



Seção temática: Ética e integridade na investigação nas Ciências Humanas e Sociais na era da IA

**Democracia, epistocracia e o papel da inteligência artificial
na melhoria das decisões democráticas***

**Democracy, epistocracy, and the role of artificial intelligence
in improving democratic decisions**

**Democracia, epistocracia y el papel de la inteligencia artificial
en la mejora de las decisiones democráticas**

Steven S. Gouveia**

<https://orcid.org/0000-0003-4176-7040>

Thailize Brandolt***

<https://orcid.org/0000-0002-3713-8877>

Resumo: Este artigo examina criticamente os desafios epistêmicos das democracias contemporâneas e propõe o uso ético da inteligência artificial (IA) como meio de fortalecer práticas democráticas informadas, alinhando-as a certos ideais da epistocracia. Partindo da crítica clássica de Platão à democracia e da defesa contemporânea da epistocracia por autores como Jason Brennan, discutem-se as limitações cognitivas que afetam a qualidade das decisões políticas, incluindo ignorância racional, polarização, vieses cognitivos e desinformação, e como elas prejudicam tanto a legitimidade quanto a eficácia dos governos eleitos. Em vez de adotar uma epistocracia elitista e excludente, defende-se aqui uma democracia mais informada, capaz de conjugar o ideal democrático de igualdade com o acesso ampliado e assistido ao conhecimento relevante para a deliberação política. A IA é apresentada como possível aliada nesse esforço, funcionando como infraestrutura de capacitação. O artigo também analisa criticamente os riscos da IA, incluindo viés algorítmico, manipulação informacional e reprodução de desigualdades, defendendo uma abordagem atenta às consequências normativas desse tipo de tecnologia. Por fim, sustenta-se que uma democracia mais justa e informada só será possível se tecnologias como a IA forem concebidas como ferramentas de justiça epistêmica, orientadas não à substituição, mas ao fortalecimento do julgamento político autônomo dos cidadãos.

Palavras-chave: Democracia. Epistocracia. Inteligência artificial.

* Financiamento: SG: Financiado pelo Projeto CEEC Individual da FCT 2022.02527.CEECIND, no Grupo Mente, Linguagem e Ação, Instituto de Filosofia, Universidade do Porto (Faculdade de Letras, Via Panorâmica s/n, P-4150-564 Porto, Portugal). TB: Esta pesquisa obteve apoio do Programa Institucional de Internacionalização da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior/Programa de Doutorado-sanduíche no Exterior (Capes/PDSE).

** Investigador Contratado do *Mind, Language and Action Group* (MLAG), Instituto de Filosofia, Universidade do Porto, Portugal, e Professor Honorário da Faculdade de Medicina Andrés Bello, Chile. E-mail: <stevensequeira92@hotmail.com>.

*** Doutoranda em Filosofia pela Universidade de Caxias do Sul (UCS) e Pesquisadora Visitante no Departamento de Filosofia da Universidade do Porto, Portugal. E-mail: <thailize.brandolt@gmail.com>.

Abstract: This article critically examines the epistemic challenges of contemporary democracies and proposes the ethical use of artificial intelligence (AI) as a means to strengthen informed democratic practices, aligning them with certain ideals of epistocracy. Drawing on Plato's classical critique of democracy and the contemporary defense of epistocracy by authors such as Jason Brennan, the article discusses the cognitive limitations that affect the quality of political decisions—including rational ignorance, polarization, cognitive biases, and disinformation—and how these undermine both the legitimacy and the effectiveness of elected governments. Instead of adopting an elitist and exclusionary epistocracy, the article advocates for a more informed democracy, one that combines the democratic ideal of equality with expanded and assisted access to relevant knowledge for political deliberation. AI is presented as a potential ally in this effort, functioning as an enabling infrastructure. The article also critically analyzes the risks of AI, including algorithmic bias, informational manipulation, and the reproduction of inequalities, advocating for an approach attentive to the normative consequences of this technology. Finally, it argues that a more just and informed democracy will only be possible if technologies such as AI are conceived as tools of epistemic justice, oriented not toward replacing but rather strengthening citizens' autonomous political judgment.

Keywords: Democracy. Epistocracy. Artificial intelligence.

Resumen: Este artículo examina críticamente los desafíos epistémicos de las democracias contemporáneas y propone el uso ético de la inteligencia artificial (IA) como medio para fortalecer las prácticas democráticas informadas, alineándolas con ciertos ideales de la epistocracia. A partir de la crítica clásica de Platón a la democracia y de la defensa contemporánea de la epistocracia por autores como Jason Brennan, se analizan las limitaciones cognitivas que afectan la calidad de las decisiones políticas, incluida la ignorancia racional, la polarización, los sesgos cognitivos y la desinformación, y cómo estas debilitan tanto la legitimidad como la eficacia de los gobiernos electos. En lugar de adoptar una epistocracia elitista y excluyente, aquí se defiende una democracia más informada, capaz de conjugar el ideal democrático de igualdad con un acceso ampliado y asistido al conocimiento relevante para la deliberación política. La IA se presenta como una posible aliada en este esfuerzo, funcionando como infraestructura de capacitación. El artículo también analiza críticamente los riesgos de la IA, entre ellos el sesgo algorítmico, la manipulación informativa y la reproducción de desigualdades, defendiendo un enfoque atento a las consecuencias normativas de este tipo de tecnología. Finalmente, se sostiene que una democracia más justa e informada solo será posible si tecnologías como la IA se conciben como herramientas de justicia epistémica, orientadas no a sustituir, sino a fortalecer el juicio político autónomo de los ciudadanos.

Palabras clave: Democracia. Epistocracia. Inteligencia artificial.

Introdução: os obstáculos e as deficiências da democracia, a epistocracia como (possível) alternativa

Desde seu surgimento, no período da Grécia Antiga, a democracia tem convivido com críticas e contestações, ao mesmo tempo que, por ora, ainda não nos deparamos com outro tipo de regime político que melhor comporte a ideia de igualdade e justiça na participação política. No entanto, quando pensamos em um mundo tão desigual economicamente quanto o que vivemos hoje, principalmente nos países do sul global, e ainda mais se levarmos em consideração o grande número de desinformações espalhadas pela internet e o alto potencial da inteligência artificial (IA) de interferir cada vez mais em decisões eleitorais, como garantir que a participação democrática seja realmente aquela que permite o melhor tipo de sistema político?

Na sociedade contemporânea, a ideia de governos autoritários e excludentes tem ganhado força com a esperança de serem a solução diante do descrédito em relação aos sistemas democráticos, que, mesmo prometendo uma participação popular mais justa, acabam esbarrando nos mesmos e antigos problemas políticos: governantes eleitos que se revelam totalmente incapazes de representar os verdadeiros interesses e as necessidades daqueles que os elegeram. Mas por que isso acontece? Por que a democracia está cada vez mais ameaçada e frágil? E por que, mesmo quando ela é exercida, não é garantia de um governo mais justo?

O presente texto procura explorar o papel transformador da IA na interface entre democracia e epistocracia, um sistema em que o conhecimento é central para a tomada de decisões. Como sistema político, a democracia tradicional valoriza a igualdade na participação política, enfrentando muitas vezes limitações devido à desinformação, à polarização e às decisões tomadas sem base em evidências. A epistocracia, por sua vez, propõe um modelo em que o conhecimento especializado desempenha um papel maior, mas corre o risco de elitismo e exclusão.

Uma das críticas mais famosas à democracia foi apresentada por Platão em sua obra *A República*, na qual o personagem Sócrates faz uma minuciosa análise de como as diferentes formas de governo injustas surgem da corrupçãoumas das outras. No caso da democracia, esta seria o resultado da degeneração da oligarquia, motivada pela insaciável cobiça dos homens de quererem ser ricos a qualquer custo. Desse modo, para Platão, a democracia nasceu a partir de almas ignorantes, avarentas e ambiciosas, o que permite que quem governa a pôlis sejam pessoas dominadas por desejos e vicissitudes.¹ O governo ideal, para Platão, seria aquele regido pelo rei-filósofo, pelo sábio, pois, em sua visão, a alma da pôlis reflete a alma de seu governante, e só é possível haver um bom governo por meio de um governante que tenha conhecimento (*episteme*) e boa ordenação da *psykhe*. Podemos dizer, portanto, que Platão foi o primeiro a defender um certo tipo de epistocracia, que podemos traduzir seguramente como o governo daqueles que sabem. Um adepto da epistocracia também poderá defender que somente aqueles que têm conhecimento possam ter o direito ao voto e a eleger seus governantes, em uma espécie de união entre democracia e epistocracia.

Entretanto, se formos transpor esse conceito de uma democracia epistocrática para os dias atuais, ela se tornaria mais ou menos justa diante da alta disseminação de informações que nem sempre estão comprometidas com a verdade nos diferentes canais de comunicação? Como poderíamos garantir que a epistocracia seja construída sobre os princípios da autoridade epistêmica e da responsabilidade eleitoral diante dos avanços das IAs, controladas pelas principais empresas de tecnologia do mundo, que têm como principal motivação os interesses mercadológicos? E como garantir que a epistocracia não resulte em um sistema excludente, que limita o conhecimento apenas àqueles que já fazem parte de uma elite intelectual ou monetária? Para tentarmos responder a essas perguntas, vamos, inicialmente, compreender os principais argumentos que os teóricos contemporâneos utilizam para defender um regime epistocrático.

A epistocracia contemporânea tem como um de seus principais defensores Jason Brennan, que escreveu *Against Democracy* (Brennan, 2016). No livro, o autor apresenta nove diferentes argumentos para defender a epistocracia como um sistema – normativamente – superior à democracia. Não vamos aqui nos ater aos pormenores de cada um deles, apenas destacar alguns:

- (1) a separação em três níveis de eleitores: os *hobbits* (desprovidos de conhecimento e sustentação política, mas que têm papel central nas decisões), os *hooligans* (possuem algum tipo de formação e utilizam esse “conhecimento” para inflar e/ou deslegitimar a contraparte) e os vulcanos (aqueles que possuem conhecimento científico considerável, que conseguem perceber o fenômeno político por meio de um distanciamento intelectual que lhes permite analisar as melhores opções – mas estes são a minoria, deixando a democracia refém da polarização entre *hobbits* e *hooligans*);

¹ “Quando um jovem, criado como ainda há pouco dissemos, na ignorância e na avareza, prova o mel dos vespões e se associa com estas criaturas ferozes e astutas, que sabem como fornecer prazeres de toda a qualidade, variedade e condição, e é então que, podes crê-lo, principia para ele a transformação do oligárquico e que existe na sua alma para o democrático.” (Platão, 1993, livro VIII, 559e).

- (ii) a ignorância racional: o desinteresse pela busca de informação detalhada e fidedigna é percebido como um custo mais alto do que o benefício de se votar de forma mais consciente, já que esse voto “ignorante” individual é visto como tendo o mesmo peso que um voto individual “consciente”;
- (iii) a participação política torna os eleitores mais conflitantes do que bem informados;
- (iv) a participação democrática não amplia a autonomia dos eleitores, acaba concentrando o poder representativo a interesses individuais ou de pequenos grupos;
- (v) se é necessária licença para dirigir ou para exercer certos tipos de profissão, deveria existir também um certo tipo de formação prévia que garantisse um mínimo de preparado técnico e intelectual para poder votar, já que atuar democraticamente é atuar com escolhas que afetam a todos;
- (vi) pelo direito a um governo competente: deve haver condições tanto para os governantes, quanto para os votantes, de poder exercerem suas funções com o máximo de informação confiável e conhecimento, porque somente assim é possível haver um governo competente.

Esses são alguns dos argumentos que Brennan (2016) utiliza não apenas para defender a epistocracia, mas para mostrar que a democracia não pode ser defendida apenas pelo que expressa ou simboliza, ou simplesmente por ela mesma, mas deve ser vista como uma ferramenta útil que, muitas vezes, pode falhar ao exercer sua função.

No entanto, os argumentos apresentados por Brennan (2016) também podem ser contestados por meio de observações como: infelizmente, não dispomos de critérios objetivos para saber quais seriam as condições necessárias para um eleitor ser considerado “informado”; tampouco poderíamos garantir que esse conhecimento fosse disponibilizado de forma igualitária a todos aqueles que gostariam de se tornar votantes apropriados, visto que, principalmente nos países do sul global, a desigualdade na distribuição de conhecimento ainda é muito grande; também poderíamos questionar se essa maneira de perceber as eleições não seria mais uma forma de elitizar ainda mais as decisões políticas.

Se formos considerar o princípio fundamental da política democrática – o de tratar todos os cidadãos como livres e iguais membros da comunidade política – veremos que esse preceito de equidade ou justiça social, que tem o objetivo de elevar o nível em direção a uma definição normativa de democracia nas práticas sociais, não funciona da forma prevista. Ainda mais se considerarmos que, dentro da esfera pública, as interações comunicativas entre indivíduos e grupos ocorrem diariamente sem muita mediação normativa, em um espaço cujas opiniões alternativas interagem e o debate público se desenrola como resultado dessa troca, que influencia diretamente a dinâmica democrática.

Quando pensamos na formação de maiorias eleitorais e de governos, devemos levar em conta também a opinião pública que nasce a partir das desigualdades sociais e políticas que afetam a estrutura e o funcionamento da esfera pública (Giusti; Piras, 2021). A desigualdade social e política produz também dano e injustiça epistêmica para os membros de grupos marginalizados, pois a posição que o sujeito ocupa na esfera pública – sua troca de informações e opiniões – acaba afetando diretamente seu acesso ao conhecimento e seu reconhecimento como sujeito conchedor. Por mais que a democracia almeje os princípios de igualdade e justiça, ao propor que todos tenham o mesmo tipo de participação e peso nas decisões políticas, o que cada sujeito consegue contribuir,

em face do conhecimento que adquire ao longo da vida, esbarra nas desigualdades sociais de cada sociedade.

O projeto de uma epistocracia, tal como proposto por Brennan (2016), poderia ser o vislumbre de um regime político que valorize o conhecimento e o acesso às informações necessárias e adequadas para o bom exercício do voto. Claro, Brennan (2016) fala a partir da realidade que conhece, a norte-americana, e parece não sanar as dificuldades de acesso e aquisição de conhecimento enfrentadas por grupos marginalizados. Diante disso, podemos pensar se haveria a possibilidade de transferir suas reflexões também para outros cenários democráticos.

Hoje, felizmente, dispomos de muitas ferramentas e tecnologias que poderiam servir como alternativas para sanar essas deficiências que a epistocracia parece apresentar. Consideramos que a IA, se utilizada conforme preceitos éticos previamente estabelecidos, pode ser uma importante aliada.

Os perigos da IA para a democracia

Antes de falarmos sobre o potencial positivo da IA em relação à epistocracia, precisamos também considerar suas limitações e os cenários atuais nos quais ela exerce um papel fundamental, principalmente na disseminação de informações falsas (*fake news*) e/ou na reprodução de estereótipos preconceituosos. Uma ferramenta como o *ChatGPT* (sigla para *Generative Pre-Trained Transformer*), famoso por responder perguntas e solicitações dos mais variados temas de forma mais ou menos coerente, também auxilia na criação de áudios e vídeos realistas e na geração de imagens que parecem reais, mas não são. Essas imagens, áudios e vídeos são usados cada vez com mais frequência para influenciar eleitores em períodos eleitorais.

Dentro dessa realidade, é importante mencionarmos que o crescente impacto das tecnologias de IA na vida humana permite o fortalecimento do escopo e da escalada das campanhas de desinformação e disseminação de notícias falsas. A atuação da IA nesse campo é ambígua, pois, ao mesmo tempo que acelera a propagação de dados *online* por meio de plataformas de mídia social, também nos possibilita detectar automaticamente conteúdo falso e removê-lo com um nível relativamente alto de precisão. Consequentemente, a relação entre IA e desinformação ampliou o discurso ético sobre o uso e a implementação da IA em nossas práticas diárias.

Além disso, as ferramentas de IA também têm sido usadas para reproduzir estereótipos e preconceitos de diferentes tipos. Um exemplo pode ser observado em 2015, quando pesquisas acadêmicas e de mídia revelaram um claro preconceito de gênero no mecanismo de busca do *Google*. Durante o estudo, os principais resultados para a busca de imagens de “CEO” [sigla para *Chief Executive Officer* – Diretor Executivo] mostraram predominantemente imagens de homens e, quando o usuário era identificado como mulher, o mecanismo de busca exibiu menos anúncios para cargos executivos (Langston, 2015). Isso ilustra um preconceito histórico, já que reflete desigualdades de gênero existentes na sociedade.

Mais recentemente, ao testar um tipo de busca semelhante, mas desta vez utilizando o *ChatGPT* para criar uma imagem que representasse um cargo de liderança em uma grande empresa, o resultado foi uma imagem de um homem branco em pé, em evidência, enquanto mulheres apareciam sentadas à mesa apenas observando. O mesmo aconteceu quando solicitamos a criação de uma imagem que representasse um “bom governante”: o resultado foi a imagem de um homem branco, de cabelos claros, com uma coroa na cabeça (Tremewan, 2024). É claro que estamos tratando de ferramentas em constante aperfeiçoamento, que dependem muito do perfil dos usuários e de comandos mais específicos para aprimorar os resultados gerados.

Somados à manipulação da informação, à desigualdade na distribuição do conhecimento e ao reforço de estereótipos preconceituosos, podemos questionar também até que ponto as ferramentas de IA nos permitem exercer nossa liberdade de escolha e pensamento crítico. Afinal, os conteúdos que consumimos hoje na internet são alimentados por dados e algoritmos que têm como base comportamentos já previstos e executados pelos usuários. Dessa maneira, os algoritmos estudam nossos gostos, mapeiam nossos hábitos e sugerem o que devemos consumir (sejam produtos, sejam informações/conhecimento) com base nesses dados. Assim sendo, as opções que nos são disponibilizadas estão dentro de um limite do que já estamos predispostos a gostar, como se consumir algo fora disso não fosse permitido, ou como se nosso gosto ou desejo jamais pudesse variar, ou ainda como se nosso conhecimento a respeito do que queremos nunca pudesse evoluir.

Além disso, os algoritmos não disponibilizam todas as opções possíveis para determinado tipo de produto, serviço ou informação: mostram principalmente aquelas ligadas a anunciantes, marcas com maior engajamento e maior número de seguidores. Portanto, nossa liberdade de escolha e possibilidade de adquirir conhecimento parecem condicionadas a um círculo vicioso sem fim. Quando falamos da alimentação de dados das ferramentas de IA, estamos falando de dados e informações produzidos com base no comportamento humano majoritário – ou naquele tipo de comportamento que mais interessa à reprodução do *modus operandi* atual. Conforme Ireni-Saban e Shermans (2021, p. 46):

Danks e London (2017) destacam fontes computacionais notáveis de viés, como o viés de dados de treinamento e processamento algorítmico. Por exemplo, os dados de entrada usados podem ser tendenciosos e levar a saídas também tendenciosas para as tarefas algorítmicas. [...]. Como os algoritmos orientados por IA se tornaram altamente prevalentes em nossos processos de tomada de decisão e práticas do dia a dia, o risco de gerar saídas discriminatórias e ofensivas aumenta significativamente [...]. Além do viés e das notícias falsas, pode-se mencionar violações de justiça e privacidade como brechas éticas significativas.

Diante dessa necessidade de refletirmos com mais profundidade sobre os aspectos e os impactos ético-políticos envolvendo o uso da IA, surgem algumas questões fundamentais: De quem é a responsabilidade por atribuir limites normativos à atuação e ao uso da IA? Como estabelecer um limiar na disseminação de informações sem comprometer a liberdade de expressão? Qual é o papel regulatório das *Big Techs* – empresas de tecnologia que dominam o mercado de inovação e comunicação e que hoje são, em grande parte, responsáveis pelo uso das IAs? Muitas discussões de cunho filosófico têm ganhado espaço nos últimos anos, justamente diante da preocupação com os usos indiscriminados dessas ferramentas, mas poucas reflexões têm, de fato, se traduzido em mudanças concretas. Eticistas e filósofos morais têm tentado articular as já conhecidas vertentes éticas – sejam deontológicas ou consequencialistas – para aplicação de regras morais nos usos da IA.

Para Ireni-Saban e Shermans (2021), essas teorias ainda parecem insuficientes para explicar completamente as complexidades das considerações morais às quais o uso da IA está sujeito quando aplicado no mundo real. Para abranger os dilemas éticos contemporâneos da IA, faz-se necessária uma abordagem mais flexível e contextual. As autoras propõem, como alternativa, um sistema ético da interseccionalidade, um comprometimento crítico que leva em consideração ideais normativos voltados à justiça social. Além disso, elas salientam que o desenvolvimento de sistemas de IA evidenciou, com ainda mais clareza sem precedentes, as diferenças socioeconômicas entre diferentes identidades sociais: “Uma estrutura interseccional para a ética da IA pode ser usada para examinar os vieses e problemas algorítmicos enraizados em sistemas e aplicativos existentes baseados em IA” (Ireni-Saban; Shermans, 2021, p. 41).

O conceito de interseccionalidade tem origem no movimento feminista; mais precisamente, o termo surgiu como crítica à segunda onda feminista, acusada de privilegiar pautas e reivindicações apenas de mulheres brancas e elitizadas, sem considerar aspectos de raça e classe, por exemplo. O termo foi usado pela primeira vez por Kimberlé Crenshaw (1989), que fez uma alusão a estradas que se cruzam para iluminar diferentes níveis de opressão com base em gênero, raça e classe, que interagem entre si e criam experiências únicas de marginalização e discriminação.

Ireni-Saban e Shermans (2021) sugerem que a interseccionalidade, como metodologia, pode ir além dos debates feministas, porque encoraja pesquisadores a investigarem os efeitos multifacetados das experiências de opressão, afastando-se de metodologias tradicionais e conservadoras que buscam uma solução universal para seus problemas. Contudo, como a metodologia da interseccionalidade se relaciona com os dilemas éticos envolvendo os usos da IA? Podemos pensar em dois princípios principais que norteiam esse conceito: diversidade e inclusão. Aplicando a valorização da diversidade no desenvolvimento da IA, poderíamos diminuir os vieses atravessados ou estereotipados, em sua maioria reproduzidos quando não se observam aspectos diversos de gênero, raça, classe, cultura etc. Do mesmo modo, quando refletimos sobre inclusão, pensamos em uma proposta que estimule uma representação mais apropriada e justa do envolvimento de indivíduos marginalizados e desfavorecidos no *design* e na engenharia de sistemas de IA e de algoritmos. Assim, poderíamos garantir que a IA estaria mais próxima de reproduzir e representar a realidade social.

O desdobramento de uma investigação ético-filosófica a respeito do uso da IA, que leve em consideração os diferentes tipos de opressão e discriminação e que não esteja apenas preocupada em diminuir os prejuízos meramente formais e consequencialistas dessas ferramentas, pode servir de instrumento para a estruturação de uma sociedade em que o conhecimento necessário à formação de bons votantes esteja acessível de forma igualitária e justa.

As (potenciais) vantagens da IA para uma democracia informada (epistocracia)

A IA pode ser apresentada como uma possível ponte entre estes dois sistemas (democracia e epistocracia), capaz de corrigir falhas e potencializar a qualidade das decisões democráticas sem comprometer os princípios igualitários. Como vimos, a crise da legitimidade moral e política das democracias contemporâneas não reside apenas em uma mera falha institucional ou simbólica desse regime político particular, ou na situação de desigualdades sociais e econômicas concretas, mas, acima de tudo, em um problema cognitivo: os agentes chamados a escolher e aplicar sua liberdade política estão influenciados por uma variedade de vieses cognitivos, polarizações afetivas e/ou emocionais, alta resistência a evidências e fatos relevantes para tomar decisões democráticas, entre outros fatores, que afetam diretamente tanto a legitimidade dos governos eleitos quanto a qualidade de suas decisões políticas (Brennan, 2016; Gouveia, 2020).

Nesse cenário, a proposta epistocrática – isto é, uma tentativa de valorizar os fins de justiça institucionalmente desejados da arena política – emerge como um conjunto de tentativas mais ou menos concretas de abordar esse problema epistêmico das democracias do século XXI. Para tal, é possível fugir dos cenários epistocráticos mais tradicionais, como os sufrágios limitados ou os testes de conhecimento – que levantam algumas dificuldades normativas, especialmente em relação à sua possível exclusão elitista e paternalista (e.g., Estlund, 2008) –, e argumentar a favor de uma forma de pensar a epistocracia que procure contabilizar, por um lado, a maior importância do conhecimento individual dos cidadãos e, por outro, manter a práxis democrática atual, pensando em formas de fortalecer a direção a uma verdadeira práxis democrática, baseada na qualidade informational da população e permitindo uma participação efetiva e inclusiva de todos os membros da sociedade (e não apenas de um grupo particular).

É nesse contexto teórico que a IA pode desempenhar um papel, senão central, altamente relevante: se pensada e aplicada com sustentação ética, a IA poderá oferecer um conjunto de ferramentas promissoras para o fortalecimento de uma democracia mais informada e, como consequência, mais justa para todos, como passamos a analisar a seguir.

Mitigação da ignorância racional

A ignorância racional, conceito cunhado por Anthony Downs (1957), aponta para um fenômeno psicológico que parece estar presente nas democracias modernas. Esse fenômeno parte do seguinte pressuposto: perante a quantidade de informação e esforço cognitivo necessários para que um cidadão esteja minimamente preparado para exercer um voto informado – e levando em conta o pouquíssimo poder que cada voto individual representa em uma votação democrática –, grande parte dos cidadãos percebe que, como o voto individual não alterará o resultado de uma eleição, não vale a pena investir tempo e energia na aquisição e na compreensão desse conhecimento especializado.

Do ponto de vista psicológico, isso funciona, no fundo, como um desincentivo ao conhecimento, o que caracteriza as democracias atuais e afeta, por consequência, a qualidade epistêmica das decisões coletivas no âmbito político. Se grande parte dos eleitores decide seu voto com base em fatores não racionais – como ideologias simplistas, desinformação ou intuições pouco esclarecidas –, a idealização do processo de decisão democrática fica afetada, comprometendo, assim, a própria qualidade dos governos eleitos e, por consequência, as próprias políticas a serem executadas.

Se observarmos a complexidade dos contextos democráticos, veremos que estamos, muitas vezes, diante de decisões sobre temas altamente especializados e de grande complexidade, como alterações climáticas, política monetária, liberdade de circulação e imigração ou regulação da IA. Não é claro que não seja plausível exigir, de um ponto de vista da ética normativa, algum nível de literacia científica e informacional sobre esses e outros temas por parte de quem decide votar em uma eleição. Claro que, por outro lado, exigir que todos os cidadãos adquiram, por si só, conhecimentos relevantes para tomar decisões informadas nesses temas é um cenário implausível e até injusto diante das desigualdades educacionais e cognitivas já existentes na sociedade. É precisamente aqui que a IA pode ter um papel fundamental como aliada de uma versão aprimorada da democracia e, por isso, mais justa para todos.

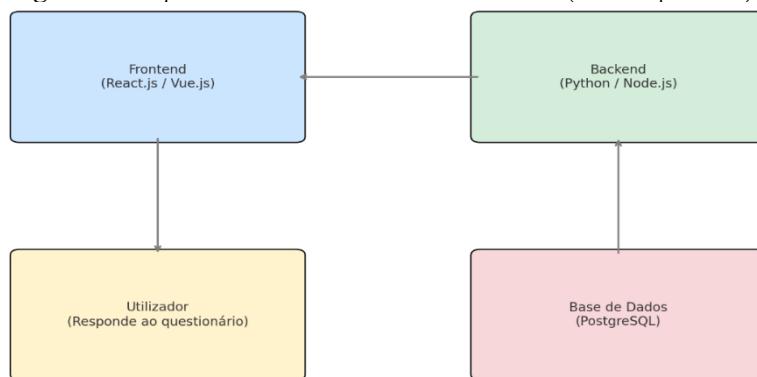
Por meio de aplicações específicas da IA, é possível mitigar pelo menos alguns dos efeitos que a ignorância racional provoca, se ela for concebida como um mecanismo de suporte cognitivo assistido, isto é, como uma plataforma cidadã capaz de oferecer aos cidadãos informações relevantes para a tomada de decisões em contexto político. A IA permite a criação de produtos que não sejam uniformes e generalizados, mas que possam ser constantemente adaptados ao nível de compreensão e ao contexto do usuário.

Pensemos, por exemplo, no uso de ferramentas como assistentes cívicos baseados em IA, que poderiam fornecer aos cidadãos uma assistência política constante: responder a dúvidas, comparar propostas políticas de diferentes partidos sobre o mesmo tema, simplificar a linguagem jurídica – muitas vezes incompreensível para a maioria das pessoas –, fornecer *fact-checking* regular, entre outras funções. Essa aplicação particular, embora não altere substancialmente os problemas de irracionalidade já identificados, poderia ao menos ajudar alguns cidadãos a tomar decisões mais informadas sobre suas próprias preferências políticas.

Exemplos ilustrativos desse tipo de tecnologia têm sido explorados em contextos experimentais ou em algumas campanhas eleitorais, como é o caso da plataforma *SmartVote*², na Suíça, ou do *VoteMatch*³, nos Países Baixos (já expandida a outros países da União Europeia). Ambas empregam algoritmos que correlacionam, por um lado, os programas partidários a serem escrutinados e, por outro, as preferências políticas dos usuários, que vão respondendo a uma série de questões de caráter político, social e econômico. No final, é oferecida não só a indicação do partido político que mais se encaixa nas preferências indicadas, como também uma explicação detalhada que justifica os porquês.

Em Portugal, a plataforma *EuVoto*⁴ foi utilizada em distintas campanhas eleitorais, como a campanha legislativa de 2015 e a campanha presidencial de 2021. Com um *design* bastante acessível, os usuários respondiam a cerca de 30 perguntas sobre tópicos centrais dessas campanhas políticas (e.g., habitação, saúde, fiscalidade, ambiente), indicando cada resposta o grau de concordância ou discordância com a temática abordada em cada questão. Ao olharmos para o lado mais técnico, resumimos, na Figura 1, a arquitetura que, provavelmente, compunha essa plataforma.

Figura 1 – Arquitetura de IA da Plataforma “Eu Voto” (modelo plausível)



Fonte: Elaborada pelos autores.

A plataforma fazia uso de uma arquitetura web modular, composta por um *frontend* interativo (e.g., *frameworks* como React ou Vue.js), um *backend* analítico (e.g., Python ou Node.js) e uma base de dados relacional que armazenava, de forma totalmente anonimizada e não persistente, tanto as posições políticas dos partidos quanto as preferências dos usuários. Utilizando métodos simples e transparentes, como a distância euclidiana ponderada, o algoritmo de correspondência comparava vetorialmente as respostas dos cidadãos com as dos programas políticos. Em alguns casos, os usuários podiam, inclusive, atribuir maior peso ou importância política e ética a determinados temas, o que reforçava a identificação partidária.

Os resultados eram apresentados em formato visual (por meio de radiais, matrizes ou gráficos de barras), o que facilitava a compreensão da afinidade política com determinado partido naquela eleição em particular, além de fornecer explicações em linguagem natural e ligações aos programas ou manifestos políticos, de modo a complementar a literacia cívica de cada participante. Isso demonstra o papel educativo desse tipo de tecnologia. É importante referir que, do ponto de vista ético, o *app* garantia o anonimato e a privacidade dos dados: não havia retenção de

² SmartVote encontra-se disponível em: <https://www.smartvote.ch/en/home>. Acesso em: 6 ago. 2025.

³ Sobre o *VoteMatch*, disponível em: https://prodemos.nl/app/uploads/sites/6/2023/03/International-Votematch_brochure-online.pdf. Acesso em: 6 ago. 2025.

⁴ *EuVoto* está disponível em: <https://www.euvoto.pt/>. Acesso em: 6 ago. 2025.

informações nem a necessidade de aceitar *cookies* personalizados, em consonância com o princípio da autodeterminação informacional avançado por Floridi *et al.* (2018).

Em suma, trata-se de um exemplo paradigmático de uma tecnologia com espírito epistocrático – pelo menos o espírito epistocrático que estamos a tentar defender neste artigo –, oferecendo ferramentas a todos aqueles que desejam tomar decisões democráticas mais informadas, permitindo o aumento da informação legítima disponível e, ao mesmo tempo, a redução do custo cognitivo da escolha política. Ademais, há uma orientação personalizada (que decorre da escolha de cada cidadão, e não de uma estrutura externa) ideologicamente neutra, o que favorece uma maior justiça epistêmica.

Por fim, corrige-se, ainda que em uma tentativa, a falta de reforço epistêmico na decisão democrática, permitindo que ao menos alguns cidadãos votem com mais conhecimento do que o fariam sem o uso dessa plataforma. Isso amplia a autonomia e a liberdade de pensamento em comparação com cidadãos que votam de forma altamente enviesada e limitada por um círculo informacional ideológico específico.

É claro que devemos sempre frisar que a introdução da IA nesse contexto deve exigir elevadas garantias de transparência, auditabilidade e neutralidade ideológica: seria muito suspeito, do ponto de vista da justiça informacional, que essas plataformas fossem desenvolvidas pelos próprios partidos, por exemplo. Assim, é necessário garantir que o desenvolvimento e a aplicação dessas tecnologias sejam conduzidos por agentes ou instituições neutras, de modo a representar uma verdadeira ferramenta de apoio cognitivo aos cidadãos, e não uma ameaça à sua autonomia democrática.

Devemos pensar em uma infraestrutura de capacitação epistêmica para todos – cuja universalização seja fator essencial para combater a falta de acesso ao conhecimento relevante, claramente presente nas democracias contemporâneas –, que permita uma deliberação pública e coletiva como apoio, e jamais como substituição, ao julgamento político.

Combate à desinformação e apoio à deliberação pública qualificada

Como vimos na subseção anterior, a ignorância racional representa um dos problemas estruturais da práxis democrática que pode e deve ser combatido. Se, por um lado, podemos atuar no aumento do conhecimento disponibilizado aos cidadãos, podemos também desenvolver, paralelamente, um projeto fundamentalmente positivo de combate aos elevados volumes de desinformação que parasitam as sociedades atuais.

Essa desinformação não é meramente acidental, mas provocada intencionalmente por agentes cujo foco é manipular e desestabilizar a opinião pública, de modo a distorcer a realidade epistêmica de determinado problema político e, ao mesmo tempo, corroer a confiança dos cidadãos nas instituições democráticas em geral. No contexto de um mundo cada vez mais digitalizado, o problema da desinformação é amplificado em sua dimensão qualitativa e quantitativa pela escala e velocidade das redes digitais, onde é possível partilhar qualquer tipo de proposição falsa sem critérios de supervisão e correção epistêmica (Wardle; Derakhshan, 2017).

Nesse sentido, a IA pode desempenhar um segundo papel – mais negativo, mas igualmente crucial – como mecanismo de identificação, contenção e correção dos fluxos desinformativos, capazes de pelo menos limitar os efeitos danosos à qualidade democrática. Modelos baseados em redes neurais e *machine learning* voltados à linguagem natural podem ser úteis na identificação de *fake news*, na remoção de contas robotizadas com o objetivo de disseminar informações falsas ou na

deteção de campanhas digitais criadas para ferir a deliberação democrática coletiva (Ciampaglia, 2018).

Tal uso pode ser considerado ilegítimo quando confrontado com o princípio da liberdade de expressão (pelo menos em seu sentido mais liberal). No entanto, é plausível, do ponto de vista do bem comum, e perfeitamente justificável restringir determinados conteúdos informacionais criados com o objetivo de desestabilizar processos democráticos, de modo a evitar suas consequências nocivas. Será importante, no entanto, garantir que essa limitação venha acompanhada de uma explicação e classificação – uma função metacognitiva que a IA também pode desempenhar – dos motivos da eliminação desses conteúdos, para que os cidadãos possam refletir, simultaneamente, sobre a qualidade epistêmica de suas próprias crenças políticas, criando-se, assim, verdadeiros cidadãos autônomos com autonomia prática, e não meramente formal (Fricker, 2007).

O caso das eleições presidenciais de 2018 e 2022 no Brasil tornou-se um dos maiores exemplos concretos desse tipo de prática desinformativa na contemporaneidade. Com uma estratégia em larga escala de desinformação, o candidato Jair Bolsonaro (e sua equipe) criou uma das infraestruturas técnicas mais influentes já utilizadas para fins políticos: da engenharia emocional a redes sociais falsas automatizadas, foram diversos os esquemas digitais destinados a influenciar o voto no candidato e, principalmente, o não-voto no candidato concorrente (Nemer *et al.*, 2025).

A plataforma *WhatsApp*, por exemplo, desempenhou um papel altamente influente na disseminação de informações falsas e no ataque a adversários políticos. Devido à sua natureza criptografada – que dificultava a identificação do emissor original – e à popularidade dessa aplicação no Brasil, tornou-se central na propagação de notícias maliciosas. Alguns estudos apontam que cerca de 40.000 grupos de *WhatsApp* estiveram ativos durante a campanha, circulando aproximadamente 28 milhões de mensagens por dia (Arnaudo, 2017). Segundo reportagem no *The Guardian*, agências independentes de informação analisaram as mensagens de campanha do candidato da direita em relação ao candidato da esquerda, identificando 3% de conteúdo enganoso ou falso no segundo, contra 42% no primeiro (Avelar, 2019).

Essa campanha de desinformação foi estruturada em três eixos centrais: (i) criação de contas falsas ou *botnets* que publicavam e partilhavam desinformação; (ii) produção de conteúdo desinformativo que viralizava em redes sociais de influenciadores digitais e grupos organizados; (iii) criação de campanhas de ódio personalizadas, que envolviam funcionários públicos, militantes e outros setores em ações de difamação (Amaral; Silveira, 2023). Essa influência foi tão clara e danosa que levou institucionalmente à exclusão de mais de 100.000 contas do *WhatsApp*, em colaboração inédita entre o Tribunal Superior Eleitoral (TSE) do Brasil e as empresas *WhatsApp* e *Google*.

Esse caso demonstra de que forma, quando somamos os níveis de ignorância racional à produção de conteúdos desregulados que exploram ao máximo os vieses cognitivos e polarizadores dos cidadãos, podemos corroer as bases epistêmicas que deveriam sustentar a democracia. Assim, o caso “Bolsonaro” constitui um exemplo prático de como, por um lado, as democracias são altamente frágeis e suscetíveis a processos de decisão enviesados e, por outro, de como a IA pode ser pensada como ferramenta de apoio à regulação e correção da informação circulante nos processos de democratização.

Considerações finais

Uma aplicação final que podemos nomear é pensar a Inteligência Artificial (IA) como uma forma de instigar a deliberação pública como mecanismo de criação de uma vontade política

intersubjetiva (Habermas, 1996). Muitos dos ambientes digitais no presente tendem a criar movimentos de câmaras de eco ou bolhas epistêmicas (Nguyen, 2020), onde os participantes apenas confirmam as suas crenças já assumidas e reforçam a sua tribalização emocional, o que os afasta da possibilidade de diálogo e espírito crítico.

No entanto, esse mecanismo pode também ser reorientado para alcançar uma reconfiguração da deliberação pública mais relevante normativamente: sistemas de IA podem ser criados para promover diálogos entre cidadãos que tenham visões opostas ou sugerir conteúdos que apoiem uma reflexão mais equilibrada sobre determinados temas políticos (algo conhecido como *epistemic nudging*; cf. Sunstein, 2016).

Ademais, várias iniciativas como *CrowdLaw*⁵, *Citizens' Assembly*⁶ ou *Deliberative Polling*⁷ já começaram a utilizar ferramentas digitais para apoiar a criação de espaços comunicacionais que permitem superar a logística dos modelos presenciais, ampliando, por um lado, a possibilidade de troca e diálogo e, por outro, o contato com diferentes visões, argumentos e opiniões. Tais plataformas podem incluir muitos cidadãos que se veem privados da pertença política nas democracias atuais, dado que grande parte do debate político ainda acontece em lugares e tempos limitados fisicamente (muitas vezes apenas nos grandes centros urbanos, onde existe maior concentração de riqueza, e longe das regiões suburbanas).

Desse modo, a IA pode ser utilizada para corrigir injustiças estruturais que se baseiam em privilégios epistêmicos de determinadas elites e no silenciamento de outros setores da sociedade: se for possível garantir maior participação na deliberação democrática, assegura-se que nem sempre as mesmas vozes (geralmente masculinas, brancas, urbanas e com elevado capital educacional) sejam novamente representadas, tomando as mesmas decisões com os mesmos enviesamentos ideológicos.

Muito mais que atuar apenas na sua qualidade “racional”, a IA pode também ajudar a aumentar os níveis de compreensão e empatia pelo Outro, transformando a comunicação em um elemento afetivo e relacional (por exemplo, ao construir pontes semânticas entre setores historicamente complexos ou ao reformular perspectivas de modo a apontar possíveis convergências entre posições diferentes).

No fundo, se bem pensada e desenvolvida, a IA pode produzir reais condições para alcançar uma democracia mais sólida, representativa e informada – algo que as democracias do século XXI não parecem estar a proporcionar, com todas as consequências trágicas que advêm desse afastamento do ideal democrático. Assim, há uma urgência ética de, pelo menos, considerar de forma exploratória essas alternativas, testando-as rigorosamente e, quiçá, introduzi-las em contextos de menor dimensão, para se compreender o impacto real nos processos de decisão democrática.⁸

⁵ Sobre *CrowdLaw*, disponível em: <https://crowd.law/?gi=fe4463677751>. Acesso em: 6 ago. 2025.

⁶ Ver *Citizens' Assembly* em: <https://citizensassembly.co.uk/>. Acesso em: 6 ago. 2025.

⁷ Consulte *Deliberative Polling* em: <https://www.involve.org.uk/resource/deliberative-polling>. Acesso em: 6 ago. 2025.

⁸ Neste artigo, utilizamos a ferramenta *ChatGPT* como apoio à elaboração de exemplos técnicos e levantamento de argumentos. Todas as formulações e interpretações são de responsabilidade dos autores.

Referências

- AMARAL, A.; SILVEIRA, F. Bolsonarismo e o fascismo na era digital. **Revista Brasileira de Estudos Políticos**, Belo Horizonte, n. 127, p. 55-94, 2023. DOI: <https://doi.org/10.9732/2023.V127.962>
- ARNAUDO, D. **Computational propaganda in Brazil**: social bots during elections. Computational Propaganda Research Project. Working Paper No. 2017.8. 2017. Oxford: Project on Computational Propaganda, 2017. Disponível em: <https://demtech.ox.ac.uk/wp-content/uploads/sites/12/2017/06/Comprop-Brazil-1.pdf>. Acesso em: 18 ago. 2025.
- AVELAR, D. WhatsApp fake news helped Jair Bolsonaro win election, analysis suggests. **The Guardian**, [s. l.], 30 out. 2019. Disponível em: <https://www.theguardian.com/world/2019/oct/30/whatsapp-fake-news-brazil-election-favoured-jair-bolsonaro-analysis-suggests>. Acesso em: 6 ago. 2025.
- BRENNAN, J. **Against democracy**. Princeton: Princeton University Press, 2016.
- CIAMPAGLIA, G. L. Fighting fake news: A role for computational social science in the fight against digital misinformation. **Journal of Social Computing**, [s. l.], v. 1, n. 2, p. 147-165, 2018. DOI: <https://doi.org/10.1007/s42001-017-0005-6>
- CRENSHAW, K. Demarginalizing the intersection of race and sex: a black feminist critique of antidiscrimination doctrine, feminist theory and antiracist politics. **University of Chicago Legal Forum**, Chicago, v. 1989, n. 1, p. 139-167, 1989. Disponível em: <http://chicagounbound.uchicago.edu/uclf/vol1989/iss1/8>. Acesso em: 19 ago. 2025.
- DOWNS, A. **An economic theory of democracy**. New York: Harper & Row, 1957.
- ESTLUND, D. **Democratic authority**: a philosophical framework. Princeton: Princeton University Press, 2008.
- FLORIDI, L.; COWLS, J.; BELTRAMETTI, M.; CHATILA, R.; CHAZERAND, P.; DIGNUM, V.; LUETGE, C.; MADELIN, R.; PAGALLO, U.; ROSSI, F.; SCHAFER, B.; VALCKE, P.; VAYENA, E. AI4People—An Ethical Framework for a Good AI Society: Opportunities, risks, principles, and recommendations. **Minds and Machines**, [s. l.], v. 28, n. 4, p. 689-707, 2018. DOI: <https://doi.org/10.1007/s11023-018-9482-5>
- FRICKER, M. **Epistemic injustice**: power and the ethics of knowing. Oxford: Oxford University Press, 2007.
- GIUSTI, S.; PIRAS, E. (ed.). **Democracy and fake news**: information manipulation and post-truth politics. New York: Routledge, 2020.
- GOUVEIA, S. S. **Homo Ignarus**: ética racional para um mundo irracional. Coimbra: Minerva, 2020.
- HABERMAS, J. **Between facts and norms**: contributions to a discourse theory of law and democracy. Cambridge: MIT Press, 1996.
- IRENI-SABAN, L.; SHERMAN, M. **Ethical governance of artificial intelligence in the public sector**. London: Routledge, 2021.

LANGSTON, J. Who's a CEO? Google image results can shift gender biases. **UW News**, Seattle, 9 abr. 2015. Disponível em: <https://www.washington.edu/news/2015/04/09/whos-a-ceo-google-image-results-can-shift-gender-biases/>. Acesso em: 19 ago. 2025.

NEMER, D.; MARKS, W.; SANFILIPPO, M. R.; OCEPEK, M. G. The human infrastructure of misinformation: a case study of Brazil's heteromated labor. In: FRISCHMANN, B.; MADISON, M.; STRANDBURG, K. J. (org.). **Governing misinformation in everyday knowledge commons**. Cambridge: Cambridge University Press, 2025. p. 148-166.

NGUYEN, C. T. Echo chambers and epistemic bubbles. **Episteme**, [s. l.], v. 17, n. 2, p. 141-161, 2020. DOI: <https://doi.org/10.1017/epi.2018.32>

PLATÃO. **A República**. Tradução: M. J. V. de Carvalho. Lisboa: Fundação Calouste Gulbenkian, 1993.

SUNSTEIN, C. R. **The ethics of influence**: government in the age of behavioral science. Cambridge: Cambridge University Press, 2016.

TREMEWAN, L. Gender bias in AI: Is there a problem with representation? **Finder**, Londres, 23 abr. 2024. Disponível em: <https://www.finder.com/uk/stats-facts/gender-bias-in-ai>. Acesso em: 19 ago. 2025.

WARDLE, C.; DERAKHSHAN, H. **Information disorder**: toward an interdisciplinary framework for research and policy making. Strasbourg: Council of Europe, 2017.

Recebido em 01/07/2025

Aceito em 11/08/2025

Publicado online em 20/08/2025