

Comuns do conhecimento e desigualdades: desafios associados a capitalismo de vigilância e IA generativa



Miguel Said Vieira*

Resumo: Apresenta a noção de comuns do conhecimento, discutindo as relações ambíguas entre tais comuns, o capitalismo de vigilância e a inteligência artificial generativa. Conclui refletindo sobre a possibilidade de recuperar o sentido contra-hegemônico de comuns do conhecimento.

No início dos anos 2000, uma série de iniciativas fundamentais relacionadas ao compartilhamento do conhecimento pipocaram no cenário global: em 2001, foi lançada a Wikipédia; em fevereiro de 2002, a Budapest Open Access Initiative (marco do movimento de acesso aberto, que defende a livre circulação de publicações científicas); e em dezembro de 2002, as licenças Creative Commons (que permitem a criadores indicar que suas obras podem ser distribuídas livremente).

Embora essas iniciativas tivessem antecedentes nas décadas anteriores — como o movimento do software livre, iniciado nos anos 1980 e já vibrante nos 1990 —, os avanços da digitalização e a popularização da internet parecem ter formado as condições perfeitas para que essa virada de milênio fosse palco para o surgimento de novas práticas de comuns.¹

Estudiosos começaram, inclusive, a identificar tais iniciativas como exemplos de um novo *tipo* de comuns: imateriais, digitais ou do conhecimento. Em contraposição aos comuns “tradicionais”, baseados no compartilhamento de recursos naturais (como estoques pesqueiros para pesca, terra para pasto e plantio etc.), estes novos comuns seriam baseados no compartilhamento de coisas imateriais: softwares, artigos científicos, músicas, fotografias... em suma, conhecimento e cultura.

*Professor de Políticas Públicas da UFABC.

Palavras-chave: comuns do conhecimento; capitalismo de vigilância; IA generativa.

1. Numa definição simples, comuns (ou bens comuns) são práticas em que uma comunidade compartilha coisas.

E enquanto os comuns anteriores compartilhavam coisas já existentes, os novos comuns compartilhavam coisas criadas pelos próprios membros de suas comunidades.

Vários desses comuns do conhecimento alcançaram grande relevância. Em 2024, a Wikipédia tem versões ativas em 339 idiomas, e a versão em inglês tem 6,8 milhões de artigos. Dos artigos científicos publicados em 2023, cerca de 46% encontram-se hoje sob acesso aberto;² e 86% dos servidores de sites web utilizam um sistema operacional baseado em software livre,³ bem como 71% dos celulares⁴.

Sob uma certa perspectiva, seria possível dizer que esses comuns — e com eles, novas formas de compartilhar conhecimento — “venceram”. Como veremos a seguir, porém, o cenário tornou-se bem mais complexo.

Capitalismo de vigilância

Nas últimas duas décadas, transcorreu uma outra transformação socioeconômica, também intimamente ligada à internet, e com efeitos provavelmente ainda mais profundos: a ascensão do capitalismo de vigilância. Para Shoshana Zuboff, que propôs essa ideia, entramos em uma fase do capitalismo em que a capacidade de prever e influenciar comportamentos a partir da coleta massiva de dados pessoais tornou-se uma das principais fontes de poder. Vejamos alguns elementos dessa transformação.

A partir dos anos 2000, a internet tornou-se cada vez mais o ponto focal da chamada “cultura participativa”: o consumo “passivo” de mídia (na forma de programas de rádio e TV, revistas etc.) cedeu espaço a uma relação mais participativa das pessoas na produção e difusão de conteúdos que lhes interessavam — por meio de espaços como fóruns, sistemas de comentários, *blogs* e redes sociais. As próprias tecnologias relacionadas à internet acompanharam essa mudança, na tendência conhecida como “Web 2.0”, que enfatizou conteúdo produzido por usuários, novas formas de interação e a possibilidade de acesso por diferentes tipos de dispositivos, incluindo celulares.

Essas mudanças, por sua vez, contribuíram para um grande avanço da publicidade online. Por um lado, tratou-se de salto quantitativo: o acesso à internet aumentou em número, frequência e engajamento; e os espaços publicitários tornaram-se abundantes e mais baratos. Mas por outro lado, o salto também foi qualitativo: um mesmo espaço publicitário online pode exibir publicidades diferentes para cada usuário; e a abundância de dados que é possível coletar sobre esses usuários — por meio de seu histórico de navegação, das postagens ou mesmo simples interações que fazem em cada site ou *app* — permitiu a criação de formas de direcionamento publicitário antes impensáveis.

A concentração da internet em um pequeno número de plataformas (controladas por megaempresas como Google, Meta e Amazon) e a popularização dos *smartphones* ampliaram as possibilidades dessa vigilância, bem como a intensidade de nossa participação num circuito contínuo de interação, coleta de dados e modulação de comportamento (não mais apenas com fins publicitários, mas também políticos); esse circuito é a base do capitalismo de vigilância.

2. Percentual verificado nos artigos indexados pela base de dados Scopus, em 2024. Ver <https://en.wikipedia.org/w/index.php?title=Open_access&oldid=1251311113#Extent>.

3. Ver <https://w3techs.com/technologies/overview/operating_system>.

4. Ver <<https://www.statista.com/statistics/272698/global-market-share-held-by-mobile-operating-systems-since-2009/>>.

Há nexos entre o capitalismo de vigilância e os comuns do conhecimento. Em primeiro lugar, há pontos de contato entre esses comuns e a “cultura participativa”: embora sejam fenômenos distintos, em ambos a criação por usuários é central; e softwares livres, especificamente, são peça fundamental na infraestrutura (majoritariamente controlada por *big techs*) que possibilitou uma web interativa e presente em todo tipo de dispositivo. Em segundo lugar, os comuns do conhecimento contribuíram para a transição de um cenário cultural marcado pela escassez a outro marcado pela abundância: em que o grosso do conteúdo que consumimos ou é produzido por usuários — seja em comuns como a Wikipédia, seja em plataformas movidas a publicidade como Instagram e Youtube —, ou é acessado em plataformas de assinatura (como Netflix e Spotify), em que se paga a *big techs* pelo acesso a um grande catálogo. (É importante notar que essas plataformas conquistaram muitos de seus clientes entre as pessoas que, antes disso, fruía de abundância similar, mas em redes P2P de compartilhamento de arquivos; redes que, embora ilegais, também podem ser entendidas como comuns — ou “acesso aberto de guerrilha”, na formulação de Aaron Swartz). À exceção do exemplo da Wikipedia, ambos os casos desse novo cenário de abundância seguem a lógica do capitalismo de vigilância: a mercadoria mais relevante deixa de ser a obra cultural, e passa a ser a modulação do nosso comportamento.

Inteligência artificial generativa

Se, até aqui, a mineração de dados fora focada em prever e modular comportamentos, o desenvolvimento de ferramentas de inteligência artificial (IA) generativa (como ChatGPT e DALL-E) demonstrou que a experiência humana também pode ser minerada para produzir obras e conteúdos sintéticos, similares aos produzidos por humanos. A IA generativa é um elemento que, embora possa alterar o cenário atual, emerge do próprio capitalismo de vigilância: como o desenvolvimento dessa tecnologia requer tanto dados como capacidade computacional em grande quantidade, a infraestrutura e os atores do capitalismo de vigilância seguem sendo centrais.

Para a previsão e modulação de comportamentos, já eram necessários muitos dados; para o treinamento eficaz de modelos de IA generativa, porém, são necessários dados — textos, imagens e outras produções humanas — em uma escala ainda maior. Empresas como Google e Meta detêm acesso privilegiado material útil para esse treinamento, como a infinidade de e-mails, postagens, áudios e vídeos produzidos por usuários de suas plataformas; e as muitas formas de interação nessas plataformas (como cliques, curtidas, denúncias, comentários), que podem ser utilizadas para refinar o treinamento desses modelos. Na prática, porém, a maioria dos conteúdos já disponibilizados na internet estão sendo usados para esse tipo de treinamento: todos os sites já publicados, todas as obras literárias já digitalizadas, e assim por diante.

Embora essa ainda seja uma questão controversa, há dúvidas sobre a legalidade do uso de boa parte desses conteúdos para treinamento: as leis de direitos autorais talvez limitem o uso de qualquer conteúdo que não esteja em domínio público, mesmo quando publicado em sites de livre acesso. E aqui, é importante notar como os comuns do conhecimento novamente podem ser instrumentalizados em benefício dos mais poderosos: em boa parte desses comuns, há a autorização expressa para qualquer tipo de uso por terceiros, eliminando a possibilidade de restrições por direitos autorais.

O software livre e os artigos científicos em acesso aberto são exemplos de comuns com obras culturais de qualidade e relevância particularmente úteis para o treinamento desses modelos. É tristemente irônico perceber que, embora esses comuns tenham sido construídos para fomentar produção comunitária e acesso não-mercantil, neste momento eles se tornam matéria-prima a beneficiar *big techs*: por mais que, em tese, comunidades também possam usar esses comuns para treinamento de IAs, só as *big techs* detêm a infraestrutura computacional e o acesso aos demais dados na escala necessária para o desenvolvimento dos modelos de IA mais avançados.

Considerações finais

Para recuperar o potencial contra-hegemônico dos comuns do conhecimento, talvez seja necessário recuperar lições dos comuns “tradicionais”, e até relativizar a distinção entre eles. É importante lembrar que comuns do conhecimento também requerem coisas materiais: da energia e minérios para que servidores funcionem, à pizza para quem programa; e, analogamente, em comuns “tradicionais” também se compartilha coisas imateriais: dos conhecimentos sobre o quê e quando plantar naquele território, aos laços afetivos em uma horta comunitária.

É fato que é mais fácil obter abundância de recursos em comuns do conhecimento, já que um software, por exemplo, pode ser facilmente copiado, e potencialmente pode ser compartilhado para muito mais pessoas do que a colheita de uma horta comunitária; no entanto, isso não elimina a possibilidade de que os comuns do conhecimento tenham regras e fronteiras delimitadas: elas não são necessárias para prevenir sobreutilização, mas podem ser úteis para favorecer comunidades em detrimento de grandes empresas, e para reduzir as desigualdades que serão ampliadas com o advento da inteligência artificial generativa.

Nesse sentido, talvez seja importante voltar a considerar a adoção de estratégias restritivas de licenciamento, tais como: o *copyleft*, que obriga quem produzir obras derivadas de um comum a compartilhá-las da mesma forma; licenças como a AGPL (Affero General Public License), que exige que software livre, que funciona através da internet, seja compartilhado com os usuários finais; as cláusulas de uso não-comercial, que proíbem o uso do bem comum em atividades com finalidade de lucro; e licenças de produção por pares, que permitem usos comerciais do comum, mas apenas para cooperativas e trabalhadores, e eventualmente para empresas que contribuam para a manutenção daquele comum.

E, para além dessas estratégias, é fundamental que as pessoas que cuidam de e contribuem para esses comuns reconheçam tais riscos de desigualdades, e mobilizem-se coletivamente contra eles: há muito a aprender aqui com movimentos sociais em geral.

