

DOSSIÊ

doi: 10.4013/fsu.2026.272.12

A vulnerabilidade informacional sob a racionalidade algorítmica: dimensões técnica, subjetiva e epistêmico-coletiva

Informational vulnerability under algorithmic rationality: technical, subjective, and collective-epistemic dimensions

Keberson Bresolin 

Universidade Federal de Pelotas – UFPEL – Programa de Pós-Graduação em Filosofia – Pelotas, RS, Brasil. E-mail: keberson.bresolin@gmail.com

RESUMO: Este artigo investiga a vulnerabilidade informacional como uma condição estrutural produzida e intensificada por ambientes organizados segundo a racionalidade algorítmica. Mediante uma investigação teórico-conceitual e interdisciplinar, articulam-se estudos sobre sistemas de recomendação e filtragem, a Teoria da Autodeterminação, as análises do poder desenvolvidas por Foucault e Han e a discussão sobre esfera pública e câmaras de eco. Sustenta-se que a vulnerabilidade informacional se forma pela interação entre três dimensões, compreendidas como a dependência técnica de sistemas opacos de seleção e predição, a reorganização subjetiva da autonomia, da competência e da relacionalidade e a fragmentação das condições epistêmicas da deliberação pública. Essas dimensões formam um circuito no qual a modulação técnica afeta a formação da agência e repercute sobre as condições coletivas de produção e validação do conhecimento. A análise oferece critérios para avaliar quando estruturas digitais ampliam relações de dependência, dificultam a contestação e comprometem a pluralidade informacional, mantendo em aberto a formulação de uma teoria normativa completa.

Palavras-chaves: vulnerabilidade informacional; racionalidade algorítmica; subjetividade digital; câmaras de eco; esfera pública.

ABSTRACT: This article investigates informational vulnerability as a structural condition produced and intensified by environments organized according to algorithmic rationality. Through a theoretical-conceptual and interdisciplinary inquiry, it brings together studies on recommendation and filtering systems, Self-Determination Theory, the analyses of power developed by Foucault and Han, and debates on the public sphere and echo chambers. It argues that informational vulnerability arises from the interaction of three

Editores responsáveis:

Inácio Helfer

Leonardo Marques Kussler

Luís Miguel Rechiki Meirelles

Evandro Barbosa

Denis Coitinho

dimensions, understood as technical dependence on opaque systems of selection and prediction, the subjective reorganization of autonomy, competence, and relatedness, and the fragmentation of the epistemic conditions of public deliberation. These dimensions form a circuit in which technical modulation affects the development of agency and reshapes the collective conditions under which knowledge is produced and validated. The analysis provides criteria for assessing when digital structures deepen relations of dependence, hinder contestation, and undermine informational pluralism, while leaving the formulation of a complete normative theory open.

Keywords: informational vulnerability; algorithmic rationality; digital subjectivity; echo chambers; public sphere.

1 Considerações iniciais

A ampliação dos fluxos de dados e das conexões digitais exige examinar como os sistemas tecnológicos interferem na formação da subjetividade, no exercício da agência e nas formas de convivência. O que está em questão não é simplesmente o acesso à informação, mas a constituição adequada da subjetividade, agência e vida comunitária.

Este trabalho parte da hipótese de que a vulnerabilidade informacional atual resulta do entrelaçamento entre a dimensão técnica, construída por sistemas algorítmicos que regulam o acesso à informação, antecipam decisões e modelam comportamentos por meio de previsão e feedback, a dimensão subjetiva, associada à frustração persistente das necessidades de autonomia, competência e relacionalidade examinadas pela Teoria da Autodeterminação, e a dimensão epistêmico-coletiva, expressa na fragmentação do espaço racional, na formação de câmaras de eco e no fechamento cognitivo que ameaça a deliberação democrática e a intercompreensão social.

Contudo, Fineman defende que o conceito de vulnerabilidade não designa uma condição unívoca porque compreende a vulnerabilidade como um aspecto universal, inevitável e duradouro da condição humana, ligado à corporeidade, à dependência e à exposição aos outros e às instituições (2008, p. 8-9). Neste aspecto, Fineman retira a vulnerabilidade do campo reservado a grupos excepcionais e contesta a imagem do sujeito autônomo e independente como condição humana ordinária. Mas sua abrangência também impõe um limite, a saber, afirmar que todos são vulneráveis ainda não explica por que certas relações sociais e institucionais expõem alguns sujeitos mais intensamente a formas específicas de dependência, manipulação ou perda de controle.

Mackenzie, Rogers e Dodds procuram responder a esse problema distinguindo as fontes inerentes, situacionais e patogênicas da vulnerabilidade. As vulnerabilidades inerentes decorrem da corporeidade, da dependência e da constituição afetiva e social dos seres humanos. As situacionais são produzidas ou agravadas por circunstâncias pessoais, sociais, políticas ou institucionais. As patogênicas constituem uma modalidade especialmente problemática de vulnerabilidade situacional, pois surgem de relações disfuncionais, estruturas injustas ou intervenções que podem apresentar-se como proteção ou assistência e terminar por ampliar a dependência e reduzir a autonomia. Os autores distinguem ainda a vulnerabilidade disposicional da ocorrente, separando a suscetibilidade estruturada para determinado dano de sua manifestação concreta (Mackenzie, Rogers e Dodds, 2014, p. 7–9). Isso permite reconhecer uma condição de vulnerabilidade antes que seja possível identificar um dano consumado.

Considerando isso, chama-se aqui vulnerabilidade informacional a condição situacional e estrutural na qual a formação do juízo, da vontade e do reconhecimento passa a depender de sistemas assimétricos e pouco transparentes de seleção, classificação e distribuição da informação. Ela possui caráter disposicional quando a estrutura digital estabelece uma suscetibilidade persistente à influência e à perda de controle, mesmo que seus efeitos não se realizem da mesma maneira em todos os usuários. Pode adquirir um caráter patogênico quando sistemas destinados a organizar o acesso à informação e facilitar a interação social produzem novas dependências ou intensificam vulnerabilidades já existentes. Isso não significa que o usuário deixe de agir ou que toda mediação algorítmica produza necessariamente um dano. A vulnerabilidade reside na desproporção entre a capacidade das plataformas de conhecer, prever e modular condutas e as possibilidades dos sujeitos de compreender, contestar ou revisar essas mediações. A racionalidade algorítmica não cria a vulnerabilidade humana em sentido absoluto, mas ela a reorganiza tecnicamente e a distribui de modo desigual nas dimensões subjetiva, técnica e epistêmico-coletiva.

Metodologicamente, a investigação é teórico-conceitual e interdisciplinar, articulando modelos computacionais, estudos psicológicos e expoentes filosóficos para identificar mecanismos e tendências cujo alcance precisa ser interpretado filosoficamente, já que os resultados mobilizados não demonstram uma determinação uniforme dos usuários pelas plataformas. O percurso argumentativo parte da reconstrução das formas técnicas de seleção, predição e retroalimentação, examina como essas mediações podem afetar as condições motivacionais em que a agência se forma e se exerce e chega às suas consequências para a produção coletiva do conhecimento e para a deliberação pública. A aproximação entre esses referenciais permite compreender como diferentes níveis de mediação participam da produção da vulnerabilidade informacional sem reduzir questões filosóficas a explicações computacionais ou psicológicas.

A primeira seção deste artigo examina a dimensão técnica da vulnerabilidade, discutindo como os algoritmos de recomendação, sistemas de aprendizado de máquina e recursos de personalização transformam os espaços digitais em mecanismos de controle, previsibilidade e modelagem comportamental. A segunda seção examina seus efeitos sobre a formação da vontade e da identidade. Por fim, a terceira seção ilustra como essas vulnerabilidades se combinam em mecanismos coletivos de erosão da esfera pública que causam polarização, insularidade epistêmica e a fragmentação da cultura comum.

Mais do que identificar riscos associados à captura informacional, o artigo procura esclarecer como essas mediações alteram as condições da liberdade, da convivência e do pensamento. O objetivo normativo é delimitado, pois o artigo não formula uma teoria completa de regulação das plataformas nem um modelo institucional de esfera pública digital. O diagnóstico oferece três critérios de avaliação, a saber, a possibilidade de compreender e contestar as operações das plataformas, a preservação das condições subjetivas da agência e a manutenção de um ambiente informacional aberto ao dissenso.

2 Vulnerabilidade técnica e controle algorítmico

A dimensão técnica da vulnerabilidade informacional diz respeito às formas pelas quais sistemas algorítmicos selecionam informações, antecipam comportamentos e organizam as condições de interação.

O ambiente digital contemporâneo deixou de ser um mero meio de circulação informacional para converter-se em uma estrutura de predição e controle, estruturada por dispositivos algorítmicos que antecipam, monitoram e retroalimentam condutas. A centralidade dos dados, combinada com personalização extrema e filtragem comportamental, constitui o núcleo da vulnerabilidade técnica, a saber, uma exposição estrutural aos imperativos da automação decisória.

Essa vulnerabilidade funda-se em modelos preditivos de aprendizado de máquina que, ao observar comportamentos passados dos usuários, constroem teorias sobre seus desejos futuros, moldando seus percursos informacionais e afetivos. Pariser analisa a capacidade preditiva desses filtros ao afirmar que a nova “geração de filtros da internet observa as coisas que você parece gostar [...] e tenta extrapolar. Eles são motores de predição, constantemente criando e refinando uma teoria sobre quem você é e o que fará e desejará a seguir” (2011, p. 10). Forma-se, portanto, de um regime de antecipação subjetiva, no qual os próprios dados dos usuários servem como matéria-prima para sua modelagem futura.

Pariser descreve a adesão involuntária do sujeito ao afirmar que “você não escolhe entrar na bolha” (2011, p. 11). A bolha também funciona como mecanismo de reforço comportamental, pois “cada clique que você dá nesse ciclo é mais uma ação de autojustificação” (Pariser, 2011, p. 71). A vulnerabilidade aparece aqui na dificuldade de escapar aos padrões que o próprio comportamento alimenta, sobretudo quando o usuário dispõe de poucos recursos para compreender ou alterar a lógica de filtragem.

Do ponto de vista computacional, essa lógica é descrita por modelos de aprendizado por reforço e pela Teoria do Feedback Social (TFS). Essa teoria postula que as pessoas modificarão suas atividades para obter recompensas simbólicas recebidas em espaços digitais, como curtidas, compartilhamentos e comentários positivos. Assim, o comportamento é moldado menos por convicções pessoais profundas e mais pela expectativa de aprovação social. Banisch *et al.* apontam que na TFS o comportamento dos usuários é “moldado pelas recompensas sociais ou pelo feedback que recebem de outros, tornando-se um arcabouço ideal para entender como as plataformas influenciam a expressão de opiniões e as escolhas” (2024, p. 2).

Esse modelo incorpora o parâmetro μ , que representa a tensão entre a busca por aprovação social e a preservação da diversidade de opiniões. Quanto menor esse valor, maior a tendência dos usuários a ajustarem seus discursos à norma percebida do grupo, renunciando à pluralidade. Banisch et al. explicam que “o parâmetro μ modela a troca entre a busca por aprovação social e a diversidade de opinião” (2024, p. 4). Quando μ assume valores baixos, o sistema converge para uma homogeneização opinativa, e mesmo diante de exposições contrárias, os sujeitos tendem a permanecer encapsulados em bolhas informacionais. Eles concluem que, “mesmo nesse cenário, equilíbrios subótimos de segregação online permanecem um resultado estável do modelo (para $\mu < 0,79$)” (Banisch et al., 2024, p. 6).

O alcance desse resultado, contudo, precisa ser delimitado, pois se trata de um modelo baseado em agentes, construído a partir de parâmetros e comportamentos formalizados e não de uma descrição imediata da conduta de todos os usuários ou do funcionamento de todas as plataformas. Embora o modelo mostre que a segregação pode emergir e permanecer estável mesmo quando os agentes manifestam alguma preferência pela diversidade, ele não autoriza concluir que esse resultado seja inevitável nos ambientes digitais ou que os sujeitos estejam inteiramente submetidos às determinações algorítmicas. A interpretação proposta aqui parte justamente dessa diferença, uma vez que o aspecto relevante da simulação não está na demonstração de uma determinação absoluta, mas na possibilidade de que certas estruturas transformem preferências e escolhas individuais que não pretendiam produzir segregação em resultados coletivos que nenhum dos participantes escolheu deliberadamente.

Ora, as dinâmicas de aprovação social digitalizadas transformam a interação em um sistema de reforço adaptativo, no qual a expectativa de aceitação e reconhecimento passa a exercer um peso significativo sobre o comportamento do indivíduo dentro de microcomunidades informacionais. Cada ação, como um clique, um compartilhamento ou um comentário, é interpretada e recompensada segundo padrões algorítmicos que favorecem a repetição de conteúdos homogêneos, gerando um ciclo de retroalimentação da conformidade. A exposição à diferença, nesse contexto, pode envolver custos sociais e cognitivos, pois aumenta o risco de isolamento, invisibilidade ou penalização simbólica. A vulnerabilidade técnica decorre da própria estrutura do ambiente, que eleva os custos sociais e cognitivos da dissidência e favorece condutas previsíveis, mesmo quando os usuários preservam alguma capacidade de contestação.

A responsabilidade por esses efeitos não pode ser atribuída exclusivamente aos sujeitos individuais, pois o design das plataformas e a moderação algorítmica exercem papel decisivo na formação desses resultados (Banisch et al., 2024, p. 11). A vulnerabilidade técnica é, nesse sentido, uma consequência estrutural dos ambientes digitais contemporâneos, e não uma falha moral dos usuários.

Essa forma de sujeição algorítmica opera pela organização prévia dos percursos informacionais, que preserva possibilidades de escolha ao mesmo tempo que favorece trajetórias compatíveis com os padrões identificados pelas plataformas. O sujeito continua decidindo sobre os conteúdos que consome, embora o campo disponível para essas decisões seja selecionado por sistemas de personalização e rastreamento comportamental. Como observa Sunstein, “você não precisa criar um ‘*Daily Me*’. Outros estão criando por você” (2018, p. 4).

Essa arquitetura de controle se impõe como um novo modo de liberdade aparente: “todos poderiam desfrutar de uma arquitetura de controle. Cada um de nós estaria totalmente encarregado do que vê e ouve” (Sunstein, 2018, p. 1). No entanto, essa liberdade conduz menos à autonomia e mais à reafirmação de crenças previamente estabelecidas, tornando os sujeitos menos propensos à dúvida, à escuta e à abertura ao outro. Em vez de ampliar horizontes, essa liberdade fabricada estreita o campo de visão. Sunstein considera a hipótese de que pessoas com determinada convicção política passem a conhecer cada vez mais autores que compartilham a mesma visão (2018, p. 34). A exposição reiterada a conteúdos que reforçam predisposições anteriores favorece a homogeneização informacional, dificulta a emergência do dissenso e transforma o ambiente digital em um espaço de confirmação contínua. A consequência direta dessa homogeneização da experiência cognitiva é o empobrecimento do espaço público, que se fragmenta em nichos ideológicos isolados, onde não há diálogo entre perspectivas distintas, mas apenas a reverberação autorreferente de certezas consolidadas.

O resultado é a erosão da esfera pública e a dissolução de referenciais comuns de realidade e de legitimidade discursiva: “a proliferação de nichos prejudicará aspectos da cultura compartilhada e promoverá a fragmentação, especialmente ao longo de linhas políticas” (Sunstein, 2018, p. 151). A vulnerabilidade técnica aparece, nesse caso, na dificuldade de encontrar conteúdos capazes de interromper padrões já consolidados de atenção e julgamento. Quando a filtragem reduz o contato com perspectivas divergentes, a experiência informacional torna-se mais previsível e oferece menos ocasiões para revisão crítica.

Os dados passam a operar como recursos de classificação e previsão que permitem às plataformas inferir preferências e intervir sobre os percursos informacionais dos usuários. Quando Han afirma que os *big data* podem tornar legíveis desejos dos quais o próprio sujeito não está consciente, sua crítica aponta para a assimetria entre a capacidade técnica de inferência das plataformas e o conhecimento que os usuários possuem sobre essas operações (2022, p. 88).

Essa transformação insere a própria estrutura do eu no campo do cálculo e da previsão algorítmica. Han afirma que “a própria pessoa se positiviza em coisa [...] os big data anunciam o fim da pessoa e do livre-arbítrio” (2018b, p. 23). A formulação de Han é radical e, por isso, não deve ser incorporada como uma descrição literal do desaparecimento da agência, pois a transformação da pessoa em um perfil calculável significa que ela passa a ser reconhecida pelas plataformas a partir daquilo que pode ser registrado, classificado e previsto, mas não que sua subjetividade se esgote nos dados ou que sua capacidade de agir tenha sido inteiramente eliminada. Essa crítica ganha maior consistência quando se dirige às condições em que a agência se forma e se exerce, sobretudo à opacidade dos processos de inferência, à assimetria entre plataformas e usuários e à possibilidade de intervir sobre a escolha antes mesmo que ela alcance um nível plenamente reflexivo. O “curtir” pode funcionar como mecanismo de adesão e de produção de dados, embora não possua sempre o mesmo significado nem constitua isoladamente uma forma completa de submissão, uma vez que seu poder não decorre do gesto individual, mas da estrutura que reúne, interpreta e recompensa continuamente essas interações.

Neste regime, a comunicação deixa de ser espaço de compartilhamento intersubjetivo para tornar-se uma operação logística, uma transmissão de sinais codificáveis, calculáveis e governáveis. Han advoga que “as ações se tornam transparentes quando se transformam em operacionais, subordinadas a um processo passível de cálculo, governo e controle” (2017, p. 11). A transparência, nesse caso, não é clareza ética, mas vigilância operativa. A hipercomunicação, longe de promover a pluralidade, enrijece o fluxo do sentido e neutraliza a alteridade: “o que assegura a transparência não é o isolamento, mas a hipercomunicação” (Han, 2017, p. 60).

Essa hipercomunicação reforça uma lógica de vigilância distribuída na qual a exposição voluntária dos usuários amplia a disponibilidade de dados para classificação e previsão. A vigilância contemporânea não é mais imposta de fora, mas internalizada como necessidade subjetiva. “A entrega dos dados não acontece por coação, mas a partir de uma necessidade interna” (Han, 2018a, p. 16). O sujeito, ansioso por visibilidade e pertencimento, voluntariamente se inscreve na lógica da exposição, desconhecendo que essa entrega contínua configura sua própria captura. Assim, a vulnerabilidade técnica passa a integrar as condições estruturais nas quais o sujeito digital forma preferências, busca reconhecimento e exerce sua agência. Sua subjetividade não é esvaziada ou substituída pelos dados, mas torna-se crescentemente dependente de sistemas que a representam como superfície mensurável, responsiva e governável.

Esse diagnóstico é aprofundado pelo conceito foucaultiano de disciplina, entendido como um conjunto de técnicas minuciosas de controle e adestramento que moldam os corpos para que se tornem úteis e dóceis dentro de determinadas racionalidades de poder. A era moderna, segundo Foucault, “descobriu o corpo como objeto e alvo do poder [...] um corpo dócil, que pode ser submetido, usado, transformado e aperfeiçoado” (1995, p. 136). Essa docilidade é produzida por meio de uma “anatomia política do detalhe” (Foucault, 1995, p. 139), isto é, um regime que opera na escala do micro, nos gestos, nos ritmos, nas posições e nos hábitos, a fim de extrair o máximo rendimento com o mínimo de desordem.

No entanto, a mudança digital do dispositivo disciplinar não extingue sua lógica, mas a reconverte, ou seja, as técnicas que antes visavam o corpo físico agora incidem sobre o sujeito informacional. Não é mais o posicionamento do corpo no espaço, mas a distribuição do comportamento no tempo e a previsibilidade da ação. Como tal, a disciplina contemporânea visa a modelagem dos dados e das condutas, criando subjetividades previsíveis, automatizáveis e governáveis por sistemas algorítmicos.

Há de se dizer, contudo, que essa aproximação não implica identidade exata entre a disciplina descrita por Foucault e o funcionamento das plataformas. A disciplina clássica opera por instituições relativamente delimitadas, distribuições espaciais e mecanismos reconhecíveis de exame. Já o poder algorítmico é privado, probabilístico, distribuído e dependente da participação contínua dos usuários. A referência foucaultiana é utilizada, portanto, para compreender a produção de normas e a classificação diferencial das condutas, não para afirmar que o ambiente digital seria apenas um panóptico transferido para as telas.

Ao identificar padrões de comportamento, os algoritmos participam da produção de normas ao distribuir visibilidade, legibilidade e relevância, favorecendo condutas compatíveis com seus critérios e reduzindo o alcance daquelas que deles se afastam. A disciplina digital, portanto, não precisa mais impor-se através da força. Sua eficácia reside justamente na invisibilidade de sua operação e na voluntariedade de sua reprodução. O disciplinamento passa a ser reproduzido pelo próprio sujeito, que internaliza as métricas de desempenho, engajamento e aprovação fornecidas pelas plataformas e as utiliza como critérios de autoavaliação.

A subjetividade, assim, torna-se o novo corpo dócil. O controle alcança os comportamentos observáveis e também as condições em que o indivíduo pensa, sente, interage e decide. Nas redes sociais, a disciplina descrita por Foucault assume uma dimensão cognitiva, pois a tecnologia passa a incidir sobre os modos de percepção e sobre os mecanismos envolvidos na decisão subjetiva. O que antes era uma arte de distribuir corpos no espaço, agora é uma ciência de modular condutas no tempo informacional. O espaço da disciplina migrou do visível ao invisível, do físico ao psicológico, do corpo ao dado.

A vulnerabilidade técnica, nesse contexto, consiste na exposição contínua a dispositivos que operam sob uma lógica de modulação comportamental silenciosa, progressiva e personalizável, baseada em inferências probabilísticas que organizam o fluxo da atenção, delimitam o horizonte de escolhas e interferem nos próprios processos de formação da vontade. Agora está nos cliques, nas pausas diante da tela, nos padrões de digitação, nos tempos de permanência e nas interações mínimas, convertidas em metadados e transformadas em *índices de valor*, pertinência e normalidade.

O poder disciplinar se atualiza, assim, nas engrenagens do capitalismo de dados, substituindo o adestramento do corpo pela captura dos processos mentais, substituindo a vigilância visível pela normatização invisível da conduta. As plataformas, ao recompensarem certas interações e punirem silenciosamente outras, instauram um regime de seletividade que forma preferências, filtra perspectivas e esculpe a subjetividade por meio da repetição e da retroalimentação. Nesse sentido, o sujeito contemporâneo já não precisa ser vigiado: ele se inscreve voluntariamente num sistema de disciplina algorítmica contínua, que, ao capturar o dado, captura também o desejo e a racionalidade.

Han relaciona vigilância e mercado ao afirmar que “*Big Brother e big deal* se aliam [...] as pessoas são tratadas e comercializadas como pacotes de dados” (2022, p. 90). Essa afirmação sintetiza a fusão entre vigilância e mercado, entre controle e monetização, que caracteriza a forma atual de sujeição digital. A observação dos sujeitos integra um processo mais amplo de conversão de suas ações em unidades de valor econômico inscritas em fluxos de dados, estatísticas e perfis de consumo. Trata-se de um novo regime de subjetivação, que Han descreve como uma reorganização da alma humana: “a sociedade positiva está em vias de reorganizar a alma humana [...] no curso e empuxo de sua positivação” (Han, 2017, p. 15).

Ao tornar experiências e condutas mensuráveis, o processo de positivação altera a maneira como os sujeitos se relacionam com o mundo e interfere nos critérios pelos quais verdade, identidade e valor são reconhecidos. A vulnerabilidade técnica constitui uma manifestação estrutural de uma racionalidade que torna os sujeitos cada vez mais legíveis a sistemas de classificação, previsão e intervenção. A disciplina já não se impõe por meio de grades ou muros, mas por meio de interfaces, notificações e algoritmos que extraem valor da subjetividade exposta.

O poder algorítmico amplia a capacidade de intervir sobre desejos, afetos e modos de conhecer ao transformar ações cotidianas em dados continuamente interpretáveis. A vulnerabilidade técnica decorre dessa exposição prolongada a sistemas cuja operação permanece pouco compreensível e cuja influência dificilmente pode ser contestada pelo usuário individual.

Essa racionalidade algorítmica regula o acesso à informação e alcança os mecanismos de motivação, desejo e identidade, participando das condições em que a agência se forma e se exerce.

3 Vulnerabilidade subjetiva no ambiente algorítmico

A estrutura algorítmica, descrita anteriormente como estrutura de controle e predição, exerce efeitos profundos sobre os estados motivacionais e relacionais dos sujeitos. Se, na seção anterior, o foco recaiu sobre os dispositivos técnicos que organizam o comportamento e estruturam as plataformas digitais como ambientes de controle e retroalimentação, a presente seção busca aprofundar as consequências subjetivas dessa configuração. Analisar-se-á, aqui, como essa base técnica engendra uma forma de vulnerabilidade subjetiva, centrada na frustração das necessidades psicológicas básicas e na erosão da autonomia, da competência e da relacionalidade. Parte-se aqui da hipótese de que a racionalidade algorítmica, ao capturar e modelar dados, interfere nas condições em que o desejo, a identidade e a agência se formam, contribuindo para uma reconfiguração da experiência subjetiva.

Na psicologia contemporânea, a Teoria da Autodeterminação (*Self-Determination Theory* – SDT), desenvolvida por Edward Deci e Richard Ryan, fornece um referencial para apreender a composição da vulnerabilidade disposicional subjetiva. A SDT postula que os seres humanos têm uma propensão inerente à integração psíquica, ao crescimento e à autorregulação, desde que sejam atendidas as necessidades psicológicas de autonomia, competência e relacionalidade, entendidas respectivamente como a experiência de agir como autor das próprias ações, de sentir-se eficaz e de estabelecer vínculos com os outros (Ryan; Deci, 2020, p. 2). Essas necessidades são consideradas universais e fundamentais, e a qualidade das interações sociais exerce papel central em sua satisfação ou frustração.

Quando tais necessidades são sistematicamente frustradas, como ocorre em contextos de vigilância excessiva, exigência de desempenho constante ou relações interpessoais instáveis, os indivíduos tendem a desenvolver estados de desorganização interna, alienação motivacional e sofrimento psicológico crônico. A vulnerabilidade, nesse sentido, decorre de ambientes que minam as condições fundamentais da autorregulação e do bem-estar, em vez de revelar uma insuficiência pessoal. Pesquisas associam a frustração dessas necessidades a sintomas de ansiedade, esgotamento, baixa autoestima e desregulação emocional (Vansteenkiste; Ryan, 2013, p. 264; Ryan; Deci, 2020).

A Teoria da Autodeterminação não foi formulada como uma teoria das plataformas digitais e os estudos citados sobre frustração das necessidades psicológicas não demonstram, por si mesmos, que sistemas de recomendação produzam necessariamente ansiedade, baixa autoestima ou desregulação emocional. A aplicação ao ambiente digital constitui uma inferência teórica deste artigo. Ela se justifica na medida em que a SDT permite examinar se determinadas características do design favorecem ou dificultam experiências de autonomia, competência e relacionalidade. A teoria fornece, assim, critérios de análise e não uma prova causal geral sobre os efeitos das plataformas.

Por sua vez, ambientes digitais podem, em determinadas circunstâncias, ampliar a autonomia, facilitar aprendizagens e sustentar vínculos significativos. O problema surge quando a captura da atenção reduz o controle do sujeito sobre seu tempo, quando métricas públicas convertem a competência em comparação incessante e quando o reconhecimento passa a depender de recompensas intermitentes e quantificadas. Não é a mediação digital em si que frustra as necessidades psicológicas, mas uma forma de organização voltada prioritariamente à retenção, ao desempenho e à previsibilidade comportamental. A vulnerabilidade subjetiva designa essa dependência das condições externas de validação e não uma fragilidade psicológica inerente ao usuário.

A vulnerabilidade adquire uma dimensão infraestrutural porque também emana das lógicas sistêmicas que organizam os ambientes de interação e subjetivação. Assim, o próprio design das plataformas digitais, em sua busca por engajamento, retenção e eficiência comportamental, pode configurar um novo regime de produção da vulnerabilidade, estruturando uma subjetividade constantemente exposta, modulada e fragilizada em seus alicerces motivacionais.

Essa transformação amplia o conceito de vulnerabilidade informacional para além da exposição de dados pessoais e o situa no plano estrutural em que a construção do conhecimento, da percepção e da vontade depende de mediações algorítmicas. A vulnerabilidade informacional reduz as condições em que o sujeito discrimina informações, forma juízos e delibera criticamente diante de fluxos mediados por algoritmos, dificultando o acesso ao dissenso e a conteúdos inesperados.

Para que isso fique claro, é necessário também refinar o conceito de “informação” que aqui está em jogo. No ambiente digital, a informação ultrapassa a condição de dado factual ou mensagem transmitida porque sua circulação passa a ser modelada por critérios de relevância algorítmica. Sua mensuração, hierarquização e distribuição frequentemente privilegiam o engajamento, a retenção e a previsibilidade em detrimento de critérios epistemológicos ou de veracidade. Nesse contexto, conteúdos compatíveis com o perfil preditivo do sujeito tendem a ganhar visibilidade, enquanto os demais podem perder alcance. Nesse sentido, o próprio acesso à pluralidade informacional é comprometido.

Como advoga Pariser, a personalização excessiva leva à formação de bolhas informacionais que limitam a diversidade de perspectivas e restringem a autonomia do pensamento. O sujeito é conduzido por trajetórias previsíveis, delimitadas por suas escolhas anteriores, e alimentado por um circuito fechado de confirmação de crenças e gostos (Pariser, 2011, p. 14). Nessas condições, a liberdade de escolha pode ser estreitada por mecanismos de retroalimentação que reforçam conteúdos familiares e reduzem a presença do divergente. A vulnerabilidade informacional consiste na diminuição das condições de autonomia epistêmica, o que dificulta a formação de juízos e decisões diante de perspectivas plurais.

A consequência disso é a modelagem algorítmica da subjetividade, desdobramento direto da vulnerabilidade informacional. A informação, nesse contexto, torna-se um vetor de controle, de modo que aquilo que corresponde ao perfil preditivo do sujeito ganha visibilidade, enquanto o que se afasta desse perfil tende a ser invisibilizado. Pariser sustenta que os meios digitais transformam a experiência informacional ao ajustar suas ofertas por meio da modulação pouco perceptível das preferências, dos desejos e das identidades dos usuários (2011, p. 63).

A exposição contínua a fluxos personalizados favorece respostas previsíveis e reforça padrões já estabelecidos, reduzindo as ocasiões em que conteúdos inesperados podem interferir na experiência do usuário. Essa lógica reduz o controle que o sujeito exerce sobre suas escolhas e permite que perfis estatísticos interfiram na formação de sua vontade, condicionando a agência a padrões que foram constituídos a partir de seus comportamentos anteriores. Perfis estatísticos passam, assim, a participar da formação da vontade ao condicionar o campo em que as escolhas são realizadas, embora não determinem integralmente a resposta do sujeito.

Esse regime utiliza mecanismos de recompensa e punição simbólica baseados na aprovação social e na distribuição desigual de visibilidade. Estudos mostram que os indivíduos ajustam suas condutas com base no retorno que recebem das interações e que a tendência à repetição aumenta conforme cresce o reconhecimento recebido (Banisch, Gaisbauer e Olbrich, 2022, p. 2–3). Essa lógica produz um ciclo de retroalimentação comportamental no qual o sujeito, em busca de aprovação e visibilidade, alinha sua expressão a padrões previamente validados. O comportamento torna-se, assim, responsivo a estímulos ambientais que reforçam a conformidade e silenciam a dissidência.

Essa dinâmica expressa uma forma atualizada do que Foucault descreveu como o “olhar normalizador (*normalizing gaze*)”. Para ele, “o exame combina as técnicas de uma hierarquia observadora e as de um julgamento normalizador. É um olhar normalizador, uma vigilância que torna possível qualificar, classificar e punir” (Foucault, 1995, p. 184). No contexto digital, esse exame não é mais realizado por instituições disciplinares tradicionais, mas por sistemas algorítmicos distribuídos que monitoram constantemente o comportamento e devolvem feedback imediato sob a forma de likes, métricas de engajamento e ranqueamentos. Trata-se de uma vigilância contínua que captura, avalia e regula a conduta por meio de normas estatísticas que definem o que é relevante, aceitável ou invisível.

O medo da desaprovação ou do isolamento, potencializado pela lógica da exposição permanente, leva muitos sujeitos à autocensura e à neutralidade estratégica, como indicam os modelos baseados no conceito de *Spiral of Silence* (Gaisbauer et al., 2020, p. 1). A partir da teoria formulada por Elisabeth Noelle-Neumann, eles sustentam que indivíduos tendem a silenciar suas opiniões quando percebem que elas são minoritárias ou socialmente rejeitadas, por receio de marginalização. No ambiente digital, esse mecanismo é intensificado pela visibilidade constante e pelos sistemas algorítmicos que quantificam e retroalimentam a aceitação pública das ideias. O resultado pode ser uma maior homogeneização discursiva, pois opiniões que recebem validação rápida tendem a ganhar visibilidade, enquanto posições dissidentes podem perder alcance ou deixar de ser expressas (Gaisbauer et al., 2020, p. 7).

Nesse cenário, o sujeito torna-se simultaneamente o objeto e o agente de sua própria normatização, operando sob uma racionalidade técnica que orienta suas ações segundo os critérios de visibilidade e aceitabilidade definidos por sistemas algorítmicos. Esse processo não se dá por imposição direta, mas por uma adaptação contínua às respostas do ambiente digital, que privilegia aquilo que é legível, replicável e recompensável. A normalização resulta da interação entre vigilância externa e autoconformação, formada na intersecção entre reconhecimento simbólico e cálculo algorítmico.

Embora esse mecanismo se assemelhe ao que Foucault chamou de *exame*, sua atualização nas plataformas digitais se distingue pela ausência de uma instância claramente identificável de julgamento. A qualificação, classificação e punição a que se refere Foucault (1995, p. 184) ainda estão presentes, mas são agora mediadas por métricas e interfaces, operando por meio de um controle distribuído e automatizado. Assim, o sujeito aprende a se adequar não pela repressão, mas pela internalização de padrões que lhe prometem aceitação e visibilidade. A sanção, nesse contexto, não é a exclusão formal, mas a invisibilidade algorítmica. A adaptação a esse regime pode levar o sujeito a reproduzi-lo ativamente, orientando sua expressão pelas recompensas intermitentes e pela lógica de desempenho que organizam a visibilidade nas plataformas.

Sob essa lógica, instala-se uma nova forma de sujeição que não mais se impõe pela coação externa, mas pela conversão da liberdade em performance e da expressão em mercadoria. Han descreve o sujeito contemporâneo como um “servo absoluto” que se acredita livre (2018b, p. 10). Essa liberdade possui caráter funcional e permanece formalmente assegurada, embora seja concretamente orientada por exigências de visibilidade, aprovação e engajamento. A busca por reconhecimento, longe de ser um movimento espontâneo do desejo humano, é explorada sistematicamente pelas plataformas como o motor de uma economia afetiva baseada na autoexploração. No neoliberalismo digital, até mesmo as emoções tornam-se capitalizáveis: “o regime neoliberal emprega as emoções como recursos para alcançar mais produtividade e desempenho” (Han, 2022, p. 65). As emoções passam, assim, a ser tratadas pelas plataformas como unidades operacionais quantificáveis e reconfiguráveis segundo a lógica da eficiência.

Esse novo regime de subjetivação articula-se intimamente com o modelo foucaultiano de disciplina, mas o transcende ao integrar controle e desejo. A disciplina, segundo Foucault, é “a técnica específica de um poder que forma os indivíduos como objetos e instrumentos de seu exercício” (Foucault, 1995, p. 170). No ambiente digital, essa técnica alcança as condições em que decisões, afetos e identidades se formam, ampliando o grau de interiorização do controle. Sob vigilância constante, o sujeito tende a internalizar os critérios de visibilidade oferecidos pelas plataformas e passa a ajustar sua própria conduta segundo padrões que lhe prometem reconhecimento e aceitação.

Han analisa a gamificação das interações sociais como um processo que traduz a busca por reconhecimento em métricas como curtidas, seguidores e comentários, nas quais a visibilidade funciona como recompensa. O autor apresenta o smartphone como um “objeto devocional” integrado a práticas de interação continuamente avaliadas por critérios de engajamento e legibilidade algorítmica (2018b, p. 24). A repetição dessas interações pode tornar o reconhecimento e o pertencimento mais dependentes das métricas oferecidas pelas plataformas.

Essa vulnerabilidade informacional também alcança a esfera pública e interfere nas condições cognitivas da deliberação. Em um ambiente onde os algoritmos filtram e hierarquizam conteúdos com base em perfis e preferências anteriores, a exposição ao contraditório é drasticamente reduzida. Sem contato com opiniões divergentes, os sujeitos deixam de acessar informações que poderiam beneficiar coletivamente a sociedade (Sunstein, 2018, p. 149). O efeito alcança a esfera pública ao erodir as condições para a construção de uma racionalidade comum, além de produzir bolhas informacionais no plano individual. A ignorância informacional, portanto, é também uma forma de injustiça relacional, ou seja, ao privar os indivíduos de visões alternativas, compromete-se o ecossistema cognitivo necessário à democracia e mina-se a possibilidade de constituir comunidades discursivas abertas à alteridade.

Em última instância, a mediação algorítmica participa da formação da subjetividade ao delimitar as ofertas disponíveis e organizar os contextos nos quais as escolhas são realizadas. Han descreve essa redução do horizonte decisório ao afirmar que “a livre escolha é extinta em prol de uma livre seleção entre as ofertas disponíveis” (2018b, p. 27). A agência permanece, embora seja exercida em condições previamente configuradas por sistemas de seleção e recomendação.

Essa reconfiguração subjetiva, entretanto, não se dá de forma isolada. Ao afetar o modo como os indivíduos se relacionam com o mundo, com os outros e consigo mesmos, ela alimenta processos coletivos de fechamento cognitivo, polarização e isolamento epistêmico. A subjetividade vulnerabilizada constitui, assim, o elo entre as dinâmicas técnicas de controle e as manifestações coletivas de desintegração comunicativa e deliberação pública comprometida.

4 Fragmentação social, polarização e isolamento epistêmico

As mediações algorítmicas interferem na produção técnica do comportamento, na experiência subjetiva e nas condições sociais de pertencimento, reconhecimento e construção do conhecimento. Forma-se uma estrutura informacional marcada pela fragmentação da esfera pública, pela polarização ideológica e pelo isolamento epistêmico. Esses fenômenos não são resultados necessários de toda comunicação digital, mas possibilidades recorrentes em ambientes que combinam homofilia, filtragem seletiva, personalização e incentivos voltados ao engajamento.

O conceito de câmaras de eco torna-se central para a compreensão das dinâmicas contemporâneas de fragmentação do espaço público. Jamieson e Cappella definem a câmara de eco como um “espaço de mídia fechado e delimitado, com potencial de amplificar as mensagens ali veiculadas e isolá-las de qualquer forma de refutação” (2008, p. 76). Para Samaržija, câmaras de eco são estruturas epistêmicas sociais que minam a confiança dos membros em fontes externas de informação. Sua característica essencial reside no fato de que, dentro delas, só se escuta o eco monótono de convicções já assumidas e das crenças de pares ideologicamente semelhantes (2023, p. 70).

É crucial distinguir essas estruturas das chamadas bolhas epistêmicas. Enquanto estas se caracterizam por uma omissão accidental de vozes externas, sendo dissolvíveis por meio de exposição a novas informações, as câmaras de eco são resilientes à refutação, pois operam ativamente na deslegitimação das fontes alternativas. Seu funcionamento consiste em convencer os membros de que o mundo exterior é movido por interesses torpes, ignorância ou má-fé, o que apenas fortalece a narrativa interna da câmara (Samaržija, 2023, p. 70-71).

A eficácia dessas estruturas reside na maneira como oferecem explicações simplificadas para experiências marcadas por ambiguidade e complexidade, acomodando ansiedades cognitivas e favorecendo a permanência de sujeitos predispostos a interpretar informações externas como ameaças às crenças do grupo. A resiliência epistêmica dessas câmaras decorre do fato de que toda evidência externa é reinterpretada a partir da premissa de que as crenças internas são inquestionáveis, ao passo que os interlocutores externos são vistos como moral e intelectualmente suspeitos (Samaržija, 2023, p. 72).

A influência das estruturas digitais não elimina inteiramente a responsabilidade epistêmica dos sujeitos, embora limite as condições concretas em que ela pode ser exercida. Samaržija argumenta que a permanência em câmaras de eco pressupõe uma adesão deliberada a sistemas de crenças autorreferentes. Sair de uma câmara de eco é, portanto, possível em tese, mas extremamente difícil na prática, já que a simples exposição a informações ou evidências contrárias não costuma ser suficiente para dissolver a adesão ideológica (Samaržija, 2023, p. 73).

As câmaras de eco participam ativamente da fragmentação do espaço público ao consolidar enclaves ideológicos autorreferentes. O ambiente digital, longe de configurar uma esfera pública aberta e racional, como idealizado por Habermas (2001, p. 116), converte-se em inúmeras comunidades epistêmicas isoladas, em que a reafirmação identitária prevalece sobre a deliberação e o confronto de ideias.

A persistência desses enclaves ideológicos pode ser compreendida pelo modelo das espirais reforçadoras, desenvolvido por Jamieson e Cappella (2008, p. 83) a partir da teoria de Slater (2007). Trata-se de um processo circular no qual identidades e convicções são continuamente confirmadas e, por isso mesmo, reforçadas por exposição seletiva a conteúdos que reiteram preconceitos e expectativas do grupo.

E quanto mais uma pessoa ouve mensagens que se alinham com suas crenças existentes, mais ela vai descartar o que não se alinha, reforçando a necessidade de pertencer a um grupo e tornando-se ainda mais seletiva sobre o conteúdo que confirma o que já está inclinada a acreditar. Em sistemas comunicacionais perfeitamente fechados, como câmaras de eco consolidadas, os efeitos dessas espirais atingem seu máximo potencial, produzindo estreitamento cognitivo, encapsulamento ideológico e intolerância. A consequência é o colapso gradual da capacidade de escuta, da abertura ao outro e da deliberação pública. A redução da visão de mundo à repetição de conteúdos validados internamente compromete a diversidade informacional e enfraquece as condições necessárias à construção de uma esfera pública democrática. Trata-se de um processo em que o sujeito, enquanto agente epistêmico, vai sendo progressivamente absorvido por uma lógica de confirmação, que mina sua autonomia crítica e sua responsabilidade deliberativa, reforçando o tribalismo digital e a fragmentação social.

Esse processo tem implicações diretas na constituição da identidade, da afiliação e da coesão grupal. Jamieson e Cappella argumentam que as câmaras de eco podem promover engajamento e coerência ideológica ao mesmo tempo que produzem polarização, desqualificação retórica do adversário e fragmentação deliberativa (2008, p. 247). Trata-se de uma coesão construída pela exclusão sistemática do outro, pela ridicularização do dissenso e pela promoção de formas de *tribalismo digital* que sacrificam a complexidade do debate em favor da pureza ideológica.

Os modelos computacionais de Banisch, Gaisbauer e Olbrich permitem representar algumas dessas dinâmicas ao simular como a homofilia e a autoinsulação informacional podem produzir estados estáveis de segregação cognitiva. Os chamados “quartos semipúblicos de expressão” (*semipublic rooms for expressing opinion*) funcionam como microambientes ideológicos que permitem a exteriorização de opiniões socialmente controversas ou marginalizadas. Esses espaços, embora possam romper parcialmente a *espiral do silêncio* (*spiral of silence*), ao conferir segurança para a expressão de ideias rejeitadas no discurso dominante, acabam paradoxalmente por consolidar zonas discursivas homogêneas, contribuindo para o aprofundamento da polarização (Banisch; Gaisbauer; Olbrich, 2022, p. 8).

Habermas analisa esse processo como uma fragmentação da esfera pública em múltiplas esferas semipúblicas em competição, denominadas *Halböffentlichkeit* (2022, p. 63–64). Essa transição rebaixa a esfera pública política, convertendo-a em mais uma entre várias arenas de visibilidade estratégica, sujeita à lógica da competição simbólica e à erosão dos critérios de validade comunicativa. Quando a troca crítica de argumentos se torna impraticável, o pluralismo perde suas condições de exercício e a racionalidade democrática torna-se progressivamente mais frágil.

A fragmentação da esfera pública não é intrinsecamente antidemocrática, pois Fraser questiona a superioridade de uma esfera pública única e sustenta que, em contextos de desigualdade estrutural, os contrapúblicos subalternos (*subaltern counterpublics*) permitem que grupos excluídos elaborem e difundam interpretações próprias sobre suas identidades, interesses e necessidades (Fraser, 1990, p. 66-68). Embora possam reproduzir exclusões e conformidades internas, esses contrapúblicos ampliam a contestação discursiva e reduzem parcialmente as vantagens dos grupos dominantes.

Mesmo que Fraser não examine os ambientes digitais, sua concepção permite distinguir os contrapúblicos das câmaras de eco, uma vez que os primeiros procuram fazer seus discursos circularem em esferas mais amplas e estabelecer relações de contestação com outros públicos (Fraser, 1990, p. 67-70). As comunidades digitais perdem esse potencial quando se tornam impermeáveis, desqualificam fontes externas e impedem a circulação de argumentos, de modo que o problema democrático não está na multiplicidade dos públicos, mas no isolamento que impossibilita a contestação entre eles.

Como argumenta Bresolin, o problema central da deformação da esfera pública contemporânea não está apenas no acúmulo de desinformação ou de *fake news*, mas na crescente incapacidade dos sujeitos de reconhecer essas informações como falsas. Trata-se de uma falência epistêmica que resulta não da ausência de dados, mas da erosão dos critérios partilhados de validação e da perda de confiança nos mecanismos coletivos de verificação (2023, p. 779). O fechamento epistêmico favorecido pelas plataformas digitais empobrece o debate público e compromete os fundamentos normativos da deliberação democrática.

Os modelos de Banisch, Jacob, Willaert e Olbrich (2024) mostram que configurações segregadas podem permanecer estáveis sob determinadas condições, inclusive quando os agentes manifestam alguma preferência pela diversidade. O resultado não autoriza, entretanto, a postular um “determinismo algorítmico emergente”. Além de se tratar de uma simulação baseada em premissas específicas, o próprio modelo admite mais de um equilíbrio possível. O importante aqui é mostrar como interações localmente orientadas pela aprovação podem gerar resultados coletivos indesejáveis e não em demonstrar que os usuários sejam conduzidos invariavelmente à clausura.

A polarização e o isolamento epistêmico devem, portanto, ser compreendidos como efeitos estruturais possíveis e recorrentes de certas estruturas sociotécnicas. Sua intensidade depende das *affordances* da plataforma, das redes de interação, das práticas dos usuários e das instituições que regulam a circulação pública da informação.

As plataformas digitais não operam, portanto, como meios neutros de expressão, mas como ambientes projetados segundo *affordances* que favorecem o fechamento epistêmico. Willaert e Olbrich observam que essas *affordances*, possibilidades de ação estruturadas pelo design da plataforma, estão estreitamente ligadas aos conceitos de câmaras de eco e bolhas de filtro. Nesse contexto, emergem identidades coletivas digitais ancoradas em símbolos de pertencimento e sinais de coesão ideológica, como *hashtags*, bandeiras e imagens de perfil (2024, p. 14). Esses marcadores simbólicos operam como dispositivos de diferenciação, segmentando os usuários em comunidades discursivas autorreferentes, cuja adesão se dá mais pela identificação afetiva do que pela racionalidade argumentativa.

A filtragem informacional extrema pode surgir como efeito estrutural de plataformas cujo design privilegia a personalização e o engajamento, reduzindo a diversidade cognitiva e favorecendo a conformidade ideológica. Sunstein sustenta que a exposição repetida a autores e conteúdos que reforçam convicções políticas semelhantes estreita o horizonte cognitivo e fragmenta a cultura comum (2018, p. 34). Essa “proliferação de nichos”, segundo o autor, compromete os elementos compartilhados da vida democrática e promove uma fragmentação deliberada ao longo de linhas políticas (Sunstein, 2018, p. 151).

Essa dinâmica repercute na própria experiência da aprendizagem e da construção do saber. Pariser afirma que, dentro das câmaras de eco, há menos espaço para os encontros fortuitos que geram insight e descoberta, a saber, “um mundo construído a partir do familiar é um mundo no qual não há nada a aprender” (2011, p. 14). A perda da serendipidade cognitiva compromete a plasticidade intelectual, reduz a abertura à diferença e empobrece a experiência democrática.

A fragmentação da esfera pública pode manifestar-se em diferentes experiências de segmentação informacional, com intensidade variável conforme as plataformas e as práticas dos usuários. Pariser pergunta se um escândalo como o Rathergate ainda seria possível em um ambiente informacional profundamente fragmentado (2011, p. 2). O horizonte democrático, que exige abertura à alteridade e a disposição dialógica, cede lugar à construção de bolhas identitárias que funcionam como filtros de confirmação e territórios de reafirmação afetiva. Pariser relaciona esse fechamento à perda de perspectiva recíproca ao afirmar que “a democracia exige que os cidadãos vejam o mundo sob o ponto de vista uns dos outros, mas estamos cada vez mais enclausurados em nossas próprias bolhas” (2011, p. 8–9).

O fechamento comunicacional reduz a exposição à alteridade quando a comunicação digital privilegia a circulação do semelhante e dificulta o encontro com perspectivas divergentes, como sugere Han ao afirmar que “a comunicação alcança sua velocidade máxima ali onde o igual responde ao igual [...] a negatividade da alteridade e do que é alheio ou a resistência do outro atrapalha” (2017, p. 12). A multiplicação de microesferas identitárias e o enfraquecimento de critérios comuns de validação tornam a contestação recíproca mais difícil e fragilizam as condições da deliberação pública.

Quando os grupos passam a operar com fontes de informação, critérios de legitimidade e normas de pertencimento inteiramente separados, a construção de uma vontade pública comum torna-se mais difícil. O problema surge da perda de instâncias confiáveis e compartilhadas de mediação epistêmica, que transforma o espaço público em uma justaposição de discursos com pouca capacidade de contestação recíproca. A erosão da esfera pública alcança, assim, os próprios critérios pelos quais uma afirmação pode ser reconhecida como verdadeira, legítima ou argumentativamente válida, produzindo uma crise de confiança no tecido comunicativo que sustenta a convivência democrática.

Bresolin afirma que o problema ultrapassa a circulação de *fake news* e alcança a erosão da capacidade coletiva de reconhecê-las como tais (2023, p. 779). A deformação da esfera pública reside na perda dos dispositivos comuns de reconhecimento do erro, da mentira e da manipulação. Nesse quadro, deterioram-se a qualidade da informação e as condições necessárias à formação de uma vontade pública informada, indispensável a qualquer democracia.

A análise foucaultiana da disciplina oferece um vocabulário para compreender essa tendência ao examinar as técnicas de ordenação espacial dos sujeitos, a saber, “a disciplina parte da distribuição dos indivíduos no espaço. Cada indivíduo tem seu lugar; e cada lugar, seu indivíduo” (Foucault, 1995, p. 141–143). Essa racionalidade disciplinar, cujo objetivo é organizar, separar e tornar os corpos legíveis e úteis, encontra no ambiente digital uma reatualização técnico-informacional. Os enclaves informacionais constituem atualmente zonas de controle e previsibilidade semelhantes aos espaços fisicamente compartimentados do dispositivo disciplinar. A demarcação algorítmica, os sistemas de categorização por perfil e a separação de usuários com base em interesses e crenças atualizam, em outro contexto, técnicas de classificação e distribuição analisadas por Foucault. A disciplina digital alcança os fluxos de dados, os posicionamentos discursivos e os padrões de engajamento, acrescentando à ordenação física dos corpos uma forma de segmentação informacional.

Esses novos dispositivos organizam a administração da informação enquanto delimitam as fronteiras daquilo que pode ser visto, dito, compartilhado e reconhecido socialmente. O espaço digital converte-se, assim, em um campo de visibilidade hierarquizada no qual a disciplina opera por meio da segmentação informacional, reforçando o isolamento epistêmico e a compartimentalização da experiência social.

5 Considerações finais

A vulnerabilidade informacional forma-se por um circuito de relações que ultrapassa a ação de um único mecanismo e a suposta fragilidade dos usuários, uma vez que os sistemas de seleção, predição e recompensa organizam o campo das interações possíveis, os sujeitos aprendem a orientar sua expressão segundo métricas de visibilidade e aprovação e a repetição dessas condutas contribui para ambientes coletivos mais homogêneos e menos permeáveis à contestação. Nesse processo, a mediação técnica afeta a experiência subjetiva, que, por sua vez, retroalimenta as estruturas epistêmicas e sociais nas quais a própria tecnologia opera.

A articulação entre as dimensões técnica, subjetiva e epistêmico-coletiva mostra que a vulnerabilidade designa uma relação assimétrica entre os sujeitos e as estruturas informacionais, reunindo efeitos tecnológicos, psicológicos e políticos que costumam ser examinados separadamente. A vulnerabilidade informacional nomeia a condição disposicional em que a formação do juízo, da vontade e do reconhecimento depende de mediações opacas cuja compreensão, contestação e revisão permanecem fora do alcance proporcional dos sujeitos. Essa dependência pode anteceder a ocorrência de um dano identificável e se distribui de maneira desigual entre indivíduos e grupos.

Uma estrutura digital amplia a vulnerabilidade informacional quando suas operações relevantes permanecem opacas ou dificilmente contestáveis, quando seus incentivos tornam a autonomia, a competência e o pertencimento dependentes de métricas de desempenho e quando sua organização reduz a circulação entre perspectivas distintas e enfraquece os critérios compartilhados de validação. Esses critérios oferecem um horizonte normativo para a avaliação das plataformas e delimitam os problemas que uma resposta regulatória ou institucional precisa enfrentar, embora ainda deixem em aberto quais medidas seriam mais adequadas para cada ambiente digital.

Uma subjetividade anterior e supostamente imune à mediação técnica serviria como uma referência inadequada para enfrentar esse problema, já que sujeitos e espaços públicos sempre foram formados por instituições, linguagens e relações de poder. A especificidade da racionalidade algorítmica encontra-se na escala, na opacidade e na capacidade de individualizar suas intervenções enquanto seus efeitos coletivos permanecem pouco visíveis. A crítica da vulnerabilidade informacional deve incidir sobre essa assimetria e preservar as condições para que sujeitos e públicos possam compreender, contestar e transformar as mediações técnicas que participam da formação de sua experiência.

Referências

- BANISCH, S.; GAISBAUER, F.; OLBRICH, E. 2022. Modelling Spirals of Silence and Echo Chambers by Learning from the Feedback of Others. *Entropy*, **24**(10): p. 1484.
- BANISCH, S.; JACOB, D.; WILLAERT, T.; OLBRICH, E. 2024. *A Dynamical Model of Platform Choice and Online Segregation*. [S.l.]: arXiv:2411.04681.
- BRESOLIN, K. 2023. Liberdade de expressão, esfera pública, plataformas digitais e mídias sociais. *Ethic@*, Florianópolis, **22**(2): p. 761–790.
- FINEMAN, M. A. 2008. The vulnerable subject: anchoring equality in the human condition. *Yale Journal of Law & Feminism*, New Haven, **20**(1): p. 1–23.
- FOUCAULT, M. 1995. *Discipline and Punish: The Birth of the Prison*. 2. ed. New York: Vintage Books.

- FRASER, N. 1990. Rethinking the public sphere: a contribution to the critique of actually existing democracy. *Social Text*, (25/26): p. 56-80.
- GAISBAUER, F.; OLBRICH, E.; BANISCH, S. 2020. Dynamics of opinion expression. *Physical Review E*, **102**(4): p. 042303.
- HABERMAS, J. 2001. *Strukturwandel der Öffentlichkeit*: Untersuchungen zu einer Kategorie der bürgerlichen Gesellschaft. Frankfurt am Main: Suhrkamp.
- HABERMAS, J. 2022. Überlegungen und Hypothesen zu einem erneuten Strukturwandel der politischen Öffentlichkeit. In: HABERMAS, J. *Ein neuer Strukturwandel der Öffentlichkeit und die deliberative Politik*. Berlin: Suhrkamp, p. 9–68.
- HAN, B.-C. 2017. *Sociedade da transparência*. Tradução de Enio Paulo Giachini. Petrópolis: Vozes.
- HAN, B.-C. 2018a. *No exame*: perspectivas do digital. Tradução de Lucas Machado. Petrópolis: Vozes.
- HAN, B.-C. 2018b. *Psicopolítica*: o neoliberalismo e as novas técnicas de poder. Tradução de Maurício Liesen. Belo Horizonte: Âyiné.
- HAN, B.-C. 2022. *Infocracia*: digitalização e a crise da democracia. Tradução de Gabriel S. Philipson. Petrópolis: Vozes.
- JAMIESON, K. H.; CAPPELLA, J. N. 2008. *Echo Chamber*: Rush Limbaugh and the Conservative Media Establishment. New York: Oxford University Press.
- MACKENZIE, C.; ROGERS, W.; DODDS, S. 2014. Introduction: what is vulnerability, and why does it matter for moral theory? In: MACKENZIE, C.; ROGERS, W.; DODDS, S. (ed.). *Vulnerability: new essays in ethics and feminist philosophy*. New York: Oxford University Press. p. 1–29.
- PARISER, E. 2011. *The filter bubble*: What the Internet is hiding from you. New York: Penguin Press.
- RYAN, R. M.; DECI, E. L. 2020. Intrinsic and extrinsic motivation from a self-determination theory perspective: definitions, theory, practices, and future directions. *Contemporary Educational Psychology*, **61**: p. 101860.
- SAMARŽIJA, H. 2023. Echo Chambers and Epistemic Responsibility. In: KRASNOV, Y.; LU, S. (Eds.). *The Epistemology of Digital Media*. London: Routledge, p. 67–83.
- SLATER, M. D. 2007. Reinforcing Spirals of Mutual Influence of Media Selectivity and Media Effects and Their Impact on Individual Behavior and Social Identity. *Communication Theory*, **17**(3): p. 281–303.
- SUNSTEIN, C. R. 2018. *#Republic*: Divided Democracy in the Age of Social Media. Princeton: Princeton University Press.
- VANSTEENKISTE, M.; RYAN, R. M. 2013. On psychological growth and vulnerability: Basic psychological need satisfaction and need frustration as a unifying principle. *Journal of Psychotherapy Integration*, Washington, DC, **23**(3): p. 263–280.
- WILLAERT, T.; OLBRICH, E. (com contribuições de BANISCH, S.). 2024. *Platform Affordances*. Bruxelas: Vrije Universiteit Brussel; Leipzig: Max Planck Institute for Mathematics in the Sciences.

Declaração de Disponibilidade de Dados:

Todo o conjunto de dados que dá suporte aos resultados deste estudo foi publicado no próprio artigo.

Submetido: 02 de outubro de 2025.

Aceito: 24 de agosto de 2026.