inteligente. Yale J. Law Technol. 20:103–76

Brayne S, Christin A. 2020. Tecnologias de previsão de crimes: a recepção de algoritmos na polícia e nos tribunais criminais. Soc. Probl. 2020:spaa004

Breckenridge K, Szreter S, eds. 2012. Registro e Reconhecimento: Documentando a Pessoa na História Mundial. Oxford, Reino Unido: Br. Acad.

Breiman L. 2001. Modelagem estatística: as duas culturas. Stat. Sci. 16(3):199–215

Bridges KM. 2017. A Pobreza dos Direitos de Privacidade. Stanford, CA: Stanford Univ. Press

Baixado de www.annualreviews.org. Convidado (guest) IP: 191.177.164.206 Em: Ter, 17 Mar 2026 21:02:23

Brown A, Chouldechova A, Putnam-Hornstein E, Tobin A, Vaithianathan V. 2019. Rumo à responsabilidade algorítmica nos serviços públicos: um estudo qualitativo das perspectivas da comunidade afetada sobre a tomada de decisão algorítmica nos serviços de bem-estar infantil. In Anais da Conferência CHI 2019 sobre Fatores Humanos em Sistemas de Computação, p. 41. Nova York: Assoc. Comput. Mach.

Buolamwini J, Gebru T. 2018. Tons de gênero: disparidades de precisão interseccionais na classificação comercial de gênero de gênero. PMLR 81:77–91

Burrell J. 2016. Como a máquina 'pensa': entendendo a opacidade em algoritmos de aprendizado de máquina. Big Data Soc 3(1). **https://doi.org/10.1177/2053951715622512**

Cahn AF. 2019. A primeira tentativa de regulamentar a IA foi um fracasso espetacular. Fast Company, 26 de nov. **https://www.fastcompany.com/90436012/the-first-effort-to-regulate-ai-was-a-spectacular-failure**

Callon M. 1998. A inserção dos mercados econômicos na economia. Sociol. Rev. 46(Suppl. 1):1–57

Caruana R, Lou Y, Gehrke J, Koch P, Sturm M, Elhadad N. 2015. Modelos inteligíveis para saúde: prevendo risco de pneumonia e readmissão hospitalar em 30 dias. In Anais da 21ª Conferência Internacional ACM sobre Descoberta de Conhecimento e Mineração de Dados, pp. 1721–30. Nova York: Assoc. Comput. Mach.

Chasalow K, Levy K. 2021. Representatividade em estatística, política e aprendizado de máquina. In Anais da Conferência ACM 2021 sobre Justiça, Responsabilidade e Transparência, pp. 77–89. Nova York: Assoc. Comput. Mach.

Chouldechova A, Benavides-Prado D, Fialko O, Vaithianathan R. 2018. Um estudo de caso de tomada de decisão assistida por algoritmo nas decisões de triagem de linha direta de maus-tratos infantis. In Anais da Conferência ACM 2018 de Justiça, Responsabilidade e Transparência, pp. 134–148. Nova York: Assoc. Comput. Mach.

Chouldechova A, Roth A. 2018. As fronteiras da justiça no aprendizado de máquina. arXiv. http://arxiv.org/abs/**1810.08810**

Churchman CW. 1967. Problemas perversos. Manag. Sci. 14(4):B-141–B-142

Citron DK. 2008. Devido processo tecnológico. Wash. Univ. Law Rev. 85:1249–313

Coglianese C, Lehr D. 2019. Transparência e governança algorítmica. Adm. Law Rev. 71:1–56

Crawford K. 2013. Os vieses ocultos em big data. Harvard Business Review, 1 de abr. https://hbr.org/2013/04/**the-hidden-biases-in-big-data**

Crump C. 2016. Formulação de políticas de vigilância por meio de aquisição. Wash. Law Rev. 91:1595–662

Davidson NM. 2016. Direito administrativo localista. Yale Law J. 126:564–634

Dawes RM, Faust D, Meehl PE. 1989. Julgamento clínico versus atuarial. Science 243(4899):1668–74

De-Arteaga M, Fogliato R, Chouldechova A. 2020. Um caso para humanos no circuito: decisões na presença de pontuações algorítmicas errôneas. In Anais da Conferência CHI 2020 sobre Fatores Humanos em Computação Sistemas, Nova York: Assoc. Comput. Mach.

Dencik L, Hintz A, Redden J, Warne H. 2018. Pontuações de dados como governança: investigando usos da pontuação de cidadãos em serviços públicos. Relatório, Data Justice Lab., Cardiff Univ., Cardiff, Reino Unido. https://datajustice.files.wordpress.com/2018/12/data-scores-as-governance-project-report2.pdf

Dettlaff A, Rivaux S, Baumann DJ, Fluke JD, Rycraft JR, James J. 2011. Desvinculando a fundamentação: a influência da raça, renda e risco na decisão de fundamentação no bem-estar infantil. Child. Youth Serv. Rev. 33(9):1630–37

Didier E. 2020. América pelos Números: Quantificação, Democracia e o Nascimento das Estatísticas Nacionais. Cambridge, MA: MIT Press

Dietvorst B, Simmons JP, Massey C. 2016. Superando a aversão a algoritmos: as pessoas usarão algoritmos imperfeitos se puderem (mesmo que minimamente) modificá-los. Manag. Sci. 64(3):1155–70

Early DE. 2016. Ferramenta de 'triagem' do Vale do Silício calcula os casos de sem-teto mais necessitados. Mercury News, 17 de fev. **https://www.mercurynews.com/2016/02/17/silicon-valley-triage-tool-calculates-the-neediest-homeless-cases/**

Ehrenkranz M. 2019. Tecnologia de reconhecimento facial da Amazon mais uma vez identifica políticos como criminosos. Gizmodo, 13 de ago. **https://gizmodo.com/amazons-face-recognition-tech-once-again-pegs-politicia-1837215790**

Engstrom DF, Ho DE, Sharkey CM, Cuéllar M-F. 2020. Governo por algoritmo: inteligência artificial em agências federais agências administrativas. Relatório, Adm. Conf. US, Washington, DC. https://www-cdn.law.stanford.edu/wp-content/uploads/2020/02/ACUS-AI-Report.pdf

Ensign D, Friedler SA, Neville S, Scheidegger C, Venkatasubramanian S. 2018. Ciclos de feedback descontrolados em policiamento preditivo. PMLR 81:160–71

Baixado de www.annualreviews.org. Convidado (guest) IP: 191.177.164.206 Em: Ter, 17 Mar 2026 21:02:23

Espeland W, Sauder M. 2007. Rankings e reatividade: como medidas públicas recriam mundos sociais. Am. J. Sociol. 113(1):1–40

Eubanks V. 2018. Automatizando a Desigualdade: Como Ferramentas de Alta Tecnologia Perfilam, Policiam e Punem os Pobres. Nova York: St. Martin's

Fink K. 2018. Abrindo as caixas-pretas do governo: liberdade de informação e responsabilidade algorítmica. Inform. Commun. Soc. 21(10):1453–71

Flood J. 2010. The Fires: Como uma Fórmula Computacional, Grandes Ideias e as Melhores Intenções Incendiaram Nova York City—E Determinaram o Futuro das Cidades. Nova York: Riverhead Books

Gama J, Žliobaitė I, Bifet A, Pechenizkiy M, Bouchachia A. 2014. Uma revisão sobre adaptação a mudanças de conceito. ACM Comput. Surv. 46(4):44

Garip F. 2020. O que o fracasso em prever resultados de vida pode nos ensinar. PNAS 117:8234–35

Garvie CA. 2019. Lixo entra, lixo sai: reconhecimento facial com dados falhos. Relatório, Georgetown Cent. Priv. Technol., **Washington, DC. https://www.flawedfacedata.com/**

Gilman M. 2020. Lawgoritmos da pobreza: um guia do advogado de pobreza para combater danos de decisões automatizadas em comunidades de baixa renda. Relatório, Data Soc. Res. Inst., NY. https://datasociety.net/wp-content/uploads/ 2020/09/Poverty-Lawgorithms-20200915.pdf

Goodman EP. 2019. O desafio da mudança algorítmica equitativa. The Regulatory Review, Fev. **https://www.theregreview.org/wp-content/uploads/2019/02/Goodman-The-Challenge-of-Equitable-Algorithmic-Change.pdf**

Green B. 2019. A Cidade Suficientemente Inteligente: Colocando a Tecnologia em Seu Lugar para Recuperar Nosso Futuro Urbano. Cambridge, MA: MIT Press

Guay JP, Parent G. 2018. Pernas quebradas, ajustes clínicos e risco de reincidência: uma análise das decisões de ajustar os níveis de risco com o LS/CMI. Crim. Justice Behav. 45(1):82–100

Hadden SG, Feinstein JL. 1989. Simpósio: sistemas especialistas. Introdução aos sistemas especialistas. J. Policy Anal. Manag. 8(2):182–87

Hamilton M. 2015. Aventuras no risco: prevendo reincidência violenta e sexual na lei de sentenças. Ariz. State **Law J. 47. https://arizonastatelawjournal.org/wp-content/uploads/2015/06/Hamilton_Final.pdf**

Harcourt BE. 2018. A falácia dos sistemas: uma genealogia e crítica da política pública e análise de custo-benefício. J. Legal Stud. 47(2):419–47

Harkinson J. 2016. Será que este algoritmo do Vale do Silício pode escolher quais pessoas sem-teto recebem moradia-ia? **Mother Jones, 29 de junho. https://www.motherjones.com/politics/2016/06/homelessness-data-silicon-valley-prediction-santa-clara/**

Hartzog W, Selinger E. 2015. Vigilância como perda de obscuridade. Washington Lee Law Rev. 72(3):1343–87

Hofman JM, SharmaA, Watts DJ. 2017. Predição e explicação em sistemas sociais. Science 355(6324):486–88

Hogue C. 2013. Relatório sumário da organização governamental: 2012. Relatório, US Census Bur., Washington, DC

Inst. Cyber Law Policy Secur. 2021. Força-tarefa de Pittsburgh sobre algoritmos públicos. **https://www.cyberinstitute.pitt.edu/algorithms**

Jacobs AZ, Wallach H. 2021. Medição e justiça. In Anais da Conferência ACM 2021 sobre Justiça, Responsabilidade e Transparência, pp. 375–85. Nova York: Assoc. Comput. Mach.

Johnson RA, Rostain T. 2020. Ferramenta de vigilância ou holofote sobre a desigualdade? Big data e a lei. Annu. Rev. Law Soc. Sci. 16:453–72

Kaplan A. 1964. A era do símbolo: uma filosofia da educação em biblioteconomia. Libr. Q. 34(4):295–304

Kaye K. 2021. Oakland, CA reconsiderará a tecnologia ShotSpotter "totalmente não comprovada". Smart Cities Dive, 9 de fev. **https://www.smartcitiesdive.com/news/oakland-ca-to-reconsider-totally-unpro shotspotter-tech/594718/**

Kelion L. 2020. Excel: por que o uso da ferramenta da Microsoft fez com que resultados da Covid-19 fossem perdidos. BBC, 5 de out. https:// **www.bbc.co.uk/news/technology-54423988**

Kellogg KC, Valentine MA, Christin A. 2020. Algoritmos no trabalho: o novo terreno contestado do controle. Acad. Manag. Ann. 14(1):366–410

Kitchin R. 2017. Pensando criticamente e pesquisando algoritmos. Inform. Commun. Soc. 20(1):14–29

Kleinberg J, Ludwig J, Mullainathan S, Obermeyer Z. 2015. Problemas de políticas de predição. Am. Econ. Rev. 105(5):491–95

Baixado de www.annualreviews.org. Convidado (guest) IP: 191.177.164.206 Em: Ter, 17 Mar 2026 21:02:23

Kleinberg J, Ludwig J, Mullainathan S, Sunstein C. 2018. Discriminação na era dos algoritmos. J. Legal Anal. 10:113–74

Kleinberg JM, Mullainathan S, Raghavan M. 2017. Compromissos inerentes na determinação justa de pontuações de risco. No 8º Innovations in Theoretical Computer Science Conference, ed. CH Papadimitrou, art. 43. Saarbrücken, Alemanha: Dagstuhl Publ.

Koepke JL, Robinson DG. 2018. Perigo à frente: avaliação de risco e o futuro da reforma da fiança. Wash. Law Rev. 93:1725–807

Koh PW, Sagawa S, Marklund H, Xie SM, Zhang M, et al. 2020. WILDS: um benchmark de mudanças de distribuição no mundo real-**de distribuição. arXiv. https://arxiv.org/abs/2012.07421**

Koops B. 2021. O conceito de function creep. Law Innov. Technol. 13(1):29–56

Krafft PM, Young M, Katell M, Huang K, Bugingo G. 2019. Definindo IA na política versus prática. In Proceedings da 2020 AAAI/ACM Conference on AI, Ethics, and Society (AIES). Nova York: Assoc. Comput. Mach.

Kroll JA, Huey J, Barocas S, Felten EW, Reidenberg JR, et al. 2016. Algoritmos responsáveis. Univ. Pa. Law Rev. 165:633–705

Ku M, Gil-Garcia JR. 2018. Pronto para análise de dados? Coleta e criação de dados em governos locais. In Proceedings da 19ª Conferência Internacional Anual sobre Pesquisa em Governo Digital, ed. A Zuiderwijk, CC Hinnant, art. 36. Nova York: Assoc. Comput. Mach.

Kube A, Das S, Fowler PJ. 2019. Alocando intervenções com base em resultados previstos: um estudo de caso sobre serviços para pessoas em situação de rua-de serviços. Proc. AAAI Conf. Artif. Intel. 33(1):622–29

Lee MK, Kusbit D, Kahng A, Kim JT, Yuan X, et al. 2019. WeBuildAI: estrutura participativa para governança algorítmica na Proceedings da ACM Conference on Human-Computer Interaction (CSCW), art. 181. Nova York: Assoc. Comput. Mach.

Lerman J. 2013. Big data e suas exclusões. Stanford Law Rev. Online 66:55–63

Levy K, Johns DM. 2016. Quando dados abertos são um Cavalo de Troia: a instrumentalização da transparência na ciência **e governança. Big Data Soc. 3(1). https://doi.org/10.1177/2053951715621568**

Lipsky M. 1980. Street Level Bureaucracy: Dilemas do Indivíduo nos Serviços Públicos. Nova York: Russell Sage Found.

Lipton ZC. 2018. O mito da interpretabilidade de modelos. Queue 16(3):31–57

Lum K, Chowdhury R. 2021. O que é um "algoritmo"? Depende de quem você pergunta. MIT Technology Review, fev. **26. https://www.technologyreview.com/2021/02/26/1020007/what-is-an-algorithm/**

Lum K, Isaac W. 2016. Prever e servir? Significance 13:14–19

MacKenzie D. 2008. Um Motor, Não uma Câmera: Como Modelos Financeiros Moldam Mercados. Cambridge, MA: MIT Press

Madden M, Gilman M, Levy K, Marwick A. 2017. Privacidade, pobreza e big data: uma matriz de vulnerabilidades para americanos pobres. Wash. Univ. Law Rev. 95:53–125

Malomo F, Sena V. 2017. Inteligência de dados para governos locais? Avaliando os benefícios e barreiras ao uso de big data no setor público. Policy Internet 9(1):7–27

McKelvey F, MacDonald M. 2019. Inovações em políticas de inteligência artificial no governo federal canadense-do Canadá. Can. J. Comm. 44(2):43–50

Metcalf J, Moss E, Watkins EA, Singh R, Elish MC. 2021. Avaliações de impacto algorítmico e responsabilidade: A co-construção de impactos. In Proceedings da 2021 ACM Conference on Fairness, Accountability, and Transparency, pp. 735–46. Nova York: Assoc. Comput. Mach.

Mitchell S, Potash E, Barocas S, D'Amour A, Lum K. 2021. Justiça algorítmica: escolhas, suposições e definições. Annu. Rev. Stat. Appl. 8:141–63

Molnar P, Gill L. 2018. Bots no portão: uma análise de direitos humanos da tomada de decisão automatizada no sistema de imigração e refúgio do Canadá. Res. Rep. 114, Citiz. Lab Int. Hum. Rights Progr., Fac. Law, Univ. Tor., Tor., Can. **https://citizenlab.ca/wp-content/uploads/2018/09/IHRP-Automated-Systems-**Report-Web-V2.pdf

Mormann F. 2021. Além dos algoritmos: em direção a uma teoria normativa da regulação automatizada. Boston Coll. Law Rev. 62:2

Mulligan DK, Bamberger KA. 2019. Aquisição como política: processo administrativo para aprendizado de máquina. Berkeley Technol. Law J. 34:773–851

Baixado de www.annualreviews.org. Convidado (guest) IP: 191.177.164.206 Em: Ter, 17 Mar 2026 21:02:23

Mulligan DK, Joshua AK, KohliN, Wong RY. 2019. Esta coisa chamada justiça: confusão disciplinar ao perceber um valor na tecnologia. Em Proceedings of the 2019 ACM Conference on Human-Computer Interaction, pp. 119–35. Nova Yorque: Assoc. Comput. Mach.

Nissenbaum H. 2010. Privacidade em Contexto: Tecnologia, Política e a Integridade da Vida Social. Stanford, CA: Stanford Univ. Press

Obermeyer Z, Powers B, Vogeli C, Mullainathan S. 2019. Dissecando o viés racial em um algoritmo usado para gerenciar a saúde de populações. Science 366(6464):447–53

O'Brien DT, Hill NE, Contreras M, Phillips NE, Sidoni G. 2018. Uma avaliação da equidade na política de atribuição baseada em domicílio das escolas públicas de Boston Doc., Boston Area Res. Initiat., Boston, MA. https://news.northeastern.edu/wp-content/uploads/2018/07/BPSHBAP.pdf

Ohm P. 2009. Promessas quebradas de privacidade: respondendo ao surpreendente fracasso da anonimização. UCLA Law Rev. 57:1701–77

Pasquale F. 2015. A Sociedade da Caixa Preta: Os Algoritmos Secretos que Controlam o Dinheiro e a Informação. Cambridge, MA: Harvard Univ. Press

Passi S, Barocas S. 2019. Formulação de problemas e justiça. Em Proceedings of the 2019 ACM Conference on Fairness, Accountability, and Transparency, pp. 39–48. Nova Yorque: Assoc. Comput. Mach.

Pozen DE. 2018. O desvio ideológico da transparência. Yale Law J. 128:100–65

Pozen DE, Schudson M, eds. 2018. Transparência Problemática: A História e o Futuro da Liberdade de Informação. Nova Yorque: Columbia Univ. Press

Quiñonero-Candela J, ed. 2009. Mudança de Conjunto de Dados em Aprendizado de Máquina. Cambridge, MA: MIT Press

Rabanser S, Günnemann S, Lipton Z. 2019. Falhando em voz alta: um estudo empírico de métodos para detectar mudança de conjunto de dados shift. In Advances in Neural Information Processing Systems, ed. MI Jordan, Y LeCun, SA Solla. Cambridge, MA: MIT Press

Reisman D, Schultz J, Crawford K, Whittaker M. 2018. Avaliações de impacto algorítmico: uma estrutura prática para responsabilização de Rep., AI Now Inst., Nova Yorkhttps://ainowinstitute.org/aiareport2018. agências públicas. pdf

Richardson R, ed. 2019. Enfrentando caixas pretas: um relatório paralelo do Sistema Automatizado de Decisão da Cidade de Nova York Task Force Rep., AI Now Inst., Nova Yorkhttps://ainowinstitute.org/ads-shadowreport-2019.html

Richardson R, Schultz JM, Sutherland VM. 2019. Litigando algoritmos 2019 Relatório dos EUA: novos desafios para o governo no uso de sistemas de decisão algorítmica. Rep., AI Now Inst., Nova Yorque. https://ainowinstitute.org/litigatingalgorithms-2019-us.pdf

Salganik MJ, Lundberg I, Kindel AT, Ahearn CE, Al-Ghoneim K, et al. 2020. Medindo a previsibilidade de resultados de vida com uma colaboração científica em massa. PNAS 117:8398–403

Salsburg D. 2001. A Dama Provando Chá: Como a Estatística Revolucionou a Ciência no Século XX. Nova Yorque: W.H. Freeman

Schrock AR. 2016. Hacking cívico como ativismo e defesa de dados: uma história da publicidade ao governo aberto dados. New Media Soc. 18(4):581–99

Scott JC. 1998. Vendo como um Estado: Como Certos Esquemas para Melhorar a Condição Humana Fracassaram. New Haven, CT: Yale Univ. Press

Selbst AD. 2017. Impacto desigual na polícia de big data. Ga. Law Rev. 52:109–95

Selbst AD, Barocas S. 2018. O apelo intuitivo das máquinas explicáveis. Fordham Law Rev. 87:1085–139

Selbst AD, Boyd D, Friedler SA, Venkatasubramanian S, Vertesi J. 2019. Justiça e abstração em sistemas sociotécnicos . In Proceedings of the Conference on Fairness, Accountability, and Transparency, pp. 59–68. Nova York: Assoc. Comput. Mach.

Simon J. 1988. Os efeitos ideológicos das práticas atuariais. Law Soc. Rev. 22(4):771–800

Singleton RA Jr., Straits BC. 2005. Abordagens à Pesquisa Social. Oxford, Reino Unido: Oxford Univ. Press

Skeem J, Scurich N, Monahan J. 2020. Impacto da avaliação de risco na justiça dos juízes ao sentenciar réus relativamente pobres. Law Hum. Behav. 44(1):51–59

Skitka LJ, Mosier KL, Burdick M. 1999. A automação influencia a tomada de decisão? Int. J. Hum. Comput. Stud. 51(5):991–1006

Sloane M, Moss E, Awomolo O, Forlano L. 2020. Participação não é uma solução de design para aprendizado de máquina. arXiv. **https://arxiv.org/abs/2007.02423**

Baixado de www.annualreviews.org. Convidado (guest) IP: 191.177.164.206 Em: Ter, 17 Mar 2026 21:02:23

Smith A. 2018. Atitudes em relação à tomada de decisões algorítmicas. Pew Research Center, 16 de nov. https://www.pewresearch.org/internet/2018/11/16/attitudes-toward-algorithmic-decision-making/
Smith A. 2019. Mais da metade dos adultos dos EUA confia que as autoridades usem reconhecimento facial de forma responsável. Pew Research Center, 5 de set. https://www.pewresearch.org/internet/2019/09/05/more-than-half-of-u-s-adults-trust-law-enforcement-t
Solow-Niederman A, Choi Y, Van den Broeck G. 2019. A vida institucional da avaliação de risco algorítmica. Berkeley Technol. Law J. 34:705–44
Spielkamp M, ed. 2019. Automatizando a Sociedade: Um Balanço da Tomada de Decisão Automatizada na UE. Berlim: AlgorithmWatch. https://www.ivir.nl/publicaties/download/Automating_Society_Report_2019.pdf
Spivack J, Garvie C. 2020. Uma taxonomia de abordagens legislativas para reconhecimento facial nos Estados Unidos. In Regulating Biometrics: Global Approaches and Urgent Questions, ed. A Kak, pp. 86–95. Nova York: AI Now Inst.
Stevenson M. 2018. Avaliando a avaliação de risco na prática. Minn. Law Rev. 103:303–84
Stevenson MT, Doleac JL. 2019. Avaliação de risco algorítmica nas mãos de humanos. Discuss. Pap. 12853, Inst. Labor Econ., Bonn, Alemanha. http://ftp.iza.org/dp12853.pdf
Stone HS. 1971. Introdução à Organização de Computadores e Estruturas de Dados. Nova York: McGraw-Hill
Surden H. 2007. Direitos estruturais na privacidade. SMULaw Rev. 60:1605–29
Suresh H, Guttag JV. 2020. Uma estrutura para entender as fontes de danos ao longo do aprendizado de máquina ciclo de vida. arXiv. https://arxiv.org/abs/1901.10002
Toros H, Flaming D. 2018. Priorizando a assistência a pessoas em situação de rua usando algoritmos preditivos: uma abordagem baseada em evidências Cityscape 20(1):117–46
Valverde M, Flynn A, eds. 2020. Cidades Inteligentes no Canadá: Sonhos Digitais, Projetos Corporativos. Toronto, Can.: James Lorimer & Co.
Veale M, Brass I. 2019. Administração por algoritmo? A gestão pública encontra o aprendizado de máquina no setor público-ing. In Algorithmic Regulation, ed. K Yeung, M Lodge, pp. 121–49. Oxford, Reino Unido: Oxford Univ. Press
Veale M, Van Kleek M, Binns R. 2018. Necessidades de design para justiça e responsabilidade no suporte algorítmico em tomada de decisão de alto risco no setor público. In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, pap. 440. Nova York: Assoc. Comput. Mach.
Vogl T. 2020. O impacto da evolução da tecnologia da informação nas formas de conhecimento no setor público serviço social: exemplos do Canadá e do Reino Unido. In Proceedings of the 53rd Hawaii International Conference on System Sciences, ed. TX Bui, pp. 2083–92. M¯anoa: Hawaii Int. Conf. Syst. Sci.
Vogl TM, Seidelin C, Ganesh B, Bright J. 2020. Tecnologia inteligente e o surgimento da burocracia algorítmica-cracia: inteligência artificial em autoridades locais do Reino Unido. Public Adm. Rev. 80:946–61
Waldman AE. 2019. A falsa promessa da lei de privacidade. Wash. Univ. Law Rev. 97:773–834
Weed J. 2021. Dicas para escrever currículos e passar pelos filtros de IA. New York Times, 21 de março. https://www.nytimes.com/2021/03/19/business/resume-filter-articial-intelligence.html
Wexler R. 2018. Vida, liberdade e segredos comerciais: propriedade intelectual no sistema de justiça criminal. Stanford Law Rev. 70:1343–429
Whittaker M, Crawford K, Dobbe R, Fried G, Kaziunas E, et al. 2018. Relatório AI Now 2018. Rep., AI Now, Nova York. https://ainowinstitute.org/AI_Now_2018_Report.pdf
Wildasin DE. 2010. Transferências intergovernamentais para governos locais. In Municipal Revenues and Land Policies, ed. GK Ingram, Y-HHong, pp. 47–76. Cambridge, MA: Lincoln Inst. Land Policy
Winston A. 2014. Conselho Municipal de Oakland revoga o Domain Awareness Center. East Bay Express, 5 de março. https://eastbayexpress.com/oakland-city-council-rolls-back-the-domain-awareness-center-1/
Wood C. 2013. Domain Awareness Center pode trazer policiamento proativo para Oakland. Government Technology, 17 de set. https://www.govtech.com/em/safety/Domain-Awareness-Center-Policing-Oakland.html
Yang Q, Steinfeld A, Rosé C, Zimmerman J. 2020. Reexaminando se, por que e como a interação humano-IA é especialmente difícil de projetar. In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems. Nova York: Assoc. Comput. Mach.
Yeung K. 2017. Regulação algorítmica: uma análise crítica. Regul. Gov. 12(4):505–23
Young M, Katell M, Krafft PM. 2019b. Regulação da vigilância municipal e responsabilidade algorítmica. Big Data Soc. 6(2). https://doi.org/10.1177/2053951719868492

Baixado de www.annualreviews.org. Convidado (guest) IP: 191.177.164.206 Em: Ter, 17 Mar 2026 21:02:23

YoungM, Magassa L, Friedman B. 2019c. Rumo a um design de políticas tecnológicas inclusivas: um método para grupos sub-representados para fortalecer documentos de políticas tecnológicas. Ethics Inform. Technol. 21(2):89–103

YoungM, Rodriguez L, Keller E, Sun F, Sa B, et al. 2019a. Além de aberto vs. fechado: equilibrando a privacidade individual e a responsabilidade pública no compartilhamento de dados. In Proceedings of the ACM Conference on Fairness, Accountability, and Transparency, pp. 191–200. New York: Assoc. Comput. Mach.

Zhang B, Dafoe A. 2019. Inteligência artificial: atitudes e tendências americanas. Public Opin. Rep., Cent. Gov. AI, Fut. Hum. Inst., Univ. Oxford, Oxford, Reino Unido

Ziewitz M. 2016. Governando algoritmos: mito, confusão e métodos. Sci. Technol. Hum. Values 41(1):3–16

Zuboff S. 2019. A Era do Capitalismo de Vigilância: A Luta por um Futuro Humano na Nova Fronteira do Poder. Nova York: Public Aff.

Baixado de www.annualreviews.org. Convidado (guest) IP: 191.177.164.206 Em: Ter, 17 Mar 2026 21:02:23