

Pesquisa-Ação Participativa em Estudos Críticos de Dados: Analisando IA a partir de uma Abordagem Sul-Norte¹

Participatory Action Research in Critical Data Studies: Analyzing AI from a South-North Approach

Investigación-acción participativa en estudios de datos críticos: análisis de la IA desde un enfoque Sur-Norte

Andrea MEDRADO²
Pieter VERDEGEM³

Resumo

Inspirados na pesquisa-ação participativa (PAP) e no trabalho de pensadores latino-americanos como Freire e Fals Borda, propomos um fluxo Sul-Norte visando desafiar a forma como a centralidade do Norte é tida como certa em relação às epistemologias, experiências e entendimentos da IA. A partir de oficinas realizadas com um grupo diversificado de estudantes, profissionais da área de tecnologia e ativistas argumentamos que a PAP pode não apenas capacitar comunidades marginalizadas no Sul Global, é importante aprendermos mais sobre sua relevância e utilizações em contextos do Norte Global. Nossa análise se aprofunda em três conceitos específicos em torno da IA e da (in)justiça de dados: *autonomia, empatia e diálogo*, o que provocou, entre outras reflexões, lembrar que a natureza humana é *sentipensante* (Borda, 2003).

Palavras-chave: IA; Estudos de Dados Críticos; Justiça de Dados; Sul-Norte; Autonomia.

¹Uma versão em inglês deste texto foi publicada na *Revista Big Data & Society* em março de 2024. <https://doi.org/10.1177/20539517241235869>. Tradução de Luíza Ferreira. (Foram feitas algumas adaptações formais em função dos critérios de publicação da Revista Mídia e Cotidiano).

² Doutora em Comunicação Social, Professora Departamento de Comunicação Drama e Filme (Department of Communication Drama and Film), Universidade de Exeter, Reino Unido. É co-vice-presidente da Associação Internacional de Pesquisa em Mídia e Comunicação, IAMCR. E-mail: a.medrado@exeter.ac.uk. ORCID: [0000-0002-9408-9688](https://orcid.org/0000-0002-9408-9688)

³ Professor da CAMRI, Escola de Mídia e Comunicação de Westminster (Westminster School of Media and Communication), Universidade de Westminster, Reino Unido. E-mail: p.verdegem@westminster.ac.uk. ORCID: orcid.org/0000-0001-7906-002X



Abstract

Inspired by participatory action research (PAP) and the work of Latin American thinkers such as Freire and Fals Borda, we propose a South-North flow aiming to challenge the way in which the centrality of the North is taken for granted in relation to epistemologies, experiences and understandings. of AI. Based on workshops held with a diverse group of students, technology professionals and activists, we argue that PAP can not only empower marginalized communities in the Global South, it is important that we learn more about its relevance and uses in contexts of the Global North. Our analysis delves into three specific concepts around AI and data (in)justice: autonomy, empathy and dialogue, which provoked, among other reflections, remembering that human nature is sentient (Borda, 2003).

Keywords: AI; Critical Data Studies; Data Justice; South North; Autonomy.

Resumen

Inspirándonos en la investigación acción participativa (PAP) y el trabajo de pensadores latinoamericanos como Freire y Fals Borda, proponemos un flujo Sur-Norte con el objetivo de desafiar la forma en que se da por sentada la centralidad del Norte en relación con las epistemologías. experiencias y comprensiones de la IA. Basándonos en talleres realizados con un grupo diverso de estudiantes, profesionales de la tecnología y activistas, sostenemos que el PAP no sólo puede empoderar a las comunidades marginadas del Sur Global, es importante que aprendamos más sobre su relevancia y usos en contextos del Norte Global. Nuestro análisis profundiza en tres conceptos específicos en torno a la IA y la (in)justicia de los datos: autonomía, empatía y diálogo, lo que provocó, entre otras reflexiones, recordar que la naturaleza humana es sensible (Borda, 2003).

Palabras clave: AI; Estudios de Datos Críticos; Justicia de datos; Sur norte; Autonomía.

Introdução

Após uma longa pausa devido à pandemia, estávamos em uma sala para participar de uma mesa em uma conferência. Organizada por uma universidade europeia, o tema da conferência era a “descolonização da Internet”. Nosso artigo refletia sobre as possibilidades de aplicar a pesquisa-ação participativa (PAP) latino-americana aos estudos de dados críticos e inteligência artificial (IA). Citando autores como Paulo Freire e Orlando Fals Borda em uma versão inicial deste trabalho, e ao responder perguntas dos participantes, em sua maioria europeus, a autora 1, Andrea Medrado, não pôde evitar a síndrome de impostora: Estaríamos fazendo justiça ao



trabalho desses autores? Faria sentido usar o PAP, uma abordagem radical que surgiu nas décadas de 1970 e 1980, para analisar a IA, que é considerada o futuro?

Enquanto Andrea lidava com suas perguntas internas, chegou o momento de perguntas e respostas. Uma das perguntas era sobre o perfil dos participantes. O trabalho tinha como base oficinas realizadas em Londres com um grupo composto predominantemente por estudantes universitários, profissionais da área de tecnologia e ativistas de classe média. A pergunta era: “Já que essas abordagens foram desenvolvidas para trabalhar com grupos oprimidos (para usar um termo freireano), será que seriam relevantes para pessoas que não tem tanta vivência de marginalização social?” A pergunta foi útil e convidou-nos a refletir sobre a adequação do uso de abordagens que decorrem de contextos de marginalização e a empregá-las em diferentes contextos. Simplificando, que lições podem ser aprendidas com a aplicação, no *Norte Global*, de abordagens metodológicas enraizadas no *Sul Global*?

- De que maneira a tradição latino-americana de PAP (pesquisa-ação participativa) pode nos inspirar a analisar questões acerca da IA e (in)justiça de dados?

As perguntas adicionais são:

- Quais estruturas epistemológicas são oferecidas pela PAP para nos ajudar a compreender questões críticas em torno da IA e das assimetrias de poder entre o chamado *Norte Global* e o *Sul Global*?
- Quais prioridades de pesquisa emergem ao abraçar fluxos epistemológicos e metodológicos *Sul-Norte* para estudos críticos de IA e dados?
- Quais insights as abordagens de PAP latino-americanas podem oferecer quando aplicadas em contextos do *Norte Global* com participantes que não são tão afetados por questões de marginalização?

Como estudiosos de dados *críticos*, nossa tarefa vai além de fazer as melhores perguntas *críticas*. Também se faz necessário buscar as formas *críticas* mais adequadas para respondê-las. Assim, à medida que buscamos respostas sobre questões de IA e (in)justiça de dados, argumentamos que elas devem decorrer das realidades daqueles que são precisamente os mais afetados pelo capitalismo extrativista (digital): os grupos marginalizados no *Sul Global*. Desta forma, convidamos os estudiosos de dados críticos a estarem mais abertos ao que chamamos de fluxos epistemológicos e metodológicos *Sul-Norte*. Ao fazer isso, desafiamos o tratamento dos contextos do *Sul*



(como o contexto latino-americano), como meros campos cobiados para a aplicação de modelos teóricos e metodológicos do Norte Global. Ecoando perspectivas de estudiosos que se concentram em dados, tecnologia e decolonialidade (Aguilar, 2020; Aouragh e Chakravartty, 2016), pretendemos reverter essa dinâmica epistemológica (colonial).

Aqui, fazemos uma pausa para algumas explicações conceituais: o que queremos dizer com o *Sul Global*? O que queremos dizer com *justiça de dados*? E o que estamos chamando de *estudos críticos de dados*? Começamos pela primeira pergunta. O termo *Sul Global* é problemático porque aglomera o outro, reforçando a alteridade dos outros. Seria simplista impor semelhanças entre um país latino-americano e um país africano, por exemplo, simplesmente porque eles têm uma economia mais vulnerabilizada do que, digamos, um país da Europa Ocidental. Assim, ao adotar o termo poderíamos contribuir indiretamente para homogeneizar o que constitui o verdadeiro *Sul*. Isto é ainda mais complicado pelo fato de existirem *Suis nos Nortes* e vice-versa (como bolsões de pobreza no Norte Global e bolsões de riqueza no Sul Global).

Ao mesmo tempo, o termo *Sul Global* pode ser útil e estratégico, e ajudar a conectar realidades marginalizadas, embora elas sejam, obviamente, muito distintas. Referindo-se a mais do que uma localização geográfica, o *Sul* aparece como uma metáfora para opressão, desigualdades sociais e sofrimento humano. Para citar Santos (2016), o Sul Global trata de traçar uma *linha abissal* que separa aqueles considerados humanos (localizados na zona MetrÓpole, o *Norte*) daqueles considerados sub-humanos (localizados na zona *Colonial*, o *Sul*). É importante ressaltar que o Sul é de natureza conectiva (Medrado e Rega, 2023). O termo pode ser um início de conversa, desencadeando intercâmbios sobre legados coloniais, opressão e marginalização. Transmite o significado de um projeto político de solidariedade⁴, particularmente entre ex-colônias submetidas à violência. Expressa também uma abertura para ouvir o *outro* e uns aos outros como *outros*. Comunica a vontade de transformar as coisas para melhor, de corrigir o que está errado, de unir forças e combater as injustiças.

Passamos à segunda pergunta, o que seria *justiça de dados*? Nas sociedades contemporâneas, a vasta disponibilidade de dados – “como um subproduto do uso de

⁴ Esses insights foram compartilhados em um webinar organizado pela Northwestern University. Yasemin Y. Celikkol, Jessica Winegar, Pablo Boczkowski e Anto Mohsin resumiram alguns dos principais debates que giram em torno do uso do termo. <https://www.qatar.northwestern.edu/news/articles/2022/3-ias-critical-conversations-2.html>



dispositivos e serviços tecnológicos pelas pessoas” (Taylor, 2017, p. 1) – tem implicações na forma como as pessoas são tratadas pelo Estado e pelo setor privado. Portanto, o poder dos dados para produzir categorizações e intervenções sociais no mundo deve estar conectado a uma agenda de justiça social. A justiça de dados pode ser definida como “equidade na maneira como as pessoas são tornadas visíveis, representadas e tratadas como resultado de sua produção de dados digitais” (Ibid.). Além disso, Taylor (2017) propõe três pilares que devem funcionar como base para a justiça de dados: (in)visibilidade, (des)engajamento com a tecnologia e antidiscriminação (Ibid.). O primeiro aspecto trata de representação e privacidade; o segundo, de atenção para compartilhar os benefícios dos dados e autonomia nas escolhas tecnológicas; e o terceiro aspecto inclui a capacidade de desafiar preconceitos.

Vinculada a uma agenda de justiça social, a justiça de dados é um conceito crucial nos estudos *críticos de dados* (ECD – Critical Data Studies em inglês). Kitchin e Lauriault (2014) argumentam que o ECD precisa desvendar os conjuntos sociotécnicos nos quais os dados são produzidos, circulados e usados, incluindo como eles moldam a política e a regulamentação. Poder e política são temas-chave do ECD, que é informado pela teoria social crítica, incluindo feminismo, economia política, pós-colonialismo, CTS etc. O estudo da inteligência artificial (IA) também faz parte do ECD. Não é possível tratar de assuntos ligados à IA sem falar sobre dados. A ligação entre eles é ilustrada nesta definição básica de IA como “programação de computador que aprende e se adapta aos dados” (Verdegem, 2021, p. 5). No contexto da IA, os dados não são apenas uma mercadoria, mas também uma fonte de criação de valor (Sadowski 2019). A *IA estreita* (ou fraca), o tipo de IA de que tratamos neste artigo, funciona analisando conjuntos de dados, identifica padrões e probabilidades e os codifica em um modelo (Broussard 2018). Isso significa que problemas e desafios relacionados aos dados estão presentes na IA e vice-versa.

Ao analisarmos a IA utilizando a PAP, somos inspirados por perspectivas críticas focadas na economia política, no colonialismo e nos estudos raciais críticos. A economia política analisa como as sociedades capitalistas produzem mídia e comunicação como mercadorias. A IA faz parte do *capitalismo de dados* (Sadowski 2019), que se caracteriza pela centralidade da mercantilização e extração de dados, bem como por um cenário industrial concentrado (Verdegem, 2022). Apenas algumas empresas – *GAFAM* (Google/Alphabet, Apple, Facebook/Meta, Amazon e Microsoft) no mundo ocidental e *BAT* (Baidu, Alibaba e Tencent) na China – têm acesso



simultâneo a grandes quantidades de dados e podem investir em uma estrutura computacional de alto desempenho necessária para executar modelos de IA. Porém dados e IA também fazem parte de um projeto de supremacia branca. Observar como o poder corporativo das grandes tecnologias se desenvolve e sustenta as desigualdades é o tema o qual Benjamin (2019) chama de *estudos de códigos críticos raciais*. Estudos como esse reпреndem a afirmação de que a tecnologia é neutra. Noble (2018, p. 4) chama de *opressão algorítmica* o racismo estrutural e o sexismo que acontecem online, que se baseiam em "falhas de dados orientadas por algoritmos que são específicas para pessoas de cor e mulheres". Desenvolvidos no complexo militar-industrial-acadêmico em grande parte branco, os dados são frequentemente capturados por instituições baseadas na supremacia branca, reproduzindo a forma da "branquitude como ideologia" (Katz, 2020, p. 8).

Abraçando essas perspectivas críticas, realizamos quatro oficinas em Londres em junho e julho de 2022 (descritos nas seções a seguir). Com suas hierarquias de produção de conhecimento menos rígidas, as abordagens PAP foram vitais para ampliar a consciência crítica dos participantes (ou *conscientização*, no vocabulário freireano). Além disso, com suas dimensões políticas explícitas, com foco especialmente em questões de assimetrias de poder na IA e na (in)justiça de dados, tal conscientização se manifestou em três eixos principais – *empatia, autonomia e diálogo*. Em um primeiro momento, inspirados pelos princípios da PAP, os participantes começaram a problematizar o que chamaram de interpretação vazia da *empatia*, estabelecendo paralelos com a dinâmica transnacional do capitalismo de dados, que prejudica as comunidades marginalizadas no Sul Global. Em seguida, a PAP ofereceu uma lente crítica para analisar questões de IA e *autonomia* de maneiras menos individualistas, mais coletivas e politicamente engajadas. Em terceiro lugar, o espírito dialógico da PAP permitiu que os participantes localizassem várias interseções entre a IA e o *diálogo*. Ao criticar a ideia de uma IA superior, os participantes lembraram-se das possibilidades oferecidas pela inteligência humana e a combinação de pensar, fazer e sentir ou o que Fals Borda (2003) chama de nossa natureza *sentipensante*.

Revisitando as Raízes Latino-Americanas da Pesquisa-Ação Participativa (PAP)



Em um contexto em que a própria tecnologia digital está por trás de danos sociais como discriminação e racismo (Benjamin, 2019; Noble, 2018), aumentar a participação é uma tentativa de entender as perspectivas e necessidades das comunidades historicamente marginalizadas, contribuindo para uma IA mais inclusiva. No entanto, embora o aumento da participação tenha potencial para a concepção de tecnologia emancipatória, há menos concordância sobre o que seria um nível significativo de participação (Birhane et al, 2022).

Enquanto tentam identificar boas práticas de participação em um estudo sobre IA como um catalisador para o progresso social, Bondi et al. (2021) ressaltam a necessidade de incluir como parceiros iguais membros de comunidades que os sistemas de IA provavelmente impactarão. Os autores apresentam uma lista de perguntas que podem servir como princípios orientadores, tais como: Como podemos entender e incorporar pontos de vista de muitos grupos de partes interessadas? Como as comunidades impactadas podem ser identificadas e representadas? Esse processo inclui grupos marginalizados? Como são abordadas as preocupações da comunidade? Quais são os planos de longo prazo para a manutenção do trabalho (Bondi et al., 2021)? Essas questões oferecem um conjunto valioso de diretrizes para pesquisas socialmente conscientes no campo da IA e, de fato, em qualquer campo de estudo. No entanto, dado o que sabemos sobre a desigualdade entre como os gigantes da tecnologia no Norte Global e as comunidades marginalizadas no Sul Global se beneficiam da IA, essas expectativas de relações iguais são confiáveis? A pensadora indígena Yásnaya Aguilar escreve sobre como o mito ocidental do crescimento perpétuo prejudica aqueles que são mais vulneráveis no sistema capitalista global. Para citar Aguilar (2020), esse crescimento avança por meio de um sistema digestivo que utiliza a tecnologia como um de seus principais componentes para mastigar pessoas marginalizadas, transformando nosso planeta em um lugar inóspito. Não é exagero afirmar que isto representa uma nova ordem colonial.

Qual seria então a nossa tarefa como estudiosos de dados críticos que investigam questões de IA e (in)justiça de dados? Fazer perguntas críticas que se relacionam com as assimetrias de poder entre os chamados Norte e Sul Globais é um bom ponto de partida, mas não é suficiente. Sugerimos que, além de fazer as perguntas certas (críticas), precisamos buscar as maneiras certas (também críticas) de respondê-las. Argumentamos que as abordagens metodológicas e epistemológicas mais adequadas para esse empreendimento são justamente as que decorrem das realidades daqueles



que são excluídos no ecossistema tecno-colonialista global (Madianou, 2019). Aqui, enquanto os entusiastas afirmam que a IA é o futuro e não há como voltar atrás, sugerimos um retorno a um passado radical. Pretendemos fazer isso revisitando as raízes latino-americanas da pesquisa-ação participativa (PAP) e particularmente o trabalho de autores como Paulo Freire (1972) e Orlando Fals Borda (1987).

A partir de uma perspectiva mais ampla do Sul Global, a PAP surgiu em diferentes regiões do mundo como uma resposta às abordagens desenvolvimentistas de cima para baixo. Ao contrário da tradição norte-americana de pesquisa-ação (PA) (Whyte, 1994), estas vertentes da PAP estão ligadas à necessidade de novas metodologias de pesquisa para responder às necessidades dos movimentos sociais do século XX, em particular a reforma agrária e a luta anticolonial. A PAP foi desenvolvida simultaneamente em vários países, como Tanzânia, Índia, Brasil, Chile e Colômbia (Rahman, 2006; Vio Grossi, 1982). Apesar das diferenças contextuais, evoluiu através de processos ascendentes (de baixo para cima) co-desenvolvidos com grupos marginalizados.

Na América Latina, a PAP foi uma tentativa de combinar teoria e prática para enfrentar realidades sociais desiguais. Uma de suas características essenciais foi a rejeição às assimetrias, que muitas vezes integravam paradigmas de dependência e processos de pesquisa. Para a PAP, não há distinções entre sujeitos e objetos de pesquisa – todos os participantes são sujeitos ativos. Quando pesquisas sobre IA participativa, por exemplo, apontam para a necessidade de tratar todos os atores envolvidos como parceiros iguais e igualitários, demonstrando compromisso com o que Freire (1972) chamava de *pedagogia crítica*. O objetivo dessa abordagem é desafiar a forma *bancária* de construção do conhecimento em que os alunos (ou os pesquisados) são considerados *folhas em branco* a serem preenchidas pelos professores (ou pelos pesquisadores). Ao adotar essa abordagem, podemos desafiar as relações de poder em que professores/pesquisadores são considerados conhecedores, enquanto os alunos/pesquisados são considerados sem conhecimento.

Além disso, autores como Fals Borda (2003) têm questionado outras formas de demarcação, nomeadamente as distinções entre fazer, pensar e sentir (ou mãos, cérebros e corações) na construção do conhecimento. O autor desenvolveu ideias em torno de abordagens *sentipensantes* (pensar-sentir). Derivada da combinação das palavras em espanhol *sentimiento* e *pensamiento*, *sentipensante* refere-se a combinar a mente com o coração para guiar a jornada no caminho certo, suportando seus muitos



contratempos (Fals Borda, 2003: 9). Este é um apelo para que uma abordagem diferente seja adotada por pesquisadores, educadores, ativistas, intelectuais: uma abordagem que se opõe diretamente à atitude fria e supostamente neutra do cientista eurocêntrico positivista. Desta forma, o conhecimento não é obtido a partir do fazer, em oposição ao pensar, ou em oposição ao sentir. Em vez disso, todos os conhecimentos são *sentipensantes* e podem ser adquiridos por meio de diálogos respeitosos e empáticos.

Alinhado às perspectivas feministas e decoloniais, a PAP reconhece explicitamente que a construção do conhecimento não ocorre em um vácuo. Na verdade, o conhecimento é situado (Haraway, 1988), exigindo uma prática de posicionamento que atenda às relações de poder em jogo nos processos de produção de conhecimento. Por esse motivo, as *vivências*, ou as experiências vividas, são cruciais para a PAP. *Vivências* podem ser definidas como “experiências plenas de um evento com todas as suas possibilidades”. Em outras palavras, as *vivências* não podem ser observadas (de forma objetiva); “elas só podem ser vividas, sentidas e vivenciadas” (Glassman e Erdem, 2014, p. 212). Quer sejam vivenciadas pelos participantes, pelos pesquisadores ou por ambos (pois as fronteiras entre os dois se confundem), o que importa é que essas *vivências* podem ser transformadoras de maneiras *sentipensantes* (Fals Borda, 2003). Essa natureza transformadora está ligada a outra característica fundamental: o compromisso com a mudança social.

Além disso, segundo Freire (1972), um processo de participação coletiva só pode ocorrer se vier acompanhado de um processo de *conscientização*. O objetivo é que a comunidade supervisione esse processo, sabendo operacionalizá-lo, entendendo-o de forma lógica e avaliando-o criticamente (Mendez e Sanabria, 2003, p. 123). A PAP não é uma abordagem realizada por um especialista. Pelo contrário, requer a formação de equipes que permitam que as pessoas descubram suas capacidades de ação. Os processos de ação-reflexão devem ser participativos e desenvolvidos coletivamente em torno de noções de solidariedade (Mendez e Sanabria, 2003, p. 124).

Aqui, somos referidos ao nosso entendimento do Sul Global como um projeto de solidariedade política. De fato, autores como Fals Borda e Freire são perspicazes porque posicionam explicitamente a pesquisa participante de forma política. Se a pesquisa participante consiste em obter uma pluralidade de vozes e ecoá-las em níveis iguais, a PAP latino-americana abraça o ato de tomar partido. Trata-se de tomar o lado dos colonizados, ou oprimidos, em termos freireanos. Se aplicarmos essas ideias às



preocupações deste artigo – analisar a IA através de um fluxo Sul-Norte - a participação deve partir do reconhecimento de que ignorar padrões de opressão e privilégio sistêmicos contribui para sistemas de IA opacos e injustos. A seguir, discutimos como essas questões foram analisadas em uma série de oficinas (workshops) realizadas em Londres.

3. Metodologia

Inspirados pela PAP latino-americana, organizamos quatro oficinas durante junho e julho de 2022⁵. Os participantes (ver Tabela 1) consistiram em 20 pessoas em quatro categorias principais: 11 estudantes de graduação; 5 estudantes de pós-graduação; 2 profissionais de tecnologia e 2 ativistas de direitos humanos. O grupo foi composto por 11 mulheres e 9 homens (todos cisgêneros) com idades variando de 20 a 43 anos. Obtivemos aprovação do comitê de ética em pesquisa da universidade antes da realização das oficinas.

N	Perfil	Idade	Gênero	Nacionalidade
1	Estudante de pós-graduação e profissional de mídia	43	F	Britânica
2	Estudante de pós-graduação e profissional de mídia	38	F	Polonesa
3	Profissional de tecnologia	34	M	Libanesa
4	Estudante de graduação	20	F	Indiana
5	Ativista tecnológico(a)	36	M	Egípcia
6	Estudante de graduação	21	F	Libanesa
7	Estudante de graduação	22	M	Chinesa
8	Profissional de tecnologia	33	F	Peruana
9	Estudante de pós-graduação	35	F	Chilena
10	Estudante de graduação	20	F	Britânica
11	Estudante de graduação	27	M	Indonésia
12	Estudante de pós-graduação	30	M	Italiana
13	Ativista tecnológico(a)	35	F	Brasileira
14	Estudante de graduação	20	F	Tcheca
15	Estudante de graduação	20	F	Indiana
16	Estudante de graduação	33	M	Chinesa
17	Estudante de pós-graduação	35	M	Britânica

⁵Salientamos que as oficinas ocorreram antes da introdução do *ChatGPT*, o que explica por que o tema não surgiu em nossas discussões.



18	Estudante de graduação	21	F	Chinesa
19	Estudante de graduação	22	M	Búlgara
20	Estudante de graduação	22	M	Chinesa

Tabela 1: Perfis dos participantes da oficina

Como discutido anteriormente, nos inspiramos nos princípios da *vivência*, *conscientização*, *práxis/ação* e reflexão crítica (Freire, 1972; Fals Borda, 1987; Glassman e Erdem, 2014). Ao reunir contribuições dos participantes em nossa primeira reunião, projetamos as oficinas da seguinte forma:

- **Dia 1 – Vivências da IA**

Compartilhamos histórias de IA no cotidiano de trabalhadores invisíveis no Sul Global. Isso desencadeou uma conversa sobre IA, (in)justiça de dados e desigualdade social, à medida em que também apresentamos ao grupo a tradição latino-americana da PAP. Ao final, foi decidido que os participantes documentariam seus próprios *dias com a IA* produzindo um texto, vídeo ou ilustração.

- **Dia 2 – My day in AI (Meu dia com a IA)**

Os participantes retornaram duas semanas depois, tendo produzido uma peça criativa para documentar suas experiências diárias em e com a IA. Com base nas *vivências* da IA, este exercício foi alinhado com pesquisas que enfocam a dataficação tal como é vivida, sentida e vivenciada no nível do cotidiano (Kennedy, 2018, p. 27). Os artefatos foram disponibilizados a todos os participantes para que pudessem se envolver com o conteúdo uns dos outros antes da segunda oficina. Durante esta sessão, os participantes foram questionados sobre quais conteúdos produzidos por outras pessoas lhes despertava empatia e por quê.

- **Dia 3 – Discussão sobre a intervenção à IA**

Os participantes passaram algum tempo articulando problemas em termos de IA que não contribuem para a justiça social, tais como: os inputs da IA vêm de uma fonte predominante (perspectiva ocidental, branca, masculina) (empresas poderosas no Norte Global); os sistemas de recomendação orientados por IA interferem em nossa autonomia; os modelos de negócios não apoiam o diálogo com empatia; temos pouca informação sobre os grandes players/atores da IA e quais são suas agendas. Todos esses problemas, que estavam ligados a questões críticas, foram compartilhados em um Padlet.



• **Dia 4 – Desenho da Intervenção à IA**

Dividimos o grupo maior em dois grupos menores. No brainstorming coletivo, os participantes propuseram duas intervenções para enfrentar os problemas acima mencionados. A primeira era referente às reportagens e representações não críticas da IA pela mídia (seja com exageros ou tropos narrativos distópicos de Hollywood). A solução do grupo seria propor um reality show onde os participantes passariam uma semana com dispositivos de outras pessoas, como tablets ou telefones, interagindo com as recomendações que recebem dos mecanismos de busca e das mídias sociais. Com base nessas recomendações, eles seriam convidados a imaginar como é viver a vida dessa outra pessoa. Durante este exercício, eles também seriam solicitados a ter empatia com essa pessoa e usar sua consciência crítica (*conscientização*) para identificar se (de que maneiras) a pessoa está sujeita à discriminação impulsionada pela IA. A segunda intervenção relacionou-se a questões de políticas de mídia. Os participantes propuseram o desenvolvimento de uma plataforma chamada *Você também ouviu falar ('Have you also heard')*? O objetivo era permitir que as plataformas de mídia social apontassem aos usuários perspectivas diversas sobre o mesmo problema com a lógica de: suas opiniões são... mas você também ouviu...? O grupo acreditava que isso ajudaria a reduzir a radicalização de pontos de vista que pode ser causada pela lógica algorítmica de expor os usuários aos mesmos tipos de conteúdo. A Figura 1 fornece uma visão geral de como os três principais conceitos de PAP – empatia, diálogo e autonomia – emergiram das oficinas.

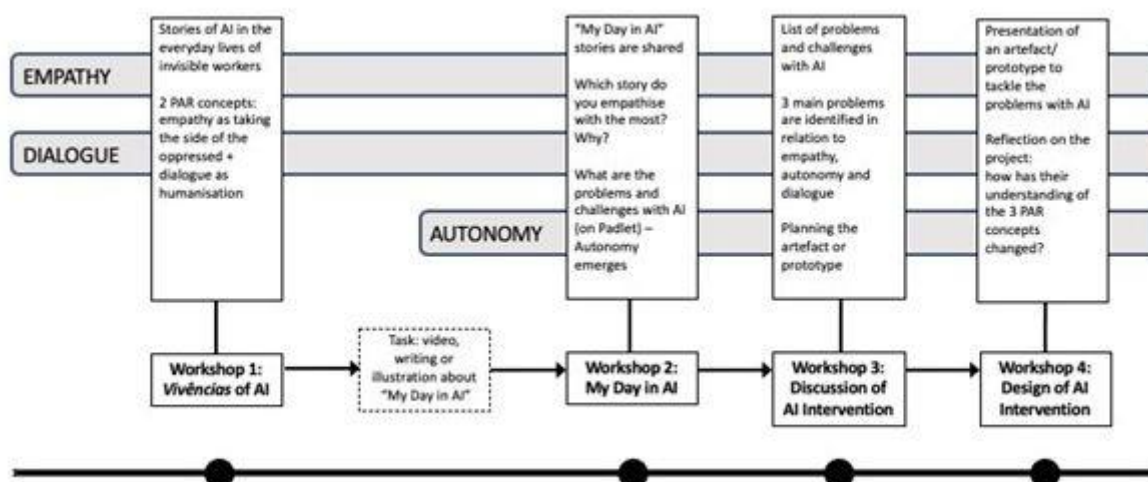


Figura 1: Cronograma Visual das Oficinas



Apresentamos as estruturas das quatro oficinas para demonstrar como nossa abordagem PAP funciona na prática. No entanto, não temos espaço aqui para nos aprofundarmos na viabilidade dessas intervenções propostas. Ao nos concentrarmos nas oficinas como nosso local empírico, argumentamos que, apesar de termos participantes de classe média que não eram tão afetados por questões de marginalização, a PAP nos levou a conversas profundamente infundidas com preocupações em torno da IA e da (in)justiça de dados.

Outra ressalva precisa ser feita. Embora tenhamos sido inspirados pela PAP latino-americano, não conseguimos permanecer tão fiéis ao seu ethos de agenda aberta como gostaríamos. Existem desafios significativos em termos de aplicação da PAP, tal como é entendida por estudiosos latino-americanos radicais (Freire, 1972; Fals Borda, 1987), no contexto do sistema de Ensino Superior do Reino Unido. O projeto foi financiado por uma rodada de financiamento inicial de subsídios de impacto dentro de nossa instituição. Mesmo que indiretamente, a agenda de impacto coloca o pesquisador em uma posição hierárquica mais alta do que os destinatários de sua experiência, pois o impacto consiste em apresentar histórias de estudos de caso bem-sucedidas (Wróblewska, 2021).

Finalmente, como discutimos anteriormente, os participantes da oficina eram em sua maioria estudantes universitários de classe média de diversos países de origem. Nesse caso, indiscutivelmente, ao contrário dos participantes de origens economicamente vulneráveis, havia menos preocupações em termos de relacionamentos potencialmente (economicamente) exploradores, mesmo que se aplicassem os princípios habituais de reciprocidade (humana, humanitária). No entanto, uma questão que emergiu foi: a partir de uma perspectiva da PAP, será que esses participantes ainda seriam capazes de oferecer informações valiosas sobre questões de IA e (in)justiça de dados, uma vez que não são tão afetados por questões de desigualdade e marginalização? Nas próximas seções, demonstramos que nossos participantes estavam claramente lidando com questões de IA, extração e desigualdades, mesmo que não tenham sido inicialmente identificadas como prioridades de pesquisa.

Empatia e IA: Compartilhar sentimentos ou tomar o lado dos oprimidos?



Ao final do primeiro dia de oficina, decidimos documentar nossas *vivências* de IA com um exercício criativo chamado *My Day in AI*. Peças de escrita criativa, vídeos e ilustrações foram produzidas e enviadas para um drive compartilhado. Retornamos para a segunda oficina, duas semanas depois. Este segundo encontro começou com uma reflexão sobre como nos sentimos ao documentar o papel que a IA desempenhou no nosso cotidiano. Em seguida, compartilhamos os escritos, vídeos e ilustrações novamente e perguntamos: ao olhar para esses artefatos, você pode nos dizer algum pelo qual você tenha empatia? Por quê? Em seguida, reproduzimos uma peça de escrita (em estilo diário) com a qual muitos participantes se identificavam.

My Day in AI (Meu dia com a IA)

Uma das coisas que preciso fazer é pagar o depósito do meu quarto, mas por algum motivo não consigo. Preciso de ajuda e dependo do chat ao vivo no portal. Tento obter respostas de um bot, que infelizmente não pode me ajudar. Mas mesmo que o bot pudesse me ajudar, ele realmente não se importa com minhas dificuldades financeiras para arcar com os custos da universidade. Os bots realmente não se importam se estamos endividados. (Trecho da redação do Participante 11, 19/06/2022).

Começamos a discussão questionando o que significava *empatia*. Como esperado, as primeiras respostas foram sobre a empatia como algo que nos permitia nos colocar no lugar dos outros. A participante 6, uma estudante de graduação, acrescentou que a empatia também exigia níveis de previsão e imaginação: “trata-se de ter a capacidade de entender e compartilhar os sentimentos de outra pessoa, e isso pode acontecer imaginando-se na situação dessa outra pessoa” (Notas da oficina 2, 23/06/2022). Então, a probabilidade de nos encontrarmos em uma situação semelhante à descrita pelo Participante 11 explicaria por que nos identificamos com o que escreveu? A participante 2, uma estudante de pós-graduação com especialização em relações públicas, observou que a maioria das pessoas podia se identificar com essa experiência pois as deixava desconfortáveis: “Muito na vida do participante 11 pode depender de uma decisão tomada por um bot que, obviamente, é incapaz de se importar”. (Notas da oficina 2, 23/06/2022).

A empatia foi um conceito que emergiu organicamente das atividades da oficina. Compartilhar nossos artefatos criativos significava que poderíamos nos conectar em nossas experiências diárias de/com a IA de forma empática. Além disso, os participantes também foram capazes de identificar pontos de interseção entre IA e empatia. Isto foi articulado pelo Participante 3, um profissional de tecnologia:



Hoje ouvimos com frequência que existem muitos problemas com o mundo porque não temos empatia, então a empatia é tida como uma solução para os males sociais. E ouvimos muitas coisas semelhantes sobre IA. Então, se é esperado que nossa vida se torne muito melhor por causa da IA, devemos agora esperar que esses sistemas de alguma forma tenham uma capacidade de empatia construída dentro deles? (Notas da oficina 2, 23/06/2022).

As observações do Participante 3 desencadearam um debate com os participantes que levantou a seguinte pergunta: que tipos de expectativas temos em relação à IA e aos sistemas baseados em dados? Se não for hoje, haverá um tipo de IA empática no futuro? Como a IA e a empatia podem ser vistas à luz do capitalismo de dados? Aqui, os participantes também identificaram um *vazio* em ambos os discursos de empatia e IA. Começamos pelo primeiro. A participante 8, uma profissional de tecnologia, observou que “a empatia, assim como muitas outras palavras, como descolonização, foi apropriada e se tornou uma espécie de marketing ou sinalização virtual. E realmente não significa muito, então está totalmente vazia” (Notas da oficina 2, 23/06/2022). Este ponto ecoa perspectivas críticas sobre empatia que foram levantadas em estudos decoloniais e feministas. Em uma crítica à comunicação no campo humanitário, Chouliaraki (2013, p. 15) observa como a empatia e a solidariedade são encenadas como “projetos individualistas de valores contingentes e ativismos consumistas”. Como nossos participantes disseram, esse tipo de empatia vazia tem menos a ver com política e mais com recompensar a si mesmo.

Dentro da mesma vertente da literatura crítica, Pedwell (2016, p. 2) questiona: “como podemos entender as ligações complexas entre empatia e relações transnacionais de poder?”. Por *transnacional*, o autor se refere aos “processos inter-relacionados e mutáveis de colonialismo, escravidão, diáspora, migração, desenvolvimento, globalização, neoliberalismo e mídia global” (Ibid). Ao colocá-la em um ecossistema transnacional, Pedwell (2016: 13) demonstra que a empatia representa uma tecnologia eficaz para maximizar a competitividade econômica dentro dos circuitos do capital. Ela acrescenta que as mobilizações de empatia dependem de distinções de gênero, raça, classe e geopolíticas. Tais distinções também determinam quem pode capitalizar a empatia como um modo de capital neoliberal e quem se limita a realizar trabalho emocional não reconhecido (Ibid). Simplificando, quando a empatia é entendida como política transnacional, seus efeitos desiguais são revelados, sendo ligados a distinções e exclusões geopolíticas.



Aqui, podemos estabelecer vínculos com discursos vazios de IA compartilhados pelas empresas que dominam o que é chamado de *capitalismo de dados* (Sadowski, 2019). O panorama industrial da IA caracteriza-se pela centralidade da mercantilização e extração de dados, bem como pela concentração da capacidade computacional e do talento da IA (Verdegem, 2022). Apenas algumas grandes empresas de tecnologia têm acesso simultâneo a grandes quantidades de dados e podem investir em uma estrutura computacional de alto desempenho necessária para executar modelos de IA. Em suma, como *motor principal* da acumulação de capital (Brevini, 2022), a IA é usada para aumentar a acumulação de capital no Norte Global, ao mesmo tempo que agrava as desigualdades no Sul Global. Assim, as explícitas exclusões geopolíticas que Pedwell (2016) critica em compreensões vazias de empatia emergem em relação à dinâmica transnacional do capitalismo de dados. Com esse enquadramento crítico em mente, começamos a discutir outro resultado criativo do *My Day in AI*, uma ilustração acompanhada de um comentário produzido pela Participante 2, uma estudante de pós-graduação.



Figura 2: Ilustração intitulada *Um processo não linear de devir contínuo*, produzida pela Participante 2, 18/06/2022)

Apenas alguns indivíduos criam softwares de IA, mas bilhões a utilizam diariamente. Não se extrai a consciência, mas ela pode ser compartilhada e utilizada como Inteligência Aumentada. A simbiose entre humanos e máquinas se fortalece. Quem se beneficiará disso? Os poucos ou os bilhões de nós? (Trecho do comentário da Participante 2 sobre sua ilustração, 18/06/2022).



A escrita da participante 2 levou os participantes a se concentrarem na questão sobre quem exatamente se beneficia da IA. Isto demonstra como a tradição latino-americana de PAP inspirou os participantes a adotar uma lente crítica para investigar questões de IA e (in)justiça de dados. Aprender mais sobre as *vivências* de IA uns dos outros (com o exercício *My Day in AI*) forneceu aos participantes as informações necessárias para um exercício de empatia. À medida que o debate evoluiu, os participantes passaram a problematizar uma interpretação vazia e um tanto apolítica da empatia, estabelecendo paralelos com questões de IA e desigualdade social em escala global. Assim, a PAP permitiu que os participantes se guiassem através de um processo de *conscientização* – de uma compreensão de sinalização apolítica e virtual a uma compreensão da empatia que está atenta às dinâmicas de poder transnacionais, desde tomar como garantidos os benefícios sociais da IA até questionar o desequilíbrio em termos de quem lucra com isso (grandes empresas de tecnologia no Norte Global e não grupos marginalizados no Sul Global). Assim, utilizar a empatia como ferramenta para interagir com as *vivências* uns dos outros nos aproximou das interpretações decoloniais e feministas do conceito (Chouliaraki, 2013; Pedwell, 2016).

De forma complementar, a empatia adquiriu níveis mais profundos do que simplesmente conectar-se aos sentimentos dos outros; em vez disso, tornou-se associada à disposição de ouvir os oprimidos (Freire, 1972). A empatia surgiu como ponto central para alcançar uma consciência crítica (*conscientização*), incluindo uma compreensão das contradições sociais, políticas e econômicas. Em outras palavras, a empatia poderia ser vista como uma práxis política que permitia que as pessoas se identificassem com grupos oprimidos, não apenas para encontrar coisas em comum, mas sim para transformar as realidades que os oprimiam (Manyozo, 2022, p. 43). A seguir, analisamos como a empatia funcionou como uma lente socialmente engajada para olhar para outra questão crítica, a da IA e da autonomia.

5. De uma autonomia individualista a um sentido coletivo de *Autonomia*

Na segunda oficina, os participantes revisitaram outra parte da escrita do Participante 12. Este estudante de pós-graduação estava trabalhando em diferentes shows em Londres para obter recursos financeiros antes de retornar para a Itália. A redação descreveu a seguinte experiência com algoritmos de IA:

Uma nova notificação apareceu no meu iPhone. Eram mensagens de texto do LinkedIn: ‘ei, você conhece ...?’. Fiquei surpreso porque a



peessoa era uma ex-namorada. Como o aplicativo propôs o contato dela se não temos conexão direta no LinkedIn? Mais tarde naquela manhã, recebi um telefonema da minha mãe. Disse a ela que no sábado estaria ocupado me preparando para o show do Harry Styles. Ela não tem ideia de quem ele é e percebi que também não sei muito sobre ele. Depois do almoço, os algoritmos do TikTok começaram a me oferecer uma quantidade - antes inexistente - de vídeos do Harry Styles. (Notas do Participante 12 em My Day in AI, 20/06/2022).

Muitos participantes não perceberam que essas recomendações algorítmicas, os sistemas de IA que conhecemos e usamos todos os dias, exemplificam o que chamamos de *IA estreita* (ou fraca). A *IA estreita* é distinta da *AGI*, inteligência artificial geral ou *IA forte*. A *AGI* é o que Broussard (2018, p. 32) chama de “o tipo de IA de Hollywood”, referindo-se à (questionável) promessa de que as máquinas têm habilidades cognitivas que poderiam (potencialmente) superar as capacidades intelectuais humanas. Em contraste, a *IA estreita* é o tipo de IA que muitas vezes passa despercebida em nossas vidas diárias. O que ela faz é identificar padrões e probabilidades em grandes conjuntos de dados e codificar esses padrões em um modelo computacional. Embora possam ser poderosos, esses sistemas são limitados na gama de tarefas que podem executar.

À medida em que os participantes se identificavam com a experiência do Participante 12 de receber conteúdos forçados, o foco estava em quão falhos e invasivos eram esses sistemas. Afinal, “o sistema de recomendação não poderia realmente dizer que o Participante 12 não estava interessado em Harry Styles” (Participante 18, notas da oficina 2, 23/06/2022). Outra participante disse que se sentiu incomodada pelo depoimento do Participante 12 porque “a IA está dizendo ao Participante 12 com quem se reconectar, como se informar melhor para uma conversa com sua mãe...”. Ela então perguntou: “isso significa que não temos mais autonomia sobre nossas interações, o que sabemos e o que dizemos?” (Participante 2, Oficina 2, 23/06/2022).

Inicialmente, as interpretações de autonomia dos grupos pareciam girar em torno de uma noção individualista, que é comum nas compreensões do termo no Ocidente/Norte Global. Os participantes pareciam incomodados com o que percebiam como tecnologia reduzindo o poder das decisões individuais. De acordo com Floridi et al (2018, p. 698), no contexto da IA, a conceituação de autonomia está relacionada com as decisões das pessoas de adotar a IA e “ceder voluntariamente parte de seu poder de decisão às máquinas”. A autonomia – na ética da IA – trata, portanto, de como salvaguardar a autonomia dos seres humanos e manter seu controle sobre máquinas/sistemas de computador. Isso nos leva a uma discussão fundamental nos



círculos de IA, e o chamado *problema de controle*, ou a necessidade de manter os seres humanos no controle *sobre* as máquinas e sistemas de IA (Russell, 2019). Da mesma forma, nossos participantes da oficina criticaram os sistemas de recomendação por sua insubordinação às nossas verdadeiras necessidades e desejos (humanos). "Esses sistemas de recomendação", eles perguntaram, "têm muito controle sobre o que dizemos, fazemos e sentimos?" (Notas da oficina 2, 23/06/2022).

O Participante 12 complexificou ainda mais o debate ao vincular recomendações automatizadas a sistemas automatizados de tomada de decisão. Em resposta às críticas dos membros do grupo sobre o que consideraram “recomendações estúpidas que obviamente não refletem o que gostamos e queremos”, ele observou que “pelo menos elas são apenas irritantes, mas não necessariamente prejudiciais” (Notas da oficina 2, 23/06/2022). “*As coisas ficam mais sérias*”, acrescentou, “quando os sistemas começam a tomar decisões por nós. Eu moro em uma aldeia remota na Itália e, quando faço pedidos online, esses sistemas decidiram que meu endereço não funciona. Sou um cidadão de menor valor porque minha família mora em uma aldeia?” (Participante 12, Notas da oficina 2, 23/06/2022).

Aqui, novamente, a PAP contribuiu para o desenvolvimento desses debates de forma consistente com um processo de *conscientização* (Freire, 1972). Ao falar sobre como os sistemas classificavam alguns endereços como reconhecíveis e, portanto, merecedores de entregas em domicílio, enquanto outros endereços e pessoas não eram merecedores, os participantes passaram a abordar dilemas éticos sobre IA e autonomia. Como perguntamos em relação à empatia, o que realmente significa autonomia na perspectiva do oprimido, usando um termo freireano?

O participante 5, um ativista de tecnologia, levantou a questão dos dados e das políticas de segurança preditiva. Ele lembrou ao grupo que a prevenção do crime era muitas vezes uma justificativa para decisões discriminatórias baseadas em dados, visando pessoas de cor e cidadãos da classe trabalhadora (Participante 5, notas da oficina 2, 23/06/2022). Autoras como Ricaurte (2022) escreveram sobre políticas preditivas. Tais políticas, argumenta ela, são problemáticas porque significam que as tecnologias de IA são utilizadas geopoliticamente como parte de um complexo industrial militar e como uma ferramenta estratégica para a vigilância dos pobres globais.

Ao discutir tais exemplos, os participantes perceberam que a autonomia pode ter diferentes significados e implicações, dependendo se estamos falando de um



contexto privilegiado do Norte ou de um contexto marginalizado do Sul. Eles também concluíram que as implicações da perda de autonomia podem adquirir dimensões mais drásticas a partir de uma perspectiva do Sul Global. Incluindo graves violações dos direitos humanos. Esse foi o caso dos refugiados *rohingya* em Bangladesh, que tiveram seus dados biométricos compartilhados de maneiras que poderiam colocá-los ainda mais em perigo (Madianou, 2021). Assim, combinar a PAP com a nossa reflexão nas oficinas revela como a IA pode fazer recomendações sobre com quem você interage, o que você vê, o que você fala e como, mas também quem pode ser percebido como humano, quem pode ser livre ou encarcerado, quem pode viver ou morrer.

Além disso, podemos inferir que a PAP ofereceu uma lente crítica para analisar questões de IA e autonomia de maneiras menos individualistas e mais coletivas e politicamente engajadas, tomando o lado dos oprimidos, como foi o caso das interpretações freireanas e feministas decoloniais da empatia (Freire, 1972; Chouliaraki, 2013; Pedwell, 2016). Nessa linha, Couldry e Mejias (2019, p. 155) definem a *autonomia* não como um autogoverno, mas sim como uma integridade socialmente fundamentada sem a qual não podemos reconhecer a nós mesmos ou aos outros como indivíduos. Segundo os autores, “a autonomia é crucialmente entendida como algo que não podemos negociar, uma integridade mínima da vida humana – é disso que desistimos quando desistimos da autonomia” (Ibid).

Assim, somos lembrados de como é importante resgatar o conceito de integridade mínima da vida humana como autonomia. Quando uma pessoa é rotulada como um futuro criminoso devido à sua infância por meio de um sistema de decisão algorítmica, tal integridade é severamente desrespeitada e, portanto, não há autonomia. Autores como Arturo Escobar (2018, p. 172) acrescentam que a autonomia adquire uma camada adicional de densidade política. Usando a palavra em sua forma espanhola, *autonomia* passa a ser uma “forma completamente diferente de regra ancorada na vida das pessoas, uma luta pela libertação e por um novo tipo de sociedade em harmonia com outras pessoas e culturas”. Além disso, as abordagens da PAP em nossas oficinas nos levaram a uma compreensão da *autonomia* como um horizonte político que orienta a prática política baseada na interexistência e no interser.

Durante nossas discussões sobre empatia e autonomia na segunda oficina, percebemos que nenhum desses conceitos poderia ser analisado isoladamente. Em vez disso, todos eles estavam interconectados de maneiras complexas e precisavam ser mapeados. Decidimos então gerar um esquema visual em um quadro Padlet, criando



colunas que identificassem *temas/problemas*, *questões comuns* e *perguntas* em relação à IA e à (in)justiça de dados. Cada participante foi solicitado a adicionar suas contribuições em um Padlet compartilhado (veja abaixo) em seu próprio tempo.

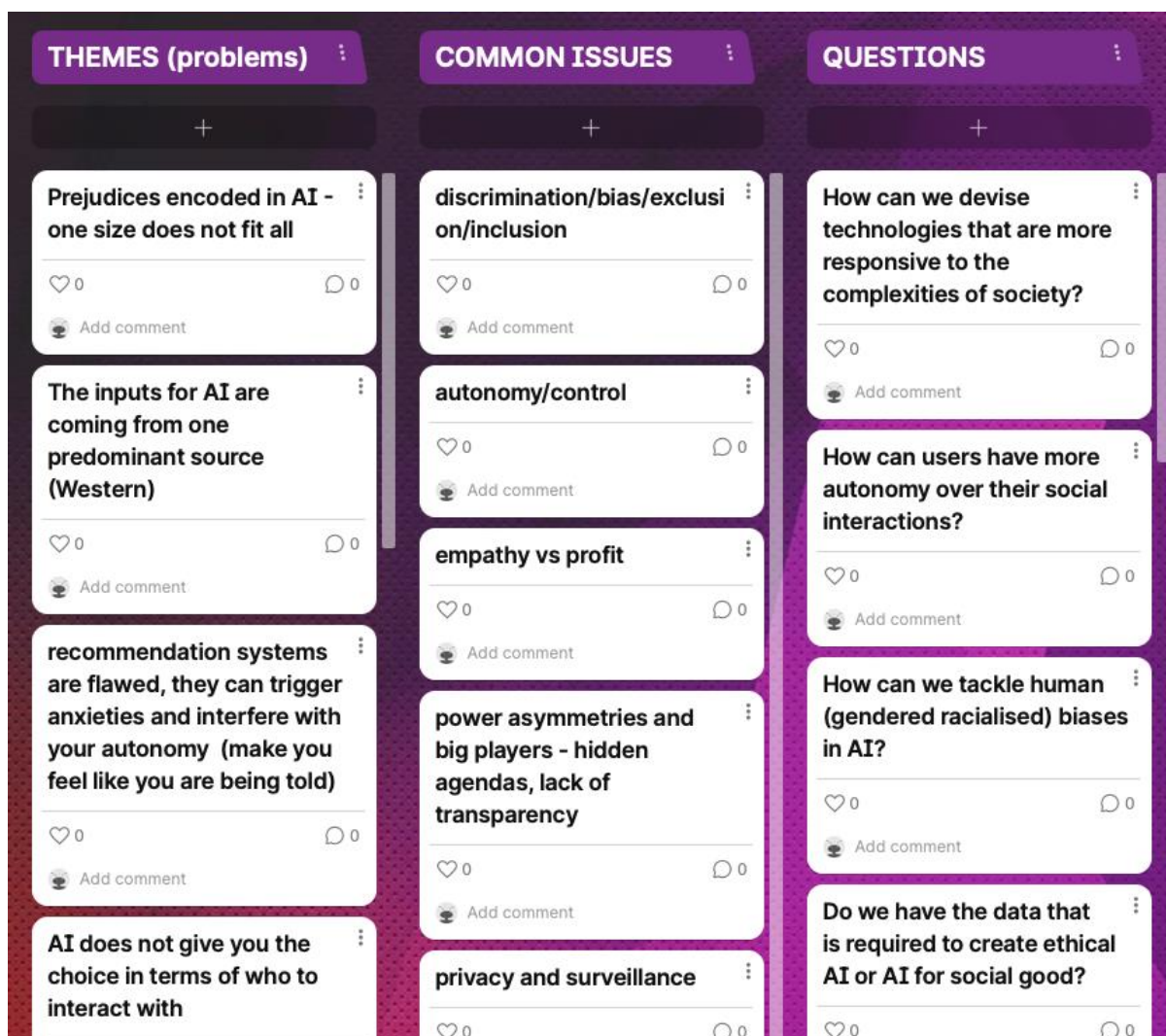


Figura 3: Captura de tela de um Padlet criado pelos participantes da oficina para identificar problemas, questões comuns e perguntas em relação à IA e à (in)justiça de dados (primeira metade da tela)⁶

⁶ Optamos por manter a imagem original (em inglês), em função de ser uma captura de tela. Além disso, é feita uma discussão no texto que permite sanar quaisquer dúvidas em relação às tabelas incluídas no artigo (*nota dos Editores*).

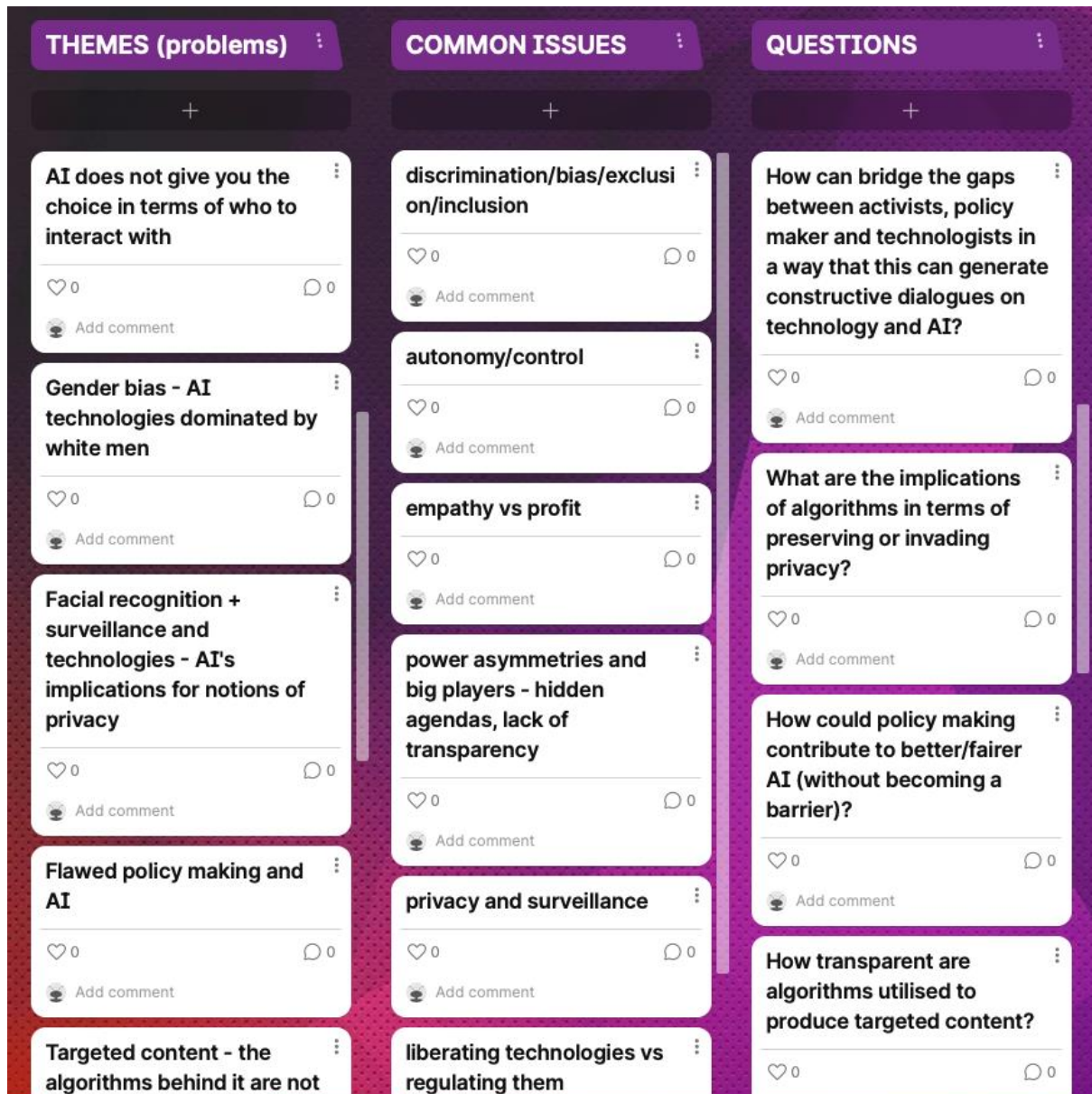


Figura 4: Captura de tela de um Padlet criado pelos participantes da oficina para identificar problemas, questões comuns e perguntas em relação à IA e à (in)justiça de dados (segunda metade da tela)

Em seguida, revisitamos o Padlet e nossas discussões sobre autonomia na terceira oficina. Fomos confrontados com uma série de problemas, questões e perguntas inter-relacionadas, tais como: (1) “os preconceitos são codificados na IA, um tamanho não serve para todos” (coluna 1 na Figura 3), o que leva a “discriminação, preconceito, exclusão/inclusão” (coluna 2 na Figura 3) e à pergunta “como podemos conceber tecnologias que respondam melhor às complexidades da sociedade?” (coluna 3 na Figura 3); (2) “os sistemas de recomendação são falhos, podem desencadear ansiedades e interferir em sua autonomia (fazer você se sentir como se estivesse



recebendo ordens)” (coluna 1 na Figura 3), que está ligado à “questão da autonomia/controle” (coluna 2 na Figura 3) e à pergunta “como os usuários podem ter mais autonomia sobre suas interações sociais” (coluna 3 na Figura 3) e (3) “preconceito de gênero – tecnologias de IA dominadas por homens brancos” (coluna 1 na Figura 4), ligada à questão de “assimetrias de poder e grandes atores – agendas ocultas, falta de transparência” (coluna 2 na Figura 4) e à pergunta “como podemos preencher as lacunas entre ativistas, formuladores de políticas e tecnólogos de forma que isso possa gerar diálogos construtivos sobre tecnologia e IA” (coluna 2 na Figura 4)? As perguntas, particularmente a última, foram escolhidas para serem abordadas por meio de uma intervenção criativa em nossa quarta e última oficina.

Diálogo: Desafiando hierarquias de inteligência (artificial)

Uma semana após a segunda oficina, estávamos de volta à sala da universidade – todos olhando para o Padlet projetado em uma tela e coçando a cabeça. O participante 3, um profissional de tecnologia, quebrou o silêncio:

Sim, eu me pergunto se essa ponte de diálogo é de fato possível”. Eu trabalho na área de tecnologia: Os modelos de negócios que temos em tecnologia não foram construídos para apoiar o diálogo. Há muito mais lucro a ser obtido com a polarização do que com o diálogo. (Participante 3, notas da Oficina 3, 30/06/2022).

Antes de voltarmos nossa atenção para as questões que giram em torno da IA e do diálogo, há alguns pontos críticos a serem feitos sobre a conceituação de diálogo de Freire (Freire e Macedo, 1995) e como eles são fundamentais para as abordagens PAP. Os diálogos são multicamadas e as jornadas de diálogo são compostas por várias etapas. A humanização representa uma dessas etapas. Não se trata apenas de reposicionar os oprimidos em melhores condições. Trata-se de redesenhar a relação *oprimido-opressor* e abordar a assimetria de poder dessa relação (Suzina e Tufte, 2020). Freire também destacou pontos essenciais sobre a *humildade* como um caminho na jornada do diálogo. De fato, a humildade aparece como um requisito para reconhecer que as pessoas – quaisquer pessoas – têm conhecimento. A humildade também sugere que a verdade autêntica e, de fato, a *inteligência* autêntica, um termo central neste artigo sobre IA e (in)justiça de dados, não residem em nenhum indivíduo ou grupo, nem podem ser impostas a um grupo por outro.



Nossos participantes foram inspirados pelos princípios dialógicos da PAP em um sentido que começaram a questionar: se todos nós devemos nos beneficiar da IA, coletivamente como sociedade, por que as comunidades marginalizadas no Sul Global oferecem tão pouca contribuição quando se trata de desenvolvimentos de IA? Neste contexto, nós (os organizadores) mencionamos processos de desenvolvimento de diretrizes éticas de IA. Referindo-nos à literatura neste campo, observamos como o desenvolvimento da ética em IA está concentrado na América do Norte, na UE, no Japão e em um punhado de outros países do Norte Global (Jobin, Ienca e Vayena, 2019). Essa sub-representação de grandes partes do mundo – África, América Latina, Sudeste Asiático, entre outras áreas – significa que o Sul Global está excluído de ter voz e contribuir para esse debate. Além disso, é preciso questionar quem exatamente está envolvido nas iniciativas que desenvolvem diretrizes éticas e se elas são representativas da sociedade. A participante 8, uma profissional de tecnologia, acrescentou ao debate: “é fácil identificar a falta de diversidade. Tudo o que precisamos fazer é analisar a participação de painéis de especialistas e grupos de trabalho e como eles são dominados por especialistas brancos do sexo masculino” (Participante 8, notas da Oficina 3, 30/06/2022). De fato, os participantes escreveram em nosso Padlet compartilhado que “as tecnologias de IA são dominadas por homens brancos” e isso foi identificado como um dos problemas fundamentais da IA.

Assim como aconteceu com a IA e a empatia, os participantes começaram a estabelecer paralelos críticos entre a IA e o diálogo. Primeiro, os participantes problematizaram a ideia de uma IA abrangente, onipresente e onisciente, bem como a premissa de que “os sistemas de IA corresponderam ou excederam as capacidades do pensamento humano” (Katz, 2020, pp. 94-95). Como discutimos anteriormente em relação à AGI, isso se baseia na promessa questionável de que as máquinas podem fazer, pensar e até sentir coisas, não apenas como seres humanos, mas também de formas superiores às capacidades humanas. A participante 8, uma profissional de tecnologia do Peru que estava familiarizada com a PAP e o trabalho de Freire, fez a observação precisa de que “essa ideia está em oposição direta aos princípios de humildade, diálogo”. Ela também perguntou: “para começar, por que a indústria de tecnologia está tão obcecada em classificar a inteligência?” (Participante 8, Notas da oficina 3, 30/06/2022).



A pergunta da participante 8 referiu-se a um tropo comum nos discursos de IA: uma preocupação com a categorização de diferentes tipos de inteligência artificial em termos de inovações tecnológicas. A literatura sobre IA é útil aqui. Huang, Rust e Maksimovic (2019), por exemplo, distinguem entre sistemas de IA que são *mecanicamente inteligentes* e projetados para realizar tarefas repetitivas; sistemas que estão *pensando* de forma inteligente e que são projetados para aprender e se adaptar de forma autônoma a partir de dados; e sistemas de IA que podem adquirir uma inteligência de *sentimentos* e que são projetados para interagir empaticamente com as pessoas (2019, p. 45). Nossos participantes reconheceram essas diferenças, mostrando preocupações sobre as máquinas substituírem os humanos em um nível emocional (o que significa que as máquinas poderiam então superar totalmente os humanos) ou mostrando um ceticismo em relação às máquinas inteligentes.

Quando discutimos a escrita do Participante 11 sobre bots que não se importavam, outros participantes foram rápidos em dizer coisas como: “vamos lá, a IA simplesmente não pode ter empatia” (Notas da Oficina 2, 23/06/2022). Curiosamente, essa discussão sobre distinções de inteligência levou alguns participantes a refletirem sobre o que nos torna humanos. O que nos torna humanos, eles disseram, é precisamente o fato de que pensar, criar e fazer não são mutuamente exclusivos. Em outras palavras, nos humanos, cérebros, mãos e corações podem coexistir. “Por que continuamos nos preocupando tanto com máquinas super inteligentes em vez de nos concentrarmos nas possibilidades do talento humano? Podemos sentir enquanto pensamos”, disse a participante 20, uma estudante de graduação. Estas observações nos lembram da importância das abordagens *sentipensantes* de Fals Borda (2003) para questionar a demarcação rígida e os binarismos do Ocidente/Norte Global.

Aqui, o grupo também nos ofereceu o que chamaram de perspectiva pragmática. Quando questionado sobre as implicações problemáticas da exclusão das comunidades marginalizadas do Sul Global, o Participante 19, um estudante de graduação, disse sem rodeios que “incluir a todos é simplesmente idealista demais” (Notas da Oficina 3, 30/06/2022). Outra observação relacionada à inevitabilidade da IA. “A IA veio para ficar”, disse o Participante 18, outro estudante de graduação, “então é melhor encontrarmos algo de bom nela” (Notas da Oficina 3, 30/06/2022). Nós, os organizadores, perguntamos: “devemos apenas aceitar o fato de que a ascensão da IA é inevitável, mesmo que isso ocorra às custas da exclusão de grandes quantidades da população global (e mesmo que grandes exclusões sejam inerentemente não propícias



ao diálogo)? A resposta foi quase unânime, com os participantes dizendo que a IA é de fato inevitável, exigindo uma atitude pragmática de apenas aproveitar “o melhor destas tecnologias para que elas possam beneficiar o maior número possível de pessoas”, como disse o Participante 18 (notas da Oficina 3, 30/06/2022).

Aqui, pode-se argumentar que a aceitação da inevitabilidade da IA como uma tecnologia de exclusão não é consistente com as perspectivas da PAP sobre: (1) *empatia*, no sentido de tomar o lado daqueles que são oprimidos (Freire, 1972), (2) *autonomia*, como a luta por um novo e mais justo horizonte social (Escobar, 2018); e (3) *diálogo*, como a criação de relações horizontais em que todas as realidades importam (Freire e Macedo, 1995). Ao mesmo tempo, os diálogos empáticos representam processos complexos e contínuos. Eles não envolvem apenas acordo, mas também conflito, negociação. (Pedwell, 2016, p. 20). De qualquer forma, em nossas oficinas, a PAP latino-americana funcionou como uma ferramenta para compartilhar reflexões, fazer perguntas críticas e imaginar realidades alternativas.

7. Considerações finais

Neste artigo, nos concentramos na IA e em algumas de suas implicações problemáticas para a (in)justiça de dados, colocando-a no campo emergente de estudos críticos de dados (ECD). Ao abraçar um fluxo *Sul-Norte*, pretendemos inspirar alternativas às perspectivas epistemológicas e metodológicas que decorrem do Norte Global. Na introdução, perguntamos como a PAP, uma abordagem radical desenvolvida no passado, poderia nos inspirar ao abordar a IA e a (in)justiça de dados no presente e no futuro. Também questionamos como uma abordagem focada em fortalecer os oprimidos pode ser aplicada a pessoas e comunidades que não experimentaram tanta marginalização social. Depois de conduzir e analisar os dados das oficinas, descobrimos que se pode aprender muito com a aplicação da PAP no Norte Global, em contextos onde as pessoas enfrentam diferentes lutas e desigualdades.

Compartilhamos uma jornada na qual a tradição latino-americana de PAP nos inspirou de várias maneiras a aprofundar questões de IA e (in)justiça de dados. A primeira inspiração veio do foco nas *vivências*, facilitando o compartilhamento das experiências vividas com a IA. Esperamos ter demonstrado como os limites fluidos da PAP entre sujeitos e objetos, pesquisadores e pesquisados, professores e alunos, conduziram a um ambiente de perguntas críticas que são perspicazes para o ECD. Aqui,



os processos menos hierárquicos de construção de conhecimento coletivo da PAP contribuíram para tornar os participantes mais atentos às dinâmicas de poder. Após quatro dias de oficinas de *fazer* (artefatos criativos), *pensar* e *sentir* sobre IA e (in)justiça de dados de maneira *sentipensante* (Fals Borda, 2003), três principais estruturas habilitadas para PAP surgiram para entender questões críticas que giram em torno da IA: *empatia*, *autonomia* e *diálogo*.

Examinar esses três elementos de formas politicamente e socialmente comprometidas, em vez de *vazias*, emergiu como uma prioridade de pesquisa. Aqui, a *empatia* tornou-se mais alinhada com suas interpretações decoloniais e feministas (Chouliaraki, 2013; Pedwell, 2016), bem como com a noção freireana (1972) de que a empatia deveria significar tomar o lado dos oprimidos. Refletindo sobre *autonomia*, os debates também mudaram de direção – de preocupações sobre sistemas de recomendação e controle individual sobre máquinas para tomada de decisão automatizada e suas implicações para quem pode ser percebido como humano, quem pode ser livre ou encarcerado, quem pode viver ou morrer. Em suma, a *autonomia* passou a ser compreendida menos como um fenômeno individualista e mais como uma luta coletiva por realidades sociais mais justas (Escobar, 2018). Por fim, os participantes perceberam que humanizar todos os atores - colocando-os todos em igualdade de condições - era um pré-requisito para o *diálogo*. Também constituía uma forma de *inteligência humanizadora*, que se opunha à ideia de *inteligência artificial superior*.

Esses três elementos – *autonomia*, *empatia* e *diálogo* – estão profundamente entrelaçados, um contribuindo para o outro. Do ponto de vista do Sul Global, para grupos marginalizados, a *autonomia* tem a ver com autodeterminação ou a capacidade de ter mais voz sobre os destinos da própria vida. Esta possibilidade é algo que está no centro das lutas dos oprimidos. Assim, se almejamos a mudança social, devemos mostrar empatia em um sentido de solidariedade política. Além disso, ao nos envolvermos em diálogos abertos e horizontais, podemos humanizar e ter empatia com o outro.

Embora a PAP latino-americana nos oferecesse lições valiosas, nossa abordagem encontrou limitações. As restrições na aplicação da PAP estão relacionadas à forma como a maior parte das bolsas de estudos é organizada no Norte Global. Muitas vezes, isso deve partir de objetivos claros para gerar impacto máximo, o que se opõe à agenda aberta e à abordagem ascendente da PAP. Embora a geração de impacto tenha



como objetivo gerar um impacto positivo nas partes interessadas não acadêmicas, a retórica da cocriação e do desenvolvimento é muitas vezes vazia, pois é um desafio envolver as comunidades marginalizadas no processo de pesquisa (pois elas precisam priorizar seu tempo e energia). Concordamos com o apelo de Lehuedé (*a ser lançado*) por uma reflexividade crítica radical, para evitar uma dupla hélice de extração: extração de dados pela indústria de tecnologia e extração de conhecimento por pesquisadores e suas instituições.

Por fim, nos inspiramos na PAP latino-americana, situando-a dentro das abordagens do Sul Global, reconhecendo a força dos diálogos Sul-Norte e Sul-Sul como ferramentas para potencializar projetos de solidariedade política entre os oprimidos (Medrado e Rega, 2023). Não reivindicamos uma espécie de universalismo latino-americano proveniente do Sul. Em vez disso, nosso objetivo é chamar a atenção para a necessidade de tratar os *Suis* como fontes de inspiração em humildes tentativas de aprender uns com os outros em ECD. Uma direção futura seria estabelecer diálogos com os meios de comunicação e outros parceiros em todo o Sul e Norte Global para analisar a IA e a (in)justiça de dados. Poderá esta abordagem dialógica (inspirada na América Latina) mudar nossa atenção dos mitos da IA sobre inteligência superior e resolução de problemas de cima para baixo, para fazer mais perguntas sobre IA, *autonomia, empatia e diálogo*? Sim, é verdade, essas questões são complexas. No entanto, embora nem sempre possamos lhes fazer justiça, isto não deve nos impedir de sermos guiados por uma preocupação com a justiça (de dados) em nossas jornadas compartilhadas de construção de conhecimento.

Referências

AGUILAR GIL, Yasnaya. Elena. A modest proposal to save the world. **The Rest of World**, 2020. Disponível em <https://restofworld.org/2020/saving-the-world-through-tequiology>. Acesso em 25 maio 2024.

AOURAGH, M.; CHAKRAVARTTY, P. Infrastructures of empire: Towards a critical geopolitics of media and information studies. **Media, Culture & Society**, v. 38, n. 4, p. 559-575, 2016.

BIRHANE, A. et al. **Power to the people?** Opportunities and challenges for participatory AI. 2022. DOI: <https://doi.org/10.1145/3551624.3555290>.

BREVINI, B. **Is AI good for the planet?** Cambridge: Polity, 2022.



- BROUSSARD, M. **Artificial unintelligence**: How computers misunderstand the world. Cambridge, MA: MIT Press, 2018.
- BONDI, E.; XU, L.; ACOSTA-NAVAS, D.; KILLIAN, J. Envisioning communities: A participatory approach towards AI for social good. In: **Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society (AIES '21)**, May 19-21, 2021, Virtual Event, USA. ACM, New York, NY, USA, 12 pages. DOI: <https://doi.org/10.1145/3461702.3462612>. Acesso em 25 maio 2024.
- CHOULIARAKI, L. **The ironic spectator**: Solidarity in the age of post-humanitarianism. Cambridge/Malden: Polity Press, 2013.
- COULDRY, N.; MEJIAS, U. Data colonialism: Rethinking big data's relation to the contemporary subject. **Television & New Media**, v. 20, n. 4, p. 336-349, 2019.
- ESCOBAR, A. **Designs for the pluriverse**: Radical interdependence, autonomy and the making of worlds. Durham; London: Duke University Press, 2018.
- FALS BORDA, O. The application of participatory action-research in Latin America. **International Sociology**, v. 2, n. 4, p. 329-347, 1987.
- FALS BORDA, O. **Ante la crisis del país**: Ideas acción para el cambio. 1ª ed. Bogotá: Panamericana, 2003.
- FLORIDI, L.; COWLS, J.; BELTRAMETTI, M. et al. AI4People – An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. **Mind and Machines**, v. 28, p. 689-707, 2018.
- FREIRE, P. **Pedagogy of the oppressed** (Pedagogia do oprimido). New York: Continuum, 1972.
- FREIRE, P.; MACEDO, D. A dialogue: Culture, language, and race. **Harvard Educational Review**, v. 65, n. 3, p. 379-390, 1995.
- GLASSMAN, M.; ERDEM, G. Participatory action research and its meanings: Vivencia, praxis, conscientization. **Adult Education Quarterly**, v. 64, n. 3, p. 206-221, 2014.
- GURUMURTHY, A. Freirean 'humility' in the age of proliferating spectacle: A reflection. In: SUZINA, A. C.; TUFTE, T. (Eds.). **Freire and the perseverance of hope**. Amsterdam: Institute of Network Cultures, 2022. p. 65-70.
- HARAWAY, D. Situated knowledges: The science question in feminism and the privilege of partial perspective. **Feminist Studies**, v. 14, n. 3, p. 575-599, 1988. DOI: doi.org/10.2307/3178066. Acesso em: 25 maio 2024.
- HUANG, M.-H.; RUST, R.; MAKSIMOVIC, V. The feeling economy: Managing in the next generation of artificial intelligence (AI). **California Management Review**, v. 61, n. 4, p. 43-65, 2019.
- JOBIN, A.; IENCA, M.; VAYENA, E. The global landscape of AI ethics guidelines. **Nature Machine Intelligence**, v. 1, p. 389-399, 2019.
- KATZ, Y. **Artificial whiteness**: Politics and ideology in artificial intelligence. New York: Columbia University Press, 2020.



KENNEDY, H. Living with data: Aligning data studies and data activism through a focus on everyday experiences of datafication. **Krisis: Journal for Contemporary Philosophy**, v. 1, p. 18-30, 2018.

KITCHIN, R.; LAURIAULT, T. Towards critical data studies: Charting and unpacking data assemblages and their work. In: THATCHER, J.; ECKERT, J.; SHEARS, A. (Eds.). **Thinking big data in geography**. New regimes, new research. Lincoln/London: University of Nebraska Press, 2014. p. 83-94.

LEHUEDÉ, S. The Double Helix of Data Extraction: Radicalising Reflexivity in Critical Data Studies. **Technology and Regulation**, v. 2024, p. 84-92, 2024. DOI: 10.26116/techreg.2024.009.

MADIANOU, M. Nonhuman humanitarianism: When 'AI for good' can be harmful. **Information, Communication & Society**, v. 24, n. 6, p. 850-868, 2021.

MADIANOU, M. Technocolonialism: Digital innovation and data practices in the humanitarian response to refugee crises. **Social Media and Society**, v. 5, n. 3, p. 1-13, 2019.

MANYOZO, L. Empathy for the other, a Freirean perspective. In: SUZINA, A. C.; TUFTE, T. (Eds.). **Freire and the perseverance of hope**. Amsterdam: Institute of Network Cultures, 2022. p. 42-45.

MEDRADO, A.; REGA, I. **Media activism, artivism and the fight against marginalisation in the global south: South-to-south communication**. London: Routledge, 2023.

MENDEZ, M. G. C.; SANABRIA, J. C. D. Reflexión metodológica sobre la aplicación concreta de la investigación acción participativa (IAP) en contextos rurales del estado de Colima. **Estudios sobre las Culturas Contemporáneas Época II**, Colima, v. 11, n. 17, 2003.

NOBLE, S. **Algoritmos da Opressão** (Algorithms of Oppression). Nova Iorque: New York University Press, 2018.

PEDWELL, C. Decolonising empathy: Thinking affect transnationally. **Samyukta: A Journal of Gender and Culture**, v. 1, n. 1, 2016.

RAHMAN, A. Roots of action research and self-reliance thinking in Rabindranath Tagore. **Action Research**, v. 4, p. 231-245, 2006.

RICAURTE, P. Ethics for the majority world: AI and the question of violence at scale. **Media, Culture & Society**, v. 44, n. 4, p. 726-745, 2022.

RUSSELL, S. **Human compatible: Artificial intelligence and the problem of control**. Londres: Viking, 2019.

SADOWSKI, J. When data is capital: Datafication, accumulation, and extraction. **Big Data & Society**, v. 6, n. 1, 2019.

SANTOS, B. S. **Epistemologias do Sul: Justiça contra o Epistemicídio** (Epistemologies of the South: Justice against Epistemicide). Londres e Nova Iorque: Routledge, 2016.

SUZINA, A. C.; TUFTE, T. Freire 's vision of development and social change: Past experiences, present challenges and perspectives for the future. **International Communication Gazette**, v. 82, n. 5, p. 411-424, 2020.



TAYLOR, L. What is data justice? The case for connecting digital rights and freedoms globally. **Big Data & Society**, jul-dez, p. 1-14, 2017.

VERDEGEM, P. Introduction: Why we need critical perspectives on AI. In: VERDEGEM, P. (Ed.). **AI for everyone? Critical perspectives**. Londres: University of Westminster Press, 2021.

VERDEGEM, P. Dismantling AI capitalism: The commons as an alternative to the power concentration of Big Tech. **AI & Society**, abril, 2022. DOI: doi.org/10.1007/s00146-022-01437-8

VIO GROSSI, F. Peasant participation, adult education, and agrarian reform in Chile. In: HALL, B.; GILLETTE, A.; TANDON, R. (Eds.). **Creating knowledge: A monopoly**. Nova Deli, Índia: Society for Participatory Research in Asia, 1982. p. 153-174.

WHYTE, W. **Participant observer: An autobiography**. Ithaca, NY: ILR Press, 1994.

WRÓBLEWSKA, M. N. Research impact evaluation and academic discourse. **Humanities and Social Sciences Communications**, v. 8, n. 1, 2 mar. 2021. DOI: doi.org/10.1057/s41599-021-00727-8

★

Este é um ARTIGO publicado em acesso aberto (*Open Access*) sob a licença *Creative Commons Attribution*, que permite uso, distribuição e reprodução em qualquer meio, sem restrições, desde que o trabalho original seja corretamente citado.