



UNIVERSIDADE FEDERAL DO RIO DE JANEIRO
FUNDAÇÃO OSWALDO CRUZ
UNIVERSIDADE DO ESTADO DO RIO DE JANEIRO
UNIVERSIDADE FEDERAL FLUMINENSE

DÉNIZARD CUSTODIO DOS SANTOS JÚNIOR

INJUSTIÇA PROGRAMADA:

A Tendência Colonialista e os (Ab)usos da Inteligência Artificial no Campo da Saúde

Rio de Janeiro

2024

DÉNIZARD CUSTODIO DOS SANTOS JÚNIOR

INJUSTIÇA PROGRAMADA:

A Tendência Colonialista e os (Ab)usos da Inteligência Artificial no Campo da Saúde

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Bioética, Ética Aplicada e Saúde Coletiva, de instituições de ensino superior associadas, como requisito parcial à obtenção do Título de Mestre em Bioética orientado pelo Professor Doutor Rodrigo Siqueira-Batista e coorientado pelo Professor Doutor Eugênio da Silva.

Rio de Janeiro

2024

C987 Custodio, Dénizard.

Injustiça programada: a tendência colonialista e os (ab)usos da inteligência artificial no campo da saúde / Dénizard Custodio dos Santos Júnior. – Rio de Janeiro, 2024.

101 f.: il.; 30 cm.

Orientador: Rodrigo Siqueira-Batista.

Coorientador: Eugênio da Silva.

Dissertação (Mestrado) - UFRJ/UFF/UERJ/FIOCRUZ. Programa de Pós-Graduação em Bioética, Ética Aplicada e Saúde Coletiva, 2024.

Referências: f. 92-101.

1. Inteligência artificial. 2. Informática médica. 3. Colonialismo. 4. Bioética.
I. Siqueira-Batista, Rodrigo. II. Silva, Eugênio da. III. Universidade Federal do Rio de Janeiro. IV. Título.

CDD 171.7

FOLHA DE APROVAÇÃO

DÉNIZARD CUSTODIO DOS SANTOS JÚNIOR

INJUSTIÇA PROGRAMADA:

A Tendência Colonialista e os (Ab)usos da Inteligência Artificial no Campo da Saúde

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Bioética, Ética Aplicada e Saúde Coletiva, em associação UFRJ-FIOCRUZ-UERJ-UFF, como requisito parcial à obtenção do título de Mestre em Bioética Aplicada e Saúde Coletiva.

Aprovada em: 26 de junho de 2024.

Dr. Rodrigo Siqueira-Batista (Orientador)
Universidade Federal do Rio de Janeiro

Dr. Eugênio da Silva (Coorientador)
Universidade do Estado do Rio de Janeiro
Centro Universitário Serra dos Órgãos

Dra. Andréia Patrícia Gomes
Fundação Oswaldo Cruz

Dr. Wanderson Flor do Nascimento
Universidade de Brasília

AGRADECIMENTOS

A conclusão desta dissertação marca o fim de uma longa jornada que não teria sido possível sem o apoio e a contribuição de diversas pessoas e instituições, às quais necessito expressar minha mais profunda gratidão.

Primeiramente, gostaria de agradecer aos professores do programa PPGBIOS, especialmente aos meus orientadores, Prof. Dr. Rodrigo Siqueira-Batista e Prof. Dr. Eugênio da Silva, cujas indicações e vasto conhecimento foram essenciais para o desenvolvimento desta pesquisa, muito obrigado pelas correções. Agradeço também aos membros da banca examinadora, Prof^a. Dra. Andréia Patrícia Gomes e Prof. Dr. Wanderson Flor do Nascimento por suas valiosas contribuições e sugestões que certamente virão. Especialmente agradeço a paciência do professor Flor, pois muitas foram as madrugadas de auxílio sem as quais esta dissertação jamais iria para frente.

Agradeço demais aos funcionários das bibliotecas, “bandex” e demais instituições em que fui recebido. Um agradecimento extraordinário ao bibliotecário Roberto Unger, chefe da biblioteca do Instituto de Estudos em Saúde Coletiva, o qual dedicou um pouco de seu rigor para essa investigação.

Aos meus amigos, agradeço pela paciência, apoio e pela ajuda prestada ao longo desta caminhada. Aos amigos do programa de mestrado do PPGBIOS, como Endi Kirby, Carolina Esteves e Nayhara Gentil, obrigado! Em especial, agradeço demais por todo suporte e eventos, Inês Nin. E um imenso obrigado para Mariana Coutinho pelas reclamações e felicidades compartilhadas em nossas reuniões. Esse sucesso dedico aos meus!

Aos meus pais e irmãos e a minha família, meu mais profundo agradecimento pelo carinho, incentivo e suporte. Vocês foram e são minha base para essa e outras pelejas e sem vocês, esta conquista não seria possível nem um pouco. Particularmente, obrigado mãe por sempre me alertar aos perigos das tecnologias e pai por me fazer desconfiar de tudo que é muito fácil.

Uma gratidão singular merece ser expressada à minha companheira, Aisha Martins, a quem agradeço por estar ao meu lado durante as diversas noites de surto e incertezas. Sua paciência e apoio constante foram cruciais para que eu pudesse alcançar e findar esse objetivo.

Por fim, expresso também meu reconhecimento ao sistema econômico brasileiro, que, por meio de políticas públicas e instituições de fomento à pesquisa, permite que investigações

científicas como esta sejam realizadas de forma gratuita, beneficiando a comunidade acadêmica, fortalecendo o ensino superior e a sociedade como um todo. Sigamos por mais!.

Em vez de entificar, “endeusar”, trata-se, antes, de politizar o tema da tecnologia; politizar, aqui, significa fazer emergir como problemático aquilo que, para muitos, é visto como resolvido ou como a solução de todos os males

Gildemarks Costa e Silva

RESUMO

CUSTODIO, Dénizard. **Injustiça programada:** a tendência colonialista e os (ab)usos da inteligência artificial no campo da saúde. Dissertação (Mestrado em Bioética, Ética Aplicada e Saúde Coletiva) – PPGBIOS, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2024.

Dado o alto grau de dependência tecnológica no qual nos encontramos, podemos dizer que vivemos em uma sociedade digital, com aspectos tecnológicos a todo instante. A implementação de sistemas de inteligências artificiais (IA) a nossa vida cotidiana se tornou comum. Como as inteligências artificiais influenciam o uso das novas tecnologias na área da saúde? De que forma essas questões podem ser analisadas à luz da colonialidade, da bioética de intervenção e da bioética de proteção? A IA e demais tecnologias têm sido apontadas como solução para diversos problemas em nossa sociedade. No entanto, ao supô-la como o arauto do futuro, se esquece que suas bases surgiram no passado. Os modelos de IA já estão amplamente em uso no setor da saúde, o que levanta questões éticas sobre a utilização deles em ambientes tão sensíveis, pois é imprescindível garantir decisões isentas de preconceitos e práticas discriminatórias, que, efetivamente, são cotidianas no mundo. Esse estudo explora as interseções entre a colonialidade e os sistemas de inteligências artificiais, escrutinando como a IA tem sido uma ferramenta de poder e controle que reflete e reforça a estrutura colonial da sociedade agora em níveis antes não alcançáveis. Por meio de uma pesquisa bibliográfica – analisada à luz da bioética de intervenção, bioética da proteção e do colonialismo de dados – identificou-se um uso acrítico da IA. Como resultado das buscas nota-se que é possível utilizar de melhores processos, práticas e metodologias para um desenvolvimento ético desses modelos inteligentes no capitalismo neoliberal, que podem mitigar os danos observados durante o trabalho. Destacam-se também iniciativas e preocupações pertinentes ao desenvolvimento e utilização de IA por países subdesenvolvidos. Por fim, transforma-se este trabalho em um alerta e espera-se que tal pesquisa e métodos motivem outros investigadores e demais envolvidos na utilização de IA na saúde a expandir o campo de estudos.

Palavras-chave: inteligência artificial; bioética; colonialidade; informática médica.

ABSTRACT

CUSTODIO, Dénizard. **Injustiça programada:** a tendência colonialista e os (ab)usos da inteligência artificial no campo da saúde. Dissertação (Mestrado em Bioética, Ética Aplicada e Saúde Coletiva) – PPGBIOS, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2024.

Since the high degree of technological dependency, we find ourselves in, we can say that we live in a digital society, with technological aspects present at all times. The implementation of artificial intelligence systems (AI) into our daily lives has become commonplace. How do artificial intelligences influence the use of new technologies in healthcare? How can these issues be analyzed in light of coloniality, intervention bioethics, and protection bioethics? AI and other technologies have been presented as solutions to various problems in our society. However, in considering them as heralds of the future, we forget that their foundations emerged in the past. AI models are already widely used in the healthcare sector, raising ethical questions about their use in such sensitive environments. It is crucial to ensure decisions are free from biases and discriminatory practices, which are indeed commonplace in the world. This study explores the intersections between coloniality and artificial intelligence systems, examining how AI have been tools of power and control that reflect and reinforce the colonial structure of society, now at levels previously unattainable. Through bibliographic research in the fields of intervention bioethics, protection bioethics, and data colonialism, an uncritical use of AI was identified. As a result of these searches, it was observed that it is possible to use better processes, practices, and methodologies for the ethical development of these intelligent models in neoliberal capitalism, which can mitigate the observed damages during the work. Initiatives and concerns about the development and use of AI by underdeveloped countries are also highlighted. Finally, this work becomes an alert, and it is hoped that this research and methods will motivate other researchers and others involved in the use of AI in healthcare to expand the field of studies.

Keywords: artificial intelligence; bioethics; coloniality; medical informatics.

RESUMEN

CUSTODIO, Dénizard. **Injustiça programada:** a tendência colonialista e os (ab)usos da inteligência artificial no campo da saúde. Dissertação (Mestrado em Bioética, Ética Aplicada e Saúde Coletiva) – PPGBIOS, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2024.

Puesto el alto grado de dependencia tecnológica en el que nos encontramos, podemos decir que vivimos en una sociedad digital, con aspectos tecnológicos presentes en todo momento. La implementación de sistemas de inteligencia artificial (IA) en nuestra vida cotidiana se ha vuelto común. ¿Cómo influyen las inteligencias artificiales en el uso de las nuevas tecnologías en el ámbito de la salud? ¿Cómo pueden ser analizados estos problemas a la luz de la colonialidad, la bioética de intervención y la bioética de protección? La IA y otras tecnologías se han presentado como soluciones a varios problemas en nuestra sociedad. Sin embargo, al considerarlas como heraldos del futuro, olvidamos que sus bases surgieron en el pasado. En todo caso, los modelos de IA ya se están utilizando ampliamente en el sector de la salud, lo que plantea cuestiones éticas sobre su uso en entornos tan sensibles. Es fundamental garantizar decisiones libres de prejuicios y prácticas discriminatorias, que de hecho son cotidianas en el mundo. Este estudio explora las intersecciones entre la colonialidad y los sistemas de inteligencia artificial, examinando cómo la IA ha sido una herramienta de poder y control que refleja y refuerza la estructura colonial de la sociedad, ahora en niveles antes inalcanzables. Mediante una investigación bibliográfica en los ámbitos de la bioética de intervención, la bioética de protección y el colonialismo de datos, se identificó un uso acrítico de la IA. Como resultado de estas búsquedas, se observa la posibilidad de utilizar mejores procesos, prácticas y metodologías para un desarrollo ético de estos modelos inteligentes en el capitalismo neoliberal, los cuales podrían mitigar los daños observados durante el trabajo. También es posible mirar iniciativas y preocupaciones pertinentes sobre el desarrollo y uso de IA por parte de países subdesarrollados. En última instancia, este trabajo se convierte en una alerta y se espera que esta investigación y métodos motiven a otros investigadores y a aquellos involucrados en la utilización de IA en la salud a ampliar el campo de estudios.

Palabras clave: inteligencia artificial; bioética; colonialidad; informática médica.

LISTA DE ILUSTRAÇÕES

Figura 1- Definições de inteligência artificial, organizadas em quatro categorias.....	23
Figura 2 - Utilização da IA, aprendizagem de máquina e profunda e seus usos na área da saúde	26
Figura 3 - Casos selecionados de manifestação de viés algorítmico em visão.....	42
Figura 4 - Algumas aplicações e preocupações relativas à IA	53
Figura 5 - O uso e dependência de IA em diferentes áreas da medicina.....	54
Figura 6 - Classificação e definição de riscos do uso da IA.....	83
Quadro 1 - Maneiras e práticas seguras para implementação de IA	80

LISTA DE ABREVIATURAS E SIGLAS

ADPPA	Projeto de Lei Federal Americana de Privacidade e Proteção de Dados
AM	Aprendizagem de máquina
ANAHP	Associação Nacional de Hospitais Privados
ANPD	Autoridade Nacional de Proteção de Dados
APP	Aplicativo usado em dispositivos móveis
BI	Bioética de Intervenção
BP	Bioética de Proteção
BVS	Biblioteca Virtual em Saúde
DeCS	Descritores em Ciências da Saúde
EUA	Estados Unidos da América
GDPR	Regulamento Geral de Proteção de Dados (União Europeia)
LGPD	Lei Geral de Proteção de Dados - Lei 13.709/2018 (Brasil)
IA	Inteligência artificial
KNN	K Vizinhos Mais Próximos
LR	Regressão Logística
MTST	Movimento dos Trabalhadores Sem-Teto
NBC	Classificação Bayesiana Ingênuas
NHS	National Health Service
PEC	Prontuários Eletrônicos Dos Cidadãos
PLS	Mínimos Quadrados Parciais
PROCON	Programa de Proteção e Defesa do Consumidor
RF	Floresta Aleatória
RNA	Rede Neural Artificial
SOBRASP	Sociedade Brasileira para a Qualidade do Cuidado e da Segurança do Paciente
SUS	Sistema Único de Saúde
UE	União Europeia

SUMÁRIO

1 APRESENTAÇÃO	12
2 QUESTÃO, HIPÓTESE E OBJETIVOS	14
2.1 QUESTÃO NORTEADORA	14
2.2 HIPÓTESE DE PESQUISA	14
2.3 OBJETIVOS	15
2.3.1 OBJETIVO GERAL	15
2.3.2 OBJETIVOS ESPECÍFICOS	15
3 METODOLOGIA	16
3.1 REVISÃO DA LITERATURA	16
3.2 ANÁLISE CRÍTICA DE TEXTOS E PROPOSIÇÃO DE FERRAMENTAS TEÓRICAS	17
4 INTRODUÇÃO: A CORRIDA PERIGOSA	19
4.1 ANTROPOMORFIZAÇÃO DA IA E A NEUTRALIDADE DOS ALGORÍTMICOS ...	28
5 IA DATIFICAÇÃO DA VIDA?	32
5.1 A ESTRUTURA OCULTA	32
5.2 LÓGICAS QUE GOVERNAM.....	36
5.3 DISCRIMINAÇÃO ALGORÍTMICA E OUTRAS OPRESSÕES COLETIVAS	40
5.4 A GUERRA PERDIDA PARA O CAPITALISMO INFORMACIONAL?.....	48
6 O (AB)USO DA IA NA SAÚDE	53
6.1 POR UMA VIDA NÃO PROGRAMADA	55
6.2 A SAÚDE COMO MERCADORIA	58
7 DEVIR CURUPIRA: O PROBLEMA DA HERANÇA EM IA	70
8 CONCLUSÃO: A LUTA CONTINUA	78
REFERÊNCIAS	92

1 APRESENTAÇÃO

A inteligência artificial (IA) tem sido apontada como solução para muitas adversidades em diferentes âmbitos da sociedade. No entanto, ao supor essa tecnologia como resolvedora dos problemas, esquecemos que a ferramenta também pode causar danos. Nesse contexto, o tema deste estudo é explorar as interseções entre a inteligência artificial e a colonialidade, analisando em que aspectos a IA tem atuado como uma ferramenta que reproduz as estruturas coloniais históricas em níveis antes humanamente inalcançáveis. Em suma, discutiu-se IA, colonialidade, colonialismo de dados, solucionismo tecnológico, dataísmo, bioética de proteção e a bioética de intervenção almejando o objetivo de investigar como a inteligência artificial pode reforçar formas de dominação que perpetuam problemas coloniais, especialmente nos usos da IA na área da saúde. Com isso, o estudo está ordenado da seguinte maneira:

Na quarta seção serão analisadas as definições de inteligência artificial, aprendizagem de máquina e se discorrerá sobre o surgimento do campo de estudo. Da mesma forma, se discute o impacto da ferramenta ChatGPT e sua influência na mídia, principalmente apresentando a “corrida perigosa” de certas empresas para alcançar a soberania sobre a IA mais potente para oferecê-la ao mercado global.

Já no quinto segmento, será introduzido o fenômeno da datificação da vida, em que nossos dados pessoais são coletados, armazenados e utilizados para diversos fins, tornando pessoas e coisas em dados, tirando daí o valor nessa nova fase do capitalismo. Diversos autores identificam e associam essa prática como um novo tipo de colonialismo, já que os dados são controlados e explorados por instituições, muitas vezes sem o consentimento ou noção dos sujeitos, o que faria parte de um “colonialismo de dados”. Os conceitos de dataísmo e solucionismo tecnológico também entrarão no debate, que poderiam ser resumidos como uma espécie de “embasbacamento” a respeito da tecnologia, em que “supostamente, o ser humano, por meio da tecnologia, irá construir uma vida feliz para todos” (Silva, 2013, p. 848).

Os impactos éticos dessas ideologias e práticas institucionais em vigor na sociedade contemporânea são abordados durante todo o trabalho, mas com um detalhamento maior na sexta seção, na qual se enfrentarão os desafios munidos das contribuições advindas da bioética de proteção (BP), proposta por Fermin Roland Schramm e por Miguel Kottow, apresentada na presente dissertação principalmente pela lupa de Rodrigo Siqueira-Batista e também os aportes da bioética de intervenção (BI), criada por Volnei Garrafa e Dora Porto, sobretudo pela lente de Wanderson Flor. Tais colaborações são combinadas para avaliar a condição dos usos e

abusos da IA no campo da saúde, entendendo como certos autores analisam o uso de dados médicos e o que seriam exatamente “dados médicos” na era digital em que a saúde representa um grande e valoroso mercado para as empresas.

A última parte se limitará a discorrer sobre a expansão da definição de *devir curupira*, conceito criado por Elen Nas (2020) que se refere ao fato da IA apontar para o futuro, mas sua base epistemológica, cultural e social ser guiada pelo passado. Como a IA é desenvolvida, treinada principalmente por poucos países, e é utilizada em outros lugares distantes da sua criação, potencialmente há a chance de perpetuar assim as próprias visões e interesses dos criadores. Se discorrerá também a respeito das aplicações que seriam viáveis caso o *devir curupira* fosse adotado como princípio no desenvolvimento de IA. Estas possibilidades serão apresentadas visando reduzir e mitigar os problemas elencados durante o presente estudo, isto é, a implementação e perpetuação de vieses e preconceitos históricos em inteligência artificial que controlam grande parte das operações em uma sociedade capitalista informacional.

2 QUESTÃO, HIPÓTESE E OBJETIVOS

2.1 QUESTÃO NORTEADORA

A partir dos pressupostos anteriormente apresentados, proponho a seguinte questão de pesquisa:

De que maneira a inteligência artificial influencia o uso das novas tecnologias em saúde? Como esses problemas podem ser analisados à luz da colonialidade e da bioética de intervenção e da bioética da proteção?

2.2 HIPÓTESE DE PESQUISA

A tese orientadora desta pesquisa foi uma preocupação pelo que nos tem acercado de todos os lados: diversas tecnologias e várias partes da sociedade se integram em inteligências artificiais, indicando que a IA é o futuro, e que devemos utilizá-las para auxiliar em tudo que for possível. Mas o problema nessa suposição é justamente que se pressupõe a IA como sendo neutra e imparcial, até mesmo quando envolvidas em erros massivos. A neutralidade dessas tecnologias seria uma ideia popularmente disseminada, e o conceito de neutro seria baseado em que a “tecnologia em si não é nem boa nem má, é a maneira como ela é utilizada que determina seu valor” (Alves, 1968, p. 16). Essa definição, no entanto, conforme será detalhada no decorrer do trabalho não se sustenta dado que as tecnologias são produtos sociais e, portanto, carregam valores atrelados em sua própria concepção e criação. Com isso, não é de se espantar que no decorrer das leituras uma tendência se destacou, isto é, apesar da Inteligência Artificial ser considerada uma tecnologia neutra, ela pode perpetuar e potencializar desigualdades e problemas sociais já presentes na sociedade.

Utilizando da bioética de intervenção e a bioética da proteção e as possibilidades de abordagem a partir da perspectiva decolonial para analisar inteligência artificial, esse estudo busca elucidar algumas características que podem ser criticadas e divulgadas para fortalecer e ampliar a criação de IAs éticas e que sirvam para ajudar a todos, universalizando a igualdade e equidade e não a discriminação. Portanto, a partir da literatura acerca do tema, foram levantadas as seguintes hipóteses:

1. Socialmente, a IA é tida como um produto tecnológico imparcial e isento de vieses. Mas pode conter e reforçar vieses sociais que impactam seu uso.

2. O uso da IA, principalmente no campo da saúde, pode afetar negativamente os pacientes e cidadãos.
3. O devir curupira como princípio e outros métodos podem auxiliar no desenvolvimento de sistemas de IA mais equitativos e adequados aos contextos em que serão utilizados.

2.3 OBJETIVOS

2.3.1 Objetivo Geral

Averiguar em que medida a IA está aprimorando mecanismos de dominação, que acabam por perpetuar problemas oriundos da colonialidade, principalmente avaliando os ab(usos) da IA na saúde, utilizando para análise o olhar do colonialismo de dados e bioéticas de intervenção e de proteção.

2.3.2 Objetivos Específicos

- a) Conceituar inteligência artificial (IA) e analisar o seu estado atual na sociedade;
- b) Conceituar colonialismo de dados e datificação e aplicar tais definições para entender as escalas de reprodução da colonialidade por meio da IA na vida das pessoas;
- c) Identificar os usos e abusos da utilização da IA no campo da saúde;
- d) Revisar o conceito de devir curupira e ampliar sua definição e relevância no desenvolvimento e estudo de inteligência artificial;
- e) Utilizar os pressupostos teóricos das bioéticas de intervenção e da proteção para embasar as percepções, buscando encontrar um caminho sem opressões automatizadas para trilhar.

3 METODOLOGIA

3.1 REVISÃO DA LITERATURA

A presente dissertação representa um trabalho baseado em análise teórica, ou seja, de cunho bibliográfico analítico, que envolveu leitura de livros, teses, dissertações e artigos, sob o método estrutural de análise de textos.

O estudo trilhou um caminho de fomento que envolveu a revisão narrativa da literatura, que é adequada para o estudo por ser uma metodologia de revisão de literatura que permite uma abordagem temática mais abrangente se comparada a outras revisões de literatura e não tão rígida quanto à busca, segundo Unger (2023, p. 133). Essa característica se dá por conta que “esse tipo de revisão promove a análise e a sumarização dos dados advindos das publicações mais relevantes dentro do tópico em estudo, a fim de inspirar novas pesquisas ou reconciliar aquelas existentes” (Carvalho, 2029, p. 917). No entanto, mesmo que o método não tenha uma padronização de busca rigorosa, utilizou-se uma estratégia de busca definida com base nos descritores, que foram identificados nos Descritores em Ciências da Saúde (DeCS) da Biblioteca Virtual em Saúde (BVS). Segundo Lopes (2002, p. 42) um descritor faz parte de um “conjunto limitado de termos autorizados para uso na indexação e busca de documentos” em bases de dados com o intuito da recuperação lógica de documentos), sendo essencial para uma pesquisa científica. Com base nisso, houve a escolha de quais descritores seriam relevantes para o trabalho.

Os descritores selecionados foram: 1) “Inteligência Artificial”; 2) “Bioética”; 3) “Informática Médica”; 4) “Direitos Civis”; 5) “Populações minoritárias, vulneráveis e desiguais em saúde”; 6) “Populações Vulneráveis”. Colonialidade e colonialismo (nos respectivos idiomas) foram palavras chaves usadas ao buscar alguma correlação com os descritores acima nos portais e em repositórios digitais como o Google Acadêmico.

Em seguida, realizou-se a busca nas bases de dados PubMed, Literatura Latino-Americana e do Caribe em Ciências da Saúde (Lilacs) e Scientific Electronic Library Online (SciELO), combinando os descritores indicados acima, mantendo entre as variações o descritor “Inteligência Artificial”, objeto principal deste estudo. A busca apresentou poucos resultados relacionados com o escopo da pesquisa encontrados ao buscar dentro dos portais.

Os poucos resultados de busca se mantiveram mesmo levando em consideração a publicação de artigos em três idiomas português, inglês e espanhol, que foram usados como os

idiomas do critério de inclusão. Outros critérios de inclusão foram os autores e portais exibirem os textos completos de maneira disponível, com um recorte temporal de 1990 a 2024.

Para não haver problemas no decorrer do texto, é necessário um esclarecimento sob a noção de conceito, pois é preciso levar em conta que todo conceito (i) aglutina uma multiplicidade de elementos – e, portanto, podem ser ambíguos – e (ii) geralmente remetem a um problema. Conforme afirma Abbagnano (1998, p. 164) em seu Dicionário de Filosofia, o conceito é

[...] todo processo que torne possível a descrição, a classificação e a previsão dos objetos cognoscíveis. Assim entendido, esse termo tem significado generalíssimo e pode incluir qualquer espécie de sinal ou procedimento semântico, seja qual for o objeto a que se refere, abstrato ou concreto, próximo ou distante, universal ou individual, etc.

Desta forma, ao ler os textos, conforme se avançava no entendimento dos conceitos, foram utilizados fichamentos e demais técnicas para fixar o conhecimento obtido. Conforme elucidam Jean e Dominique (2006) o fichamento é uma técnica útil para combater o esquecimento de textos, portanto, empregá-la é criar um vestígio escrito de um trabalho penoso: “as fichas são, portanto, absolutamente indispensáveis. É a partir delas que se pode ter uma ideia precisa, ao mesmo tempo global e detalhada, dos textos filosóficos e de seus autores” (Folscheid; Wunenburger, 2006, p. 24).

3.2 ANÁLISE CRÍTICA DE TEXTOS E PROPOSIÇÃO DE FERRAMENTAS TEÓRICAS

Ao perpassar sobre os textos disponíveis, se impõe a necessidade de compreendê-los criticamente, isto é, passando pelo crivo do entendimento. O trabalho de entendimento dos conceitos e explicação dos mesmos durante a elaboração desse texto “liga desde o início a produção das noções, sua análise e a revelação da argumentação, mesmo se somos obrigados a fazê-lo de maneira parcelar e anárquica. A ordem da preparação é, portanto, a desordem provisória do vaivém” (Folscheid; Wunenburger, 2006, p. 126).

No entanto, o uso desses textos e conceitos será plenamente baseado na interpretação que o filósofo Michel Foucault tinha de sua obra: ela serve para ser usada de acordo com alguma finalidade de igual maneira uma ferramenta. Não dispondo nesse processo de exclusividade de uso ou mesmo de fidelidade a tal autor/teoria. Nas palavras do filósofo francês, ele definiu que sua obra, ou seja, seu discurso é:

Meu discurso é, evidentemente, um discurso de intelectual e, como tal, opera nas redes de poder em funcionamento. Contudo, um livro é feito para servir a usos não definidos por aquele que o escreveu. Quanto mais houver usos novos, possíveis, imprevistos, mais eu ficarei contente. Todos os meus livros seja História da loucura seja outros podem ser pequenas caixas de ferramentas. Se as pessoas querem mesmo abri-las, servirem-se de tal frase, tal ideia, tal análise como de uma chave de fenda, ou uma chave-inglesa, para produzir um curto-circuito, desqualificar, quebrar os sistemas de poder, inclusive, eventualmente, os próprios sistemas de que meus livros resultam, pois bem, tanto melhor! (Foucault, 2006, p. 52).

Assim, pretendo utilizar dos autores dispostos aqui como uma caixa de ferramentas que auxiliam para propor contribuições ao real problema bioético que se enfrenta nesse texto, a saber, a colonialidade presente na inteligência artificial. De igual maneira, almejo contribuir com tal arsenal o tornando mais amplo e dessa forma ajudar em um eventual curto-circuito no colonialismo de dados. Com todas as ferramentas expostas à mesa, podemos nos voltar ao trabalho principal, isto é, o resultado final desse caminho metodológico. Em verdade, antes de chegarmos às conclusões, precisamos partir da introdução e ir aprofundando mais na problemática para entendermos o desfecho dessa revisão narrativa. Dito isso, vamos começar do começo, tratando da temática no início.

4 INTRODUÇÃO: A CORRIDA PERIGOSA

Um tópico tem dominado de forma massiva as redes sociais e estado presente em muitos noticiários, apavorando alguns, alegrando outros e surpreendendo a todos. O fato é que a Inteligência Artificial (IA) é uma inovação que há pelo menos quinze anos tem muitos holofotes, mas apesar disso, o campo já tem mais de meia centena de anos.

De acordo Russel e Norvig (2013) o primeiro trabalho hoje reconhecido como IA foi realizado pela dupla composta por Warren McCulloch e Walter Pitts, que em 1943, descreveram a criação de um neurônio artificial, a partir de alguns fundamentos de neurofisiologia, da lógica proposicional de Bertrand Russell e Whitehead e da teoria da computação de Turing (Russel; Norvig, 2013, p. 41). Alan Turing é considerado o “pai da computação” porque sua pesquisa foi fundamental para impulsionar os estudos nessa área, como o da dupla mencionada anteriormente. Turing já vinha produzindo artigos extremamente vitais para o desenvolvimento do campo desde 1936, como no artigo¹ em que ele explicava o funcionamento do que seria a “máquina universal”, isto é, um modelo abstrato de computador. Mas foi em um artigo intitulado “*Computing Machinery and Intelligence*”, de 1950, que ele alcançou maior popularidade ao lançar um paradigma², que ficou conhecido como o teste de Turing. Ainda em 1950, Minsky e Edmonds “construíram o primeiro computador de rede neural” (Russel; Norvig, 2013, p. 41). Segundo Siqueira-Batista e Silva (2019), logo após, em 1956, houve a realização de um seminário organizado por John McCarthy, Marvin Minsky, Claude Shannon e Nathaniel Rochester, que pode ser considerado um marco para a área, demarcando e consolidando a IA enquanto campo de estudos, e tendo justamente pela primeira vez, utilizado o termo inteligência artificial.

Durante os anos seguintes, até 1962, a empregada robô Rosie e o computador companheiro Rudi, dos Jetsons, faziam com que todos desejassem ter uma casa inteligente e um assistente que cuidasse de todas as tarefas. Nisso se dá um dos possíveis inícios da propagação³ da ideia de IA no público em geral. Stanley Kubrick, exímio diretor de cinema, pouco tempo depois consolidou o processo que continuaria a conquistar a mídia ao longo dos anos. Em 1968, ele lança o filme “*2001: Uma Odisseia no Espaço*”, obra que leva aos

¹ Intitulado “*On Computable Numbers, with an Application to the Entscheidungsproblem*”, o texto era bem técnico e teórico, dificultando a compreensão por não especialistas.

² O teste de Turing criou um paradigma na computação, por ter estabelecido “realizações científicas que, durante algum tempo, fornecem problemas e soluções modelares para uma comunidade de praticantes de uma ciência” (Kuhn, 1998, p. 13).

³ Ao que tudo indica, o primeiro longa a utilizar o conceito foi *Metrópolis*, de 1927 de Fritz Lang, apesar do filme ser bastante impopular, motivo pelo qual não é citado diretamente no texto.

telespectadores o conceito de uma inteligência artificial: Hal 9000, o computador inteligente que controla a nave. Dessa forma, por conta do grande sucesso do filme, outras representações vieram a habitar o imaginário sociocultural da população mediado pelas telas e, com o passar do tempo, novas roupagens foram utilizadas: um replicante, uma cópia indistinguível de um humano no premiado “*Blade Runner*”; uma IA caçadora incansável que tenta aniquilar a humanidade, tal máquina foi criada por uma empresa chamada Skynet, no “*O Exterminador do Futuro*”; uma releitura de Pinóquio ou então um filme em que uma IA no corpo de uma criança⁴ busca ser amado, como vemos no “*A.I. - Inteligência Artificial*” de Spielberg. Com o passar do tempo, o desenvolvimento da própria IA e da cultura midiática, novas representações foram surgindo e adicionando cada vez mais camadas a esse até então estranho conceito abstrato futurista e distante da inteligência artificial. Por conta desse entrelaçamento com a cultura, podemos pensar que a IA, com certeza, permanecerá por muito tempo conosco. Ou nem tanto, já que recentemente, alguns bilionários e alguns funcionários de grandes empresas do ramo tecnológico decidiram escrever uma carta aberta⁵ voltada aos projetos de inteligência artificial, pedindo uma pausa de seis meses de pesquisas maiores ou similares ao ChatGPT⁶. Essa atitude poderia ser interpretada como uma possível ameaça à IA em si ou a certo modelo ou uso dela.

O caso é surpreendente, pois mesmo após já estar se tornando comuns carros sem motoristas e dezenas de outros estabelecimentos funcionando sem funcionários humanos e muitos outros problemas intensificados pela IA, somente agora a elite parece que interpretou com seriedade esse assunto⁷. Os autores da carta pedem uma pausa na “corrida perigosa”⁸ de modelos de inteligência artificial, para que se possa criar protocolos para o desenvolvimento e design de IA. Aparentemente, ao vislumbrar que uma IA também pode servir como força de trabalho intelectual, que é habitualmente dominado pela elite, a substituição dos autores por uma IA nesse local de trabalho indicaria uma perturbação na ordem “natural” das coisas. Ou

⁴ Sob essa descrição se poderia incluir também o filme “*EVA*” de Kike Mafllo, de 2011.

⁵ Pause Giant AI Experiments: An Open Letter. Future of Life Institute. Disponível em: <https://futureoflife.org/open-letter/pause-giant-ai-experiments>. Acesso em: 2 abr. 2023.

⁶ A ferramenta foi banida da Itália por não se adequar às leis de proteção de dados. Veja em: Itália bloqueia ChatGPT após suspeita de violação de regras de coleta de dados. G1, Tecnologia, 31/03/2023. Disponível em: <https://encurtador.com.br/CLR02>. Acesso em: 6 abr. 2023.

⁷ Os trabalhos autômatos não são um recurso oriundo apenas da IA, na realidade são bem mais antigos, mas se intensificam por conta dela. Enquanto estava em um âmbito de trabalhos físicos, não pareceu haver tanta urgência em cuidar do desaparecimento de alguns empregos.

⁸ Houve uma corrida armamentista entre empresas de tecnologia como Microsoft e Google, por exemplo, que no ano de 2022 lançaram vários produtos pautados em IA. E objetivando oferecer a IA mais potente, não se importando em entender, prever ou controlá-las de forma confiável.

talvez com a possibilidade de um subalterno artificial para ser explorado as coisas fiquem demasiadamente sintéticas, a dominação de classe perderia o sabor.

No final de maio de 2023, surgiram novamente preocupações por parte de alguns autores (mais de quinhentos especificamente, entre eles alguns cientistas de IA e milionários da indústria tecnológica), que foram além das questões de ética de perigo ou não, e levantaram até mesmo a possibilidade de risco de extinção humana. Sem nenhuma fundamentação, e em apenas um parágrafo de 22 palavras, sugerem que a IA poderia ter o mesmo potencial destrutivo de pandemias ou guerras nucleares⁹. É necessário salientar que muitos dos autores da declaração são aqueles que estão diretamente envolvidos no desenvolvimento e construção dessas poderosas ferramentas. Isso pode demonstrar uma espécie de ação contrastante e mesmo injusta, já que ao mesmo tempo em que financiam pesquisas com essa tecnologia, ao surgir um produto de uma empresa rival, possivelmente mais tecnologicamente refinado e avançado, esses cientistas alertam a população e pedem uma pausa nas pesquisas¹⁰. Dessa maneira, essa aparente contradição levanta uma questão intrigante: por que esses mesmos indivíduos, que são responsáveis por impulsionar o avanço da IA, expressam preocupações tão alarmantes? A resposta pode estar, em parte, na própria natureza ambivalente da IA. Por um lado, a IA oferece inúmeras oportunidades e benefícios, impulsionando avanços tecnológicos, automação e otimização de tarefas. Por outro lado, também enfrentamos desafios e dilemas éticos à medida que a IA se torna cada vez mais qualificada e autônoma. O ChatGPT impactou de forma decisiva o debate público a respeito do uso de IA, pois como alega João Texeira (2023) “se escreveu tanto sobre esse GPT3 que ele virou uma espécie de Pedra Filosofal”.

Mas, antes de tudo, o que é o ChatGPT que serviu de parâmetro e impulsionou esse tumulto sobre o tema? Sobre isso, dois pesquisadores estadunidenses se debruçaram para entender essa ferramenta. Sobre a definição do sistema, vejamos:

O ChatGPT é uma ferramenta pública desenvolvida pela OpenAI que é baseada na tecnologia do modelo de linguagem GPT. É um chatbot altamente sofisticado capaz de atender uma ampla variedade de solicitações baseadas em texto, incluindo responder a perguntas simples e realizar tarefas mais avançadas, como gerar cartas de agradecimento e orientar indivíduos em discussões difíceis sobre problemas de produtividade. O ChatGPT consegue isso aproveitando seus extensos bancos de dados e design eficiente para entender e interpretar as solicitações do usuário e, em seguida,

⁹ CAIS, Center for AI Satety. Statement on AI Risk. Disponível em: <https://www.safe.ai/statement-on-ai-risk>. Acesso em: 31 maio. 2023.

¹⁰ Essa informação é relevante para observar uma possível espécie de controle de narrativa por parte dos autores. Já que a premissa na carta é de que IA é criada por alguém, e não por eles mesmos envolvidos no financiamento e criação de IAs. E ao evocar a extinção da humanidade, se desvia a atenção de problemas mais cotidianos da IA, como uso de dados sem consentimento, vieses e usos indevidos. O’Neil (2020) demonstra que é mais fácil a IA extinguir a democracia do que a própria humanidade.

gerar respostas apropriadas em linguagem quase natural (Lund; Wang, 2023, p. 3, tradução nossa).

Essa inteligência artificial generativa é um sucesso em parte devido ao seu modelo de linguagem, que é particularmente eficaz em responder as perguntas feitas pelos usuários, se adaptando ao estilo de escrita. Mas também outro elemento faz dele o mais utilizado, que é a acessibilidade: o ChatGPT não exige que o usuário tenha conhecimentos prévios em computação e pode ser ativado por perguntas simples em uma caixa de texto¹¹. O modelo fez sucesso em sua terceira versão de aprimoramento, e segue sendo atualizado. Devido a sua capacidade adaptativa, o especialista em linguística computacional, Robert Dale (2021, p. 113) afirma que o GPT tem a habilidade de gerar textos particularmente difíceis de serem distinguidos de textos criados por humanos, da seguinte maneira:

Os conceitos por trás do GPT são refinados através de um processo de duas etapas: pré-treinamento gerativo e não supervisionado usando dados não rotulados e ajuste fino discriminativo e supervisionado para melhorar o desempenho em tarefas específicas. Durante a fase de pré-treinamento, o modelo aprende naturalmente, assim como uma pessoa poderia aprender em um novo ambiente, enquanto a fase de ajuste fino envolve um refinamento mais orientado e estruturado pelos criadores (Lund; Wang, 2023, p. 3, tradução nossa).

Assim, esse método de linguagem prioriza a resposta mais natural para o receptor da mensagem e permite que a IA responda de forma quase indistinguível da resposta de um ser humano. Isso poderia ser suficiente para resolver e engavetar alguns dos dilemas mais clássicos da computação como o teste de Turing¹² e pôr dúvidas sobre o dilema da sala chinesa¹³, devido a tamanha aparente compreensão que o modelo atua, mesmo com textos incompletos, errados ou carregados de informalidades.

No entanto, até o momento, escrevi¹⁴ como se o conceito de inteligência artificial fosse compreendido plenamente por todos. No geral, apesar de diferentes definições encontradas no

¹¹ De maneira geral, a IA generativa é a mais recente a dominar a mídia e com suas múltiplas possibilidades, ela também é uma das que apresenta maior risco, já que será mais acessada e utilizada por todos.

¹² O teste de Turing é um teste proposto por Alan Turing, e nele se avalia a possibilidade e capacidade da máquina se passar por humano. O juiz entrevista a máquina e um humano, tentando descobrir qual é qual. Se a máquina enganar o juiz, fazendo-o pensar que é um ser humano, ela passa no teste de Turing.

¹³ É um experimento mental de John Searle, filósofo da linguagem, que critica a ideia de que uma máquina pode entender a linguagem e ter consciência. Na sala, alguém sem conhecimento de chinês recebe símbolos e regras para combinar, podendo se comunicar sem entender, assim como a máquina, não consciente de como está agindo.

¹⁴ A utilização da primeira pessoa ao expor as observações e a terceira pessoa do plural visando apenas incluir o leitor no texto, se dá por conta da possibilidade de violência epistêmica ao evocar a dita universalidade da ciência. Segundo Grosfoguel (2008, p. 3), os paradigmas eurocêntricos hegemônicos que inspiraram a filosofia e as ciências ocidentais assumem um ponto de vista universalista, neutro e objetivo, mas diversos estudiosos indicam que “falamos sempre a partir de um determinado lugar situado nas estruturas de poder, dessa forma,

campo, elas frequentemente partilham da ideia de que a IA é um campo computacional que se concentra na criação de sistemas que podem se comportar e agir de forma inteligente. Elas “pensariam” por reproduzirem habilidades e ações que demandariam inteligência para serem executadas. Conforme apontam Stuart Russell e Peter Norvig (2013, p. 25) ao compilar a definição dada por oito estudiosos, os consensos sobre o que é a IA podem ser divididos em raciocínio e comportamento, em que os sistemas de IA seriam aqueles que:

A) Pensam e agem como humanos

B) Pensam e agem racionalmente

A categoria A se refere à “fidelidade ao desempenho humano”, isto é, o quão similares são as ações da máquina comparada às ações humanas. Já a categoria B mede o “sucesso comparando-o a um conceito ideal de inteligência, chamado de racionalidade” (Russel; Norvig, 2013), então racional é fazer o certo com as informações que dispõe. Na Figura 1 podemos ver as citações de cada uma das oito definições dentro das duas distintas categorias:

Figura 1- Definições de inteligência artificial, organizadas em quatro categorias

Pensando como um humano	Pensando racionalmente
<p>“O novo e interessante esforço para fazer os computadores pensarem (...) <i>máquinas com mentes</i>, no sentido total e literal.” (Haugeland, 1985)</p> <p>“[Automatização de] atividades que associamos ao pensamento humano, atividades como a tomada de decisões, a resolução de problemas, o aprendizado...” (Bellman, 1978)</p>	<p>“O estudo das faculdades mentais pelo uso de modelos computacionais.” (Charniak e McDermott, 1985)</p> <p>“O estudo das computações que tornam possível perceber, raciocinar e agir.” (Winston, 1992)</p>
Agindo como seres humanos	Agindo racionalmente
<p>“A arte de criar máquinas que executam funções que exigem inteligência quando executadas por pessoas.” (Kurzweil, 1990)</p> <p>“O estudo de como os computadores podem fazer tarefas que hoje são melhor desempenhadas pelas pessoas.” (Rich and Knight, 1991)</p>	<p>“Inteligência Computacional é o estudo do projeto de agentes inteligentes.” (Poole <i>et al.</i>, 1998)</p> <p>“AI... está relacionada a um desempenho inteligente de artefatos.” (Nilsson, 1998)</p>

Fonte: Russel; Norvig, 2013, p. 25.

Assim, apesar de o conceito de inteligência artificial utilizado não ser um absoluto consenso, irei trabalhar com uma síntese das definições encontradas no livro de Russell e Norvig, em que a IA pode ser definida como sistema computacional que se distingue por

“nossos conhecimentos são, sempre, situados”, o que justifica a evidenciação do ‘eu’ autor no decorrer do texto. Essa opção não atrapalha as possíveis contribuições científicas deste estudo senão indica e aprofunda o cuidado que se deve ter lidando com as estruturas de poder.

realizar tarefas, pensando e agindo com inteligência, que pode ser conseguida por meio de vários métodos.

No entanto, a ideia de que a racionalidade deve seguir os parâmetros humanos é bastante criticada. Essa concepção pressupõe que apenas as ações e pensamentos humanos são considerados racionais, e aplicá-la à IA pode ser problemático¹⁵, principalmente porque nos próximos anos é possível de estarmos lidando com uma nova criatura. Rodrigo Siqueira-Batista (2022), define essa criatura como um autopsicobot, isto é, um ser autoconsciente que age autonomamente e que seria completamente indistinguível de um *homo sapiens*, que “nesse momento são teóricos, mas que podem se estabelecer no futuro”. Conforme apontam Nas, Silva e Siqueira-Batista (2018, tradução nossa): “o uso da IA traz novos contornos [...], principalmente quando se considera o recente status de ‘celebridade’ alcançado por robôs que ganharam cidadania, como Sophia”¹⁶. Pensando na possibilidade de as decisões e sugestões tomadas pela IA poderem ser questionadas, o campo de Inteligência Artificial Explicável (XAI)¹⁷ surge, considerando alguns fatores como a interpretabilidade e a explicabilidade, como fundamentais para um funcionamento responsável de qualquer inteligência não natural, garantindo certa transparência do sistema independente da forma de aprendizado empregado.

Os sistemas mais avançados utilizam métodos de aprendizagem de máquina (AM) para atingir seus objetivos. No entanto, não é toda máquina que aprende, há formas de produzir um comportamento considerado inteligente sem que haja a necessidade de recorrer a métodos de aprendizagem. As chamadas inteligências artificiais simbólicas, que de acordo com Badia (2015, p. 97, tradução nossa) são tipos de IA que operam “com representações abstratas do mundo real que são modeladas usando linguagens de representação baseadas principalmente na lógica matemática e suas extensões”, não aprendem, elas já são programadas baseadas em respostas que o sistema precisa dar conforme os estímulos (comandos) recebidos, ofertando apenas respostas objetivas a situações já cadastradas antes. Elas não aprendem novos elementos, mas exibem os elementos previamente inseridos, ou seja, o conhecimento embutido que é acionado por determinado estímulo é deduzido e é exibido como resultado. Assim, o máximo de “aprendizado” dessa IA seria responder a novas situações caso sejam classificadas

¹⁵ Na diferenciação das definições de Russel e Norvig, o “ideal de inteligência” tem como base a inteligência humana. Ver mais sobre a discussão de superioridade racional humana e colonialismo em CUSTODIO, Denizard; BROEDEL, Larissa. A pseudossuperioridade humana numa perspectiva Bioética: Do ceticismo antigo aos estudos animais contemporâneos. VIII Congresso Mundial de Bioética e Direito Animal, 2022.

¹⁶ STONE, Zara. Everything You Need To Know About Sophia, The World’s First Robot Citizen. Forbes. Disponível em: <https://www.forbes.com/sites/zarastone/2017/11/07/everything-you-need-to-know-about-sophia-the-worlds-first-robot-citizen/?sh=50a9d3e346fa>. Acesso em: 3 abr. 2023.

¹⁷ XAI é um campo de pesquisa dentro do macro de IA, que tem como finalidade criar modelos e aplicações que expliquem o funcionamento de técnicas e métodos de funcionamento das “caixas pretas” da IA.

como novas condições para conhecimentos e elementos já registrados no sistema anteriormente.

Mas falando sobre algumas das formas mais utilizadas de aprendizagem de máquina (*machine learning*) são elas:

A) aprendizagem supervisionada, que ocorre quando o algoritmo é treinado com dados previamente rotulados, onde cada rótulo representa uma classe, e assim seu objetivo após aprender, é classificar corretamente novos dados.

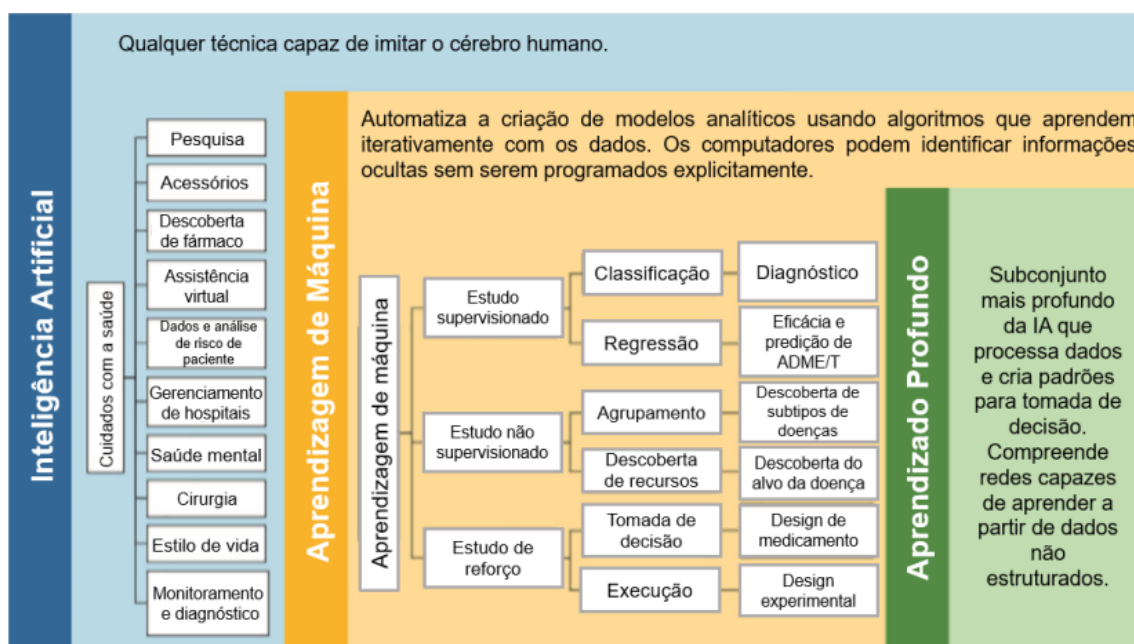
B) Aprendizagem não supervisionada os dados usados no treinamento não são rotulados e a IA se encarrega de reconhecer padrões nesses dados de forma a agrupá-los segundo algum critério de semelhança.

C) Aprendizagem por reforço, que introduz uma recompensa ao sistema quando realiza bem uma tarefa, logrando com que o sistema busque novas recompensas, realizando plenamente a ação.

E dentro desses subgrupos se incluem as diversas técnicas utilizadas, como por exemplo a classificação bayesiana ingênua (NBC), regressão logística (LR), floresta aleatória (RF), vizinhos mais próximos (KNN), mínimos quadrados parciais (PLS) e rede neural artificial (RNA)¹⁸. Arabi (2020) traduz uma imagem que ilustra bem a utilização e versatilidade desta tecnologia quando aplicada ao campo da saúde, os mais distintos usos justificam em parte o motivo do investimento das empresas neste campo. Na Figura 2 podemos visualizar com clareza que a assistência à saúde será fortemente influenciada pela IA, com recomendações que vão desde um simples auxílio na pesquisa médica, até o monitoramento, diagnóstico e sugestões de estilos de vida mais saudáveis para pacientes, bem como o desenvolvimento de medicamentos ou realização de cirurgias.

¹⁸ Para acompanhar uma explicação mais detalhada, levando em consideração até mesmo diferenças entre a perspectiva simbólico-dedutiva e a perspectiva conexionista-indutiva no funcionamento da inteligência artificial, ver capítulo dois “o que as máquinas aprendem?” de Silva, 2022.

Figura 2 - Utilização da IA, aprendizagem de máquina e profunda e seus usos na área da saúde



Fonte: Mak, *et al.* 2018 *apud* e adaptado por Arabi, 2020, p. 7.

A dificuldade, no entanto, é que quanto mais complexo o algoritmo, independentemente da técnica utilizada, pior será sua explicabilidade¹⁹. Devido a isso, mesmo com diversos princípios e contribuições ao tema, as criações ainda apresentam severas vicissitudes. Tais problemas levaram Geoffrey Hinton, um pesquisador pioneiro em pesquisas de AI super modernas que utilizam essas aprendizagens, a abandonar seu emprego na Google em maio de 2023 para expressar suas preocupações com o desenvolvimento dessas tecnologias, conforme foi relatado por Cade Metz, repórter do New York Times (2023). Hinton deixou a Google porque não concordava com a direção que a empresa estava tomando em relação à inteligência artificial, pois “é difícil evitar que a utilizem para coisas ruins”. “É assustador” para ele pensar no avanço da tecnologia nos próximos cinco anos, principalmente no campo da desinformação, em que em breve o cidadão comum já “não será mais capaz de saber o que é verdade”. Em contrapartida, o cientista-chefe do Google, Jeff Dean, disse em um comunicado que continuam comprometidos com uma abordagem responsável à IA. A equipe está “aprendendo a entender os riscos emergentes, ao mesmo tempo em que inovamos com

¹⁹ Levando isso em conta, o governo brasileiro criou diretrizes para gerir a IA, a Estratégia Brasileira de Inteligência Artificial (EBIA). A diretriz, no eixo ético, diz prezar por produções que visem os “campos de equidade/não-discriminação (fairness), responsabilidade/prestação de contas (accountability) e transparência (transparency), conhecidas como a matriz FAT” (Paiva, 2021).

ousadia”. Cade Metz, na mesma reportagem da NYT (2023, tradução nossa, grifo nosso), acrescenta que Hinton diz que:

Até o ano passado, [...] a Google atuava como um “administrador adequado” da tecnologia, **com cuidado para não liberar algo que pudesse causar danos**. Mas agora que a Microsoft aumentou seu mecanismo de busca Bing com um chatbot - desafiando o negócio principal do Google - o Google está correndo para implantar o mesmo tipo de tecnologia. **As gigantes da tecnologia estão presas em uma competição que pode ser impossível de parar**, disse Hinton.

Dessa forma, com as empresas de tecnologia competindo para implementar e desenvolver inteligências artificiais mais potentes e funcionais, torna-se mais fácil de a IA se tornar mais potencialmente perigosa, pois o lucro obtido pelas empresas é potencializado por práticas exploratórias que, por vezes, desconsideram a ética (Cassino, 2021). No entanto, no atual momento de discussão dessa tecnologia, as teorias de que a IA exterminarão a humanidade estão sendo mais uma barreira do que auxiliando no desenvolvimento de sistemas éticos de IA. Segundo Goldman (2023), Kyunghyun Cho, professor da universidade de Nova York, afirma que essas narrativas condenatórias sobre a IA estão desviando a atenção das questões que são realistas, sejam elas negativas ou positivas. Isto é, exagerando e alucinando sobre as possibilidades do que poderia acontecer, acaba-se afastando da realidade de problemas palpáveis que já ocorrem e suas possíveis soluções.

De forma palpável, Cathy O’Neil (2020) elabora em seu livro como as novas tecnologias estão ameaçando a democracia, implementando *algoritmos de destruição em massa*, ou ADMs como chama, mas diferente do que se pode pensar como algo elaborado, são tecnologias relativamente “simples” que atingem uma enormidade de pessoas. Por exemplo, em um dos capítulos O’Neil explica “como as ADMs fazem a peneira de candidatos, boicotando alguns e ignorando muitos outros. [...] o software muitas vezes traz preconceitos venenosos no código, aprendendo com registros passados a como ser injusto”, isto é, em um único software de seleção de candidatos que funciona em diferentes empresas podemos ver vieses racistas, sexistas, econômicos, geográficos e demais filtros que excluem da maneira mais economicamente viável a população, tornando a seleção mais rápida e eficiente, mas muitas vezes com base em certos preconceitos. De toda maneira, uma posição muito similar à opinião de Cho é compartilhada por muitos pesquisadores de IA e especialistas em ética com quem a Revista Nature (2023) conversou, e eles se diziam frustrados com o tom apocalíptico das narrativas observadas. Isso chegou a tal ponto de vários cientistas da popular revista Nature se pronunciarem, pedindo para diminuir as narrativas alarmistas sobre a IA, pois

isso não era construtivo para o debate em geral sobre o uso da IA e estava muito longe da realidade. Eles dizem que esse tipo de narrativa só fomenta a corrida armamentista em busca dessa “máquina superpoderosa” e incentiva a indústria ao mesmo tempo em que enfraquece os argumentos que buscam a regulação. Eles ressaltam que apenas alguns grupos realmente participam e dominam essa conversa sobre IA, deixando muitos grupos excluídos.

Essa narrativa de extinção é uma grande história esse é o fascínio da humanidade desde o início dos tempos [...] O cavalo de Tróia [que parece inofensivo, mas é malicioso] – essa é uma história semelhante, certo? [...] O lamentável é que os cientistas, bem como os formuladores de políticas, as pessoas que estão tomando decisões ou criando esses avanços, estão apenas sendo positiva ou negativamente entusiasmados por tais avanços, não sendo críticos sobre isso. Mas, no momento, a inteligência artificial geral é uma espécie de varinha mágica que eles estão apenas tentando balançar para hipnotizar as pessoas para que as pessoas não sejam críticas sobre o que está acontecendo (Goldman, 2023, tradução nossa).

É importante que haja um equilíbrio entre o avanço tecnológico e a segurança, para garantir que a IA faça o bem para todos, e não perpetue preconceitos ou amplifique problemas sociais já preexistentes. Mas, para isso, é necessário que a discussão seja conduzida e baseada em fatos e não em fantasmas ou narrativas alarmistas como a levantada pelos autores das cartas citadas no começo da seção.

Não podemos esquecer que os problemas fazem parte do mesmo contexto que permite e possibilita a criação da IA. Para Feenberg (2002), a tecnologia nada mais é que um artefato sociocultural e que, por isso, não está livre de influências históricas, políticas, culturais²⁰, e por vezes, Feenberg tentou “mostrar que a tecnologia que nos é apresentada como politicamente neutra, eterna, a-histórica, sujeita a valores estritamente técnicos e, portanto, não permeada pela luta de classes, é uma construção histórico-social” (Novaes; Dagnino, 2004, p. 191).

4.1 ANTROPOMORFIZAÇÃO DA IA E A NEUTRALIDADE DOS ALGORÍTMICOS

No decorrer do trabalho, um pensamento que pode ocorrer ao leitor é uma certa espécie de culpa por parte da própria IA nos processos aqui descritos. Tal observação é compartilhada por Álvaro Vieira Pinto (2005) que, ao analisar as relações históricas dos homens com as máquinas, percebe que há uma tendência em substantivar a técnica. A substantivação da técnica, seria um tipo de ideologização da tecnologia, a tornando um ser, que geralmente assume um caráter de “ser fantástico”, e “ao se assumir a técnica como um ser autônomo,

²⁰ Dessa forma, também está envolta na luta de classes, como disposto no próximo capítulo.

também aqui, passa-se a falar dela como se não existisse um ser humano na condição de agente central e, principalmente, como se não houvesse grupos sociais com interesses diversos no processo tecnológico” (Silva, 2013, p. 851). Isto é, ao se conceber a IA como autônoma, no sentido pleno do termo, se exime a culpa dos humanos. Portanto, é preciso elucidar que, neste trabalho, a premissa utilizada é que ainda não há IA forte²¹ o suficiente para ser realmente autoconsciente de si e, diante disso, também não poderia agir com arbitrariedade para ser culpabilizada. Independentemente de passar por certos testes, se há pensamento consciente ou apenas emulação de pensamento, conforme debatem os filósofos da mente sobre o tema, neste estudo partirei do pressuposto de que os diferentes tipos de IA ainda não são seres com autonomia.

O professor de direito Ryan Calo prefere utilizar outro termo para denominar essa espécie de autonomia limitada, que ele chama de comportamento emergente, justamente para evidenciar que a discussão nesse caso não é exatamente autonomia, mas sim um comportamento adquirido que é derivado do código escrito, agindo de acordo com as interações com o ambiente (*inputs*). Por isso é necessário não confundir com a definição de autonomia que comumente temos ao pensar em autonomia humana. Como o autor esclarece:

Eu uso o termo “emergente” em vez de “autonomia” intencionalmente. A autonomia sugere que os robôs estão de alguma forma tomando a decisão de agir de uma determinada maneira. Pouco se ganha e muito se perde, fingindo que os robôs contemporâneos exibem algo parecido com intenção (Calo, 2015, p. 539, tradução nossa).

Verdadeiramente, as inteligências artificiais atuais ainda não exibem comportamentos dignos de serem consideradas autônomas no sentido lato do termo. O que se pode perder ao não considerar essa diferença é acreditar na noção de que os problemas relacionados às IAs são de autoria delas ou que não podem ser evidenciados, em último caso, os criadores delas como negligentes ou mesmo criminosos, dependendo do ocorrido. E apesar de este estudo ainda utilizar o termo autonomia, é preciso evidenciar essa diferenciação. Dessa forma, tais inteligências são autônomas, porém de maneira limitada mediante as inserções algorítmicas adquiridas. Assim, quaisquer impressões que direcionem para uma interpretação que a IA age por “maldade” ou em alguma outra espécie de antropomorfização da IA serão decorrentes dos próprios vícios habituais e inerentes à sociabilização humana. Ou seja, discordo neste trabalho que exista algum sistema que tenha “a autonomia decisória desenvolvida pela inteligência

²¹ Simplificadamente, “IA forte” é um termo usado para definir inteligências que seriam autoconscientes. Sobre o debate entre IA forte e fraca, consultar o capítulo 26 do livro de Russell e Norvig (2013).

artificial [que] rompe padrões sociais” (Ribeiro, 2020). Tal inteligência plenamente autônoma talvez possa existir em um breve momento, mas isso está além do escopo deste texto²², o que não incompatibiliza as reflexões aqui desenvolvidas de serem aplicadas em uma IA autoconsciente, quando esses modelos existirem.

Dessa maneira, a IA atuará conforme os dados e linhas de código que são inseridos em seus algoritmos. Algoritmos esses que são virtualmente e literalmente neutros em princípio, já que não podem ser postos diante a sistemas éticos de valores por suas consequências, pois essas consequências são apenas resultados de má execução na escrita dos códigos, isto é, produtos da humanidade.

Mas essa neutralidade por princípio é questionada por diversos autores, pois as tecnologias apresentam características do contexto em que foram criadas. A ideia comumente explanada de que a tecnologia é neutra tal qual uma faca, e pode ser usada para o bem ao cortar alimentos e salva a todos, ao mesmo tempo em que pode fazer o mal por degolar inocentes é uma suposição instrumental e de certa forma ingênua. Segundo Rubens Alves, a tecnologia é o resultado de relações sociais de produção dentro da sociedade e, portanto, a máquina não poderia ser outra coisa senão “um instrumento que perpetua um mundo em que os donos dos meios de produção exploram os deserdados” (Alves, 1968, p. 14). Conforme esclarecem Novaes e Dagnino (2004), a ingenuidade contida no conceito de neutralidade técnica descrito acima seria considerar a tecnologia em termos de meios, como se a ferramenta fosse “ser ativada a cada momento por decisões livres e sempre novas do homem”. Mas de acordo com Alves (1968), “o que caracteriza a tecnologia é que os meios são, em si mesmos, os fins”, isto é, seu objetivo final. Logo, a ênfase nas mais diversas possibilidades que compõem as ações possíveis de uma tecnologia ou ferramenta não deve sobressair ao fato de que todas elas são possíveis pela própria existência da ferramenta. É o funcionamento dos meios, e não o produto o que realmente conta. Dessa forma, o entendimento de que a tecnologia seria produto de uma a racionalidade técnica, sendo influenciada também pela racionalidade política da sociedade que permitiu a sua criação, abre margens para uma posição crítica sobre os temas aqui trabalhados. Como estabelece Feenberg (2002), os valores de um sistema social específico e os interesses da classe dominante se instalam no desenho das máquinas e em outros supostos procedimentos racionais.

²² Para aprofundamento no tema e entender porque ainda não é viável sustentar a existência de uma singularidade tecnológica ou uma IA forte, ver artigo de Joelma Marques (2021). Também, no bojo desta discussão, poderá ser verificado o conceito de *autopsicobots*, formulado por Siqueira-Batista e colaboradores.

De toda maneira, retomando a questão da culpabilização, os possíveis erros decorrem geralmente por conta da má limpeza dos dados usados ou da não criação de implementações algorítmicas justas no código da máquina, em concordância com a maioria dos estudos já feitos na área, conforme apontam os pesquisadores Peres *et al.* (2021). Por outro lado, caso quisermos passar a IA por algum sistema ético consequencialista, teríamos que assumir que as máquinas estão reproduzindo, espelhando e reverberando valores, práticas e crenças induzidas por seus criadores humanos. Nesse caso, a dita neutralidade tecnológica seria, decerto, avariada, e a culpa não poderia recair sob os atores da ação (IA), mas sim da instituição/equipe envolvida na criação da máquina. A suposta neutralidade e imparcialidade não se funda em princípio com a suposição ingênua nem nos desdobramentos, por conta das consequências. Assim, embora a IA ainda não possa ser culpabilizada diretamente, seus efeitos danosos já são sentidos por certos grupos, com um novo instrumento perpetuando antigos problemas. Essa situação por si já justifica a necessidade de se estudar sobre a colonialidade presente em inteligências artificiais.

5 IA DATIFICAÇÃO DA VIDA?

Neste capítulo será explorada a confluência entre conceitos fundamentais dos estudos anticoloniais e os novos aspectos do capitalismo contemporâneo buscando evidenciar singularidades das práticas tecnológicas no campo da saúde. Essa unidade tem papel fundamental pois lança as bases teóricas para compreensão aprofundada do tema central deste estudo.

Ao abordar o conceito de colonialidade, explorarei características estruturantes que foram implementadas no período colonial e se mantém até os dias atuais, conservando opressões e dinâmicas de poder na sociedade, segundo autores Quijano e Mignolo. Logo após, examinaremos os conceitos de dataísmo e solucionismo tecnológico, que incentivam a crença na tecnologia como solução viável para quaisquer desafios e que acabam ignorando e promovendo problemas. Uma das questões que será evidenciada com maior propriedade será o racismo algorítmico, que será outro foco, destacando preconceitos raciais perpetuados através das tecnologias, e que no âmbito da saúde exageram uma disparidade já existente. Esses elementos aparentemente sem conexão direta, parecem se aglutinar e convergir no que de Laymert dos Santos chama de virada cibernética para o capitalismo informacional, mudança econômica que valora e rege os dados incentivados pelos demais fenômenos citados acima.

Tratar sobre os conceitos de colonialidade, dataísmo tecnosolucionismo e racismo algorítmico, não obstante da virada cibernética e do desenvolvimento do capitalismo informacional, ajuda a compreender melhor como tais fenômenos influenciam e moldam em certa medida a execução da medicina na contemporaneidade. Com isso combinado, podemos dar um real início ao detalhamento e explicação de cada um dos itens supracitados, procedendo conforme o descrito.

5.1 A ESTRUTURA OCULTA

O sociólogo Anibal Quijano estudou a fundo certos problemas sociais persistentes e notou que é algo que atinge diversos países, a partir da construção geopolítica do globo desde a modernidade. Quijano (2005, p. 124) afirma que a invasão e colonização da América causou uma “mudança histórica verdadeiramente enorme e que não afeta somente a Europa, mas o conjunto do mundo. [...] Trata-se da mudança do mundo como tal”. O resultado dessa mudança, segundo o autor, foi uma estruturação do sistema-mundo baseado em “três elementos centrais que afetam a vida cotidiana da totalidade da população mundial: a

colonialidade do poder, o capitalismo e o eurocentrismo²³”. Assim, a colonialidade, conforme define Mignolo (2005, p. 38), “é constitutiva da modernidade, e não derivativa” dela, sendo o lado obscuro e nem tanto percebível quando nos referimos a essa época geralmente recordada nos livros por tantos avanços para a “humanidade”, sendo uma espécie de estrutura de poder oculta. Decerto, a humanidade retratada nos livros não é a mesma que passou pelas muitas faces da colonização, como elucida Lino Neves (2015, p. 282):

A colonização tem muitas faces, e muitos nomes, genocídio, extermínio, epistemicídio, violência, agressão, violação, invasão, destruição, escravidão, evangelização, aculturação, assimilação, integração, subordinação, subalternização, discriminação, dizimação, anulação da expressão cultural, desestruturação social, desestruturação política, desestruturação étnica, perda de identidade, perda de integridade, perda de território, sofrimento sem fim, preconceitos sempre renovados.

O conceito de colonialidade nomeia, de acordo com Mignolo (2017, p. 2) a “lógica subjacente da fundação e do desdobramento da civilização ocidental desde o Renascimento até hoje, da qual colonialismos históricos têm sido uma dimensão constituinte, embora minimizada”, lógica esta que transcendeu ao colonialismo e não desapareceu com a independência dos países ex-colônias. Segundo Mignolo (2005, p. 36) a colonialidade é “o eixo que organizou e continua organizando a diferença colonial”, naturalizando certos padrões nas relações de poder, como hierarquias raciais e culturais, implementando dominações de diversos tipos. Ou seja, é um termo que qualifica as dominações e opressões persistentes e remanescentes desde o início da exploração das metrópoles e é condição fundacional do capitalismo moderno, um “dos elementos constitutivos e específicos do padrão mundial do poder capitalista” (Quijano, 2010, p. 73). Grosfoguel e Bernadino-Costa (2016) afirmam que, a partir do século XVI, o colonialismo não só orquestrou a distribuição e organização do novo sistema mundo entre os conquistadores e conquistados em prol de uma construção conceitual da própria Europa, como também a modernidade em si e as enunciações decorrentes sobre controle do ser, da produção do conhecimento, do trabalho e do Estado. Assim, de acordo com os estudiosos do campo, a colonialidade pode ser vislumbrada ao observar as relações de poder, saber e ser de uma sociedade.

No entanto, ademais dessa conceituação de colonialidade, é preciso destacar que de acordo com alguns autores, estamos passando por um processo de dominação similar ao colonialismo visto séculos atrás, porém atualmente estaríamos vivendo em um colonialismo de

²³ De acordo com Quijano (2015, p. 126), eurocentrismo é uma perspectiva de conhecimento hegemônica e dominante, que suplanta outras formas de saber pré-existentes, tanto na Europa quanto no resto do mundo.

dados. Esse processo substitui as formas de controle fundamentadas apenas em poder bélico por outras formas exploração e controle, já que esse novo sistema “combina as práticas extrativistas predatórias do colonialismo histórico com os métodos abstratos de quantificação da computação” (Couldry; Mejías, 2018, p. 2, tradução nossa). Os autores ainda afirmam que todo o sistema capitalista se adaptou para a extração de dados, utilizando a estrutura construída pelo colonialismo histórico como alicerce e aprimorando seus tentáculos. Dessa forma, a correlação com o colonialismo não é apenas uma metáfora, mas sim uma nova forma de colonialismo, pois haveria uma união de várias forças, com a finalidade de se apropriar da vida humana (capturando e extraindo dados de forma massiva), erguendo para isso uma potente

infraestrutura para extração de dados (tecnológica, ainda em expansão); uma ordem (social, ainda emergente) que liga os humanos a essa infraestrutura; um sistema (econômico) construído sobre essa infraestrutura e ordem; um modelo de governança (social) que se beneficia dessa infraestrutura, ordem e sistema e trabalha para vincular cada vez mais os humanos a eles; uma racionalidade (prática) que dê sentido a cada um dos outros níveis; e, finalmente, um novo modelo de conhecimento que redefina o mundo como aquele em que essas forças juntas abrangem tudo o que há para ser conhecido da vida humana. Os dados, em suma, são o novo meio de refazer o mundo à imagem do capital (Couldry; Mejías, 2019 *apud* Damasceno; Angeli Neto, 2022. p.160).

Isto é, com a potencialização oriunda dessas seis diferentes fontes, o colonialismo de dados se apresenta como uma real ameaça quando pensamos que os dados pessoais estão sendo extraídos e transformados em valor ao ponto de mover toda essa estrutura para esta finalidade. Da mesma forma que Quijano, Damasceno e Neto (2022, p. 161) afirmam que “o colonialismo e a colonialidade de dados são instrumentos de dominação ligados à colonialidade do poder, do ser e do saber”, exemplificando a correlação entre os conceitos de colonialidade e colonialismo de dados. Dessa forma, com o avançar do texto, quando o termo colonialismo de dados for mencionado, pode-se ter como certo que estarão presentes facetas da colonialidade. Isso significa que serão tidos como termos correlatos, mesmo não tendo exatamente o mesmo significado, por sua significação ser próxima o suficiente deverá ter sua interpretação coligada. Nada obstante, podemos entender que a colonialidade é um aspecto mais amplo, que moldou o que Grosfoguel (2008) chama de sistema-mundo capitalista/patriarcal/ ocidentalocêntrico/ cristão-cêntrico/ moderno/ colonial para atender essa nova funcionalidade do capitalismo informacional: extração, conversão, armazenamento e utilização de dados.

A essa extração, exploração, mercantilização e conversão massiva de coisas e pessoas em dados, damos o nome de datificação, que se apresenta como resultado do sistema

capitalista atual, segundo os autores. Nesse mesmo sentido, a datificação da vida é a conversão de dados pessoais em mercadoria, comercializados com diversas finalidades e extraídos muitas vezes sem o consentimento do sujeito. Em outras palavras:

Nosso ponto é que, por meio da datificação, as interações sociais em todas as suas formas tornam-se um domínio no qual o poder de mercado pode ser exercido e o valor pode ser extraído, às vezes por meios extensos de trabalho e outras vezes sem qualquer atividade aparente por parte dos dominados, mas sempre sob a rubrica de uma espécie de apropriação e exploração. Apropriação e exploração são as razões pelas quais acreditamos que ainda é relevante falar sobre colonialismo na era dos dados. Não são características ultrapassadas de um sistema antigo, mas relações que definem recentemente nossa era atual (Couldry; Mejías, 2019 *apud* Damasceno; Angeli Neto, 2022, p. 163).

Nesse mesmo sentido, Feenberg (2002), ao analisar a história social da tecnologia (de maneira geral) no modo de produção capitalista, argumenta que sua construção não foi feita de forma democrática. Ele acredita que a vitória das elites capitalistas ignorou e continua a ignorar a participação de muitos grupos sociais na definição do desenho tecnológico. Dessa forma, conforme Novaes e Dagnino (2004, p. 193) a tecnologia não é neutra, na realidade “é um importante veículo para dominação cultural, controle social e concentração do poder industrial”. Com essa dominação, até seria possível transformar valores e interesses em regras; tornar procedimentos, equipamentos e artefatos em moldadores das rotinas cotidianas, pelas quais o poder e a hegemonia das elites dominantes se tornam naturais²⁴ (Feenberg, 2002, p. 15). Se há dominação, logicamente precisa-se de pelo menos dois grupos, ou duas classes: os dominantes e os dominados. Dessa forma, é uma sociedade dividida, com a opressão sendo exercida no meio dessa luta de classes entre opressores e oprimidos; senhores e serviçais; elite aristocrata e povo ralé; ou simplificada em ricos e pobres.

Dentro dessa lógica colonial, a hegemonia ocupa um papel seminal tanto no ser quanto no conhecimento produzido²⁵, o que fomenta questões principalmente aos estudos que são feitos sabendo dessa estrutura de dominação. Essa parte pode inclusive ajudar a elucidar ao leitor o motivo pelo qual o texto é escrito com o enunciador presente. Para tornar isso mais claro, Ramón Grosfoguel declara que:

Os paradigmas eurocêntricos hegemônicos que ao longo dos últimos quinhentos anos inspiraram a filosofia e as ciências ocidentais do ‘sistema-mundo patriarcal/capitalista/colonial/moderno’ assumem um ponto de vista universalista,

²⁴ Isso pode ser percebido no uso dos celulares hoje em dia. As redes sociais não ironicamente se tornaram uma dependência muito grande e com isso criaram uma demanda de consumo “orgânica”.

²⁵ Respectivamente chamados de colonialidade do ser e do saber pelos estudiosos do campo.

neutro e objetivo. Algumas intelectuais feministas chicanas e negras e também alguns estudiosos do Terceiro Mundo, [...] vieram recordar-nos que falamos sempre a partir de um determinado lugar situado nas estruturas de poder. Ninguém escapa às hierarquias de classe, sexuais, de género, espirituais, linguísticas, geográficas e raciais do ‘sistema-mundo patriarcal/capitalista/colonial/moderno’. Esta questão não tem a ver apenas com valores sociais na produção de conhecimento nem com o facto de o nosso conhecimento ser sempre parcial. O essencial aqui é o lócus da enunciação, ou seja, o lugar geopolítico e corpo-político do sujeito que fala (Grosfoguel, 2008, p. 4).

Dessa forma, de acordo com Grosfoguel (2008), a produção de conhecimento é sempre restrita e “nas ciências ocidentais, aquele que fala está sempre escondido, oculto, apagado da análise. [...] O lugar epistémico étnico-racial/sexual/de género e o sujeito enunciador encontram-se, sempre, desvinculados”. Ao deixar claro que as observações são minhas e não me ocultar nas entrelinhas durante o texto, permito que o leitor compreenda melhor o principal empenho deste estudo: lidar com a questão da colonialidade. Mas, a colonialidade e suas facetas não serão as únicas lógicas que veremos por aqui.

5.2 LÓGICAS QUE GOVERNAM

O dataísmo é a crença de que tudo pode ser convertido e reduzido a algoritmo e que o processamento de dados pode solucionar todos os problemas. Além disso, como indica o pesquisador Jose Van Dijck (2014, p. 204): “O dataísmo pressupõe confiança na objetividade de métodos quantificados, bem como na independência e integridade das instituições que utilizam esses métodos – sejam plataformas corporativas, agências governamentais ou pesquisadores acadêmicos”. Isto é, independente de quem será o utilizador dos métodos quantificados tecnológicos, quais dados serão usados ou quem criou o algoritmo, há uma pressuposição de integridade e boa moderação. O interessante dessa premissa é que de igual maneira à modernidade, o dataísmo tem um lado obscuro que geralmente fica oculto e pouco se discute sobre: o lado prejudicial dessas tecnologias e serviços de dados. Claro que há ressalvas em temas que são geralmente resgatados, mas somente para certos grupos²⁶, não como um problema capaz de afetar milhões de pessoas diferentes. O filósofo brasileiro Álvaro Viera Pinto, na década de 1970, bem antes das criações de IAs robustas, já observava uma atitude comum de “embasbacamento” a respeito da tecnologia, em que “supostamente, o ser

²⁶ O grupo das crianças é geralmente usado como primeiro grupo a ser afetado, mesmo para produtos que não visem crianças como consumidores.

humano, por meio da tecnologia, irá construir uma vida feliz para todos”²⁷ (Silva, 2013, p.848). Dessa forma, na visão do autor, começa-se a usar máquinas como objetos de adoração, e não um instrumento de transformação. Esse embasbacamento acabaria levando a população a um endeusamento e santificação moral tanto das tecnologias quanto dos “dos processos adotados pelos países centrais e, [...] nisso se abre a possibilidade de fazer da tecnologia a forma atualmente mais eficaz de dominação” (Silva, 2013, p. 848). Para ele, o necessário é alinhar a tecnologia com a realidade concreta, mas isso não é o que geralmente acontece. Dessa forma, segundo o autor, a tecnologia se converte em valor moral, sempre positivo. Isto é, de acordo com esse posicionamento do autor, todas as tecnologias (incluindo a IA) são sempre boas e não há espaço para questionar os erros cometidos por conta do uso dessas ferramentas, pois tais erros não são decorrentes das tecnologias, mas sim das circunstâncias.

Aliado a essa ideologia de santificação dos processos tecnológicos, há também o solucionismo tecnológico, observado e conceituado por Evgeny Morozov (2018, 2020), que impulsionado por uma política neoliberal normalizou como resposta padrão a todos os problemas encontrados na sociedade a ideia de que “como não há alternativas (ou tempo ou dinheiro), o melhor que podemos fazer é colocar curativos digitais sobre os danos” (Morozov, 2020) “os problemas devem ser resolvidos por meio de aplicativos, sensores e ciclos infinitos de retroalimentação – todos fornecidos por startups” (Morozov, 2018, p. 91). Essas soluções digitais, dentro dessa ideologia, seriam a única resposta possível diante de qualquer problema, seja ele “da desigualdade ao crime, do enfrentamento às doenças às mudanças climáticas” (Amaral; Salles, 2020, p. 11). Como explicam Amaral e Salles (2020, p. 10) “o problema de acreditar que a tecnologia resolverá todos os problemas humanos é que os gurus desse setor presumem que os algoritmos são a expressão da verdade e jamais se equivocam, abdicando de métodos para reforma ou retificação dos dados”. Morozov (2020) ainda diz que “é mais fácil usar a tecnologia solucionista para influenciar o comportamento individual” do que abordar as complexas questões políticas que causam os problemas e as crises.

Esses dois grandes eixos, de um lado os “tecnológicos” dataísmo, com sua presumida/santificada imparcialidade, e o solucionismo com seus recursos engenhosos e de outro a “antiga” colonialidade reproduzindo opressões, integram um dos principais aspectos da realidade capitalista contemporânea e também dessa pesquisa. Esses eixos permitem compreender como a realidade tem sido moldada de forma que a dominação política seja

²⁷ Para Pinto a tecnologia estava sujeita à algumas formas de ideologização, dentre elas, algumas almejavam mostrar que a tecnologia desenvolvida nos países centrais é benéfica para toda a humanidade, incluindo as regiões mais pobres, e que essas devem recebê-la de forma positiva.

exercida de forma mais sutil e menos evidente, pois se há máquinas no processo, não há parcialidade envolvida. A conscientização da coleta e gerenciamento de dados da população mundial é crucial para entendermos como uma IA pode acabar afetando milhões de pessoas no capitalismo neoliberal contemporâneo. Podemos assim, conceber “uma sociedade algorítmica [...] onde procedimentos do tecnoliberalismo – como classificação, pseudo-meritocracia, mensuração e vigilância – são automatizados e consolidados em caixas-pretas que permitem poucos desvios” (Silva, 2020, p. 17).

Sergio Amadeu (2021, p. 40) insiste que no caso brasileiro, há um “exemplo cabal do reforço mútuo da relação entre o neoliberalismo e a colonialidade”, que foi a entrega dos dados do MEC (Ministério da Educação) à empresa Microsoft. Parte da sutileza da colonialidade é tornar certos problemas em não questões, gerindo o que pode ser questionado ou tratado como verdade a ponto de ser questionado e o que não²⁸. O caso é apresentado pelo autor assim:

A direção do MEC decidiu entregar os dados do SiSU para serem processados na nuvem da Microsoft, chamada Azure. Ou seja, hospedou os dados do desempenho escolar de milhões de estudantes brasileiros para serem tratados na plataforma estadunidense. O principal argumento foi o do alto custo em manter esses dados em um data center próprio. [...] Dados dos estudantes que cursaram o ensino médio, como a renda familiar bruta mensal de cada um, os valores recebidos em diversos programas sociais, a nota no Enem, as médias populacionais relacionadas à cor declarada e a deficiências, entre outras informações sensíveis, foram entregues à plataforma Microsoft Azure (Amadeu, 2021, p. 40-41).

Nesse sentido, Ferreira e colegas (2020) levantam o argumento que de igual forma como acontece na educação, as empresas de tecnologia geralmente dão ênfases excessivas em soluções tecnológicas e incrementadores de eficiência, potencialmente negligenciando aspectos coletivos e importantes do campo, deslocando indicadores e parâmetros próprios da área. As empresas *tech* na área da saúde tendem a recair no mesmo problema. No artigo publicado em 2020, Gisele Ferreira e outros pesquisadores comparando essa situação da “venda” de dados de estudantes brasileiros com um caso sul-africano similar analisado pelo cientista social Kwet e alegam que a semelhança entre o que acontece no Brasil e a análise de Kwet do caso sul-africano (e provavelmente de outros países do “Sul global”) é impressionante. Em particular, quando as empresas GAFA²⁹ oferecem generosamente

²⁸ Essa prática é conhecida como epistemicídio, e conforme diz o autor, “integra o regime de verdade da colonialidade” que é normalizada pelas infraestruturas de submissão baseadas na alienação técnica. Sendo “fundamentais para o ordenamento neoliberal em uma sociedade fortemente dataficada” (Amadeu, 2021, p. 36).

²⁹ Google, Amazon, Facebook, Apple. Mas a esse respeito, Tarcízio Silva *apud* Faustino e Lippold, 2023, p.148 elucida que um dos pilares do racismo algorítmico é justamente a prática de grandes empresas de tecnologia colonizarem infraestruturas tecnológicas em alguns países menos conectados, de forma a restringir o acesso desses povos ao seu monopólio. Um exemplo famoso é a oferta de internet gratuita e de baixa qualidade pelo

tecnologias a estudantes desfavorecidos, os dados são extraídos sem impedimentos e posteriormente tratados de uma forma que torna as especificidades locais desprovidas de importância.

Em 2014, a Google ofertou a diversas universidades brasileiras seu pacote educacional de graça, propagandeando uma “inovação sem custos”, um auxílio para a “nova etapa” do armazenamento e gestão de dados universitários, a adesão foi rápida e bem facilitada. O problema é que, pouco tempo depois, em 2021, o armazenamento que era “ilimitado” passou a ser cobrado, tornando as instituições reféns pois poderiam perder os arquivos. Tal mudança afetou cerca de 80% das instituições brasileiras³⁰. O caso foi contado pelo jornalista Paulo Ribeiro, que relatou como o PROCON foi acionado e entrevistado, considerando o descumprimento da oferta e prática de publicidade enganosa, além de um contrato abusivo por parte da empresa, multando a empresa em 1 milhão de reais. Segundo Henrique Parra *et al.* (2018), essas instituições deixaram de elaborar diversos projetos de desenvolvimento de sistemas próprios de armazenamento por conta de terem se tornado dependentes da Google e terem a certeza que não precisariam economizar, já que o armazenamento era ilimitado.

Em outro caso, Diogo Moyses (2021), coordenador do Programa de Telecomunicações e Direitos Digitais do Instituto Brasileiro de Defesa do Consumidor (IDEC) destaca que os dados de 223 milhões de brasileiros foram expostos, incluindo informações sensíveis como o celular, CPF e até mesmo pontuação de crédito. Esse evento foi uma clara contravenção do código de proteção de dados no Brasil (LGPD), mas de certa forma, a falta de uma punição severa contra as empresas envolvidas no vazamento indica que o governo brasileiro não está preparado adequadamente para proteger os direitos do consumidor nesses casos.

O filósofo João Teixeira (2023) compartilha da mesma opinião dos demais pensadores citados aqui, definindo que especialmente no caso brasileiro, apesar dos avanços sociais, há muito retrocesso no sentido ético e político e, mesmo com a tentativa de retomar as rédeas nacionais para um desenvolvimento geral, não é possível fazê-lo justamente por conta da dependência criada pela relação neoliberal com o colonialismo, aprofundada nos últimos anos. O autor declara:

Nós retrocedemos muito nos últimos anos em termos políticos e em termos éticos. E agora justamente que se está tentando aqui no Brasil recuperar a “história do Brasil”, eu vejo a

Google e pelo Facebook para países com baixíssima conexão, como Gâmbia e Sri Lanka; o preço cobrado é que as pessoas só podem acessar os produtos dessas mesmas empresas em vez de terem acesso ilimitado à internet.

³⁰ O projeto Educação Vigida estima que cerca de 72% das instituições brasileiras de educação superior entregaram o armazenamento de seus e-mails para o Google e 8% para a Microsoft.

tecnologia numa espécie de contramão da história, porque nós somos só consumidores de novidades, consumidores de smartphone. Mas digamos assim a engenharia de software, etc., isso tudo não recebe atenção dos nossos governos. A pesquisa não recebe atenção, é melhor comprar os pacotes de fora. Estamos reproduzindo, e reproduzindo o colonialismo (Texeira, 2023).

5.3 DISCRIMINAÇÃO ALGORÍTMICA E OUTRAS OPRESSÕES COLETIVAS

Aprofundando no quão problemático é importar as pressuposições e ideologias embutidas nas tecnologias para outros contextos, é necessário falar das discriminações algorítmicas. Essas discriminações podem ser situações menos fatais tal qual uma tradução ou rótulos que sugerem estereótipos de gêneros ou raciais em certas profissões como também pode escalar para uma opressão baseada em um erro no sistema policial de reconhecimento facial de criminosos, os associando com pessoas não brancas. A esse respeito, deve-se ressaltar que 90,5% das pessoas presas por reconhecimento facial no Brasil são negras. Esses e outros casos de opressão podem ser encontrados na linha do tempo criada por Tarcízio Silva³¹. O autor define o conceito de discriminação algorítmica como sendo

o modo pelo qual a disposição de tecnologias e imaginários sociotécnicos em um mundo moldado pela supremacia branca realiza a ordenação algorítmica racializada de classificação social, recursos e violência em detrimento de grupos minorizados. Tal ordenação pode ser vista como uma camada adicional do racismo estrutural, que, além do mais, molda o futuro e os horizontes de relações de poder, adicionando mais opacidade sobre a exploração e a opressão global que já ocorriam desde o projeto colonial do século XVI (Silva, 2022, p. 69).

A discriminação não pode ser entendida como apenas uma forma de violência a um grupo específico, mas nessa parte iremos nos concentrar em alguns exemplos específicos para reforçar a problemática tratada. Mas a discussão sobre raça perpassará toda essa sessão, pois, como afirma Mignolo (2017) a raça é uma categoria criada na modernidade para legitimar e sustentar o imperialismo e colonialismo europeu, que por conseguinte implementou o que chamamos hoje de sociedade ocidental. É comum que por ter sido fundamento para a criação, as problemáticas sociais que nos atingem hoje acusem justamente a raça como um dos principais indicados a serem solucionados. Não se pode olvidar que em uma sociedade tecnocolonialista e neoliberal que nos reduz e transforma em mercadoria, “encontramos no racismo um elemento ideológico que diferencia o preço de cada mercadoria e, sobretudo, os

³¹ Para visualizar o mapeamento temporal de discriminação algorítmica, acesse: SILVA, Tarcízio. Mapeamento de Danos e Discriminação Algorítmica. Desvelar, 2023. Disponível em: <https://desvelar.org/casos-de-discriminacao-algoritmica/>. Acesso em: 06 jan. 2024.

critérios que definem e autorizam quais delas podem ser descartadas e quais, mesmo quando supérfluas, não são passíveis de tal redução” (Faustino; Lippold, 2023, p. 147), fortalecendo a máxima de Elza Soares de que a carne mais barata do mercado é negra. Mas, conforme foi dito, há outras formas de discriminação que levam outros aspectos como definidores de opressão: identidade e orientação sexual, deficiência, idade, religião entre outros além da opressão racial.

Ainda assim, o pesquisador Tarcízio Silva se debruçou para entender a correlação entre o funcionamento da visão computacional e o racismo algorítmico. A visão computacional pode ser definida como o desenvolvimento e replicação da visão humana usando máquinas, para que elas possam fazer e automatizar trabalhos que antes só poderiam ser feitos por olhos humanos. E desde os primeiros trabalhos sobre esse campo de estudo certos pesquisadores observaram que as máquinas reagem de forma diferente ao visualizar pessoas negras e não brancas em comparação com a população branca. Esse fato, conforme já foi esclarecido, se dá por conta da falta de treinamento dessa IA com base de dados com diversidade populacional. Mas o fato é que, em dez anos de avanços, o problema persiste, como se aparentemente não importasse para as instituições representarem bem essa porção populacional de não brancos. De forma geral, pode-se afirmar que “arquiteturas algorítmicas favorecem a performance de neutralidade cultural, mas com categorias genéricas com muito mais destaque do que termos específicos [...], onde esta ideia de neutralidade é marcada socialmente como a percepção dos desenvolvedores” (Silva, 2020, p. 436-437). Resumidamente, Tarcízio elabora a Figura 3 em que indica os problemas e vieses de casos mais famosos envolvendo a visão computacional. É importante destacar que por conta da Figura 3, podemos observar que a ausência de testes e o conjunto de dados insuficientes são as causas da maior parte dos casos analisados, indicando que essa temática não é uma das preocupações das equipes de desenvolvimento, mesmo com dez anos de diferença entre o primeiro caso analisado e o último.

Figura 3 - Casos selecionados de manifestação de viés algorítmico em visão computacional

Caso / Reação	Data	Problema/Viés Algorítmico	Causa/problema “técnico”
Reconhecimento de faces de computador da HP não reconhece usuário negro	12/2009	Desumanização; Invisibilidade	<ul style="list-style-type: none"> • Base de dados insuficiente • Ausência de testes
Google marca pessoas negras como gorilas	07/2015	Representação e associação racista; desumanização	<ul style="list-style-type: none"> • Base de dados insuficiente • Base de dados com associações racistas intencionais • Ausência de testes
Robôs interagentes não encontram rosto de mulher negra	03/2017	Desumanização; Invisibilidade	<ul style="list-style-type: none"> • Base de dados insuficiente • Ausência de testes
Faceapp embranquece pele para deixar “mais bonita” a selfie	04/2017	Representação eurocêntrica de beleza; desumanização	<ul style="list-style-type: none"> • Base de dados insuficiente • Ausência de testes
APIs não reconhecem gênero e idade de mulheres negras	02/2018	Representação eurocêntrica de gênero e idade	<ul style="list-style-type: none"> • Base de dados insuficiente • Ausência de testes
Kairos retira do ar aplicativo de “diversidade”	06/2018	Tipologia racial essencialista	<ul style="list-style-type: none"> • Tipologia centrada nos EUA
APIs de análise de expressões faciais associam emoções negativas a negros	01/2019	Percepção eurocêntrica; estereotipização	<ul style="list-style-type: none"> • Base de dados insuficiente • Ausência de testes
Google Vision confunde cabelo negro com peruca	02/2019	Reforço de apropriação cultural; desumanização	<ul style="list-style-type: none"> • Base de dados insuficiente • Base de dados com exemplos de apropriação estético-cultural • Ausência de testes
Carros autônomos tem mais chance de atropelar pessoas negras	03/2019	Desumanização; risco físico direto	<ul style="list-style-type: none"> • Base de dados insuficiente • Ausência de testes

Fonte: Tarcísio Silva, 2020, p. 436-437.

Contudo, se tratando de opressão algorítmica e racismo, é emblemático recordar o estudo conduzido por Safiya Noble³², que escrutinou as representações “atribuídas a identidades raciais e de gênero que, [...] em setembro de 2011, permitiu que o buscador do Google associasse a identidade de meninas negras (codificada na palavra-chave “Black Girls”) com conteúdo pornográfico” (Amaral; Elesbão, 2022, p.3). Amaral e Elesbão utilizam essa pesquisa buscando evidenciar como as narrativas de neutralidade técnica se contrapõem com

³² As obras e citações aqui feitas da autora serão reproduzidas e traduzidas de forma autônoma e por vezes serão traduções apoiadas na tradução já feita por Amaral e Elesbão, 2022.

as aparentes “falhas” expostas frequentemente no motor de busca mais utilizado no ocidente, da empresa Google. É importante notar que o que a empresa admite como “erros” refletem preconceitos já presentes na sociedade e, no estudo de Noble, foi percebido como os vieses racistas e sexistas estavam presentes pois:

[As] ‘falhas’ de dados impulsionadas por algoritmos costumam ser dirigidas às mulheres e às pessoas negras, apontando a racionalidade fundamental da lógica algorítmica de gestão de informações. Para tanto, ressaltam-se as formas sociais, político-econômicas e tecnológicas envolvidas nas práticas algorítmicas de classificação de informações empreendidas pela empresa, em especial a natureza eminentemente publicitária de seu negócio, e as (falsas) narrativas tecnicistas de neutralidade e objetividade que sustentam sua legitimidade perante os usuários. (Amaral; Elesbão, 2022, p. 3).

A autora entende que o SEO (*Search Engine Optimization*, isto é, otimização de motor de pesquisa) utilizado pela Google qualifica e classifica os sites para aparecerem como resultado de uma pesquisa de forma a se preocupar com os sites que podem pagar para entregar um SEO melhor, não necessariamente tendo a preocupação de que o resultado seja fidedigno ou represente bem o assunto buscado. E esse é um dos motivos para muitos resultados serem opressivos a certos grupos. Entendendo a Google como uma empresa que intermedia comercialmente os anunciantes e proprietários de site, Noble expõe que a empresa prioriza os interesses econômicos, não com a boa representação de diferentes identidades, pois “na verdade, a Google está no negócio de otimização de vendas” (Noble, 2018, p. 33). A conclusão da autora é que essas representações pejorativas pautadas no racismo e sexismo são expressão das próprias relações de poder da sociedade capitalista neoliberal, pois o resultado que é exposto quando usamos o motor de busca são oriundos de uma identidade tensionada “entre as lutas por justiça social organizadas em torno de identidades e histórias coletivas e a comercialização de tais identidades para vender produtos, serviços e ideologias em um esforço para acumular maiores lucros” (Noble, 2013 *apud* Amaral; Elesbão, 2022, p. 16).

A priorização de resultados baseados por investimento acaba por criar uma situação em que certos sites estarão em uma posição melhor classificada por terem maior financiamento, mesmo que tais sites contenham conteúdos tendenciosos e discriminatórios. O problema de terem suas representações nesses grandes motores de busca de forma a reproduzir as opressões já sofridas socialmente é que essa má representação será retroalimentada digitalmente, e além disso, como os melhores resultados são apresentados mediante maior investimento em SEO, esses grupos podem não ter condições de mudar isso por conta da falta de dinheiro ou falta de um agenciamento que formule ações coletivas. O problema da má representação nos meios

digitais não atinge apenas pessoas negras, podendo abranger a diversos grupos não pertencentes ao padrão normativo vigente na sociedade ocidentalocêntrica, cristocêntrica, patriarcal, capitalista, moderna colonial. Amaral e Elesbão fazem uma nova pesquisa no motor de busca trazendo outros grupos que potencialmente poderiam indicar resultados parecidos com os dos encontrados por Noble. Fazem isso para comparar e monitorar os resultados dada a diferença de tempo entre as pesquisas, pois a autora fez seu estudo em 2013, a dupla estava buscando oito anos depois, em 2021, tempo suficiente, no meio da tecnologia, para implementar mudanças caso fossem realmente um desejo da empresa.

Augusto Amaral e Ana Clara Elesbão pesquisaram os termos “*black girls*”, “*african american girls*”, “*american girls*”, “*american indian girls*”, “*asian girls*”, “*latina girls*” e “*hispanic girls*”. E foi devolvido a eles não mais conteúdos pornográficos em todas as opções, mas apenas duas delas continham propagandas de cunho sexista. Nos termos *asian girls*, “pelo menos quatro entre os dez sites listados eram de conteúdo pornográfico. Já no caso da página de resultados para ‘*hispanic girls*’ foram gerados anúncios de sites de encontros” (Amaral; Elesbão, 2022, p. 25). O resultado para a dupla é que

Em linhas gerais, o que se pode perceber é que os resultados encontrados corroboram os levantamentos feitos ao longo do estudo de caso proposto sobre a natureza comercial do mecanismo de busca e sobre as formas segundo as quais a Google mantém páginas de grande circulação e interesse publicitário no topo da lista de resultados, tendo em vista que, mesmo quando não foram encontrados sites pornográficos para os termos pesquisados, pode-se verificar a prevalência de páginas de conteúdo mercantilizável, mantendo o controle das representações e identidades das mulheres e meninas racializadas (Amaral; Elesbão, 2022, p. 25).

Nesse mesmo sentido, Granjeira e Almeida (2023) evidenciam como uma produtora midiática de teorias da conspiração e outras interpretações questionáveis, O Brasil Paralelo, se tornou a maior anunciante nacional em propagandas na Meta (Facebook, Instagram) e também está presente nas pesquisas no Google. O TSE chegou a condenar a produtora por “promover deliberada produção e difusão exponencial de notícias sabidamente falsas”, causando caos informacional no ano de 2022. Esse caso permite desvelar os funcionamentos não neutros dessas tecnologias. A conexão entre SEO e IA ocorre, pois, a ferramenta de SEO irá indicar os melhores resultados, que serão inseridos no sistema como os dados a serem interpretados. Se esses dados apontam justamente para o viés racista, sexista, LGBTfóbico, xenofóbico, capacitista, etarista e demais tipos de opressão e intolerância, a IA irá reconhecer eles como sendo correlações lógicas derivadas dos dados recebidos. Então ao conceber uma garota asiática (*asian girl*), por exemplo, não será errôneo para a máquina associá-la a estereótipos

unicamente sexuais, já que, dentro da maioria dos dados recebidos, está propriamente indicações de pornografia. Apenas para um exemplo comparativo com outro grupo que não envolve racialidade e nem gênero, o algoritmo de análise e aprovação de procedimentos de saúde do maior plano de saúde dos EUA, UnitedHealth, falhou considerando os problemas de uma população de idosos, chegando a negar 90% dos pedidos de procedimentos de maneira errônea³³. O caso foi para o tribunal e evidencia o despreparo mesmo de grandes empresas para lidar com essas minorias. Por essa razão, um maior cuidado na etapa de limpeza de dados é essencial ao almejar uma IA justa e livre de práticas discriminatórias e violentas para certos grupos.

A base de dados utilizada também é um ponto a se preocupar, e muito. Tarcízio Silva (2022, p. 72-101), ao elaborar seu livro sobre o racismo algorítmico, elucida de que forma os estereótipos preconcebidos sobre uma população ou grupo podem acabar moldando as categorias e etiquetas que são utilizadas ao se formular as bases de dados que treinam as IAs. Citando um estudo em uma base de dados visual, isto é, com imagens, o autor demonstra que muitas vezes os resultados obtidos podem ser problemáticos, ilógicos e até mesmo cruéis. Os aspectos que mais se encaixam nos padrões, são melhores representados, enquanto os grupos mais “marginais” têm sua representação feita de forma questionável ou pejorativa. O projeto de visão computacional ImageNet Roulette recebeu uma *selfie* de uma asiática e aplicou como rótulos alguns termos pejorativos comuns no inglês aos asiáticos, como por exemplo “olho puxado”. Ao passo que aspectos mais “padrões” têm uma identificação mais adequada porque

Em sociedades ocidentais, características “desejáveis”, “positivas”, “normais” e modos de existência são construídos e mantidos através de alinhamentos com a narrativa dominante, dando vantagens àqueles que se encaixam no status quo. Grupos e indivíduos nas margens, por outro lado, são frequentemente percebidos como os “pontos fora da curva” ou “desviantes”. Classificação de imagens e práticas de etiquetamento, sem as precauções necessárias e consciência destas histórias problemáticas, selecionam tais estereótipos e preconceitos e os perpetuam (Vinay Uday Prabhu, 2020, p. 6 *apud* Silva, 2022, p. 95).

Assim, o algoritmo representa e reflete de forma absolutamente justa e fidedigna o que acontece no mundo hoje, através dos dados que são incluídos. O mundo e conseqüentemente os dados retirados dele e postos nas máquinas terão sequelas coloniais. O algoritmo, em si, até pode ser totalmente neutro, considerando o modelo de aprendizado, os dados inclusos no treinamento, as suposições teóricas e assim por diante, mas essas condições são limitadas e

³³ MOLE, Beth. UnitedHealth uses AI model with 90% error rate to deny care, lawsuit alleges. **Arstechnica** 2023. Disponível em: <https://encurtador.com.br/fuwE3>. Acesso: 28 fev. 2024.

definidas por humanos, que podem e diluem seus vieses e preconceitos no resultado da análise, no cerne do AM, mesmo de forma não intencional. Não é sem razão, dado que os treinamentos muitas vezes são fomentados através de estruturas online para suprir os bancos de dados, e a:

Internet Ocidental, como estrutura social, representa e mantém a cultura branca, masculina, burguesa, heterossexual e cristã por meio de seu conteúdo. Essas ideologias são mediadas de forma translúcida pelo design do navegador [da Internet] e pelas práticas de informação concomitantes. Usuários, provedores de conteúdo, formuladores de políticas e designers anglófonos imprimem suas estruturas raciais em suas experiências na Internet, reproduzindo as dinâmicas raciais através do meio eletrônico e, ao mesmo tempo, redistribuindo recursos culturais ao longo de linhas raciais. Essas práticas recriam nitidamente a dinâmica social online que reflete os padrões offline de interação racial, marginalizando mulheres e pessoas de cor (Brock, 2011 *apud* Amaral; Elesbão, 2022, p. 6).

Essa seria uma das diversas entrelinhas que expõe a relação entre a branquitude³⁴ e o capital, o que evidencia a pouca ou nenhuma possibilidade de utilização dos modelos de IA sem a adaptação aos contextos, necessidades e limitações específicas de países que não se encaixam nos padrões ocidentalocêntricos, patriarcais, capitalistas, modernos e coloniais. A relação é possível de ser observada no Vale do Silício, já que, “nos mercados e ambientes produtivos de tecnologia de ponta [...], tanto a diversidade dos corpos quanto a das epistemes são raras, o que produz impactos materiais e simbólicos nos sistemas usados por grande parte das populações mundiais” (Silva, 2022, p. 96). O epistemicídio, isto é, a destruição ou silenciamento de outros conhecimentos não aplicados ao padrão também são problemáticos e poderiam ser explorados como uma forma de violência ainda mais impactante, dada a escassez de modelos inteligentes desenvolvidos do zero e criados em locais que efetivamente levam em conta aspectos e conhecimentos específicos do lugar.

Na expectativa de ampliar o entendimento do que pensa a respeito das discriminações algorítmicas, Mohamed, Marie-Therese e Isaac (2020) explanam sobre como as discriminações algorítmicas não ficam apenas no campo do enviesamento e na reprodução de forças coloniais por máquinas, sendo essa um pequeno elemento de um todo muito maior. Os autores definem como exploração algorítmica o uso antiético das tecnologias para explorar grupos, instituições ou indivíduos. E nessa prática podem ser diferenciados outros elementos, como os testes betas e a evasão ética para o treinamento de IA em locais em que a legislação é omissa; o trabalho fantasma feito para alimentar esses modelos inteligentes; e até mesmo a desapropriação algorítmica, que se refere ao monopólio de certas técnicas tecnológicas

³⁴ Segundo Bento, Maria A S, (2002, p. 7 *apud* Silva, 2022, p. 80), a branquitude é “um lugar de privilégio racial, econômico e político, no qual a racialidade, não nomeada como tal, carregada de valores, de experiências, de identificações afetivas, acaba por definir a sociedade”.

avançadas por parte de uma minoria econômica em detrimento de muitas outras partes populacionais.

A prática desse tipo de exploração se daria de forma bastante comum, infelizmente. O exemplo que temos maior quantidade de evidência é o trabalho fantasma (*ghost work*)³⁵, em que trabalhadores de países pobres são contratados por empresas multinacionais para rotular e classificar dados, auxiliando no treinamento de IA, o problema é o uso de mão de obra barata que reproduz as relações econômicas históricas entre os países. De acordo com a pesquisa de Mohamed e colaboradores (2020, p. 669), antigas “colônias” do EUA e dos ingleses³⁶ como Filipinas, Quênia e Índia e Uganda voltaram a estabelecer contatos “exploratórios” por meio dessa espécie de trabalho invisível que é necessário para fundamentar a inovação da IA.

Esse trabalho fantasma pode ser entendido como uma forma de evasão ética (*ethics dumping*), que significa, neste caso, a exportação empresarial de práticas antiéticas³⁷, procurando países com uma legislação mais frouxa para atuar, geralmente em populações mais pobres e marginalizadas. A execução de testes betas, o que significa as primeiras testagens de novas tecnologias, é por vezes feita com um agrupamento diferente da população-alvo almejada com o treinamento da IA. A nível de exemplo, o conhecido e supracitado caso da Cambridge Analytica realizou testes betas de seus algoritmos nas eleições nigerianas de 2015 e no Quênia de 2017, muito tempo antes de os utilizar nos EUA e no Reino Unido. Conforme indicam Mohamed e amigos (2020, p. 668) essas testagens refletem bem o tipo de tratamento histórico que o Império Britânico (e outras Metrôpoles) deram para suas colônias, as tratando apenas como laboratórios para novos medicamentos e tecnologias.

É importante observar o alinhamento dessas ações de discriminação algorítmica com as antigas práticas do colonialismo, fortalecendo a tese do colonialismo de dados. Isso configura uma falha internacional em proteger os cidadãos, apontando justamente a necessidade de implementação de legislações internacionais que se complementem. Para se proteger e protegerem os cidadãos os sistemas políticos, principalmente de países subdesenvolvidos, deveriam se atentar para essa diferenciação tão fundamental na hora de implementar tecnologias estrangeiras.

³⁵ Não estando restrito sobre a IA, mas de forma similar podemos citar as plataformas que atuam em muitos países de maneira exploratória, compartilhando uma pequena taxa dos lucros obtidos com os trabalhadores.

³⁶ Isto é, por dominarem o inglês conseguem servir bem aos propósitos expansionistas desses dois países dentro das pesquisas de IA. Ver mais em: Smink, 2023.

³⁷ Ou seja, no país de origem é considerado ilegal ou incorreto, mas em outros países já não é mais preciso respeitar.

Mas na realidade, não se pode entender sempre esses problemas como apenas oriundos do mau funcionamento da equipe criadora da IA, porque como Judy Goldsmith e Olivia Erdélyi argumentam, as razões são multifacetadas e por conta disso se aglutinam aspectos diferentes e esferas diferentes da sociedade são mobilizadas no processo para reagir às consequências de uma ação “robotizada”.

Por vezes, estes problemas resultam da imperfeição das aplicações de IA, como quando os sistemas de IA produzem preconceitos discriminatórios. Outras vezes, surgem problemas porque a IA faz o seu trabalho com perfeição, como evidenciado pela crescente ameaça à privacidade representada pelas aplicações de padrões de reconhecimento [...]Algumas instâncias de IA são eticamente questionáveis (por exemplo, robôs sexuais infantis [...]), potencialmente perigosas (por exemplo, decisões autônomas de morte por máquinas) ou levantam desafios sistêmicos mais amplos (por exemplo, questões de desigualdade e deslocamento de trabalho e a impugnação de paradigmas éticos, legais e sociais já existentes) (Goldsmith; Erdélyi, 2018, p. 246, tradução nossa).

Mas independentemente do resultado sobre quais forem as razões por trás, tais situações de violência continuam a ocorrer e conforme mais se avança na sociedade neoliberal baseada no colonialismo de dados, mais fica iminente o surgimento de outro modelo inteligente que utilize de discriminações e técnicas problemáticas em grupos sociais. Como alega Tarcízio Silva (2024), no caso brasileiro, há pesadas vulnerabilidades na defesa de direitos humanos e altos índices de desigualdade, racismo, violência de gênero e LGBTfobia, e por isso, a realidade concreta dessa violência não pode ser ignorada ao planejar sistemas e mecanismos de controle social da tecnologia. Dito isso, temos algo a fazer sobre e contra isso?

5.4 A GUERRA PERDIDA PARA O CAPITALISMO INFORMACIONAL?

Um ponto que merece especial atenção é a exposição dessas inteligências à quantidade de dados disponíveis nas redes. A internet possibilita uma infinidade de recursos, inclusive sendo eles nocivos para o AM³⁸. Se as fantasiosas saídas do solucionismo viraram possíveis foram devido a grandiosa influência que os dados e a tecnologia têm em nosso cotidiano, pois no atual período do capitalismo, os dados são o que comandam. Não à toa, alguns chamam a época de Era da informação. Efetivamente, conforme demonstra a obra de Laymert Garcia dos Santos (2013), houve uma virada cibernética na forma de funcionamento do mundo

³⁸ Nocivos no sentido de terem o comportamento esperado alterado por conta da interação com o público, por exemplo, como foi com o SimSimi, IA que foi criada para conversas saudáveis, porém começou a ameaçar sequestros e assassinatos após a interação. Mas a responsabilidade de limitar esses comportamentos é dos desenvolvedores, não da IA. Entenda: App de bate-papo SimSimi é suspenso no Brasil após “aprender” respostas impróprias. Disponível em: <https://11nk.dev/YKWr3>. Acesso em: 21 maio 2023.

contemporâneo. O conceito de virada cibernética surgiu como uma crítica ao conceito de virada cultural, mas levando em conta principalmente a influência tecnológica e científica no capitalismo contemporâneo:

Desde a década de 1970, mas principalmente nos anos mais recentes, tem se firmado a tese segundo a qual o capitalismo estaria se transfigurando ao incorporar a dimensão da cultura ao processo de produção e até mesmo ao fazer dela o motor da acumulação. [...] A discussão sobre a virada cultural operada pelo capitalismo contemporâneo é muito instigante. Tendo, porém, a achar que ela tem um alcance limitado por não considerar a centralidade das tecnologias da informação no processo em curso – mesmo conferindo importância a essas tecnologias, o pensamento sobre a virada cultural vê o papel destas como apenas um dos fatores da transformação. [...] Mais importante do que a transformação desta em mercadoria parece **ser a “virada cibernética”, que selou a aliança entre o capital e a ciência e a tecnologia**, e conferiu à tecnociência a função de motor de uma acumulação que vai tomar todo o mundo existente como matéria-prima à disposição do trabalho tecnocientífico (Santos, 2013, p. 9-10, grifos nossos).

Eduardo Mariutti (2020, p. 2), ao escrever sobre o conceito de virada cibernética, explicita que o processo é composto por uma faceta “menos nítida”, que “envolve a penetração do capital na dimensão molecular da realidade, isto é, no nível da informação”. Nesse sentido, o mundo todo se converte em mercadorias, que podem e são codificadas para serem traduzidas em informações, dados e códigos. Na mesma página, Mariutti continua afirmando que apesar dessa aliança ter desencadeado “uma avalanche tecnológica que rompe as fronteiras sociotécnicas existentes e, ao mesmo tempo, possibilita a fusão de linhagens tecnológicas que até então caminhavam separadamente”. Tal avalanche não liberou o potencial gigantesco que possuía em suas possibilidades emancipatórias dos sujeitos, na realidade essa potência é tolhida por conta de que “a virada cibernética permanece instrumentalizada e posta a serviço da valorização do capital. O esforço do capitalismo de ponta é, portanto, traduzir, controlar e se apropriar das informações que constituem o homem, a natureza e a cultura”.

Portanto, o funcionamento dessa “nova” fase do capitalismo, o que chamarei aqui de capitalismo informacional³⁹, seria novamente predatória, de forma até mais intensa, “porém quase invisível do homem e da natureza. As novas formas de pilhagem – amparadas na tecnociência e nos novos sistemas de propriedade intelectual, que adentraram no nível das informações genéticas e no patenteamento de seres vivos” (Mariutti, 2020, p. 3). Mariutti (2020, p. 5) completa para dizer que essa pilhagem apesar de mais intensa, é “uma pilhagem

³⁹ A definição de acordo com a síntese de Laymert (2013, p. 84) sobre o tema, diz que “levando a instrumentalização ao extremo, tal estratégia considera tudo o que existe ou existiu como matéria-prima a ser processada por uma tecnologia que lhe agrega valor. Nesse sentido, a única ‘coisa’ que conta na nova ordem é o que pode ser capturado da realidade e traduzido numa nova configuração. A única ‘coisa’ que conta é a informação”.

muito mais sutil do que fora a colonização do novo mundo e a conquista da África. Mas, mesmo assim, uma pilhagem”.

Assim, ao raciocinar a partir dessa interpretação, já que tudo a volta só vale pelo tanto de que se gera ao ser codificado, é fácil chegar à conclusão que a vida humana, independente de quão complexa seja, será e já está sendo tratada como um recurso, um dado. Como afirma Cassino (2021, p. 27): “Nossa vida social tornou-se um recurso que pode ser extraído e utilizado pelo capital como forma de acumulação de riquezas. Tanto populações do Norte Global quanto do Sul passaram a ser fontes de informações que alicerçam o capitalismo”. Ao fim, “tudo gera dados capturáveis” e por conta disso os dados se tornaram um recurso lucrativo. Então, a vida agora é datificada, transformada em bits e fazem parte de um sistema maior que gera lucro.

Apesar de haver exploração e predação em qualquer parte do globo, no mesmo compasso dos avanços tecnológicos, os países que detinham capital econômico para investir nas novas tecnologias capazes de gerir tanta informação continuam a dominar os demais por meio da concentração e manipulação dos dados obtidos. Em países da América Latina ou do Sul Global, bem como em países que tentam resistir ou não se submeter ao domínio euronidense⁴⁰, “pelo colonialismo, são obrigados a europeizar-se ou a americanizar-se pela tecnificação” (Corbisier, 1959). João Teixeira ainda ressalta que a globalização e a internet criaram uma falsa ideia de conectividade e participação igualitária entre os países do globo, no entanto, os polos tecnológicos são todos mantidos de forma insulares em poucas regiões e inacessíveis para o resto do mundo. O autor vê essa situação com péssimos olhos, declarando que é uma guerra perdida para nós, os periféricos do capitalismo informacional, como podemos ver abaixo:

Nós vivemos uma época que a tecnologia tá toda encapsulada, ela é totalmente insular ali naquela região do vale do silício e outras poucas regiões. E a coisa talvez mais errada que as pessoas possam dizer é que a internet integra as pessoas, não há uma coisa que exclua mais as pessoas do que a internet porque nós estamos acoplados a periferia dessas máquinas que estão lá. Nós só somos uma ponta e essa ponta se relaciona com uma tecnologia que é uma caixa preta ninguém sabe como funciona [...] Nós somos só a ponta e é uma ponta que exclui, **somos os excluídos, os periféricos, os marginalizados.** [...] **Esta é uma guerra perdida, é uma guerra que nós já perdemos,** todos os nossos dados foram transferidos para essas empresas insulares. **Foi uma guerra onde não se derramou um pingão, uma gota de sangue,** mas eu acredito que tenha sido a pior revolução, uma das piores revoluções já acontecidas durante a história da humanidade (Teixeira, 2023, grifos nossos).

⁴⁰ Europeus e estadunidenses.

Apesar da visão de Texeira poder ser taxada de pessimista por vislumbrar essa guerra como perdida, efetivamente tivemos poucos avanços em relação ao cuidado com os dados e a não reprodução de opressões por meio das máquinas. Tais avanços são ainda mais insignificantes se comparado com o salto tecnológico sofridos nos últimos cinco anos, os incidentes de discriminação algorítmica são frequentes e diversos, como evidenciado na linha do tempo supracitada (ver nota 31). Ao que parece, independente da solidificação e robustez do desenvolvimento das IAs, a resolução de problemas discriminatórios não está em pauta dentro da agenda das empresas. Um dos caminhos viáveis para responder os casos de opressão via IA seria entender como funcionam as “pontas soltas” que usam de dados inconclusos, incompletos ou mesmo preconceituosos que podem ter sido utilizados para escrever um software, o problema, no entanto, é que a complexidade escala de forma inversamente proporcional à explicabilidade da inteligência, conforme já mencionado.

Contudo, mesmo com o esforço dos estudiosos e especialistas, não são todos os modelos de IA que passam pelo processo de explicabilidade publicamente. Na realidade, é o extremo contrário, principalmente porque os sistemas de IA mais usados pelas empresas são patenteadas e têm seu “funcionamento” guardado por lei de propriedade intelectual⁴¹. No entanto, mesmo não entendendo a fundo seu funcionamento, muitos estados/empresas vendem os dados de seus cidadãos e clientes às corporações, não se importando com qual tipo de tratamento é dado às pessoas, agora, “codificadas”. A ligação dessas poderosas forças de aprendizado (IA) aos dados de milhões de pessoas, categorizando-as⁴² é premente ao se levar em conta a criação da IA e por qual motivo devemos ter cuidado e priorizar o desenvolvimento de ferramentas éticas para não impactar negativamente a vida dessas pessoas. Não devemos esquecer que a IA pode ser usada para nos vender qualquer coisa, enquanto transformam um banco de informações sensíveis em uma mercadoria valiosa⁴³.

De toda forma, dentro da sociedade as inteligências artificiais estão emergindo veloz e irreprimivelmente e estão transformando totalmente o modo de funcionamento do mundo até então conhecido. O que abre diversas possibilidades e o dobro de riscos. No campo da saúde, não difere: a IA tem a potência de auxiliar as decisões médicas e acelerar absurdamente o

⁴¹ Isto é, já que não se pode explicar plenamente o funcionamento, o “segredo” está nos dados, treinamento e a técnica que levou ao aprimoramento da IA para chegar a tais resultados.

⁴² Elen Nas (2020, p. 59) diz que: “Classificar, em si, é discriminar. Classificações são objetos de poder”.

⁴³ O armazenamento de dados não implica na sua venda, porém guardamos apenas o que nos é valioso. Como vimos, no capitalismo informacional, os dados são o que importam e as empresas sabem usar disso para lucrar. Não podemos esquecer do escândalo da Cambridge Analytica, que comprou informações coletadas de cerca de 50 milhões de usuários do Facebook, que foram usadas para criar um sistema que previa e influenciava as escolhas dos eleitores nas urnas durante a campanha de Donald Trump em 2016, “vendendo” um candidato perfeito a presidente. Ver: BBC, 2018.

atendimento e diagnóstico. A única diferença é que neste campo em específico, os dilemas e desafios éticos são intensificados, dado que envolvem diretamente o poder de salvar a população da morte ou acelerar seu encontro com ela.

6 O (AB)USO DA IA NA SAÚDE

Neste terceiro capítulo, iremos aprofundar as interseções e problemáticas já mencionadas e tensionadas anteriormente, discutindo-as na área da saúde com auxílio de duas vertentes bioéticas. A unidade começará apresentando a bioética de proteção e a bioética de intervenção, que ajudarão na explanação sobre os usos e abusos tecnológicos em contextos médicos. Esses usos serão evidenciados questionando o conceito de dados médicos, bem como o comércio da saúde como um produto qualquer.

Devido às intensas mudanças alçadas no campo da saúde pelo avanço tecnológico, alguns problemas como a monetização da saúde ficaram ainda mais evidentes, fazendo com que as disparidades econômicas se tornassem um impeditivo maior para acessar serviços que, por vezes, são facilitados pela tecnologia. Além disso, com tais avanços se tornou ainda mais fácil coletar, armazenar, utilizar e até mesmo divulgar os dados médicos de um sujeito, fragilizando uma pessoa que já está em estado de paciente e comprometendo a privacidade e autonomia do indivíduo. Assim, essa unidade busca examinar os conceitos supracitados para ampliar certas preocupações éticas relacionadas ao uso da tecnologia na saúde. Abaixo vemos, na Figura 4, de forma simplificada algumas interações entre diferentes áreas da medicina e o uso da IA, bem como as “preocupações éticas e legais e uso da IA em diferentes áreas da medicina. Deve-se notar que a aplicação da IA, as preocupações éticas e as preocupações legais podem abranger mais de uma ‘área da medicina’”:

Figura 4 - Algumas aplicações e preocupações relativas à IA

	Áreas da Medicina	Usos de IA	Preocupação ética	Preocupação jurídica
1	Radiologia diagnóstica	Suporte clínico para avaliação de imagens médicas	Algoritmos enviesados	Responsabilidade social**
2	Patologia	Precisão aprimorada	Decisão tendenciosa	Responsabilidade legal**
3	Radioterapia	Tratamento melhorado, redução do tempo de planejamento e de radioterapia	Falta de Transparência	Privacidade de dados e segurança
4	Medicina Clínica	Registro eletrônico de saúde	Acessibilidade	Propriedade intelectual
5	Pediatria	Modelagem preditiva	Viés discriminatório	Regulamentação
6	Cirúrgica	Cirurgia robótica	Segurança	Regulamentos da FDA*
7	Medicina de emergência	Gerenciamento de medicamentos	O melhor interesse da criança*	Qualidade do cuidado
8	Imunolergologia	Identificação de risco	Impacto nos trabalhadores	Adequação na HIPAA*
9	Câncer de próstata	Otimização e desenvolvimento de medicamentos	Autonomia	Regulamentação de ensaios clínicos
10	Urologia	Identificação do tratamento mais eficaz	Relação médico-paciente	Má conduta

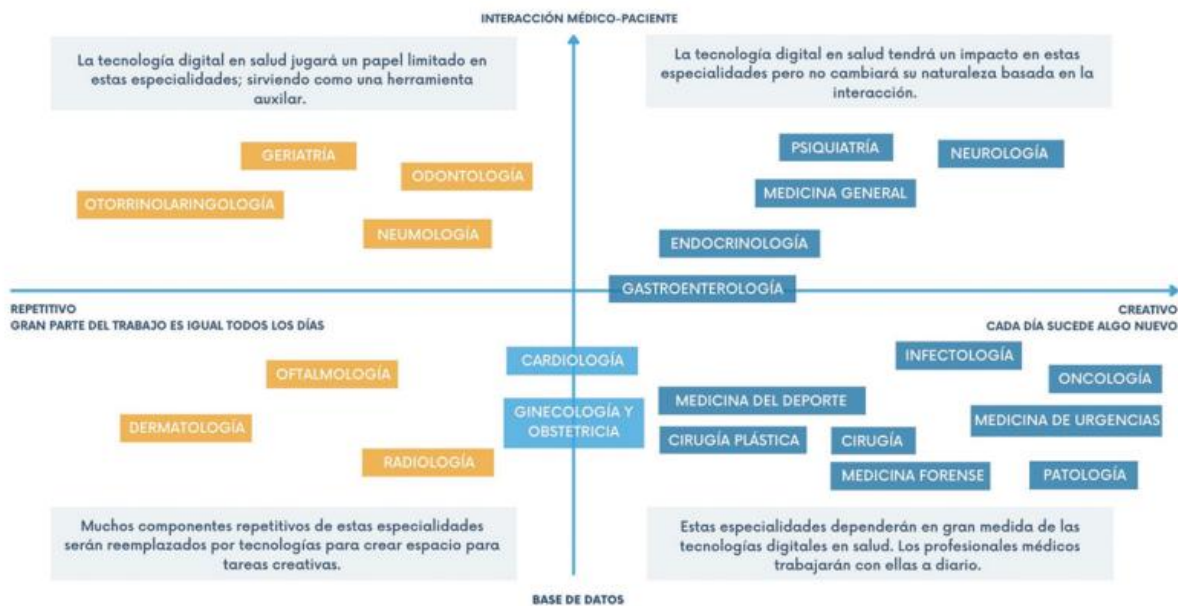
44:* 45:*:*

Fonte: Adaptado e traduzido de Debesh *et al.*, 2023, p. 2.

⁴⁴ Essas categorias são muito específicas do contexto estadunidense, mas tem seus correlatos nacionais, portanto, em uma análise crítica desse quadro poderíamos entender que o respeito à autonomia do paciente, a atenção as regras contidas na ANVISA, ANS e a LGPD, trariam um pouco do contexto brasileiro para a imagem.

⁴⁵ Essas traduções podem alterar um pouco do sentido original. No português não temos uma distinção tão marcante como acontece no inglês, com diferenças substanciais entre *Accountability* e *Liability*, respectivamente.

Figura 5 - O uso e dependência de IA em diferentes áreas da medicina



Fonte: Lanzagorta-Ortega, *et al.*, 2022, p. 19.

É importante notar os textos contidos em cada quadrante da Figura 5, explicitando cada uma das apostas dos autores para as especialidades médicas. De acordo com essas descrições, poderíamos entender o quadro dividido da seguinte forma: as especialidades na parte inferior da Figura 5 sofrerão maior influência e adotarão em maior quantidade o uso de IA, que se diferenciam das especialidades da parte superior. A divisão de lados pode ser interpretada como a quantidade de novos desafios nas áreas específicas, isto é, os campos à esquerda enfrentam os mesmos desafios na maior parte do tempo, já os profissionais da direita encaram novos problemas corriqueiramente.

Um aspecto curioso que pode ser ressaltado é que, de acordo com os autores, áreas como a oftalmologia, dermatologia e radiologia sofrerão um intenso impacto das tecnologias de inteligência artificial, fazendo com que os profissionais futuramente possam até mesmo disputar lugar com esses sistemas modernos. Por outro lado, a oncologia, a infectologia e a cirurgia plástica terão apenas um suporte da IA, implementando o uso desses sistemas cotidianamente. Essa indicação de que a IA poderia substituir médicos devido à repetitividade das ações médicas contraria a opinião da maioria dos pesquisadores do campo de IA na saúde. Esses especialistas veem os sistemas de IA apenas como um apoio ao trabalho médico, sem a intenção de eliminar os profissionais de saúde ou comprometer a relação médico-paciente em situações em que o paciente interaja apenas com um robô.

6.1 POR UMA VIDA NÃO PROGRAMADA

Como tratarei diretamente com desafios bioéticos nessa etapa da pesquisa, contarei com o auxílio de dois sistemas bioéticos: Bioética de Intervenção (BI) e Bioética de Proteção (BP) para embasar as observações.

A princípio, a Bioética de Intervenção pode ser definida como uma “ferramenta de denúncia, reflexão e busca de alternativas para a solução de problemas (bio)éticos que aparecem em um contexto típico das desigualdades registradas no hemisfério Sul, especialmente na América Latina, sobretudo os macroproblemas” (Flor; Garrafa, 2011, p.288). Sua utilização neste trabalho serve para aprofundar e auxiliar as análises feitas no campo da saúde, com o arcabouço já explorado da bioética, que neste caso leva em conta o contexto geopolítico e a situação diferenciada das vidas exploradas pela colonialidade, que por vezes, mediante aos parâmetros clássicos utilizados nas linhas de pesquisa da bioética tradicional, acabam tendo como resultado uma pesquisa um pouco deslocada⁴⁶. Serrano e Rampazzo (2017, p.262) esclarecem que apesar da bioética inicialmente ter sido considerada um “movimento dos Estados Unidos”, a disciplina foi modificada durante sua adoção pelos países latino americanos. Assim, “a Bioética passou a ser enciclopédica por definição, etimologicamente ‘ética da Vida’” (Serrano; Rampazzo, 2017, p. 269). E levando isso em conta, a bioética principialista de Beauchamp e Childress parece insuficiente lidando com problemas cotidianos “enfrentado pela maioria da população dos países latino-americanos, com níveis significativos de exclusão social” (Tealdi⁴⁷, 2008, p. 162). Dessa forma, utilizando de práticas, metodologias e autores contra hegemônicos, tenta-se aplicar uma intervenção em contextos em que os direitos da população esteja “em risco”. BI tem uma estrutura de direitos que:

incorpora o discurso da cidadania ampliada, cujos direitos estão além das garantias asseguradas pelo Estado. Assim, a intervenção deve ocorrer para garantir a todos os seres humanos:

- a) direitos de primeira geração (relacionados ao reconhecimento da condição de pessoa como requisito universal e exclusivo para a titularidade de direitos);
- b) direitos de segunda geração (que significam o reconhecimento dos direitos econômicos e sociais que se manifestam na dimensão material da existência), e
- c) direitos de terceira geração (que se referem principalmente à relação com o meio ambiente e à preservação dos recursos naturais) (Tealdi, 2008, p. 164).

⁴⁶ Claramente, a bioética precisa levar o contexto em consideração em suas análises. No entanto, quando aplicada ao contexto brasileiro, por exemplo, teve de ser adaptada. Segundo pesquisa de Ramos *et al.* (2019), a bioética brasileira é mais horizontalizada, mais politizada e crítica do que suas contrapartes, evidenciando o ponto de partida da Bioética da Intervenção.

⁴⁷ Todos os textos e trechos citados de Tealdi serão necessariamente traduzidos por autoria própria.

Por essa razão, podemos dizer que a bioética de intervenção (BI) é bastante alinhada com a mesma crítica feita por Foucault quando cunhou o termo biopolítica, pois ambas “estão interconectadas em seus propósitos de entender os mecanismos de poder sobre a vida que se fundam em torno de uma imagem de vida” (Flor; Garrafa, 2011, p. 292). Quijano (2005, p. 130) evidencia que toda sociedade é uma estrutura de poder, e que por conta disto, o poder vigente determina a sociedade, dado que esta última é também seu produto. Por isto, a sociedade replica os controles que foram implementados durante a colonização, isto é, a colonialidade. E assim sendo, os conceitos que temos por base e utilizamos para nossas elaborações estão ancorados e envoltos de colonialidade e por conseguintes as opressões aos vulneráveis. Portanto, precisa-se de uma bioética que esteja atenta para cenários periféricos do capitalismo global, algo característico da “bioética de intervenção em sua vinculação teórico-política com a saúde pública, da qual se origina a preocupação com as demandas de saúde relacionadas com temáticas de situações persistentes e emergentes” (Feitosa; Flor, 2015, p. 280).

Por outro lado, para Kottow e Schramm (2008, p. 20), uma característica universal de seres vivos é a vulnerabilidade, ou seja, a “condição ontológica de qualquer ser vivo”⁴⁸. Ao interpretar dessa forma, eles criaram a bioética de proteção, “com a perspectiva de proteger indivíduos e populações vulneradas e suscetíveis” (Siqueira-Batista; Possamai, 2022, p.11), tendo essa visão de erguer um princípio da proteção focado nos “indivíduos e populações de vulnerados e suscetíveis que não podem tomar decisões sozinhos, e não aos simplesmente vulneráveis, que, de fato, são todos os seres vivos” (Siqueira-Batista, 2021, p. 264). Tealdi (2008, p. 165) indica que “a ética de proteção se concebe naturalmente além da bioética”, pois “reconhece as reais necessidades dos seres humanos existentes, para os quais não há consolo na filosofia, mas na assistência”. Desse jeito, tentam partir para a ação visando estabelecer bases protetivas que permita “os excluídos, os fracos, os deficientes receberem a proteção necessária para desenvolverem as suas capacidades em liberdade”.

Há uma proposta baseada na BP que parece ser bastante adequada às investigações das inteligências artificiais, já que até mesmo incluí a IA forte e autônoma⁴⁹: a (Bio)ética para todos os seres. Este modelo, proposto por Rodrigo Siqueira-Batista, parte do conceito de quiddidade, que diz respeito à situação no mundo de todos os seres, submetido ao vir-a-ser e ao

⁴⁸ A vulnerabilidade que a BP tenta proteger é a chamada suscetibilidade, que impede os indivíduos ter uma vida digna e de qualidade; não essa condição existencial de ser potencialmente vulnerável.

⁴⁹ Contudo, esta possibilidade não será explorada neste estudo, como já dito, a IA forte está longe do escopo do texto.

deixar-de-ser; ou seja, tudo o que existe estaria inscrito em um mesmo plano de imanência de tempo e de espaço e, portanto, daí se derivaria a ideia de totalidade e de interdependência (Siqueira-Batista, 2020; Siqueira-Batista; Gomes, 2021). A quididade põe em xeque um aspecto conceitual das bioéticas clássicas que estendem sua preocupação aos diferentes grupos humanos. Nesse sentido, essa corrente bioética – outrora denominada preliminarmente como uma “bioética quântica” – abrange seu acolhimento a todos os seres, e está intimamente relacionado com a “superação do discurso ético tradicional que é marcado pelo antropocentrismo irresponsável, e à crítica do consumismo desenfreado, que é afim à lógica neoliberal do capitalismo tardio” (Siqueira-Batista, 2020, p. 272).

Devido às inúmeras contribuições dessas duas correntes bioéticas ao tema desse estudo e também à formação do bioeticista que lhes escreve, ambas serviram de apoio ao desenvolver esta dissertação, especialmente ao delimitar certas influências no transcorrer de todo trabalho, isto é, não apenas basearam as análises sobre as implicações éticas da IA na saúde, mas também guiaram uma compreensão aprofundada da problemática em voga.

As observações propositadamente limitadas em situações alarmantes e derivadas na América Latina e sul global tiveram sua origem justamente no reconhecimento de que se tratando de tecnologia, estes locais serão os últimos a terem a chance de usá-los, a menos que seja para servir como cobaias dos equipamentos mais avançados. Além de evidenciar que os conceitos criados no local podem e devem ser usados para remediar os problemas ali surgidos, assumindo a verve proativa da BI, que resultou em explorar as maneiras pelas quais os profissionais de saúde e outros colaboradores poderiam agir para mitigar os efeitos adversos da IA na saúde. Essas perspectivas foram oriundas da BI com suas constantes denúncias e reflexões buscando solucionar ou remediar “nossos” problemas reais. Mas a expansão do conceitual e preocupação com demais seres, não apenas os humanos para serem acolhidos nas pequenas soluções propostas, veio certamente da BP. A interpretação de que este trabalho poderia cruzar os limites de ser considerado apenas um produto bioético e ser usado não apenas como uma reflexão erudita acadêmica e trivial, mas como uma ferramenta para o desenvolvimento de sistemas, de forma a melhorar o aprimoramento da administração de IA fez com que esses dois eixos fossem a motriz deste trabalho.

É possível, no entanto, que nem sempre essas influências estejam evidentes no decorrer da dissertação, uma das razões para isso reside na complexidade e inovação dos tópicos abordados, que fazem com que muitas vezes se tenha que recorrer a uma abordagem multifacetada para tratar de tais tópicos. Somado a isso, deve-se também o fato de a escrita da dissertação geralmente já ter uma progressão não linear, no qual os pensamentos vão surgindo

e ganhando incorporação com o tempo, assim, as influências bioéticas podem ter ficado um tanto submergidas no decorrer do texto. Por fim, há o fato de que a temática é relativamente recente, muitos autores basilares do campo bioético não produziram muito sobre ele. Ainda assim, com olhar atento, nuances de tais influências podem ser vistas em todo o trabalho.

Contudo, cada bioética teve uma espécie de representante máximo, um interlocutor que auxiliou na criação e continuação da narrativa desse estudo. Apesar de cada bioética ter seu escopo e forma de abordagem, aqui se coadunaram para formar o que, a meu ver, são fluxos em que ambas se tangenciam e reafirmam a necessidade de atenção para alguns problemas, potenciais conflitos e pequenos bálsamos. Pelo lado da Bioética da Proteção, retomada em termos da (Bio)ética para todos os seres, Siqueira-Batista (2022) evidencia a urgência de as medidas protetivas incluírem todos os seres, até os artificialmente criados, intitulados por ele de autopsicobots, garantindo o bem estar de todos e segurança de todos os envolvidos enquanto sociedade. De outra perspectiva, temos a bioética de intervenção, que se concentra mais nos problemas persistentes oriundos do desbalanceamento de poder ocorrido durante a colonização e as grandes navegações. Nessa corrente o interlocutor será principalmente Flor (2011, 2015), que combina bem os estudos decoloniais com as problemáticas raciais e bioéticas aqui enfrentadas. Ambas figuras ajudaram com ideias e críticas no processo de elaboração desse texto.

6.2 A SAÚDE COMO MERCADORIA

Agora concentrando-se no âmbito da saúde, tópico desse capítulo, o que se pode observar é uma ampla gama dessas novas tecnologias sendo usadas e por muitas vezes melhorando a eficácia e eficiência no trabalho médico em diversos casos. E tendo diversos usos, também se possibilita os abusos na utilização da IA, com isso deve-se questionar se o atendimento mais rápido e tecnológico deve ser priorizado em detrimento do atendimento médico humanizado.

No entanto, é crucial destacar que como afirma Barreto (2017) que a saúde, sem o âmbito tecnológico, já tem um problema em ser desigual, e isso em nível global, não é uma situação que ocorre apenas em países pobres ou em desenvolvimento. Isso é, as desigualdades encontradas no campo da saúde não são desconectadas de outras vislumbradas em todos os ramos das sociedades do planeta. Se uma sociedade é desigual, a saúde que habita em uma parte dessa sociedade também será. Mas é importante destacar esse tópico por um motivo simples: se aparte dos recursos e aparatos tecnologicamente avançados e caros o campo já

apresenta dificuldades de ofertar os recursos disponíveis para toda população, com o advento de IA e demais tecnologias de aprendizagem de máquina, a possibilidade do problema se intensificar é ampliada a níveis exorbitantes. Não deve ser uma surpresa para o leitor dado o quão avançado está neste estudo, mas pode-se salientar que, neste caso, o campo da saúde sofre com a IA da mesma forma que a colonialidade, quero dizer: sendo ampliado, intensificado e aprimorado em suas virtudes e em seus defeitos e vicissitudes. Dessa maneira, é válido ressaltar o que nas palavras de Barreto é um marco para as desigualdades no campo da saúde, que é

o denominado “Black Report”, no Reino Unido. Em 1977, comissão nomeada pelo Ministro da Saúde [...], foi encarregada de analisar a existência de desigualdades em saúde, já que o sistema nacional de saúde (NHS) daquele país, que havia sido criado na década de 1940 estava fundado nos princípios de equidade e de acessibilidade universal. Uma das observações relevantes desta comissão foi de que no período desde a criação do NHS haviam ocorrido melhorias importantes nas condições de saúde da população britânica, independente de classe social [...]. Porém o encontro mais inesperado foi que os diferenciais dos níveis de saúde entre as classes sociais haviam persistido e para alguns problemas tinham se ampliado. Além disto, persistiam desigualdades no tocante à disponibilidade e uso dos serviços de saúde. [...] este documento teve um imenso impacto sobre as discussões posteriores relativas às desigualdades em saúde nos países desenvolvidos (Barreto, 2017, p. 2100).

Mauricio Barreto ainda evidencia que esse relatório foi tão influente por evidenciar um “aspecto cruel do capitalismo, mesmo no estágio avançado alcançado nestes países, momento em que seria esperado que estas sociedades fossem razoavelmente justas para com os seus cidadãos”. A justiça é um ponto central para diversas bioéticas, entretanto, com itens como o acesso à saúde, o julgamento pode sofrer alguns impactos da opinião popular. Conforme alega o autor, o problema do acesso deixou de ser pautado como um problema de responsabilidade estatal para ser cotado como uma dificuldade oriunda da incapacidade do cidadão. A responsabilidade por conseguir ser consultado ou examinado depende mais do sujeito que da governança disponibilizar meios para isso ocorrer. Posição muito próxima daquela defendida por Morozov (2020), alinhada com o solucionismo tecnológico, já que o cidadão agora teria que utilizar um app para agendar sua consulta, outro para pagar e mais um para se deslocar até lá.

Em décadas recentes, o crescimento da perspectiva neoliberal e o individualismo têm fortalecido a visão de que os acontecimentos em nível da sociedade são de responsabilidade dos indivíduos que os sofrem, retirando o seu caráter de fenômeno social e coletivo. Esta visão de mundo tem sido o fundamento para que influentes forças políticas entendam as desigualdades como fruto de problemas individuais e os reneguem como expressão das injustiças e, portanto, sem necessidade de políticas e ações governamentais para minimizá-los (Barreto, 2017, p. 2101).

Encarando toda essa situação com as desigualdades no campo da saúde, não podemos esquecer que Barreto parte de um escopo global, analisando o campo como sendo apenas um. Comparando a saúde oferecida entre países desenvolvidos com os demais, se observam diferenças gritantes, com problemas recorrentes ao longo dos anos. Quer dizer, ainda há diferenças substantivas e os problemas de acessibilidade “persistem e, em muitos casos, aumentam as desigualdades nos níveis de saúde entre nações, ou entre regiões, grupos sociais ou étnicos de uma mesma nação” (Barreto, 2017, p. 2103). Segundo o autor, essa menor desigualdade, ou uma maior excelência na acessibilidade cria espécies de “prêmios de cidadania”, tornando alguns itens tidos habitualmente como fundamentais e vitais, um produto de excelência que só está disponível para certos países, aqueles que decerto tem mais dinheiro para oferecer. A Bioética de Intervenção sobressalta que há um passado histórico de benefício econômico desses “certos países” em cima dos demais. Um dos exemplos dados por Barreto é o caso da “expectativa de vida [...], veremos que uma criança nascida em Serra Leoa, em 1990, tinha, somente pela sua condição do nascimento, a expectativa média de viver 38 anos menos que aquelas nascidas no Japão no mesmo ano (46 anos versus 84 anos)”. Com isso, antes de vermos os usos da IA, devemos ter noção de um dos principais problemas enfrentados no campo, que é justamente a persistente desigualdade ao ofertar saúde à população planetária.

Mas se os dados especificamente analisados apontam para esse problema contínuo, também tem que ser evidenciado que de maneira geral eles apontam na direção completamente oposta, ou seja, “em nível global, indicadores das condições de saúde da população mostram, em geral, tendências positivas” (Barreto, 2017, p. 2103), logo após, o autor no mesmo artigo indica possíveis soluções para lidar com o defeito persistente do sistema (de saúde). Isso indica que com o avanço tecnológico, também estamos conseguindo dar conta de suprir uma quantidade cada vez maior de pessoas, com serviços de prevenção, diagnóstico, tratamento, cura e reabilitação também melhores, justamente pela implementação de novas tecnologias. E uma das formas pelas quais a IA está sendo mais utilizada na saúde é na análise de dados médicos. A IA pode ser usada para analisar grandes quantidades massivas de informações médicas, desde registros e histórico dos pacientes até resultados imagéticos de exames, para identificar padrões e tendências, indicando a probabilidade, por exemplo, de maior risco de um grupo do que de outro, que permite com que se possa selecionar quem priorizar em meio a crises de superlotação, sendo um excelente apoio, principalmente, pós pandemia de covid-19⁵⁰.

⁵⁰ As superlotações não decorrem apenas dos eventos pandêmicos, mas com a urgência da covid, sistemas foram adaptados para esquematizar priorizações e treinados para tentar diminuir as filas em unidades médicas,

Segundo Debesh *et al.* (2023), por conta da praticidade e acurácia da análise de imagens e diagnóstico de IA, em países em desenvolvimento onde o número de médicos especializados seja escasso, esses serviços podem tornar acessíveis, trazendo assim uma equidade no acesso à saúde.

Os modelos inteligentes podem ainda auxiliar os médicos a diminuir o erro médico durante o processo de tratamento. Estima-se que apenas nos Estados Unidos ocorram pelo menos 5 milhões de erros médicos a cada ano, representando a terceira causa mais comum de morte no país, como evidenciado pela pesquisa da Gaceta Médica de México, liderada por Dioselina, Diego e Raúl (Lanzagorta-Ortega; Carrillo-Pérez; Carrillo-Esper, 2022, p. 57). Segundo a SOBRASP (Sociedade brasileira para a qualidade do cuidado e da segurança do paciente), no Brasil existiam cerca de 292 mil casos com falhas na assistência à saúde no ano de 2022⁵¹. E a nível global, 2,6 milhões de pacientes morrem por conta do descuido em erros evitáveis, segundo a OMS⁵².

Além disso, a IA também pode ser usada para identificar padrões em dados de saúde pública que podem ajudar a identificar e prevenir doenças em larga escala⁵³. No LABDAPS, laboratório de Big Data e Análise Preditiva em Saúde da USP, começou a ser conduzida uma pesquisa resgatando uma das questões fundamentais da epidemiologia desde John Graunt: “É possível prever a probabilidade de quem vai morrer?”. Segundo as conclusões de Chiavegatto e colegas (2018), atualmente é algo previsível mesmo entre municípios. A pesquisa sai do clássico embate em que municípios mais ricos teriam sistemas de saúde melhor equipados e melhor expectativa de vida, conseguindo tratar os dados para indicar boas gestões municipais e catalogando essas boas práticas, como por exemplo, os municípios que vão acima do esperado na longevidade dos cidadãos investem em atenção primária aos pacientes. De acordo com Chiavegatto e parceiros (2018, p. 839), os “avanços recentes em aprendizado de máquina têm o potencial para melhorar a forma como os sistemas de saúde são fornecidos e avaliados”. Isso está em consonância com o que alegam Giovanella e Lobato (2012, p. 89): “estudar os sistemas de saúde nos ajuda a conhecer como suas estruturas estão falhando ou sendo bem-sucedidas nos objetivos de promover a saúde, garantir a melhoria das condições de saúde da

inclusive o SUS digital futuramente poderá ter esse recurso, segundo ALBUQUERQUE, Karoline. **Tecmundo**, 2023. Disponível em: <https://encurtador.com.br/jzLOX>. Acesso em: 06 dez. 2023.

⁵¹ MESIANO, Beatrice; RIBBEIRO, Leonardo. CNN. Brasil registra quase 300 mil falhas na assistência à saúde em 2022. Disponível em: <https://encurtador.com.br/jstY9>. Acesso em: 21 dez. 2023.

⁵² BASSETTE, Fernanda. **Folha de S. Paulo**. 6 milhões de pacientes morrem todos os anos por erros evitáveis na assistência à saúde, diz OMS. Disponível em: <https://encurtador.com.br/mHNOZ>. Acesso em: 21 dez. 2023.

⁵³ Pensando em níveis ainda maiores, há projetos que visam eventos de escalas nacionais e mesmo planetárias, como é o caso do projeto da Google (White, 2018) que usa AM para impedir grandes desmatamentos, impactando assim populações gigantescas.

população, cuidar das pessoas e aliviar seu sofrimento”. Essas indicações inflam a busca por utilizar a maior quantidade de dados médicos para abastecer sistemas de IA e assim garantirmos uma saúde melhor e avançar para outro patamar.

Mas o que exatamente são os dados médicos? Se são tão amplamente coletados, estudados e utilizados no capitalismo informacional, devem ser valiosos. Schneble e parceiros (2020) evidenciam que na maioria dos estatutos de proteção de dados⁵⁴ tem uma categoria especial para os dados de saúde. Mas uma coisa que os autores destacam é que o conceito de dados médicos hodiernamente é bem mais amplo do que se considerava no passado, pois a quantidade massiva de dados que são produzidos é bem maior do que tínhamos séculos atrás:

Os dados de saúde evoluíram consideravelmente. Atualmente, o chamado big data biomédico abrange desde dados produzidos por serviços de saúde, atividades de saúde pública e pesquisas biomédicas até dados que registram a exposição a fatores ambientais, como luz solar e poluição, ou dados que revelam estilo de vida, condições socioeconômicas e padrões comportamentais, como aqueles de aplicativos de bem-estar e fitness, mídia social e dispositivos vestíveis (Schneble *et al.*, 2020, p. 2, tradução nossa).

Schneble e colegas criaram distinções para evidenciar os tipos de dados que já são coletados e tratados por empresas que, como se percebe pela citação acima, muitas vezes nem são do ramo da saúde. Os separam em dados de saúde diretos, indiretos, inferidos e invisíveis. Os inferidos usam a combinação de dois elementos diferentes para saber uma informação não cedida. Já os invisíveis são dados que aparentemente não têm valor ou relevância para a saúde, mas que podem ser usados futuramente por IA ou outras máquinas, são dados a “prova de tempo”. O ponto central do artigo é que precisamos diferenciar esses tipos de dados e mudar a forma que enxergamos a concessão desses dados aos órgãos e empresas que os tratam, pois, “a maioria dos coletores de dados está obtendo acesso a grandes quantidades de dados comportamentais e de saúde de forma gratuita, sem ter que cumprir nenhuma proteção” (Schneble *et al.*, 2020, p. 6, tradução nossa).

Mas nem sempre a utilização dos dados, sejam eles diretos ou invisíveis, é danosa à população. Médicos e cientistas de dados de Illinois (Pan, 2017), se uniram para treinar máquinas capazes de prever quais mulheres grávidas ainda no pré-natal teriam maior risco de desenvolver algum tipo de adversidade no parto das crianças. Os parâmetros que definiam o parto adverso incluíam filho prematuro, baixo peso, internação na UTI e mesmo morte no primeiro ano de vida. A IA tem sido usada para desenvolver algoritmos de diagnóstico de

⁵⁴ LGPD é o regimento brasileiro, inspirado em grande parte pelo estatuto GDPR da união europeia (UE). O EUA não tem uma lei que verse sobre proteção de dados de saúde que abranja todo o país.

imagens que podem ajudar os médicos, por exemplo, a identificar lesões cancerígenas mais rapidamente e com maior precisão do que os métodos de diagnóstico tradicionais. Ou mesmo predizendo quanto tempo de vida pacientes oncológicos em estado terminal terão de qualidade de vida, oportunizando ao médico rever o tratamento recomendado. Sob outra perspectiva, um estudo recente realizado por Helen Santos e sua equipe (2020) na cidade de São Paulo revelou avanços promissores na previsão de óbitos de pacientes idosos utilizando técnicas de aprendizado de máquina. Esses usos podem levar a tratamentos mais precoces e eficazes, salvando vidas ou propriamente ampliando a qualidade de vida dos cidadãos. Por conta disso, Szlovits e Patel, em publicação de 2009, admitem que a IA em medicina está se tornando não apenas uma parte, mas um componente essencial da informática médica, sendo um recurso importante na solução de problemas em atenção à saúde. Inclusive, alguns especialistas (Lanzagorta-Ortega; Carrillo-Pérez; Carrillo-Esper, 2022, p. 55) afirmam que a IA poderia representar o “momento Gutenberg” para a medicina, dada a tamanha revolução causada pela IA no campo, comparável ao impacto da prensa de Gutenberg para a comunicação. Já Topol (2013, p. 13) afirma que “a medicina está prestes a passar por seu maior abalo na história [...] pela primeira vez, podemos digitalizar humanos”.

De toda forma, segundo Scheffer e Souza (2022), a maioria dos investimentos financeiros estrangeiros no setor de saúde no Brasil vem dos EUA. Além disso, das treze empresas estudadas pelos autores, dez são do setor financeiro e não têm a saúde como foco principal. Nos últimos oito anos, o número de investimentos mais do que dobrou⁵⁵. Até as “big techs”, isto é, empresas gigantescas no ramo da tecnologia, que não são primariamente da saúde, estão investindo nesse mercado. Como aponta Lobo (2017, p. 189):

Duas das mais importantes experiências no uso IA em vários campos, inclusive medicina, são a plataforma Watson Health, da IBM, e o Deep Mind, da Inglaterra, que processam informações armazenadas na “nuvem” de oncologia e de avaliação de risco e evolução de pacientes [...] O Deep Mind tem sido usado em medicina na avaliação de scans visuais, buscando causas de cegueira. Em 2016, estabeleceu um projeto com o Royal Free London NHS⁵⁶ Foundation Trust e com o Imperial College HealthCare NHS Trust para acessar os prontuários eletrônicos de pacientes e avaliar seus dados, o que tem gerado grande controvérsia no país. O problema da confidencialidade dos dados tem sido questionado, sobretudo pelo fato de eles serem processados livremente por uma empresa privada (Google).

Como indica o autor, ao falar do tratamento livre dos dados, nem tudo nesta nova “revolução” são benesses, na realidade é bastante o oposto: a IA tem apresentado diversos

⁵⁵ Essa desproporção poderia causar desconfiança se essas análises fossem divulgadas ao grande público.

⁵⁶ NHS, isto é, National Health Service é o sistema de saúde pública do Reino Unido, o que gera tensões devido à participação de uma empresa privada estadunidense.

problemas, tem reproduzido machismo, racismo e outros vieses que estão impregnados na sociedade e acabam afetando essas criações. Esses vieses, nada mais são do que está descrito nos outros capítulos como colonialidade. Tarcízio Silva (2019) explica em seu artigo e livro como os algoritmos afetam de forma violenta corpos não brancos, como já visto acima. Carla Vieira (2019), em entrevista, destaca que a violência pode ocorrer ao privar os rostos não brancos de reconhecimento facial.

Tais fatos levaram à criação da Liga da Justiça Algorítmica, que visa unir forças para monitorar e disseminar informações sobre a inteligência artificial. Em seu site, há uma série de perguntas questionando a todos sobre os perigos da big data aos envolvidos, principalmente cidadãos. Uma das perguntas é em que país estão sediados e sendo tratados os seus dados, questão que provavelmente uma grande parcela da população não saberá responder, criando no leitor uma indignação inicial sobre a privacidade dos seus dados. Eventualmente, mesmo tendo seus dados protegidos por uma legislação, isso não se aplica em todos os cenários, pois vivemos em um mundo interligado, mas, por vezes as legislações deixam de ter validade em outro território, como é o caso nos dados de um passageiro de um voo de um país a outro. Maria Tzanou estuda um caso teoricamente simples, mas que evidencia a falta de cuidado que há no tratamento de dados⁵⁷. Tzanou estudou a fundo os dados PNR (*Passenger Name Record*) na UE, que são efetivamente registros individuais de cada passageiro de um voo. Segundo a autora apenas com esses registros que aparentemente só dizem respeito ao âmbito empresarial de transporte, se pode ter acesso à dados mais sensíveis:

Os dados PNR (Passenger Name Record) também podem revelar informações religiosas ou étnicas (por exemplo, das preferências alimentares do passageiro), afiliação a um determinado grupo, bem como dados médicos (por exemplo, assistência médica exigida pelo passageiro, ou quaisquer deficiências ou problemas de saúde que sejam informados à companhia aérea) (Tzanou, 2017, p. 107, tradução nossa).

O estudo dos PNR revela, como apontado por Schneble *et al.* (2020), a quantidade de dados de saúdes inferidos e indiretos cedidos a essas empresas sem expresse consentimento dos passageiros. Empresas estas que podem não fazer uso de tais dados, porém também podem vendê-los a outras, como no caso da Cambridge Analytica (BBC Brasil, 2018). A falta de tratamento adequado desses dados na esfera legal demonstra o quão defasada está a discussão das novas tecnologias na área da saúde. É preocupante que, na maioria dos países, não haja

⁵⁷ Segundo ela, ao avaliar acordos UE-EUA, conclui que esses acordos “não fornecem proteção adequada de dados pessoais do outro lado do Atlântico”. Demonstrando que “as confusões conceituais de privacidade e proteção de dados [...] não são apenas teoricamente errôneas, mas também perigosas na prática” (Tzanou, 2017, p. 4, tradução nossa).

algo concreto em relação à regulação e cumprimento da lei nesse âmbito, o que já é por si uma grande dificuldade. Miguel Kottow, expoente da BP, ao analisar o funcionamento dos Estados modernos alerta que muitas vezes o pensamento liberal contemporâneo adota uma forma de Estado operando com funções mínimas necessárias para manter as condições para o capitalismo, modelo conhecido como o minarquismo ou estado vigilante noturno. E ao operar dessa forma, deixa “os cidadãos desprovidos de proteção social” exercendo uma função protetiva “mutilada cujo funcionamento residual está subordinado a interesses estrangeiros” Tealdi (2008, p. 165).

Nos últimos dois séculos, com o nome ou matrícula, era possível acompanhar toda a vida de um paciente em um hospital, averiguando todos os diagnósticos e hipóteses sobre o sujeito⁵⁸. Atualmente, apenas com o nome, podemos saber não só informações criteriosas de fora do hospital como também se pode tentar prever, com base nesses dados, se essas pessoas estão sujeitas a uma eventual morte ou algo parecido⁵⁹. Panitz Manassi (2014, p. 51) indica que apesar das fragmentações dos sistemas de informações do SUS, que teoricamente dificultam o acesso aos dados médicos no Brasil, a fragmentação dos prontuários eletrônicos dos cidadãos (PEC) reflete a fragmentação do sujeito contemporâneo. Ainda assim, Manassi (2014, p. 62) esclarece que em 2009, havia 142.771.691 de usuários cadastrados no SUS, e 86,97% dos municípios brasileiros usavam os aplicativos CadUS ou CADWEB, de acordo com dados do DATASUS, sistema de dados do SUS. E de acordo com o perfil epidemiológico da Associação Nacional de Hospitais Privados (Anahp), de 2022, 95,65% dos hospitais investigados já implantaram o PEC. Se, no âmbito privado, o PEC já é realidade da maioria, por outro lado, no contexto mais geral brasileiro, o Conselho Nacional de Saúde do ano de 2019, teve como proposta uma maior implementação e adesão ao PEC para solucionar os problemas, indicando que a tendência da quantidade de dados é apenas crescer, seja no âmbito privado ou no âmbito público⁶⁰.

Por outro lado, mesmo que o sistema econômico esteja predando os dados individuais, nem sempre o acesso a eles é tão fácil. Caso as descrições aqui encontradas fizeram o leitor pensar nisso, estaríamos interpretando essa tecnologia social de forma ingênua demais. Segundo Zuboff (2019), as empresas estão coletando, armazenando e transformando os dados pessoais no ouro da nova era, extraindo dele os valores necessários para o “sucesso” nesta era

⁵⁸ Sobre o tema, recomendo leitura de Foucault 1978 e 2012.

⁵⁹ Topol (2019) em seu capítulo 12, discute bastante o tema do uso de dados para a previsão de eventos.

⁶⁰ Apesar da “ausência, inconsistência ou falta de clareza no registro dos prontuários” (Cerqueira *et al.*, 2019, p. 13) e da diversidade de meios para acessar o PEC no meio público, os PECs tendem a aumentar com o tempo, o que pode levar a excelentes usos ou catástrofes, principalmente em relação à privacidade.

do “capitalismo de vigilância”, como chama. Rígidas e robustas legislações estão surgindo nos últimos anos em diversos países visando justamente dar conta de proteger a privacidade e os dados de seus cidadãos.

Na constituição brasileira, a privacidade é um direito garantido, mas devido aos grandes avanços tecnológicos e, por conseguinte, grandes quantidades de dados criados e gerenciados, foi criada a Lei Geral de Proteção de Dados Pessoais (LGPD) para cuidar especificamente dos critérios para manter maior proteção a dados nessa “nova era”, evitando abusos e usos ilícitos das informações coletadas. Dentro dessa lei, são aplicadas sanções para mau uso dos dados, especificamente a LGPD tipifica os seguintes dados como mais cruciais para serem protegidos.

Art. 5º Para os fins desta Lei, considera-se:

I - dado pessoal: informação relacionada a pessoa natural identificada ou identificável;

II - dado pessoal sensível: dado pessoal sobre origem racial ou étnica, convicção religiosa, opinião política, filiação a sindicato ou a organização de caráter religioso, filosófico ou político, dado referente à saúde ou à vida sexual, dado genético ou biométrico, quando vinculado a uma pessoa natural. (Lei 13.709, 2018 *apud* CFM, 2022).

Dessa forma, no campo da saúde, “a profissão médica se relaciona diretamente com certo tipo específico de dado pessoal, o qual – por sua natureza – demanda ainda mais a restrita proteção da Lei, os dados pessoais sensíveis” (CFM, 2022, p. 11). A LGPD foi muito inspirada na GDPR, lei de proteção de dados da união europeia. Ambas legislações são bastante atentas aos problemas de dados, mas ainda assim surgem diversos autores criticando certos aspectos que ficaram de fora da abrangência da lei. No entanto, no Brasil temos um órgão que cuidará da aplicação e fiscalização do cumprimento das normas, a Autoridade Nacional de Proteção de Dados (ANPD). É vital destacar que a LGPD tem abrangência extraterritorial, ou seja:

atinge todas as pessoas (físicas ou jurídicas) e atividades realizadas no Brasil que tenham relação com tratamento de dados pessoais, como a coleta, a produção, a classificação, a utilização, o acesso, a reprodução, a transmissão, a distribuição, o processamento, o arquivamento, o armazenamento, a eliminação, a avaliação ou o controle da informação, a modificação, a transferência, a difusão ou a extração de dados [...] Do mesmo modo, a lei se aplica se os dados pessoais objeto do tratamento tenham sido coletados no território brasileiro (Iramina, 2020, p. 98-99).

Essa possibilidade de ter os dados dos “nacionais” protegidos demonstra como não apenas o Brasil, mas Tailândia, Chile, Coreia do Sul, Índia, Indonésia e outros cem países estão se empenhando para lidar com os problemas desse século, apesar disso, a atenção para

problemas anteriores coadunados com os novos não parece ser posta em pauta, como foi, por exemplo, com os casos que vimos acima, do uso de dados médicos britânicos pelo projeto Deep Mind coordenado por uma empresa estadunidense e a oferta de dados de vários brasileiros pelo MEC para a Microsoft. Os exemplos reforçam a “relação entre o neoliberalismo e a colonialidade” (Amadeu, 2021, p. 40). Ambas situações envolvendo o EUA reforçam que o pioneirismo deles em assuntos tecnológicos faz com que fique evidente a necessidade dos outros países se americanizar pela tecnificação (Corbisier, 1959), ademais, é notável perceber que o país mesmo sendo um dos mais avançados nesses quesitos não tem uma regulação federal que cuide especificamente sobre os dados. Até existe um esboço de uma lei federal (American Data Privacy and Protection Act, conhecida como ADPPA), mas efetivamente não há nada em vigor além do que projetos de lei estaduais apenas. Essa aparente incompatibilidade entre ser o país a frente e ao mesmo tempo não se preocupar em legislar sobre o uso dos dados é preocupante, no entanto, com o passar do tempo mais e mais países têm se dedicado a garantir maior proteção à privacidade da população.

Contudo, mesmo com tamanho esforço, é muito difícil combater os encantos oferecidos aos cidadãos para fornecer passiva e pacificamente seus dados de forma voluntária. A grande diferença desta maneira de fornecimento para as outras criticadas no trabalho é justamente a presença do assentimento no fornecimento desses dados, como acontecem nas redes sociais, que são grandes coletoras de informações sensíveis, mas de forma consentida pelo usuário, que espontaneamente escolhe mostrar sua aparência, localização e demais dados para expressar sua personalidade nessa versão *online* da vida social. Adicionalmente, Camargo (2020, p. 68) alega que “a mesma lógica de apropriação de conteúdo acontece nas redes sociais, só que desta vez o conteúdo é cedido de modo voluntário. Importante notar que a distinção deste cenário para o dos outros serviços é a voluntariedade, não o caráter do conteúdo em si”. Observando esse fenômeno, o cientista e precursor da realidade aumentada, Jaron Lanier criou um termo para qualificar esses servidores que seduzem os usuários, ele os chama de servidores sereias⁶¹. Esses computadores e servidores têm como propósito atrair a maior quantidade de pessoas, com excelentes promoções, maravilhosos produtos ou coisas similares. Dessa forma, funcionariam como as sereias de Homero, seduzindo os navegadores digitais a chegarem mais

⁶¹ No texto original o autor utiliza o termo “*siren servers*”, e o termo siren tem duas traduções possíveis, baseadas na mitologia grega, sendo uma das opções a conhecida sereia e por outro lado a distante sirena. Apesar de ambas terem vozes sedutoras, Lanier em seu livro utiliza uma alegoria com a metade pássaro da sirena, e não algo aquático como as sereias, indicando que a melhor tradução seria servidores sirenas. Mesmo assim, manteve a tradução de servidores sereias por encontrar mais literatura em português com esse termo, principalmente Camargo (2020).

perto para provar e oferecer seus dados para acessar o que queriam. E as sereIAs nesse caso não são apenas igualmente sedutoras, mas deveras perigosas, Lanier (2013) afirma que mesmo os que criam e controlam esses servidores, podem acabar caindo em desilusão com a sereia, pois os resultados de curto prazo gerados por ela, poderiam potencialmente gerar graves consequências a longo prazo para todo um grupo ou mesmo para toda humanidade. Ou seja, mesmo os controladores dos servidores sereias “IAra”, não estariam a salvo. É importante notar, que muitas vezes o usuário atraído não é informado da maneira devida sobre o que está fornecendo, ou com o que está plenamente concordando. No campo da saúde, o problema se intensifica mais.

Outro fato que aliado com a voluntariedade corrobora o acesso aos dados é a prática de invasão de dados e o seu conseqüente vazamento. Topol (2019) indica que mesmo o sistema de saúde dos EUA, um dos países pertencentes ao pequeno grupo de criadores de tecnologias, é bastante vulnerável a vazamento de dados. Segundo ele, a criptografia usada para proteger os registros médicos dos estadunidenses foi quebrada, e logo após, pelo menos quatro milhões e meio de pacientes do Community Health System tiveram seus dados vazados, contendo até endereços e registros civis. Se o país que está à frente dos maiores desenvolvimentos tecnológicos do capitalismo tem sua segurança e privacidade ameaçada, em países onde sequer há estrutura para o desenvolvimento tecnológico básico, a privacidade e segurança dos dados torna-se um luxo quase inacessível, um direito negado. Nesse sentido, os problemas da IA quando se integram à big data⁶² aparecem no mesmo compasso que nos damos conta que são basicamente insuperáveis, já que estão em processamento agora mesmo em bilhões de redes, tratando os dados de bilhões de pessoas, para apenas alguns poucos terem acesso a eles, ou mesmo criarem algo com estes dados, manipulando a outra porcentagem. O avanço tecnológico já se tornou irreparável no sentido de deixar de ser utilizado, como alega Castells:

Na era da informação, a principal lógica das redes globais predominantes é tão difusa e penetrante, que o único modo de se livrar de seu domínio parece ser ficar fora delas e reconstruir o significado com base em um sistema de valores e crenças inteiramente distintos (Castells, 1999, p. 36).

Todavia, diferentemente da época em que Castells escreveu tal mensagem, atualmente é inimaginável um regresso tecnológico que não inclua o amplo uso de IA, ou mesmo da aceitação dela em meio aos humanos devido à inter-relação entre os países e a demanda internacional criada por essas tecnologias. Assim surge um problema antigo repaginado, porém

⁶² Big data nada mais é do que o um conjunto de dados de tamanho, velocidade massivo, que extrapola as capacidades básicas de ferramentas típicas que servem como banco de dados.

ainda sem consenso desde os tempos antigos⁶³: o limite da interação do homem com a tecnologia, homem-máquina; e principalmente a expectativa que o homem tem de que a tecnologia aponte o caminho para o futuro (Morozov, 2020; Silva, 2013). Para aclarar esse ponto, podemos usar uma metáfora com um motorista e um GPS. Imaginem o tempo avançando como se fosse uma estrada, e nós, enquanto humanidade estamos a percorrendo, e temos atualmente a tecnologia como um possível auxiliador, que nos indica possibilidades e podemos consultá-las para saber qual caminho seguir, tal qual um GPS que tão estimado e familiar. Mas esse uso é opcional, já que ainda teremos que percorrer todo o caminho, tendo atenção no percurso. De outro modo, podemos compreender que a principal diferença ao lidar com IA é que, a partir de agora, o ser humano já conseguiu desenvolver um condutor autônomo⁶⁴, ou seja, além de esperar que a tecnologia (IA) mostre o futuro, podemos relaxar e sequer dirigir ou verificar qual caminho é o mais indicado. Logo, cria-se a crença de acreditar que seremos levados ao futuro automaticamente, sem precisar monitorar o caminho e o destino pelo qual estamos sendo conduzidos. Apesar da metáfora ser fortuita, devemos lembrar que acreditar que somos apenas uma única humanidade é bastante inverdade, já que apesar de compartilharmos sim características comuns a todos os humanos, estamos todos em condições e recursos diferentes para lidar com eventuais problemas. Tanto a BP quanto a BI nos alerta dos perigos de termos como inequívoca a ideia de “A humanidade”, sendo que são diferentes grupos e estes “estão posicionados de forma muito diversa em relação aos atributos e oportunidades sociais” (Tealdi, 2008, p. 166).

⁶³ A tecnologia dos antigos é a técnica que permitiu os sapiens transformarem pedras em ferramentas, criando então a tecnologia. Cada época teve de lidar com sua tecnologia, enfrentando suas dores e delícias.

⁶⁴ Recurso que já está em uso com os carros autônomos da empresa Tesla, por exemplo.

7 DEVIR CURUPIRA: O PROBLEMA DA HERANÇA EM IA

A crença que as máquinas e a IA nos levarão para o futuro apresenta uma problemática intrínseca na própria convicção. Como criações do presente, as máquinas não estão isentas dos problemas que afligem nossa sociedade hoje, que nos impedem de caminhar para quaisquer idealizações do futuro. Ao depositarmos cegamente confiança nelas, corremos o risco de sermos conduzidos a futuros massivamente duvidosos. As máquinas são programadas por seres humanos, que estão sujeitos a influências e vieses colonialistas, conforme indicam Mehrabi *et al.* (2019). Portanto, acreditar que as máquinas nos conduzirão a um futuro livre dessas problemáticas é uma visão ingênua e bastante perigosa. É fundamental reconhecer que as máquinas são poderosas ferramentas, mas seu uso e desenvolvimento devem ser guiados por princípios éticos baseados em uma profunda reflexão sobre as questões coloniais presentes em nossa sociedade, a fim de evitar a perpetuação dessas opressões coloniais através das tecnologias.

Vieira Pinto (2005, p. 42) aponta que valorizar a tecnologia atual é também exaltar a estrutura social que a criou. O endeusamento da tecnologia nesse sentido, é um endeusamento do capitalismo moderno colonial que a tornou possível. Conforme esclarece Silva, o autor é bastante crítico de um consumo tecnológico “acrítico”.

De acordo com Vieira Pinto, nas sociedades divididas, os elementos dominantes sempre exaltam o presente, no qual eles se acham bem instalados. [...] Assim, esses elementos dominantes elevam ao plano da ideologia as condições sociais, científicas e técnicas que o caracterizam. [...] e exaltar o presente é sempre uma característica dos que se beneficiam dos instrumentos existentes. [...] já não sendo possível negar ou ignorar o futuro, tratam de domesticá-lo de antemão, descrevendo-o em termos atuais. Nesse caso, apresentam o futuro em termos do modelo tecnológico existente. Assim, o futuro deixa de ser ameaçador, perde o caráter de desconhecido, porém a questão central é que ele não representa nenhuma mudança substancial no status dos grupos dirigentes (Silva, 2013, p. 842-843).

Devemos ter em mente que as máquinas são fruto do presente e, como tal, não podem escapar dos problemas de colonialidade que afetam a sociedade em que foram desenvolvidas. Assim a crença de que elas nos conduzirão a um futuro melhor é problemática, pois as máquinas são programadas por seres humanos que carregam consigo suas perspectivas, valores e ideologias. Não existe tecnologia fora do contexto social, o que ocasiona na reprodução desses conceitos coloniais pelas máquinas.

Se épocas atrás o problema político era de que “o homem é o lobo do próprio homem”, frase atribuída à Hobbes (2006, p. 9) pela tradição, mas originalmente escrita por Plauto em

sua peça *Asinaria*, atualmente o problema já não se finda nessa afirmação. A frase tem o sentido de que os homens não conseguiriam viver em sociedade harmoniosamente, pois teriam um “instinto de se auto caçar e autodestruir” (Brandeburgo; Garcia, 2023, p. 49). E parece que as tecnologias hodiernas deram um fôlego e uma sobrevida para a teoria política de Hobbes, o tornando atual, mesmo com o passar dos séculos. Uma vez que as inteligências artificiais adquirem as características dos homens, se tornam elas o próprio lobo do homem. A IA terá esses vieses porque foram treinadas dessa forma. Ou seja, se antes, o homem já era o lobo do homem, agora as máquinas criadas por eles, continuam sendo então perseguidores de seus iguais, que apesar de terem composições sintéticas e inorgânicas, ainda mantém vivo o maldito legado hobbesiano de terem os homens como próprios males de sua própria espécie.

Em outras palavras, podemos sintetizar a argumentação em: não existe IA forte, então quem está desenvolvendo a IA é a humanidade. Ao que se nota, a humanidade não está “controlando” o desenvolvimento, abrindo margem para ferir e danar seus semelhantes. Devido a ética em vigor ser consequencialista, tais atos podem e devem ser julgados a partir dessas violências sofridas por certos grupos. E a culpabilização não pode ser feita em cima dos sistemas inteligentes, visto que não são independentes dos comportamentos emergentes. Assim, a IA está reforçando a teoria hobbesiana nesse aspecto, criando, então, lobos do homem com uma potência destrutiva que jamais foi alcançada, exceto talvez pelas bombas atômicas. Portanto, é imprescindível reconhecer a necessidade de uma abordagem crítica e reflexiva ao utilizar e desenvolver tecnologias.

Tealdi (2008, p. 163) declara que “a bioética de intervenção tem fundamento filosófico utilitário e consequencialista, defendendo como moralmente justificáveis” políticas públicas e tomadas de decisões que privilegiem o maior número de pessoas, resultando nas “melhores consequências coletivas, ainda que em detrimento de determinadas situações individuais”. Esse ponto abordado acima e durante todo o trabalho de situar a problemática da IA como um potencial nocivo é vital para se intervir no tópico, interpretando que os problemas aqui elencados se configuram como problemas de saúde coletiva. Não devemos esquecer que conforme foi determinado em 1978 na Conferência Internacional sobre Cuidados Primários de Saúde, com a Declaração de Alma-Ata, seguramente a saúde “é um direito humano fundamental, e que a conquista de um nível de saúde mais alto possível é um objetivo social sumamente importante em todo o mundo, cuja realização requer a intervenção de muitos outros setores sociais e econômicos, além do setor da saúde” (Tealdi, 2008, p. 246).

Com isso posto, podemos levantar algumas possibilidades interventivas, voltadas para o poder prático intervencionista da BI ou protecionista da BP. Na conjuntura aqui descrita, há

diversos motivos para intervir em favor dos “vulneráveis” que aqui são representados por aqueles que não estão atentos ao consumo acrítico tecnológico, iludidos pelo dataísmo, pois estes estão expostos aos perigos extrativistas e formas de violências do colonialismo digital. Se tratando desse tema, pode-se destacar o conceito de soberania digital, como define Ali Shoker (2022), se refere a um estado⁶⁵ possuir a autoridade suprema sobre todos os seus recursos digitais, devendo abranger toda a cadeia de produção de geração de dados, armazenamento, processamento e demais operações com os dados. Dessa forma, há uma necessidade estratégica de cada nação proteger os “ativos digitais”⁶⁶ gerados em seu país ou por seus cidadãos. Na legislação brasileira, foi suscitada pelo Marco Civil da Internet de 2014, sendo um dos documentos pioneiros do conceito, e o Brasil foi um dos primeiros países do mundo a insistir num enquadramento robusto dos direitos digitais. Para Morozov (2018, p. 9-10), inclusive, esta foi uma iniciativa que “ainda que inconclusa, é uma manobra importante, sobretudo agora que, cada vez mais, as plataformas digitais buscam nos atrair para seus impérios digitais acenando com serviços gratuitos e convenientes”. O problema, porém, foi que a “iniciativa” não deu novos avanços desde essa época. Belli (2024) afirma que “por anos o Brasil foi referência mundial, oferecendo uma visão de como o software pode ser enxergado como ferramenta libertadora, em vez de um instrumento de extração de dados e de colonização digital”, isto é, desde 2003 o governo federal dava preferência para a utilização do software livre nas repartições, no entanto, em 2016 isso foi descontinuado. Devemos recordar que a companhia siderúrgica nacional, que foi vital para o desenvolvimento brasileiro só surgiu depois de muita reivindicação; A Petrobras seguiu o mesmo processo, sendo fundada após o movimento “O petróleo é nosso”. Com isso, quero expor em consonância com o Silveira (2023) que é necessário começar uma reivindicação severa contra o colonialismo de dados, que pode nos livrar de muitos dos males aqui observados. Mais à frente, para finalizar, veremos algumas outras soluções práticas de casos bem sucedidos de como a inteligência coletiva aliada com a tecnologia superou os perigos do colonialismo digital.

Então, em concordância com tudo o que vimos aqui, é necessário termos noção de que apesar dos avanços e benefícios possíveis que a IA nos traga, trará também problemas e em escalas antes inimagináveis. A capacidade de administração e processamento dos agentes

⁶⁵ Segundo a tradição na literatura decolonial o Estado não é um aliado nessa luta, já que basicamente seria o resultado do poder opressor consolidado, uma espécie de lobo mor, usando a leitura hobbesiana. No entanto, mesmo diante dessa situação e os riscos que envolvem essa situação (que não são poucos), essa possivelmente seria uma estratégia melhor do que desconsiderar totalmente esse grande acúmulo de poder fora da equação.

⁶⁶ O termo original aqui é *digital asset*, mas nesse caso parece fortuito não traduzir como recurso digital, mas sim como ativo digital para transmitir o valor atribuído ao dado. Não se pode esquecer que o ouro da nova era tem um valor extremamente alto.

autônomos pode gerar um erro ou injustiça reproduzida em níveis catastróficos. Reproduzida, pois aqui entendo que na própria criação e desenvolvimento (aprendizados) dessas inteligências se incluem legados malditos que acabam por deturpar o julgamento potencialmente matematicamente justo e imparcial. Legados, pois, se não há IA forte o suficiente, necessariamente quem está por trás das ações dos modelos de IA, é o ser humano. A herança que legamos à IA, seja ela cultural, situacional ou estrutural⁶⁷, necessariamente será repleta de vieses e por conseguinte, problemática.

A crença na neutralidade tecnológica esquece que tais tecnologias são produzidas por pessoas reais em ambientes reais e que eventualmente são impactados por preconceitos e vieses reais. Conforme afirma Elen Nas (2020, p. 62): “As soluções matemáticas e estatísticas não são capazes de lidar com preconceitos embutidos nos sistemas”, embutidos justamente por originarem desse mesmo ambiente não neutro. O dataísmo acaba por atrapalhar a própria possibilidade de resolução do problema, visto que nem sequer é considerado um problema. E não é considerado um problema por várias diferentes esferas, pois a população em geral não considera a possibilidade de a IA, em si, poder ser danosa; os cientistas da área estão preocupados demais com o desempenho técnico, não humanizado ou ético; e os políticos que discutem e legislam sobre ela parecem ter preocupações maiores. Dessa forma, a crença bastante popular do dataísmo de que as máquinas e IAs seriam teoricamente neutras e que suas decisões não teriam nenhum viés acaba por nublar toda e qualquer forma de enfrentamento aos resquícios de colonialidade encontrados nos códigos. Só que isso colide diretamente com o que é a IA e demais ferramentas tecnológicas, isto é, produtos originados na sociedade que os concebe. Como elucida Elen Nas (2020, p. 60, grifos nossos) ao criar o conceito de devir curupira:

A IA, portanto, é um ‘produto’ que aponta para o futuro, porém sua base epistemológica, cultural e social é guiada pelo passado. Como o personagem da lenda folclórica brasileira, o Curupira, a IA se volta para frente enquanto seus passos estão marcados para trás. [...] também a IA confunde seus seguidores apresentando o passado quando o que se espera dela é um caminho mais responsável e seguro para o futuro. Ela atua com características cruéis.

Elen, nesse artigo intitulado “inteligência artificial: uma discussão bioética dos sistemas de informação” se esforçou para escrutinar os possíveis problemas de uma IA sob uma análise eticamente minuciosa, expondo sobre a essência da tecnologia e como ela encanta por sua

⁶⁷ Aqui lê-se capitalista-colonial. O capitalismo é um sistema histórico que é repleto de conceitos próprios e inerentes a ele, que os dá sentido. Em outros modelos socioeconômicos é possível que novas funções e usos poderiam ser criados, caso tivessem tecnologia suficiente.

natureza trazendo à tona diversas ferramentas e conclusões para o campo, bem como denúncias de como a tecnologia é enviesada por ser produto da criação de um mundo não neutro. O atual estudo foi fortemente influenciado por conta do texto dela e sem dúvidas, esse artigo precisaria ter maior reconhecimento dos especialistas e equipes de desenvolvimento de IA, devido ao trabalho notável da criadora do conceito. No entanto, o conceito parece ter capacidade de ser estendido e ampliado para utilizações mais amplas em estudos tanto éticos quanto computacionais. Com isso posto, posso explorar ainda mais o que estaria nos entremeios do que poderíamos definir como devir curupira. Mas antes de prosseguir no tópico, preciso elucidar o que é curupira, como assim devir?

De acordo com Costa Neto, Fita e Aguiar (2023, p. 9) a ambiguidade observada nas palavras de Nas (2020) é constituinte dessa entidade. De fato, o curupira é descrito tanto como tendo ações benéficas e protetoras quanto punitivas e demoníacas. Ele pode tanto lhe indicar o caminho para sair de uma densa floresta, como pode também lhe fazer enlouquecer de tantos dias perdido na selva, isso é claro, se não quiser te matar. Além disso, essa indeterminação é observada também em sua aparência e materialização, ao que os estudiosos indicam, a sua característica principal e a única característica física que se mantém em seus relatos são seus pés retorcidos pois “trata-se de uma entidade fenotipicamente polimórfica, apresentando-se aos seres humanos em diferentes aspectos: um duende benfazejo, um demônio, gnomo ou ogro”. No entanto, a enganação também é algo marcante em suas ações, geralmente tirando suas vítimas do caminho que deveriam seguir, lançando alguma distração chamativa o suficiente para os fazer mudar a rota. “Além de enganar, ao deixar suas pegadas ao contrário, pois, quando o Curupira está entrando na mata, ele parece estar saindo, também engana, com assobios e gritos, os caçadores, lenhadores e outros viajantes, fazendo com que percam o rumo” (Costa-Neto; Fita; Aguiar, 2023, p. 10).

Essa característica de fazer com que os indivíduos percam o sentido e sigam caminhos tortuosos é similar ao que acontece com certos grupos diante da habilidade de processamento e criação de dados com rápidas respostas atingidas em velocidades absurdas da IA. Tal poderio já está sendo usado para espalhar notícias falsas em formatos antes não imaginados, trazendo um dano para pessoas e até mesmo o sistema político que é difícil de ser recuperado. Ainda que tenham alguns especialistas indicando que é necessário criar, aumentar e se fazerem cumprir regulamentações sobre a IA, o tema não parece estar recebendo devida atenção, principalmente no âmbito legislativo/político. A respeito da descrença do Curupira, Câmara Cascudo, ilustre historiador e folclorista, afirma que o Curupira é o:

Demônio da floresta, explicador dos rumores misteriosos, do desaparecimento de caçadores, do esquecimento de caminhos, de pavores súbitos, inexplicáveis. [...] Sempre os pés voltados para trás e de prodigiosa força física, engana caçadores e viajantes, fazendo-os perder o rumo certo, transviando-os dentro da floresta, com assobios e sinais falsos. [...] *A gente da cidade acredita em sua existência*, mas ela não é motivo de preocupação porque os Curupiras não gostam de lugares muito habitados. [...] Na mata os gritos longos e estridentes dos Curupiras são muitas vezes ouvidos pelo caboclo. *Também imitam a voz humana*, num grito de chamada, *para atrair vítimas*. O inocente que ouve os gritos e não se apercebe que é um Curupira e dele se aproxima perde inteiramente a noção do rumo (Cascudo, 2001, p. 172-173, grifos nossos).

É interessante essa citação de Cascudo, pois diferente das maquinarias do capitalismo contemporâneo que buscam enlouquecidamente pessoas para abastecer o colonialismo de dados, o curupira não preocupa a população geral. Talvez essa faceta pode ser interpretada, nessa analogia, como sendo a dificuldade de entendimento, conhecimento e mesmo acesso que o grande público tem desses motores artificiais inteligentes. Decerto, existem grandes projetos e muitas inteligências que podem ser consideradas públicas e auditáveis, no entanto, a parcela da população que os utiliza frequentemente ainda é baixa, mesmo que seus usos sejam acuradamente comprovados. O sociólogo Gabriel Vasconcellos (2010, p. 13) diz que o curupira é “uma fantasmagoria que recusa a ingerência do fator externo. Curupira é um corpo sem orifício, sem nenhum buraco, sem nenhuma fenda por onde pudesse ser penetrado”⁶⁸. Essa é mais uma característica que o curupira compartilha com a IA, ambos são opacos, herméticos⁶⁹, e por isso, quase inexplicáveis ou dificilmente são críveis quando se tem seu funcionamento esclarecido. Sobre o aspecto fantasioso e mágico, é importante ser destacado que

As tecnologias são comumente associadas a “utilidade” e, portanto, vistas como produtos que podem ser melhor controlados por determinados mercados. O fato da tecnologia ser também arte (techne) confunde sua ontologia, assim como as relações que se estabelecem com ela. A confusão reside no fato da tecnologia, como arte, possuir atributos criativos como a materialização de ideias. Este é um aspecto sedutor e envolvente da tecnologia, pois existe algo de mágico e lúdico em seu devir. Enquanto a arte representa visões particulares ou parciais, em consonância ou dissonância com alguns saberes, a tecnologia como instrumento da ciência agrega autoridade científica a uma criatividade delimitada por pragmatismos (Nas, 2020, p. 56).

Os sistemas agem de forma semelhante aos humanos, mas infelizmente, a IA não passou por sua época de renascimento para terem um retrato semelhante ao homem vitruviano elucidando sua constituição e ontologia. De toda forma, a estrutura fechada da IA, só aceita

⁶⁸ O curupira conforme descrito tem a capacidade de imitar sons, assoviar para atrair as pessoas, tal qual a sereIA. O que indica a possibilidade do curupira ter ao menos um orifício que emite sons, possivelmente a boca.

⁶⁹ O curupira conforme descrito Ambas também são e têm funcionamentos que na maior parte das vezes é mítico, misterioso. Isso seguindo a terceira lei de Clarke, em que “qualquer tecnologia suficientemente avançada é indistinguível da magia”.

ordens por dentro, isto é, do código, a não ser que seja programada para levar em conta o ambiente. E esse ponto merece uma diferenciação em comparação com o curupira. Ao passo que para o ente folclórico o fato de ser opaco é uma qualidade e uma força, pois torna suas limitações desconhecidas para os seus inimigos e os amedronta por não saber suas limitações, ainda é parte constitutiva de seu ser, isto é, para ele é importante ser assim. Além disso, ele é opaco para todos, seu corpo é fechado e sem orifício, independentemente de qualquer razão.

Esse mesmo argumento não se aplica as inteligências artificiais, visto que a opacidade de seu sistema é uma fragilidade limitante, que ocorre por vezes devido as patentes comerciais implicadas nos sistemas de inteligências artificiais, que se fossem públicos perderiam a efetividade e exclusividade do produto que é tão valiosa pro capitalismo. Jha Debesh e colegas (2023) argumentam que a implementação de IA no setor médico pode não apenas limitar as escolhas (mesmo que informadas e consentidas) dos pacientes, mas também levar à erosão da autonomia deles. Fora que tal opacidade se aplica somente ao público em geral, pois a equipe de desenvolvimento certamente tem maiores informações e compreende melhor o funcionamento do modelo. Caso os modelos fossem plenamente transparentes, o que estaria em jogo para as empresas que tocam os sistemas de IA seria a lucratividade e exclusividade por trás. A opacidade neoliberal é o que impede uma maior explicabilidade da IA. Nesse sentido, opacidade curupírica é mais democrática, já que independente do cargo, crença ou raça ele será aquele com os pés torcidos.

Mas voltando ao mote central da explicação do devir Curupira, temos por devir, conforme define os filósofos Deleuze e Guattari, que é nada mais do que um câmbio, uma mudança entre coisas. Mas esse é um termo, sem dúvidas, complexo:

A significação do termo “devir” não é unívoca. É usado às vezes como sinônimo de 'tomar-se'; às vezes, é considerado o equivalente de 'vir a ser'; às vezes, é empregado para designar de um modo geral o mudar. [...] é o que destaca o processo do ser, ou, se se quiser, o ser como processo (Mora, 2000, p. 707).

Teoricamente, o devir pressupõe o encontro de dois elementos para se alterarem, mas no caso que estou tratando aqui ambos seres não têm uma consciência bem definida, uma por não existir, até onde se sabe, uma IA realmente forte e completamente autônoma, e por outro lado a criatura folclórica não está disponível para ser enquadrado nos limites da razão, o que nos impede de definir se ele é consciente ou não⁷⁰. No entanto, a similaridade que define essa

⁷⁰ Nada impede de o Curupira ter algum tipo de consciência volitiva, levando em conta a complexidade de ações que são descritas em seus relatos poderíamos pressupor, no mínimo, alguma deliberação e raciocínio, já que ao

relação não é menos potente que a diferenciação que separa a IA do Curupira, pois “devir não é atingir uma forma (identificação, imitação, mimese), mas encontrar a zona de vizinhança, de indiscernibilidade ou de indiferenciação tal qual já não seja possível distinguir-se de uma mulher, de um animal ou de uma molécula” (Deleuze; Guattari, 1997, p. 43). De acordo com Ferrater Mora, as diferenças se tornam irrelevantes, isto é, “o devir, um movimento no qual ambos os termos são distintos, mas com uma espécie de diferença que, por sua vez, dissolveu-se imediatamente” (Hegel, 1831 *apud* Mora, 2000, p. 710). Dessa maneira, tanto o curupira quanto a IA se assemelham, mas também se diferenciam pelos traços elaborados acima.

Devir curupira então é a tendência que as novas tecnologias têm em serem traíçoeiras, ressuscitando e potenciando de forma intensa problemas do passado que ainda habitam nosso construto colonial capitalista e demais características enquanto sociedade. E dessa forma, a IA curupira assombra todo e qualquer dataísmo que creia na utilização da IA como neutra e em alguma parcela responsável pelo futuro humano. É importante destacar que o devir não deve ser encarado como essencialmente maléfico, mas sim ambíguo, dúbio. O Curupira surge aqui também como aquela figura que nos alerta para os perigos de se confiar em produtos semelhantes aos humanos, tal qual as artes de IA crendo de que são originados de pessoas reais. No meio da selva de pedra do capitalismo tardio, o Curupira existe como um emblema, um vulto, que pode se exibir para impor obediência às regulações, como para levar a humanidade à caminhos capciosos. O ponto central do devir curupírico é exibir essa dualidade e ambiguidade constitutiva dos sistemas inteligentes, que exibem suas excelentes facilidades enquanto abusam capturando dados e violentando vulneráveis.

8 CONCLUSÃO: A LUTA CONTINUA

De toda forma, o caminho ao aclamado “futuro” deve ser trilhado com base nas decisões humanas, que apesar de falhas, ao menos podem ser responsabilizadas sem grandes questões éticas. Nosso papel político, que decide a caminhada, tem como base buscar o bem comum.

Não é papel dos algoritmos apontar soluções que envolvam o destino humano na educação, saúde, trabalho, moradia e assim por diante. A pergunta, portanto, não é como fazer a IA mais justa, mas por que se faz necessário utilizar IA para avaliar acesso a serviços (incluindo os básicos, como saúde), periculosidade, competência e confiabilidade, classificar/discriminar indivíduos, os comparando com práticas passadas (Nas, 2020, p. 63, grifos nossos).

E o problema ao usar esse tipo de tecnologia no campo da saúde, conforme sintetiza Elen Nas (2020, p. 52-53), é que em sistemas ultramodernos que atingem o mundo todo, sendo essas tecnologias pontos de informação e mesmo referência em diversas áreas, “no caso da saúde, se os sistemas aprendem com dados que integram informações antigas e interpretações ultrapassadas do ponto de vista científico, o que se tem são novas ferramentas replicando velhas ideias”.

Mesmo outros especialistas de países subdesenvolvidos estão se preocupando com a temática. Ugochi Okengwu (2022), um pesquisador nigeriano, argumenta que a IA pode sim ajudar a resolver problemas sociais, se for desenvolvida levando em conta os aspectos específicos da região que será aplicada, e assim, poderá promover justiça e igualdade. O autor afirma que é preciso mover esforços visando a implementação de regulações convergentes sobre a IA em toda África, considerando os valores africanos para a montagem dessa estrutura unificadora. É relevante destacar que, de acordo com o índice global de IA (White; Cesareo, 2023), atualmente sete países em todo o continente africano tem regulações voltadas para a IA, sendo eles África do Sul, Tunísia, Nigéria, Quênia, Egito, Gana, Marrocos; número muito abaixo dos 54 países que compõem o continente africano. Na América Latina encontramos um cenário muito similar, pois dos vinte países componentes, há apenas seis países que de fato promoveram leis voltadas ao uso de inteligência artificial, como é o caso do Brasil, Colômbia, Argentina, Chile, Uruguai e México⁷¹. Essa perspectiva de Okengwu de implementar os valores, preocupações e até mesmo expectativas nacionais no desenvolvimento de uma IA é muito enriquecedora para nossa pesquisa, pois podemos considerar o mesmo argumento ao

⁷¹ Se diferenciarmos em nossa análise entre América do Sul, central e do Norte, restaria apenas cinco países dos doze que compõem a região abaixo dos “trópicos do pecado”.

pensar na implementação desses modelos inteligentes em países pobres como uma forma de intervenção aos problemas enfrentados nesses locais. Dessa maneira, de um aspecto bioético (principalmente da BI), fazer o eticamente correto levando em conta os diversos cidadãos desses países seria investir em pesquisa e desenvolvimento desses motores inteligentes visando a aplicação justa e ética deles para ser um suporte na resolução dos problemas nacionais. Pensando na limitação de recursos desses países, pode-se pensar que a adaptação de modelos já existentes também seria uma opção bastante fortuita para se considerar. Os únicos empecilhos a sua efetivação seriam como já vistos a dependência da empresa detentora do modelo e o enviesamento de dados que intoxicam certos modelos inteligentes, podendo afetar a população que se tentaria auxiliar, mas desde que o treinamento seja adaptado às necessidades específicas e leve em conta dados contextuais, não há como a IA devolver outra coisa que não seja justamente suporte e ajuda, se o treinamento for bem executado. Conforme já se afirmou diversas vezes, a IA só retornará e devolverá ações com base nos dados e treinamento feitos.

O uso justo e viável da IA, especificamente em saúde, sem correr riscos de impactar negativamente a sociedade é possível, principalmente se nos concentramos em avanços pequenos, mas palpáveis. A utilização dos softwares para criação e design de medicamentos, escaneamento e identificação de problemas de saúde por meio da imagem, e usos já em vigor hoje em dia, podem ser considerados “pequenos” diante da potencialidade da IA, no entanto, são usos que já otimizam e aprimoram sistemas que antes demandavam muitos recursos. A criação de pequenos blocos de legislação e vigilância do uso da IA socialmente já resolveria um tanto de problemas aqui citados. Para fins de estudo de caso, um estudo envolvendo pesquisadores de diversos países, inclusive países subdesenvolvidos, como Nigéria e Rússia, traz à tona diferentes tipos de metodologia que poderiam ser empregadas para mitigar os problemas com a IA no campo da saúde⁷². O texto é significativo, pois nos informa de maneira direta e contundente em que e para onde direcionar os esforços na implementação dessas técnicas, conforme podemos ver no Quadro 1 abaixo:

⁷² Mitigar porque eliminar ou zerar os erros seriam impossíveis, opinião compartilhada com Tarcízio Silva (2020). Apesar dessa ideia de sistemas neutros ser amplamente compartilhada, não parece ser realmente executável.

Quadro 1 - Maneiras e práticas seguras para implementação de IA

1. Privacidade e Segurança de Dados:

Estratégia para proteção dos dados dos pacientes: Implementar métodos robustos de criptografia para garantir a segurança dos dados dos pacientes durante a transmissão e armazenamento. Seguir padrões de segurança de dados reconhecidos internacionalmente, como o Health Insurance Portability and Accountability Act (HIPAA) nos Estados Unidos e o Regulamento Geral de Proteção de Dados (GDPR) na União Europeia *ou a Lei Geral de Proteção de Dados Pessoais (LGPD) brasileira e demais regulações.*

Melhor prática: Realizar auditorias regulares e atualizar os protocolos de segurança para se adaptar às ameaças de cibersegurança em constante evolução. Treinar a equipe de saúde nas melhores práticas de segurança de dados para prevenir violações causadas por erros humanos.

2. Manuseio e Armazenamento Responsável de Dados:

Estratégia de manuseio de dados: Anonimizar os dados dos pacientes sempre que possível para proteger a privacidade individual. Considerar a adoção de abordagens de aprendizagem federada que permitem que modelos de IA sejam treinados em dados descentralizados sem centralizar informações sensíveis.

Melhor prática: Desenvolver diretrizes de manuseio de dados específicas para IA e robótica na saúde, definindo procedimentos para coleta, armazenamento e compartilhamento de dados. Avaliar regularmente as práticas de manuseio de dados para garantir conformidade com os padrões éticos.

3. Viés e Justiça:

Estratégia para abordar a discriminação algorítmica: Para mitigar o viés, empregar diversos e representativos conjuntos de dados durante o treinamento do modelo de IA. Monitorar continuamente os algoritmos de IA em busca de viés e recalibrá-los conforme necessário.

Melhor prática: Colaborar com equipes interdisciplinares, incluindo *eticista, bioeticistas, filósofos, sociólogos e demais especialistas* para avaliar o potencial impacto social dos sistemas de IA e identificar vieses nos processos de tomada de decisão.

4. Garantindo a Justiça na Tomada de Decisão Baseada em IA:

Estratégia para transparência e explicabilidade: Desenvolver algoritmos de IA com recursos de transparência integrados, permitindo que os profissionais de saúde compreendam o raciocínio por trás das recomendações.

Melhor prática: Implementar técnicas de IA explicáveis que forneçam explicações claras e compreensíveis para as decisões impulsionadas por IA, aumentando a confiança e a responsabilidade.

5. Responsabilização e Transparência:

Estratégia para uma clara responsabilização pelas ações dos sistemas de IA: Estabelecer cadeias transparentes de responsabilidade para sistemas de IA e robótica, definindo os papéis e deveres de fabricantes, instituições e profissionais de saúde.

Melhor prática: Desenvolver mecanismos de denúncias de incidentes que permitam que, prontamente, os profissionais de saúde possam relatar erros ou eventos adversos do sistema de IA. Criar uma cultura de controle de ações e responsabilidade dentro das organizações de saúde.

Fonte: Adaptação feita pelo autor baseado em Elendu, C, *et al.* (2023, p. 3, tradução nossa)⁷³.

⁷³ A adaptação se deu por conta da edição e adição de pequenos trechos, indicadas no texto pelo uso do itálico, visando melhor utilização do enxerto no trabalho. Favor conferir versão original para maiores esclarecimentos.

A proposta de Elendu e colegas (2023) indica pontos muito úteis para nossa pesquisa, ofertando mitigações e saídas de problemas que vimos durante todo esse percurso intelectual. A cadeia de responsabilização entre as empresas fabricantes, as instituições e os profissionais envolvidos no campo da saúde representaria um avanço significativo pois com ela poderíamos evitar o debate sobre quem responsabilizar por atos cometidos por IA. Para liquidar de vez a possibilidade de culpabilizar as máquinas pelas ações cometidas, podemos lembrar que Bryson (2019) afirma que dar agenciamento moral para robôs poderia ser interpretado como uma ação imoral, isto é, seria completamente antiético tentar equiparar em grau de comparação ferramentas e artefatos com o humano, devido a capacidade moral humana ser indubitável.

A anonimização e a criptografia apoiadas num sistema maior de responsabilidade compartilhada entre os profissionais da saúde já instruídos em como manter e cultivar a proteção de dados parece ser o suficiente para lidar com a maioria dos problemas de vazamento de dados médicos já encontrados. A responsabilidade compartilhada também deve abranger os pacientes, já que também deverão ser informados e consentir com os processos digitais aos quais deverão ser submetidos.

Segundo Elendu e pesquisadores (2023, p. 3) diretrizes éticas devem ser criadas e atualizadas para estarem disponíveis para todos os envolvidos com os processos de saúde. Se essa iniciativa partisse do governo, é possível que tivesse maior aderência por parte das empresas criadoras de IA, pois eles deverão seguir as regras contidas no documento para assim poderem atuar no país. A união europeia já tem o seu *ethics guidelines for trustworthy AI*, que serve a esse propósito. Em países subdesenvolvidos, a situação já foi iniciada pelo governo brasileiro com a EBIA.

Aprofundando em uma análise do uso desses sistemas inteligentes para mitigar a falta de acesso e recursos e demais problemas em geral da saúde enfrentados por países subdesenvolvidos Fletcher, Nakeshimana e Olubeko (2021) tratam justamente de como os modelos de AM e IA poderiam lidar justamente com as especificidades desses países com falta de infraestrutura e que por vezes têm a desigualdade social econômica refletida também no setor da saúde. Os autores argumentam que pela falta de técnicos e especialistas nessa tecnologia de ponta, tais países poderiam ser tolhidos no processo de uso de IA, o que indica uma dificuldade para o desenvolvimento de sistemas nacionais. A falta de proteção legal e o preconceito social existente contra certos grupos nessas sociedades também podem ser fatores que corroborem para um mau uso da IA nessas regiões. Dessa forma, os autores desenvolvem três etapas indicadoras que podem favorecer a criação ética de IA, sendo eles os processos de adequação, o viés e a equidade, em tradução livre. A metodologia contribui muito com o

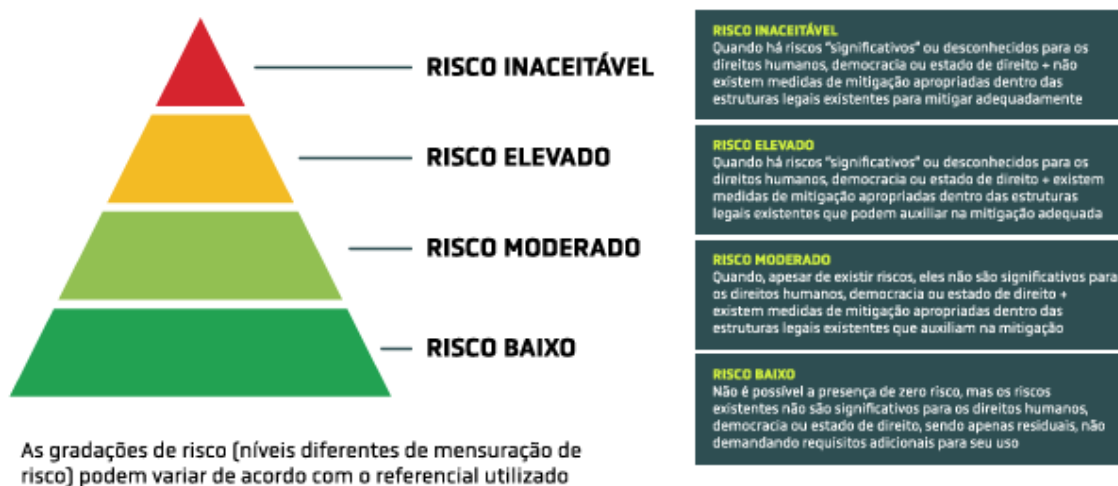
desenvolvimento do campo chamado de equidade na aprendizagem de máquina (fairness in machine learning). A adequação é o fator de conexão entre o contexto local e a população-alvo. Já a etapa de viés, se referente ao reconhecimento da tendência de um modelo para favorecer um grupo demográfico no lugar de outro, trazendo injustiça e discriminação. O fator de equidade existe especificamente para reparar isso, pois ele envolverá avaliar o impacto nos diferentes grupos sociais e optar por uma das definições matemáticas de equidade que satisfaça o conjunto de requisitos legais (jurídicos), culturais e éticos. O artigo conta ainda com hiperparâmetros matematicamente acurados que podem embasar a utilização e avaliação de sistemas de AM e IA na saúde global.

A nível global, de acordo com Belli (2020), os países do BRICS (Brasil, Rússia, Índia, China e África do Sul), um importante bloco do Sul global, abrigam 42% da população mundial, com isso quero dizer uma quantidade bastante significativa de dados sendo produzido por quase metade da população do planeta, que são abrigados por esses governos. Não sem razão, a maior parte desses países foi citada nesse estudo, por serem países com características atrativas para empresas extrativistas de dados. É preciso que esses países estejam e continuem atentos às ofertas “gratuitas” que são pagas com dados e por vezes afetem a soberania nacional. Observando os exemplos e os outros passos supracitados, bem como tendo noção do caráter dual do devir curupira, é capaz de poderem combater de forma eficaz o colonialismo de dados e o perigo dos algoritmos de destruição em massa (IA), uma das estratégias indispensáveis é uma robusta malha jurídica. A proteção legal e cobertura jurídica para os temas de IA são um dos tópicos mais discutidos na academia por conta dos desafios que a ferramenta impõe ao direito constitucional.

De forma geral, podemos entender que a maior parte das legislações atuam classificando a implementação de IAs por meio da regulação baseada em riscos, já que “essa estratégia de regulação e designação macro de faixas de risco foi adotada por diversas normativas regulando sistemas de inteligência artificial: o PL 2338/23 no Brasil, o EU AI Act, o Proyecto de ley 15869/19 do Chile e a ferramenta de avaliação de impacto do Canadá” (Bioni; Garrote; Guedes, 2023, p. 41). E apesar de ser uma técnica muito diversificada, com extensas variações entre elas a depender da legislação avaliada, ainda podem ser categorizadas por trabalharem no mesmo sentido de pautarem a IA previamente por seus potenciais impactos. As variações podem variar tanto para os nomes dados a essas faixas de risco quanto para a metodologia escolhida. A título de destaque, conforme indicado pelos autores, o PL 2338/23 brasileiro é o único que realiza uma divisão mais adaptada aos riscos maiores e mais intensos, criando as categorias de risco excessivo e alto, além das categorias dos dois primeiros

níveis (baixo ou moderado). De forma metafórica podemos entender os riscos como uma espécie de pirâmide, como a exemplar pirâmide constituída pelos pesquisadores do Data Privacy Brasil da Figura 6:

Figura 6 - Classificação e definição de riscos do uso da IA



Fonte: Bioni; Garrote; Guedes, 2023, p. 45.

O problema de todas essas regulações é pressupor uma racionalização e compreensão das consequências advindas da tecnologia. David Collingridge (1980) formulou o que ficou conhecido como o dilema de Collingridge, que basicamente põe diante dessas formulações futuras sobre a tecnologia um problema duplo:

A) Os impactos tecnológicos não podem ser facilmente previstos até que a tecnologia seja amplamente desenvolvida e utilizada⁷⁴.

B) O controle e a mudança são difíceis quando a tecnologia já está enraizada.

Ou seja, enquanto mudar os processos tecnológicos que estamos vivenciando for possível e fácil de ser executado, não enxergaremos razão para mudar; Quando a mudança se torna aparente e necessária, já estamos dependentes e, portanto, a mudança é demorada, cara e difícil. Assim, em um contexto onde a tecnologia avança exponencialmente com diversas potências em muitas direções, a legislação caminha a passos lentos, tentando prever e limitar processos danosos que o uso da IA possa causar a alguém. Sobre esse tópico na área da saúde, deve-se recordar do estudo feito por Castagno e Khalifa (2020), que afirma que mesmo profissionais da área da saúde e pacientes têm pouco conhecimento sobre o uso IA no campo.

⁷⁴ Devido à tamanha aderência social à IA, pode-se pressupor que estaríamos perto de ver todas as consequências e impactos possíveis dessas ferramentas. Mas com a IA generativa novas portas se abrem, o que teoricamente afasta da realidade essa premissa.

Assim, apenas os especialistas poderiam utilizá-las de forma consciente, o que sabemos que não é o que acontecerá na realidade.

A aposta e os riscos se intensificam ainda mais quando estamos falando de uma tecnologia cujos limites ainda estão se estabelecendo, pois, a própria ferramenta está tendo seus usos explorados e definidos no momento atual, como é a IA generativa. Nenhum dos projetos de lei tramitados no Brasil até o presente momento cita e tenta limitar a IA generativa de forma direta. Nenhuma das PL que abrangem o tema da IA (PL 5051/19; PL 21/20; PL 872/21; PL 759/23) descrevem as generativas, e a PL 2338/23 cita apenas as IA de propósito geral. Não irei elaborar aqui uma crítica de que na legislação do parlamento europeu, “por questões de segurança” a regulação não irá abranger os usos militares ou com finalidade de defesa nacional, já que essa exclusão poderia resultar na criação de todo um artigo a parte, porém, ao deixar temas tão centrais e potentes de fora da regulamentação, se permite a criação armas (e não mais apenas ferramentas), com potencial letal e de opressão. Se quando utilizadas sem essa finalidade a IA já atinge um público massivo e de forma tão violenta, quando é planejada e treinada com esse propósito sem se encaixar nas classificações de risco é uma permissividade para uma atuação brutal e devastadora contra todo indivíduo ou grupo que seja considerado de potencial risco a nação. Isso é, quando por conta de descaso ou negligência não ocorrerem massacres por um simples “erro técnico”.

Com esse dilema não quero desestimular a produção de regulamentações, mas sim indicar que talvez formulações de legislações reativas⁷⁵ pudessem ser um caminho mais interessante e propício ao lidar com uma tecnologia inovadora e geradora por definição. Conforme corroboram Bioni, Garrote e Guedes (2023, p. 121) que alegam que uma saída poderia ser “por meio do estímulo à produção de bancos de dados e sistemas de IA éticos diversos, abertos e multidisciplinares em território nacional, em combinação com o fomento à educação e capacitação”. Tal proposta elaborada parece ainda mais mitigadora de danos quando a aplicamos em países latinoamericanos ou em contexto do sul global, em situações de países pobres, com desigualdades sociais profundas e opressões estruturais abismais. Nessa situação é preciso recordar que uma “regulação de IA não seja uma importação acrítica de modelos regulatórios de outros contextos”. A título de destaque, os pesquisadores elucidam que dentre os projetos brasileiros, apenas o PL 2338/2023 avança na “tropicalização do debate”, considerando o Brasil como permeado por assimetrias e desigualdades estruturais.

⁷⁵ Segundo os autores do Data Privacy Brasil (2023, p. 67-85) as reativas seriam aquelas que tentam propor reações aos modelos. As reativas estariam em contraposição com as legislações defensivas (ou preventivas), que tentam prevenir os riscos e consequências ilegais oriundas dos sistemas inteligentes.

Para fins de encerramento, destaca-se que há uma preferência por parte da população de que esses temas, principalmente se tratando da saúde, sejam regidos pelos governos, já os cidadãos tendem a confiar mais em empresas públicas por conta do empenho com o bem público. Conforme alegam os especialistas:

Os pacientes e os pais biológicos entrevistados para este estudo expressaram a opinião de que as instituições públicas, como as universidades e o NHS, são mais confiáveis do que as empresas privadas, devido ao seu comprometimento em promover o bem público, enquanto as empresas privadas são movidas pela maximização do lucro. [...] Esta posição não é exclusiva dos participantes nesta investigação [...]; outros estudos também demonstraram o ceticismo do público sobre as motivações das empresas privadas no contexto da saúde (Dlugatch; Georgieva; Kerasidou, 2023, p. 10, tradução nossa).

Mas a gerência governamental não pode minar e dominar os esforços sociais da população em prol de uma tomada de decisão e tomada de poder coletiva, visando a saída dessa situação tão complexa. Parece um tanto utópico sugerir que essa população citada aqui no texto por vezes em situação de vulnerabilidade reaja e se engaje em busca de novas condições de vida, mas com esse leviatã⁷⁶ neoliberal-capitalista engendrado em todos os âmbitos pessoais, também se abrem frestas de poder emancipatórios dentro desses novos âmbitos.

Se é verdade que o capitalismo se estendeu a novos campos da vida, também é verdade que as lutas emancipatórias ganharam novas possibilidades abertas por essa mesma extensão. A reintrodução indireta e compulsória (não paga) no processo produtivo desse trabalhador que foi expulso da fábrica ou desse usuário comum que está navegando em seu momento de ócio faz com que as cidades, os parques e as camas componham parte da esteira produtiva estendida de produção e extração de valor. Se este diagnóstico estiver correto, a humanidade está exposta como nunca antes, mas o capital também (Faustino; Lippold, 2023, p. 186-187).

Os repórteres Roger Montti (2023) e Omer Kabir (2023) ao relatarem um documento vazado da Google, indicam que a corrida armamentista das IAs tem tido um concorrente pesado que vem de fora do meio empresarial, que pode sair vitorioso dessa disputa: a produção dos códigos abertos, isto é, uma grande e ampla comunidade de programadores que se dedicam a modificar os códigos dos produtos para liberar o acesso e melhorar a experiência para todos; na IA conseguem modificar os treinos de aprendizagem ou modificar regras anteriormente impostas que inibiam certos comportamentos. No entanto, em sua maioria atuam em cima de

⁷⁶ Monstro bíblico com vários tentáculos, foi usado por Hobbes para representar o poder “absoluto” do Estado durante a modernidade.

produtos patenteados, cometendo crime de pirataria ao fazerem isso⁷⁷. Mas se a falta de regularização e o vazio jurídico beneficia as empresas conhecidas permitindo que elas abusem de certos aspectos de seus usuários, como a lei poderia atingir milhões de indivíduos que frequentemente se unem anonimamente apenas para se ajudarem a desvendar ou criar o código almejado? A indicação dos repórteres pode servir como indicativo que apesar da corrida de *algoritmos de destruição em massa* ter contornos aparentemente definidos, suas regras podem e estão sendo subvertidas na mesma velocidade que o desenvolvimento da IA: velozmente.

Nos últimos tempos, estão surgindo diversas iniciativas que sustentam essa afirmação de que a inteligência coletiva e comunitária poderia sanar muitos dos problemas advindos do colonialismo de dados e consolidar a soberania digital não só nas mãos do Estado. Para fins de exemplo podemos observar mais recentemente, o Bibi Mob, que em 2019, em meio a calorosas discussões sobre os exploratórios trabalhos por meio das plataformas digitais, a empresa lança uma plataforma que cobra do motorista taxas mais baixas que os gigantes concorrentes da categoria, podendo atuar como franquia em prefeituras ou empresas locais. O CEO visou diminuir as extensas jornadas de trabalho e a baixa remuneração dos motoristas de aplicativo, ao mesmo tempo que promovia a segurança tanto do passageiro quanto do motorista. Souza (2022) indica que o diferencial do aplicativo, além dessas propostas de maior segurança é o valor de repasse para o trabalhador: 95% do valor fica com o motorista. O aplicativo já está presente em duzentas cidades. Esse é um dos exemplos das chamadas tecnologias de propriedade de trabalhadores, como explica Grohmann.

Rafael Grohmann (2023) também cita em outro momento, o caso de sucesso do MTST, que desde 2019 se preocupa com as tecnologias, criando até mesmo um núcleo para isso. Com isso, “um dos movimentos sociais urbanos mais fortes da América Latina compreende a tecnologia com um papel central na luta de classes”, criando até uma cartilha sobre soberania digital⁷⁸. A iniciativa do movimento popular combina a “reapropriação de tecnologias em prol da classe trabalhadora, oportunidades de renda, organização de trabalhadores de diferentes setores e luta por soberania popular”. O interessante de se observar nos projetos criados pelo núcleo de tecnologia é que são alinhados ao contexto brasileiro, já que no Brasil, para muitas pessoas, a internet se limita apenas ao *WhatsApp*, e por conta disso, os técnicos pensaram em um design já integrando essa característica na formatação do projeto *Contrate Quem Luta*, por

⁷⁷ O que acaba por ofertar mais domínio a essas empresas, sobre isso, ver: FERES, V. C. M.; OLIVEIRA, V. J.; GONÇALVES, D. D. Robin Hood às avessas: software, pirataria e direito autoral. **Revista Direito GV**, v. 13, n. 1, p. 69-94, abr. 2017.

⁷⁸ A carta pode ser acessada em: <https://nucleodetecnologia.com.br/cartilha/>. Acesso em: 30 mar. 2024.

exemplo. Assim, o MTST supera o tecnosolucionismo barato, achando que as tecnologias por si só resolveram o problema, mas sim o contexto e principalmente o uso que se fará da ferramenta digital. Oferecendo assim, “lições de uma emergente organização de trabalhadores de diferentes setores a partir de um movimento social”, com a participação do “poder popular construindo e reapropriando tecnologias, organizando e circulando lutas, e também pressionando o Estado em prol de políticas públicas que garantam direitos e dignidade [...]”. Como esclarece o autor, as pautas de economia solidária e tecnologias nunca precisaram estar tão articuladas entre si até agora, mas na atual conjuntura, esse parece ser um dos caminhos que podem trazer bons resultados dentro dessa luta computadorizada por bem estar social.

Em um outro caso, vemos que a iniciativa estatal pode sim ser muito profícua. Betim (2020) traz à tona o caso da mumbuca, moeda social criada no município carioca de Maricá, que ajuda famílias vulneráveis da cidade, incentivando o comércio local com cerca de um milhão por mês, que representa 10% da movimentação dos comércios da cidade. Durante a pandemia de COVID-19, preservou empregos e estabilidade a diversas famílias, os ajudando a ter o mínimo necessário para viver⁷⁹, tornando o banco mumbuca um farol de esperança na turbulência global. Nesse sentido, Maricá tornou-se pioneira na implementação de renda básica universal, além de ter seus ônibus urbanos sem tarifas, facilitando o acesso de seus moradores na locomoção pela sua própria cidade. Modelos como esses diminuem a influência, dependência e poder de instituições grandes e também fortalecem as relações entre os moradores, promovendo economias mais justas e solidárias. Assim, a mumbuca e as demais produções aqui citadas, deixam de ser apenas colaborações, mas se tornam um símbolo de resistência e luta contra as desigualdades oriundas de um sistema acumulador que exacerba diferenças sociais já existentes. Representa bem a ideia de que por meio de colaboração, inteligência coletiva e uso consciente da tecnologia é possível construir sistemas e modelos sociais mais sustentáveis e eticamente adequados, aliviando pressões econômicas sobre-humanas e retirando o poder das garras do neocapitalismo voraz para dar o controle nas mãos do povo. Os bioeticistas, médicos, cientistas da computação, entusiastas tecnológicos que se preocupem e demais adeptos que possam se juntar nessa luta, deveriam procurar meios de se engajar ativamente nesse processo.

A única coisa que não pode acontecer nessa luta social computadorizada, é se deixar levar por esse embasbacamento tecnológico de crer na neutralidade. Se hoje uma das formas de dominação são as tecnologias, lá também devem ser um campo de batalha em busca de

⁷⁹ O que pode parecer pouco significativo, mas em tempos de capitalismo tardio onde o mínimo não é alcançado por dezenas, torna-se vital esse tipo de ajuda. Novamente, passos pequenos, mas significativos.

libertação. Conforme recomendam os estudos fanonianos para a emancipação “nesse esforço, não se trata de refutar ou adorar a tecnologia, mas de colocar ‘a ciência e a tecnologia a serviço da emancipação’” (Faustino; Lippold, 2023, p. 180). Faustino e Lippold (2023, p. 182) esboçaram uma crítica hacker-fanoniana ao colonialismo digital, esclarecendo sobretudo que: “a tarefa colocada não é a de demonizar ou endeusar as redes e plataformas, [...] Isso implica dizer que o problema não é o aprendizado de máquinas ou a chamada inteligência artificial, em si, mas os sentidos pelos quais são projetados e, sobretudo, os usos que lhes atribuímos”. Dessa forma, descolonizar a tecnologia e confrontar a missão civilizatória “em novos moldes high-tech é, antes de qualquer coisa, colocar em xeque o caráter destrutivo do modo de produção capitalista em todas as suas dimensões sutis e declaradas”.

Esses três exemplos, sendo uma iniciativa de uma empresa, outro por mobilização popular e o terceiro sendo liderado pela própria prefeitura, demonstram que a desmobilização desses sistemas de opressão e construção de modelos tecnológicos responsáveis e éticos é algo possível e executável de diferentes formas⁸⁰. Como forma de sintetizar as medidas indicadas para alcançar a construção ética de modelos de IA, podemos levar em conta os seguintes passos:

Recomenda-se não perder de vista que os modelos inteligentes são criados em certa estrutura social, mas não podem reproduzir as mazelas que afetam essa estrutura. Com isso, o princípio curupira (ou devir curupira) deve ser adotado durante todo o processo de desenvolvimento ou implementação de sistemas inteligentes para se reconhecer que apesar de eles representarem o futuro, suas bases epistemológicas, culturais e sociais são moldadas pelo passado. Para lidar com isso, é de extrema importância que se leve em conta no desenvolvimento ou adaptação dos contextos, valores e preocupações relativas à realidade concreta e finalidade que se utilizará a IA. Treinamentos computacionais que utilizem bases de dados diversificadas que tenham seus impactos avaliados por equipes multidisciplinares plurais, são altamente desejados, visando que tais equipes plurais possam validar, se preciso for, reajustar o sistema até ele apresentar resultados satisfatórios e não danosos para algum grupo vulnerável. Nesse processo de validação ou auditoria, a priorização de mecanismos que ofereçam maior explicabilidade das decisões tomadas pelos modelos facilitaria o serviço, diminuindo a opacidade e aumentando a confiança e transparência no modelo inteligente. A opção de dividir os processos usando a metodologia de adequação, viés e a equidade pode

⁸⁰ Marcelo Faria explora as possibilidades de resistência por meio da “técnica” até mesmo com grupos mais historicamente explorados, como um grupo indígena, que criou uma moeda indígena digital chamada oyxabaten, apontando um uso territorializado da técnica, ver em Silveira; Souza; Cassino (2021, p. 191-210).

aumentar a efetividade do macroprojeto, desde o desenvolvimento até a implementação prática da IA. De toda forma, os danos causados só podem ser previstos, não há como conceber todas as utilizações possíveis, devido a isso cadeias de responsabilização devem ser implementadas⁸¹, com a utilização de elementos que permitam relatos de incidentes e denúncias. A utilização de técnicas de criptografia e anonimização no armazenamento dos dados pessoais é premente quando estamos pensando na segurança deles. Por fim, com ato político ou ação da esfera governamental, a formulação de estratégias e legislações que fomentem e apoiem elementos mais basilares da IA parecem indicar um sucesso maior na mitigação dos danos causados⁸², promovendo aspectos mais reativos para lidar com as causas dos problemas ao invés de focar apenas nos riscos possíveis trazidos pelas inteligências artificiais.

As principais considerações deste estudo se evidenciam ao contrapor de forma clara os pilares que moldaram a pesquisa e as observações colecionadas ao encerrá-la. Assim, de forma a encaminhar para o fim, devo elucidar ao leitor que as evidências mobilizadas, confrontadas com os vitais objetivos e hipóteses, nos levam às seguintes conclusões:

Ao se investigar em que medida a IA está modificando as escalas de reprodução da colonialidade socialmente, percebe-se que o fenômeno da datificação da vida impacta e escalona os danos trazidos pela colonialidade de formas antes não registradas. E tal situação não é vista como problemática, o que suscita a primeira hipótese, na qual *a IA é tida como um produto tecnológico imparcial e isento de vieses. Mas pode conter e reforçar vieses sociais que impactam seu uso.* Essa hipótese se baseia nos elementos do dataísmo e do solucionismo tecnológico vislumbrados no texto. Ao se debruçar sobre os problemas perpetuados advindos da colonialidade, foram encontradas evidências, principalmente pelas ofertas nas quais grandes empresas tecnológicas prometeram auxílio digital a diversos países subdesenvolvidos, mas estavam interessadas em explorar e gerir os dados gerados por esses países, o que impediu e dificultou a criação e desenvolvimento nacional de sistemas próprios desses países, afetando a soberania digital nacional. Essas ofertas evidenciam os problemas a nível nacional, porém os impactos a nível pessoal também são poderosos e tem seus vestígios rastreados por resultados preconceituosos em motores de busca e uso de dados inadequados, acarretando em idealizações opressoras sendo reproduzidas para um número gigantesco de usuários. Essas

⁸¹ No caso da utilização de IA na saúde, é recomendável que a cadeia de responsabilização seja mais robusta e inclua o paciente, o deixando ciente de todos os procedimentos possíveis.

⁸² Pois as preventivas já estão em processo e apesar de limitadas, já ofertam itens para lidar com os riscos trazidos pela IA; com as reativas teríamos duas frentes e uma maior possibilidade de sucesso de uma IA ética.

situações indicam uma mudança funcional no sistema econômico para uma valorização da informação como lucro, uma espécie de capitalismo informacional.

Nesse capitalismo focado na informação, os dados são essenciais, e no âmbito da saúde, uma hipótese vislumbrava um impacto considerável do uso das novas tecnologias com os pacientes, questionando como a IA poderia *afetar negativamente os pacientes e cidadãos*. Os abusos na área da saúde são indicados na maioria dos casos por vazamentos e uso indevido de dados, mas como todo o sistema econômico está se mobilizando para pregar os dados individuais, os dados médicos ficam em posição especial de risco, haja visto que atualmente se produzem muitos dados e com a complexidade dos modelos preditivos de IA, diversos dados se convertem em dados médicos indiretos, inferidos e até mesmo invisíveis. Destaca-se que mesmo sistemas de saúde avançados em países economicamente potentes, parecem não conseguir lidar bem com a invasão e vazamento de tais dados. No entanto, a hipótese foi parcialmente refutada, dado que muitos usos benéficos foram encontrados nos mesmos estudos que apontavam circunstâncias danosas aos pacientes. A IA se tornou parte essencial do âmbito médico e já é utilizada para ajudar populações inteiras desde o parto de bebês, como também apoiar escolhas de prioridades, além de atuar na prevenção de mortes a um nível municipal, promovendo uma vida mais digna e saudável aos vulneráveis.

Este estudo descreveu relações entre a colonialidade e a IA, partindo de uma análise específica baseada em autores diversos e que não formam um consenso entre si sobre o melhor termo de descrever o estágio atual de capitalismo, somente concordando em suas descrições. Conseqüentemente, a interpretação das relações desses itens pode ser drasticamente alterada mediante aos autores utilizados. As contribuições poderiam ser ampliadas e mais eficazes se o estudo se focasse em um linguajar mais técnico, voltado para o meio tecnológico que é próprio para o desenvolvimento de IA. A identificação dos ab(usos) de IA mencionada nos objetivos específicos se comprovou de forma pouco consistente, podem ter as práticas no campo da saúde ampliadas e analisadas em trabalhos futuros. Durante a escrita dessa pesquisa identificaram-se questões tangenciais que permitiriam outros estudos abordarem o fenômeno estudado para aperfeiçoar o entendimento ou para uma maior confirmação empírica dos resultados obtidos.

Os resultados sugerem que apesar de haver factualmente usos positivos de IA, não se pode omitir as evidências de má utilização ou inação para prover métodos de mitigação de danos a população. Nesse sentido, o *dever curupira como princípio e outros métodos podem auxiliar no desenvolvimento de sistemas de IA mais equitativos e adequados aos contextos em que serão utilizados*, isto é, o princípio curupira serviria como um alerta para as equipes de

desenvolvimento de softwares inteligentes para estarem cientes da tendência que as novas tecnologias têm em serem traiçoeiras, ressuscitando e potenciando de forma intensa problemas do passado que ainda habitam nosso construto colonial capitalista em nossa sociedade. Dessa maneira, ao compilar nesse estudo ferramentas teóricas e contribuições bioéticas da BI e BP, espera-se lançar luz a possíveis veredas em que se pode imaginar e principalmente vivenciar um futuro equitativo e com opressões automatizadas mitigadas, independente do grupo social, dadas as devidas limitações e especificidades dos países e populações.

Assim, muitos dos resultados deste trabalho podem ser aproveitados para se alinhar com as recentes tendências do campo da bioética de se debruçar sobre o uso de IA e também de expandir as correlações entre os estudos sobre colonialidade, IA e bioética, propondo ações efetivas em problemas reais.

REFERÊNCIAS

ABARIEGO, J.; AMARAL, A. J.; SALLES, E. B. C. (org.). **Algoritarmismos**. São Paulo: Tirant lo blanch, 2020.

ABBAGNANO, N. **Dicionário de Filosofia**. São Paulo: Martins Fontes, 1998.

ALBUQUERQUE, K. SUS Digital: como o Ministério da Saúde quer usar IA para diminuir filas. **Tecmundo**, 2023. Disponível em: <https://encurtador.com.br/jzLOX>. Acesso em: 06 dez. 2023.

ALVES, R. Tecnologia e humanização. **Revista Paz e Terra**, n. 8, 1968.

AMARAL, A. J.; ELESBÃO, A. C. S. Racismo e sexismo algorítmicos: um estudo de caso sobre o mecanismo comercial de busca do Google. **Revista Eletrônica do Curso de Direito da UFSM**, v. 17, n. 1, 2022. DOI: <https://doi.org/10.5902/1981369466455>.

AMARAL, J. A.; SALLES, E. B. C. Pandemia, vigilância e os perigos do “solucionismo tecnológico”. In: SOBRINHO, L. L. P.; CALGARO, C.; ROCHA, L. S. (org.). **Covid-19: Ambiente e Tecnologia**. Itajaí: Ed. Univali, 2020. p. 94-109.

ANAHP. Publicação anual da Associação Nacional de Hospitais Privados. **Observatório Anahp**, v. 14, 2022.

ARABI, S. A. A. **Aplicação da inteligência artificial nas indústrias farmacêuticas**. 2020. 35 f. Trabalho de Conclusão de Curso (Graduação em Ciências Farmacêuticas) – Universidade de São Paulo, São Paulo, 2020.

AUGUSTO, T.; ROSA, P. F. Tecnologia para filtros em redes sociais reforça padrão racista. **Folha de São Paulo**, Cotidiano, 2021. Disponível em: <https://encurtador.com.br/nxRZ3>. Acesso em: 5 fev. 2023.

BADIA, L. M. R. Algunas reflexiones sobre el presente y futuro de la inteligencia artificial. **Novatica**, n. 234, p. 97- 01, 2015.

BARRETO, M. L. Desigualdades em saúde: uma perspectiva global. **Ciência & Saúde Coletiva**, v. 22, n. 7, p. 2097-2108, 2017. DOI: <https://doi.org/10.1590/1413-81232017227.02742017>.

BASSETTE, F. 6 milhões de pacientes morrem todos os anos por erros evitáveis na assistência à saúde, diz OMS. **Folha de São Paulo**, 2023. Disponível em: <https://encurtador.com.br/mHNOZ>. Acesso em: 21 dez. 2023.

BBC BRASIL. O escândalo que fez o Facebook perder US\$ 35 bilhões em horas. **BBC News Brasil**, 19 mar. 2018. Disponível em: <https://www.bbc.com/portuguese/internacional-43466255>. Acesso em: 17 abr. 2021.

BELLI, L. Brasil precisa reconstruir sua soberania digital. **Estadão**, 2024.

BELLI, L. BRICS countries to build digital sovereignty. **CyberBRICS**, p.71-280, 2020.

BERNARDINO-COSTA J; GROSFUGUEL R. Decolonialidade e perspectiva negra. **Revista Sociedade e Estado**, v 31, n. 1, p. 15-24, 2016. DOI: <https://doi.org/10.1590/S0102-69922016000100002>.

BETIM, F. Maricá, no Rio, preserva empregos e negócios na pandemia e coloca a renda básica no centro do debate. **El País**, Sociedade, 2020.

BIONI, B.; GARROTE, M.; GUEDES, P. **Temas centrais na Regulação de IA: O local, o regional e o global na busca da interoperabilidade regulatória**. São Paulo: Associação Data Privacy Brasil de Pesquisa, 2023.

BRANDEBURGO, E.; GARCIA, H. As origens da crise ambiental antropocêntrica e a ofensa à dignidade do ser humano. *In: ENCONTRO DO CONSELHO NACIONAL DE PESQUISA E PÓS-GRADUAÇÃO EM DIREITO - CONPEDI*, 6., 2023, Florianópolis. **Anais eletrônicos** [...]. Florianópolis: CONPEDI, 2023. Disponível em: <http://site.conpedi.org.br/publicacoes/4k6wggq8v/9qtv6736>. Acesso em: 15 fev. 2023.

BRYSON, J. J. The past decade and future of AI's impact on society. *In: GONZALÉZ, F. et al. Towards a new enlightenment?: a transcendent decade*. Madrid: Turner, 2019.

BURKART, N.; HUBER, F. M. A survey on the explainability of supervised machine learning. **The Journal of Artificial Intelligence Research**, v. 70, p. 245-317, 2021. DOI: <https://doi.org/10.1613/jair.1.12228>.

CALO, R. Robotics and the lessons of cyberlaw. **California Law Review**, Berkeley, v. 103, n. 3, p.513-563, June 2015.

CAMARGO, G. X. **A vedação à gratuidade compulsória dos serviços digitais como forma de proteção dos dados pessoais dos usuários consumidores e mitigação do abuso de posição dominante pelas plataformas de dois ou múltiplos lados**. 2020. 211 f. Dissertação (Mestrado em Direito) – Universidade Federal de Santa Catarina, Florianópolis, 2020.

CARVALHO, Y. M. Do velho ao novo: a revisão da literatura como método de fazer ciência. **Revista Thema**, v. 16, n. 4, p. 913-928, 2019. DOI: <https://doi.org/10.15536/thema.V16.2019.913-928.1328>

CASCUDO, L. C. **Dicionário do folclore brasileiro**. 11. ed. [S. l.]: Global, 2001.

CASTAGNO, S.; KHALIFA, M. Perceptions of artificial intelligence among healthcare staff: a qualitative survey study. **Frontiers in Artificial Intelligence**, 2020. DOI: 10.3389/frai.2020.578983.

CASTELLS, M. **A sociedade em rede—volume I: a era da informação: economia, sociedade e cultura**. 8. ed. São Paulo: Paz e Terra, 1999.

CERQUEIRA, R. C. D. *et al.* **Uma análise da base de dados do Sistema de Informação Hospitalar entre 2001 e 2018: dicionário dinâmico, disponibilidade dos dados e aspectos metodológicos para a produção de indicadores sobre violência**. Rio de Janeiro: IPEA, 2019.

CONSELHO FEDERAL DE MEDICINA. **LGPD**: a Lei Geral de Proteção de Dados Pessoais e a atuação do profissional de medicina. Brasília: CFM, 2022.

CHIAVEGATTO FILHO, A. D. P. **Aula 1 - Aplicações de inteligência artificial e machine learning em saúde**. São Paulo: USP, 2018. 1 vídeo (16 min). Publicado pelo Canal USP. Disponível em: <https://www.youtube.com/watch?v=y8em7JhKwhU&list=PLAudUnJeNg4tvUFZ8tXQDoAkFAASQzOHm&index=2>. Acesso em: 7 fev. 2023.

CHIAVEGATTO FILHO, A. D. P. *et al.* Overachieving municipalities in public health: a machine-learning approach: a machine-learning approach. **Epidemiology (Cambridge, Mass.)**, v. 29, n. 6, p. 836-840, 2018. DOI: 10.1097/EDE.0000000000000919.

COLLINGRIDGE, D. **The social control of technology**. New York: St Martin Press, 1980.

CORBISIER, R. **Formação e problema da cultura nacional**. Rio de Janeiro: ISEB: MEC, 1959.

COSTA-NETO, M. E.; FITA, D.; AGUIAR, M. P. L. Curupira e Caipora: o papel dos seres elementais como guardiões da natureza. **Boletim do Museu Paraense Emílio Goeldi, Ciências Humanas**, Belém, v. 18, n. 1, 2023. DOI: <https://doi.org/10.1590/2178-2547-BGOELDI-2021-0095>.

COULDRY, N.; MEJÍAS, U. A. Data colonialism: rethinking big data's relation to the contemporary subject. **Television & New Media**, p. 1-14, 2018. DOI: <https://doi.org/10.1177/152747641879663>.

CUSTODIO, D.; BROEDEL, L. A pseudosuperioridade humana numa perspectiva bioética: do ceticismo antigo aos estudos animais contemporâneos. *In*: CONGRESSO MUNDIAL DE BIOÉTICA E DIREITO ANIMAL, 8., 2022. **Anais [...]**. [S. l.: s. n.], 2022.

DALE, R. GPT-3: What's it good for? **Natural Language Engineering**, v. 27, n. 1, p. 113-118, 2021. DOI: <https://doi.org/10.1017/S1351324920000601>.

DAMASCENO, G. P. M.; ANGELI NETO, G. O controle da informação, o colonialismo de dados e a desobediência epistêmica como forma de insurgência. **Ciências Sociais Aplicadas em Revista**, v. 23, n. 43, p. 153-169, 2022. DOI: <https://doi.org/10.48075/csar.v23i43.29078>.

DEBESH, J. *et al.* Ensuring trustworthy medical artificial intelligence through ethical and philosophical principles. **arXiv**, 2023. DOI: <https://doi.org/10.48550/arXiv.2304.11530>.

DELEUZE, G.; GUATTARI, F. **Mil Platôs: capitalismo e esquizofrenia**. Rio de Janeiro: Editora 34, 1997. v. 4.

DIJCK, J. Datafication, dataism and dataveillance: Big Data between scientific paradigm and ideology. **Surveillance & Society**, v. 12, n. 2, p. 197-208, 2014. DOI: <https://doi.org/10.24908/ss.v12i2.4776>.

DLUGATCH, R.; GEORGIEVA, A.; KERASIDOU, A. Trustworthy artificial intelligence and ethical design: public perceptions of trustworthiness of an AI-based decision-support tool in the context of intrapartum care. **BMC Medical Ethics**, v. 24, n. 1, 2023. DOI: 10.1186/s12910-023-00917-w.

ELENDU, C. *et al.* Ethical implications of AI and robotics in healthcare: a review. **Medicine**, v. 102, n. 50, e36671, 2023. DOI: 10.1097/MD.00000000000036671.

FEENBERG, A. **Transforming technology**. New York: Oxford University Press, 2002.

FEITOSA, S.; FLOR, W. A bioética de intervenção no contexto do pensamento latino-americano contemporâneo. **Revista Bioética**, v. 23, n. 2, p. 277-284, 2015. DOI: <https://doi.org/10.1590/1983-80422015232066>.

FERES, V. C. M.; OLIVEIRA, V. J.; GONÇALVES, D. D. Robin Hood às avessas: software, pirataria e direito autoral. **Revista Direito GV**, v. 13, n. 1, p. 69-94, abr. 2017. DOI: <http://dx.doi.org/10.1590/2317-6172201704>.

FERREIRA, G. *et al.* Metaphors we're colonised by? The case of data-driven educational technologies in Brazil, **Learning, Media and Technology**, 45, p. 46-60, 2020.

FLETCHER, R. R.; NAKESHIMANA, A.; OLUBEKO, O. Addressing fairness, bias, and appropriate use of artificial intelligence and machine learning in global health. **Frontiers in Artificial Intelligence**, 2021. DOI: 10.3389/frai.2020.561802.

FLOR, W.; GARRAFA, V. Por uma vida não colonizada: diálogo entre bioética de intervenção e colonialidade. **Saúde e Sociedade**, v. 20, p. 287-299, jun. 2011. DOI: <https://doi.org/10.1590/S0104-12902011000200003>.

FOLSCHEID, D.; WUNENBURGER, J-J. **Metodologia filosófica**. Tradução: Paulo Neves. 3. ed. São Paulo: Martins Fontes, 2006.

FOUCAULT, M. Gerir os ilegalismos. *In*: FOUCAULT, M. **Michel Foucault: entrevistas a Roger Pol-Droit**. São Paulo: Graal, 2006. p. 41-52.

FOUCAULT, M. **História da loucura na idade clássica**. Tradução: José Teixeira Coelho Netto. São Paulo: Perspectiva, 1978.

FOUCAULT, M. **Microfísica do poder**. São Paulo: Graal, 2012.

GIOVANELLA, L. (org.). **Políticas e Sistema de Saúde no Brasil**. 2. ed. rev. ampl. Rio de Janeiro: Editora Fiocruz, 2012.

GOLDMAN, S. Top AI researcher dismisses AI “extinction” fears, challenges “hero scientist” narrative. **VentureBeat**, 2023. Disponível em: <https://venturebeat.com/ai/top-ai-researcher-dismisses-ai-extinction-fears-challenges-hero-scientist-narrative/>. Acesso em: 8 jun. 2023.

GOLDSMITH, J.; ERDÉLYI, O. J. Regulating artificial intelligence: proposal for a global solution. *In*: AIES '18: AAAI/ACM CONFERENCE ON AI, ETHICS, AND SOCIETY, 2018. **Proceedings** [...]. [S. l.: s. n.], 2018.

GRANJEIRA, J.; ALMEIDA, R. Por dentro da máquina do Brasil Paralelo para dominar as buscas no Google. **Núcleo**, especiais, 2023. Disponível em: <https://nucleo.jor.br/especiais/2023-01-31-a-maquina-do-brasil-paralelo/>. Acesso em: 04 fev. 2024.

GROHMANN, R. Plataformas de propriedade de trabalhadores: cooperativas e coletivos de entregadores. **MATRIZES**, v. 16, n. 1, p. 209-233, 2022. DOI: <https://doi.org/10.11606/issn.1982-8160.v16i1p209-233>.

GROSGOUEL, R. Para descolonizar os estudos de economia política e os estudos pós-coloniais: transmodernidade, pensamento de fronteira e colonialidade global. **Revista Crítica de Ciências Sociais**, n. 80, 2008. DOI: <https://doi.org/10.4000/rccs.697>.

GU, J.; OELKE, D. Understanding bias in machine learning. **arXiv**, 2019. DOI: <https://doi.org/10.48550/arXiv.1909.01866>.

HOBBS, T. **Do cidadão**. São Paulo: Martin Claret, 2006.

IRAMINA, A. RGPD v. LGPD: adoção estratégica da abordagem responsiva na elaboração da Lei Geral de Proteção de Dados do Brasil e do regulamento geral de proteção de dados da União Europeia. **Revista de Direito, Estado e Telecomunicações**, v. 12, n. 2, p. 91-117, 2020. DOI: <https://doi.org/10.26512/lstr.v12i2.34692>.

KABIR, O. Can open source defeat Google and OpenAI in the AI arms race? **Ctech**, 2023. Disponível em: <https://www.calcalistech.com/ctechnews/article/s19lafwn2>. Acesso

KUHN, T. **A estrutura das revoluções científicas**. 5. ed. São Paulo: Perspectiva, 1998.

LANIER, J. **Who owns the future**. New York: Simon & Schuster, 2013.

LANZAGORTA-ORTEGA, D.; CARRILLO-PÉREZ, D. L.; CARRILLO-ESPER, R. Inteligência artificial en medicina: presente y futuro. **Gaceta médica de México**, v. 158, n. 91, p. 55-59, 2022. DOI: <https://doi.org/10.24875/gmm.m22000688>.

LOBATO, L. V. C.; GIOVANELLA, L. Sistemas de saúde: origens, componentes e dinâmica. In: GIOVANELLA, L. *et al.* (org.). **Políticas e sistemas de saúde no Brasil**. Rio de Janeiro: Editora Fiocruz, 2012. p. 89-120.

LOBO, L. C. Inteligência artificial e medicina. **Revista Brasileira de Educação Médica**, v. 41, n. 2, p. 185-193, 2017. DOI: <https://doi.org/10.1590/1981-52712015v41n2esp>.

LOPES, I. L. Uso das linguagens controlada e natural em bases de dados: revisão da literatura. **Revista Ciência da Informação**, Brasília, v. 31, n. 1, p. 41-52, jan./abr. 2002. DOI: <https://doi.org/10.18225/ci.inf.v31i1.976>.

LUND, B.; WANG, T. Chatting about ChatGPT: how may AI and GPT impact academia and libraries? ChatGPT and its impact on academia. **Library Hi Tech News**, v. 40, 2023. DOI: <http://dx.doi.org/10.2139/ssrn.4333415>.

- MANASSI, L. P. **Registro eletrônico de saúde e produção de informação da atenção à saúde no SUS**. 2014. 179 f. Dissertação (Mestrado Profissional em Saúde Pública) – Escola Nacional de Saúde Pública Sergio Arouca, Fundação Oswaldo Cruz, Rio de Janeiro, 2014.
- MARIUTTI, B. E. A “virada cibernética”: capitalismo, informação e direitos de propriedade. **Texto para Discussão**, Instituto de Economia – Unicamp, Campinas, n. 390, ago. 2020.
- MARQUES, C. J. Searle e os desafios da inteligência artificial (IA) forte. **Revista de Filosofia Reflexões**, n. 18, 2021.
- MEHRABI, N. *et al.* Survey on bias and fairness in machine learning. **arXiv**, 2019. DOI: <https://doi.org/10.48550/arXiv.1908.09635>.
- MENDES FAUSTINO, D.; LIPPOLD, W. **Colonialismo digital**: por uma crítica Hacker-Fanoniana. São Paulo: Boitempo Editorial, 2023.
- MESIANO, B.; RIBBEIRO, L. Brasil registra quase 300 mil falhas na assistência à saúde em 2022. **CNN**, 2023. Disponível em: <https://encurtador.com.br/jstY9>. Acesso em: 21 fev. 2023.
- METZ, C. The Godfather of A.I.” Leaves Google and warns of danger ahead. **The New York Times**, 2023.
- MIGNOLO, D. W. Colonialidade: o lado mais escuro da modernidade. **Revista Brasileira de Ciências Sociais**, v. 32, n. 94, e329402, 2017. DOI: <https://doi.org/10.17666/329402/2017>.
- MIGNOLO, W. D. A colonialidade de cabo a rabo: o hemisfério ocidental no horizonte conceitual da modernidade. *In*: LANDER, E. (org.). **A colonialidade do saber**: eurocentrismo e ciências sociais. Buenos Aires: Conselho Latinoamericano de Ciências Sociais, 2005.
- MIGNOLO, W. **La idea de América Latina**: la herida colonial y la opción decolonial. Barcelona: Gedisa, 2007.
- MOHAMED, S.; PNG, M. T.; ISAAC, W. Decolonial AI: decolonial theory as sociotechnical foresight in artificial intelligence. **Philosophy & Technology**, v. 33, 659-684, 2020. DOI: <https://doi.org/10.1007/s13347-020-00405-8>.
- MOLE, B. United health uses AI model with 90% error rate to deny care, lawsuit alleges. **Arstechnica**, 2023. Disponível em: <https://encurtador.com.br/fuwE3>. Acesso: 28 fev. 2024.
- MONTTI, R. Leaked Google Memo admits defeat by open source AI. **Search Engine Journal**, 2023. Disponível em: <https://www.searchenginejournal.com/leaked-google-memo-admits-defeat-by-open-source-ai/486290/>. Acesso em: 9 maio 2023.
- MORA, J. F. **Dicionário de Filosofia**: tomo I (A-D). São Paulo: Edições Loyola, 2000.
- MOROZOV, E. **Big Tech**: a ascensão dos dados e a morte da política. Tradução: Claudio Marcondes. São Paulo: Ubu Editora, 2018.
- MOROZOV, E. **Solucionismo, nova aposta das elites globais**. Tradução: Simone Paz. Ribeirão Preto: Outras Palavras, 2020.

- MOYSES, D. O libera geral dos dados pessoais, sob Bolsonaro. **Outras Mídias**, 2021. Disponível em: <https://outraspalavras.net/outrasmidias/o-libera-geral-dos-dados-pessoais-sob-bolsonaro/>. Acesso em: 9 maio 2023.
- NAS, E. Inteligência artificial: uma discussão bioética dos sistemas de informação. In: CASTRO, C. J.; NIEMEYER-GUIMARÃES, M.; SIQUEIRA-BATISTA, R. **Caminhos da bioética**. Teresópolis: Ed. UNIFESO, 2020. (Coleção FESO, v. 3).
- NASCIMENTO, C. C. E.; SILVA, E.; SIQUEIRA-BATISTA, R. The “use” of sex robots: a bioethical issue. **Asian Bioethics Review**, v. 10, n. 3, p. 231-240, 2018. DOI: 10.1007/s41649-018-0061-0.
- NATURE. Stop talking about tomorrow’s Ai doomsday when AI poses risks today. **Nature News**, 2023. Disponível em: <https://www.nature.com/articles/d41586-023-02094-7>. Acesso em: 23 out. 2023.
- NEVES, L. J. O. A necessária desconstrução da colonialidade nas Amazônias e nas Américas. **Textos e Debates**, v. 2, n. 27, 2015. DOI: <https://doi.org/10.18227/2317-1448ted.v2i27.3208>.
- NOBLE, S. U. **Algorithms of oppression**: how search engines reinforce racism. New York: New York University Press, 2018.
- NOVAES, H.; DAGNINO, R. O fetiche da tecnologia. **Org & Demo**, v. 5, n. 2, 2004. DOI: <https://doi.org/10.36311/1519-0110.2004.v5n2.411>.
- O’NEIL, C. **Algoritmos de destruição em massa**: como o big data aumenta a desigualdade e ameaça a democracia. Tradução: Rafael Abraham. Santo André: Editora Rua do Sabão, 2020.
- OKENGWU, U A. Practical implications of different theoretical approaches to AI ethics. In: CORRIGAN, C. C. *et al.* (ed.). **AI ethics in higher education**: insights from Africa and beyond. [S. l.]: Springer Briefs in Ethics, 2022.
- PAIM, J. S. *et al.* The Brazilian health system: history, advances, and challenges. **The Lancet**, v. 377, n. 9779, p. 1778-1797, 2011. DOI:[https://doi.org/10.1016/S0140-6736\(11\)60054-8](https://doi.org/10.1016/S0140-6736(11)60054-8).
- PAIVA, F. Governo cria a Estratégia Brasileira de Inteligência Artificial, EBIA. **Terra**, 09 abr. 2021. Disponível em: <https://encurtador.com.br/invER>. Acesso em: 09 abr. 2023.
- PAN, I. *et al.* Machine learning for social services: A study of prenatal case management in Illinois. **American Journal of Public Health**, v. 107, n. 6, p. 938-944, 2017. DOI: 10.2105/AJPH.2017.303711.
- PATEL, V. L. *et al.* The coming of age of artificial intelligence in medicine. **Artificial Intelligence in Medicine**, v. 46, n. 1, p. 5-17, 2009. DOI: 10.1016/j.artmed.2008.07.017.
- PERES, E. I. *et al.* **Preconceito em algoritmos de aprendizagem de máquina e suas bases de treinamento**: uma revisão sistemática de literatura. 2021. 19 f. Trabalho de Conclusão de Curso (Graduação em Computação e Informática) – Universidade Presbiteriana Mackenzie, São Paulo, 2021.

QUIJANO, A. **Colonialidade do poder, eurocentrismo e América Latina**. Buenos Aires: CLACSO, 2005. (Colección Sur).

QUIJANO, A. Colonialidade do poder e classificação social. *In*: SANTOS, B. S.; MENESES, M. P. (org.). **Epistemologias do sul**. São Paulo: Cortez, 2010.

RAMOS, F. *et al.* Desafios atuais da bioética brasileira. **Revista Bioética**, Brasília, v. 27, p.446-454, 2019. DOI: <https://doi.org/10.1590/1983-80422019273328>.

RIBEIRO, J. L. V. **Personalização da inteligência artificial: novo paradigma jurídico**. 2020. 298 f. Tese (Doutorado em Direito) – Centro Universitário de Brasília, Brasília, 2020.

RIBEIRO, P. V. Google é proibido pelo PROCON de limitar serviço oferecido como ‘ilimitado’ a universidade. **Intercept Brasil**, 2022.

RUSSELL, S.; NORVIG, P. **Inteligência artificial**. Tradução: Regina Célia Simille. Rio de Janeiro: Elsevier, 2013.

SANTOS, H. G. D. *et al.* Machine learning to predict 30-day quality-adjusted survival in critically ill patients with cancer. **Journal of Critical Care**, v. 55, p. 73-78, 2020. DOI: 10.1016/j.jcrc.2019.10.015.

SANTOS, L. G. **Revolução tecnológica, internet e socialismo**. São Paulo: Perseu Abramo, 2013.

SCHEFFER, M.; SOUZA, P. M. S. A entrada do capital estrangeiro no sistema de saúde no Brasil. **Cadernos de Saúde Pública**, v. 38, supl. 2, 2022. DOI: <https://doi.org/10.1590/0102-311X00239421>.

SCHNEBLE, C. O. *et al.* All our data will be health data one day: the need for universal data protection and comprehensive consent. **Journal of Medical Internet Research**, v. 22, n. 5, e16879, 2020. DOI: 10.2196/16879.

SCHRAMM, F. R. Bioética da proteção: ferramenta válida para enfrentar problemas morais na era da globalização. **Revista Bioética**, Brasília, v. 16, n. 1, 2008.

SERRANO, P. J.; RAMPAZZO, L. A bioética latino-americana: historicidade, atualidade e operacionalização do conceito. **Revista de Direito Brasileira**, São Paulo, v. 17, n.7, p. 261-274, maio/ago. 2017.

SHOKER, A. Digital sovereignty strategies for every nation. **Applied Cybersecurity & Internet Governance**, v. 1, n. 1, p. 1-17, 2022. DOI: <https://doi.org/10.5604/01.3001.0016.0943>.

SILVA, C. G. Tecnologia, educação e tecnocentrismo: as contribuições de Álvaro Vieira Pinto. **Revista Brasileira de Estudos Pedagógicos**, v. 94, n. 238, set. 2013.

SILVA, T. Mapeamento de danos e discriminação algorítmica. **Desvelar**, 2023. Disponível em: <https://desvelar.org/casos-de-discriminacao-algoritmica/>. Acesso em: 06 jan. 2024.

SILVA, T. Racismo algorítmico em plataformas digitais: microagressões e discriminação em código. *In: SIMPÓSIO INTERNACIONAL LATINO-AMERICANA DE ESTUDOS SOBRE VIGILÂNCIA, TECNOLOGIA E SOCIEDADE (LAVITS), 6., 2019. Anais [...]. [S. l.: s. n.], 2019.*

SILVA, T. **Racismo algorítmico**: inteligência artificial e discriminação nas redes digitais. São Paulo: Edições Sesc, 2022.

SILVA, T. Regular a inteligência artificial no Brasil pode mitigar o racismo algorítmico. **Folha de São Paulo**, 2024. Disponível em: <https://www1.folha.uol.com.br/blogs/politicas-e-justica/2023/05/regular-a-inteligencia-artificial-no-brasil-pode-mitigar-o-racismo-algoritmico.shtml>. Acesso em: 06 jan. 2024.

SILVA, T. Visão computacional e racismo algorítmico: branquitude e opacidade no aprendizado de máquina. **Revista ABPN**, v. 12, p. 428-448, 2020.

SILVEIRA, S. A. Não haverá soberania digital sem o Estado. **Outras Mídias**, 2023. Disponível em: <https://outraspalavras.net/tecnologiaemdisputa/nao-havera-soberania-digital-sem-o-estado/>. Acesso em: 06 jan. 2024.

SILVEIRA, S. A.; SOUZA, J.; CASSINO, J. F. **Colonialismo de dados e modulação algorítmica**: tecnopolítica, sujeição e guerra neoliberal. São Paulo: Autonomia Literária, 2021.

SIQUEIRA-BATISTA, R. **Regulação emocional na perspectiva da neurociência cognitiva**. Palestra 'Sobre Bioética e Neurociências' na Mesa Redonda 'Ética em Pesquisa'. I Simpósio Interdisciplinar de Cognição e Emoção, Núcleo de Estudos Interdisciplinares em Cognição e Emoção, Instituto Federal do Rio de Janeiro, 2022.

SIQUEIRA-BATISTA, R. (Bio)ética para todos os seres: proêmio. *In: CASTRO, J. C.; NIEMEYER-GUIMARÃES, M.; SIQUEIRA-BATISTA, R. Caminhos da bioética*. Teresópolis: Ed. UNIFESO, 2020. v. 3, p. 257-276.

SIQUEIRA-BATISTA, R. Quantum bioethics: ethics for all beings. **Ciência & Saúde Coletiva**, v. 13, p. 1013-1015, 2008. DOI: <https://doi.org/10.1590/S1413-81232008000300023>.

SIQUEIRA-BATISTA, R.; GOMES, A. P. (Bio)ética para todos os seres e Estratégia Saúde da Família: composições entre cuidado de si e compaixão laica. *In: CASTRO, J. C.; NIEMEYER-GUIMARÃES, M.; SIQUEIRA-BATISTA, R. Caminhos da bioética*. Teresópolis: Ed. UNIFESO, 2021. v. 4, p. 64-95.

SIQUEIRA-BATISTA, R.; POSSAMAI, R. V. Bioética da proteção de Schramm e Kottow: princípios, alcances e conversações. **Revista Bioética**, Brasília, v. 30, n. 1, p. 10-18, 2022. DOI: <https://doi.org/10.1590/1983-80422022301501PT>.

SIQUEIRA-BATISTA, R.; SILVA, E. Notas sobre os fundamentos matemáticos da inteligência artificial. **Revista de Ciência, Tecnologia e Inovação**, v. 4, n. 6, 2019.

SMINK, V. Os milhares de trabalhadores em países pobres que abastecem programas de inteligência artificial como o ChatGPT. **BBC News Brasil**, 2023. Disponível em: <https://encurtador.com.br/quyY5>. Acesso: 26 fev. 2024.

SOUZA, F. Araraquara (SP) adota 'Uber' próprio; motorista fica com 95% da tarifa. **UOL**, Economia, 2022.

STONE, Z. Everything you need to know about Sophia, the world's first robot citizen. **Forbes**, 2017. Disponível em: <https://encurtador.com.br/ikrFY>. Acesso em: 3 abr. 2023.

TEALDI, J. C. (dir.). **Diccionario latinoamericano de bioética**. Bogotá: UNESCO, 2008.

TEXEIRA, F. J. PPGLM/ UFRJ. **Workshop de Filosofia e Inteligência Artificial, João de Fernandes Teixeira (UFSCar)**. [S. l.: s. n.], 2023. 1 vídeo (86 min). Publicado pelo canal PPGLM / UFRJ. Disponível em: <https://www.youtube.com/watch?v=ANF1rzzokik>. Acesso em: 31 maio. 2023.

TOPOL, E. **Creative destruction of medicine**: how the digital revolution will create better health care. New York: Basic Books, 2013.

TOPOL, E. **Deep medicine**: how artificial intelligence can make healthcare human again. [S. l.]: Hachette Book Group, 2019.

TZANOU, M. **The fundamental right to data protection normative value in the context of counter-terrorism surveillance**. Oxford: Hart Publishing, 2017.

UNGER, R. Política de informação: poder, comunicação e controle - uma revisão narrativa. **Logeion: Filosofia da Informação**, v. 10, n. 1, p. 131-144, 2023. DOI: <https://doi.org/10.21728/logcion.2023v10n1.p131-144>.

VASCONCELLOS, G. Trinta anos depois: ideologia Curupira. **Locus**, v. 16, n. 1, 2010.

VIEIRA PINTO, Á. **O conceito de tecnologia**. Rio de Janeiro: Contraponto, 2005. 2 v.

VIEIRA, C. Inteligência Artificial: a caixa preta que prejudica as minorias. **Medium**, 08 abr. 2019. Disponível em: <https://medium.com/blogcarlavieira/intelig%C3%A2ncia-artificial-a-caixa-preta-que-prejudica-as-minorias-7bf66fa98228>. Acesso em: 7 abr. 2023.

WHITE, J.; CESAREO, S. The global artificial intelligence index. **Tortoise**, 2023. Disponível em: <https://www.tortoisemedia.com/intelligence/global-ai/>. Acesso em: 20 fev. 2024.

WHITE, T. A luta contra o desmatamento ilegal com TensorFlow. **Blog do Google Brasil**, 2018. Disponível em: <https://brasil.googleblog.com/2018/03/a-luta-contra-o-desmatamento-ilegal-com.html>. Acesso em: 5 fev. 2023.

ZUBOFF, S. **The age of surveillance capitalism**: the fight for a human future at the new frontier of power. New York: Public Affairs, 2019.